Multitasking, Incentives, and Police Officer Behavior*

James Reeves
University of Michigan

November 1, 2024
[Please click here for latest version]

Abstract

I study the consequences of incomplete contracts in the high-stakes, multitasking setting of policing. In my context, highway patrol troopers face salient traffic enforcement targets, but must balance their effort on enforcement production with completing other non-enforcement responsibilities. While the enforcement target induces trooper effort, it simultaneously distorts trooper behavior and generates a range of socially suboptimal outcomes, including lower quality enforcement, disparities unwarranted on the basis of collision risk, and delayed completion of non-enforcement responsibilities. Given significant trooper-specific heterogeneity, I develop an approach to optimally assign troopers to locations that reduces the negative externalities produced by the existing incentive scheme. In contrast, alternative personnel policies which ignore this heterogeneity improve only a subset of outcomes.

*I thank Michael Mueller-Smith, Sarah Miller, Benjamin Scuderi, and Mel Stephens for guidance and feedback throughout this project. I also benefitted from helpful conversations with Amanda Agan, Charlie Brown, Itzik Fadlon, Nicole Gandre, Felipe Goncalves, Florian Gunsilius, Sara Heller, Elisa Jácome, Andrew Joung, Logan Lee, Elizabeth Luh, Chris Malloy, Hani Mansour, Tatiana Mocanu, Emily Owens, Petra Persson, JJ Prescott, Benjamin Pyle, Nikhil Rao, Kevin Schnepel, Damián Vergara, Alice Wu, and Basit Zafar, along with seminar participants at the University of Michigan, the 2023 Midwestern Economic Association Meeting, the XIV Transatlantic Workshop on the Economics of Crime, the 2024 Western Economic Association International Meeting, and the 2024 UW-UM-MSU-UWO Labour Day Conference. In addition, I thank Sharad Goel for sharing some of the underlying data, as well as numerous individuals at the Patrol for assistance with additional data and institutional details.
Reeves: jmreeves@umich.edu.

1 Introduction

Employers throughout the economy face a core challenge of optimally managing their workforce under incomplete contracts. A long tradition in labor economics has analyzed this problem through the lens of agency theory, casting employers as principals and workers as their agents. The key tension between these contracting parties stems from a potential misalignment of incentives - principals wish to maximize output, but production is delegated to agents who must exert imperfectly observed costly effort in order to produce that output. Without sufficiently strong contractual tools to align these incentives, agents will choose to shirk (Holmström 1979; Bolton and Dewatripont 2005) or sacrifice other production margins such as quality (Baron 1972). Employers may overcome this tension through a variety of methods, including performance pay (Lazear 2000), bonuses (Kuhn and Yu 2021), quotas (Asch 1990; Chung, Narayandas, and Chang 2021), and other pecuniary or non-pecuniary incentives.[1] However, while these tools may be useful for aligning the incentives of the employer and employee in settings with rote occupations, the vast majority of workers are instead multi-tasking agents, balancing effort across multiple competing responsibilities (Applebaum, Marchionni, and Fernandez 2008; Deming and Kahn 2018), some of which may be measured more precisely than others. In such settings, correctly incentivizing agents while not encouraging adverse specialization in a single task becomes practically challenging (Holmström and Milgrom 1991; MacDonald and Marx 2001).[2]

When agents are tasked with providing public services, contracts that inadvertently encourage specialization can have substantial negative consequences for social welfare. For example, in the context of ambulance services, when governments delegate responsibility to private firms who emphasize observable and contracted measures of performance such as response times, non-contracted outcomes such as mortality risk can suffer (Knutsson and Tyrefors 2022).[3] Similarly, providing teachers strong incentives to pass marginal students may improve graduation rates, but simultaneously harm long-term human capital development (e.g., Jacob and Levitt 2003; Dee et al. 2019). Such misalignment for public sector agents can also affect the scope and distributional burden of the tax system, underscoring patterns of racial inequality and financial instability (Luh, Pyle, and Reeves 2024). In these cases, the design of compensation and incentive schemes can have consequences that extend far beyond the immediate effects on firm profits.[4]

In this paper, I study the efficacy and efficiency consequences of a salient and narrow incentive

[1]According to a recent survey, nearly forty percent of the civilian labor force faces some type of broadly defined performance pay (Gittleman and Pierce 2013). See Gibbons (1998) for an overview of incentives in organizations, Prendergast (2007) for a discussion of incentives in public sector settings, and Holmström (2017) for a review of the use of performance pay and other incentive schemes.

[2]There is an open question on why observed arrangements reflect far simpler contracts than would be predicted by theory (Chiappori and Salanié 2003). One possibility is that it's difficult for firms to design highly nuanced contracts because they are unable to appropriately assess the implied value of the contract over all possible payoffs (Hart and Holmström 1987).

[3]A rich literature has also studied the complexities and moral hazard associated with government contracting and the optimal design of such contracts (e.g., Cox et al. 1996; Lewis and Bajari 2014)

[4]Prior work has also examined the consequences of incentives on physician (Clemens and Gottlieb 2014; Shearer, Somé, and Fortin 2018; Alexander 2020) and tax collector (Khan, Khwaja, and Olken 2014) behavior.

scheme in the high-stakes setting of policing. The highway patrol troopers I examine face salient traffic enforcement targets, but must balance their performance on this measure with completing other non-enforcement actions, many of which arrive stochastically. For example, the agency may expect troopers to produce traffic enforcement, respond to collisions, and complete timely reports. However, the most well-tracked performance metric covers enforcement actions (e.g., stops, tickets), creating an environment where one aspect of troopers' responsibilities is most salient. To study these tradeoffs, I assemble a unique data infrastructure covering nearly nine years of activity in the Washington State Patrol. This data infrastructure includes administrative measures of both enforcement and non-enforcement activity, trooper schedules and assignments, along with detailed institutional documents describing both the existence and setting of enforcement targets and their role in the broader constellation of trooper performance evaluations. Together, these resources allow me to both quantify how the target encourages trooper performance as well as study and measure the degree of trooper substitution between salient and non-salient tasks.

Leveraging these institutional features and the competing nature of tasks, I develop a parsimonious multi-task principal-agent model in the spirit of MacDonald and Marx (2001) to theoretically ground the tension between the division of effort the trooper wishes to exert and the Patrol desires. A key feature of the model is that tasks are substitutes to the agent and complements to the principal. Therefore, any desirable compensation or evaluation scheme must induce the trooper to view the tasks as complements, otherwise troopers will choose to specialize in one task or shirk altogether. I use the model to derive a series of sharp empirical predictions about the nature of substitution between tasks, as well as changes to the distribution of enforcement across motorists.

Taking these predictions to the data, I begin by quantifying how responsive troopers are to changes in enforcement targets. Using trooper transfers across locations with higher or lower enforcement targets, I document near-immediate and sharp changes to enforcement volume, with no evidence of differential selection into locations on the basis of pre-transfer performance. Increasing the target raises subsequent enforcement volume by more than fifty percent of the change, a magnitude that is further corroborated in the long-run cross-sectional correlation based on first-differences.

However, these responses may simply reflect changes in the supply of law-violating motorists across areas. I therefore examine how troopers respond to missing their goals when they remain in the same location for consecutive periods. While trooper performance evaluations are a broad collection of objective and subjective measures across multiple responsibilities, internal interviews with troopers highlight the substantial saliency of the enforcement targets and how trooper daily activity is strongly influenced by its presence (Branson et al. 2016). Consequently, poor performance during one evaluation period should lead to compensating changes in trooper behavior in the following period since repeated poor performances are deleterious to career development and positive personnel reviews.

Using a regression discontinuity design comparing troopers who just make versus just miss their previous average per-shift target, I find strong evidence that troopers are responsive on a period-over-period basis. Just missing the target increases the probability of making the target in the following period by 3.2 percentage points, with larger effects for troopers who are on the margin of consecutive misses across multiple evaluation periods. This result holds both within- and across-troopers. While troopers increase per-shift enforcement activity, much of this increased enforcement arises from (generally minor) non-moving violations which include licensing, equipment, and paperwork violations. Taken together with the event-study evidence above, I conclude that enforcement targets are first-order effective at inducing additional incentivized effort from troopers in the workforce.

While this incentive scheme can induce greater enforcement volume, it may also come at the cost of an emphasis on easy rather than optimal enforcement (e.g., minor violations or inequitable effects across the motorist population). Indeed, I find evidence that marginal enforcement actions experience greater court dismissal rates. I also find that marginally missing troopers are weakly differentially more stringent towards minority motorists on a per-stop basis, a finding that is concentrated among non-moving violations where troopers have substantial discretion. These results are consistent with prior evidence highlighting differential "hassle costs" (Makowsky and Stratmann 2009) across race and income, as lower-income motorists are less likely to contest citations in court (Feigenberg and Miller 2023), a finding I confirm in my setting.[5] Given historical patterns of income inequality along racial and ethnic lines, these enforcement responses have the potential to further undermine the economic stability of low-income households (Mello 2023; Luh, Pyle, and Reeves 2024).

Is such a disparity warranted by differences in driver risk by race? I test for differences in motorist collision risk by race across the age distribution and find little conclusive evidence that collision risk is an appropriate rationale for such a disparate finding, although I am unable to fully rule out alternative explanations based on differential driving behavior.[6] Instead, I find evidence consistent with mechanisms which interact with trooper incentives and generate these socially suboptimal outcomes. For example, if non-moving violations are a tag (Akerlof 1978) for potential criminal behavior or future driving misconduct, then vehicle searches for minority motorists should be more productive with respect to contraband yield. Despite troopers being weakly more likely to search minority motorists, there is no evidence these searches are differentially more productive on this margin. Therefore, even if there are population differences in car condition, using this information as a basis for stops and searches is unwarranted on efficiency grounds. Together with the evidence on hassle costs, this result also offers an alternative systemwide explanation for previously

[5] Poverty and racial/ethnic demographics are correlated in Washington State, as they are in the United States more broadly. To the extent troopers use signals that are unobserved to the econometrician (e.g., car condition) to infer the likelihood of a contested hearing that are also correlated with race/ethnicity, then the area poverty rate should capture both race and non-race channels.

[6] Lovrich et al. (2003) calculate a "risk" score based on the seriousness of the charged offense and find that minority drivers have higher average risk scores compared to white drivers. However, a strict interpretation of this score is challenging since charged offense is also a function of trooper discretion which itself may exhibit disparate impact (Goncalves and Mello 2021). In a similar exercise, I calculate a parametric collision risk score and find that the stringency gap persists even among motorists in the same part of the predicted risk distribution, suggesting that risk alone cannot explain this gap.

documented differential stringency (Lovrich et al. 2003) that does not hinge on an interpretation of racial animus.[7] Instead, my results imply that such equilibrium outcome disparities may be the consequences of misaligned incentives.

Enforcement targets may not just impact the quality of enforcement actions, but also how other non-enforcement responsibilities may be down-weighted or shirked. To investigate this potential adverse response, I look at how troopers who marginally missed their target in the previous period respond to non-salient responsibilities (e.g., responding to collisions and paperwork completion) compared to troopers who made the target. Marginal troopers who missed the target return to enforcement actions faster than those who made the target, indicating a reallocation of effort towards enforcement production. Using a subsample of collision reports linked to incidents, I also show this effort reallocation comes at the expense of the timeliness of completing non-enforcement responsibilities, as marginally missing troopers are more likely to file collision reports more than one week after the incident and are weakly more likely to miss internal deadlines for filing collision reports which involve felony charges. I estimate the net welfare impacts of this former delay, comparing the change in net present value of the victim's insurance compensation against the social benefits of increased enforcement using a simple back-of-the-envelope calculation. Over a range of plausible parameter values, the delayed reporting is welfare-negative to welfare-neutral, suggesting that, on balance, the benefits from enforcement emphasis do not fully offset the private costs of settlement delay to the victim unless the additional enforcement deters highly severe, tail-risk events.

Can these negative externalities be mitigated? Taking the policy environment as given, alternative personnel policies may choose to target either the existing location-specific incentives or accommodate variation across troopers through reallocation. I test the importance of these channels through a variance decomposition exercise where I decompose the variation in multitasking collision responses into variation from trooper and location fixed effects. Calculating unbiased variance estimates of the focal components (Kline, Saggio, and Sølvsten 2020), I find trooper effects explain between 1.7 and 4.1 times the variation in multitasking responses as location effects, indicating that addressing trooper heterogeneity is a key factor in minimizing potential social externalities induced by enforcement targets in this setting. Put differently, while there may be average differences across locations which contribute to the observed responses, these effects are second-order relative to the variation across troopers.

Assuming that eliminating the targets or a "first-best" policy of individually tailored targets are infeasible prescriptions, a next-best approximation is to reassign troopers across locations to improve the match of trooper performance with location-specific demands across tasks.[8] I examine the scope for potential improvements by first assessing the empirical match between troopers and

[7] I implement the Antonovics and Knight (2009) test for racial bias in the context of searches and find no evidence in support of this mechanism.

[8] In state legislative hearings, Patrol leadership have argued against legislation that would ban the use of quotas in performance evaluations. Moreover, repealing quotas exacerbates the underlying imperfect monitoring problem and may heighten the presence of negative externalities (Edwards and Rushin 2023).

locations, classifying them into discrete types using k-means clustering. I find there is substantial potential for improvement, with many high-disparity generating troopers assigned to high minority share locations and low-enforcement volume troopers assigned to high target locations. I leverage tools from the optimal transport literature to compute the optimal allocation of troopers to locations while incorporating rich cross-trooper and cross-location heterogeneity. I then predict counterfactual outcomes using a simple data-driven approach which has a similar intuition to an Oaxaca-Blinder decomposition (Oaxaca 1973; Blinder 1973) and maintains a connection to the underlying principal-agent model.

I benchmark this "second-best" policy against several alternatives, including uniformly raising or lowering the target across locations. Relative to the observed equilibrium, the optimal unconstrained reassignment policy reduces the distortionary degree of enforcement emphasis along with racial disparities in response enforcement. Moreover, the policy also increases ex-ante enforcement in high traffic areas, which subsequently reduces collision volume, a key goal of the Patrol. Incorporating trooper location preferences through geographically-constrained matching still generates improvements. In contrast, simpler personnel policies which alter the enforcement target holding assignments fixed improve some, but not all of these outcomes. Thus, one solution to minimizing the welfare losses from incomplete contracts in multitask settings is to optimally match the heterogeneous workforce with heterogeneous job demands, which is in spirit, a feasible accommodation of the theoretical task grouping result from Holmström and Milgrom (1991).

The central contribution of this paper is to empirically quantify the consequences of misaligned incentives in the high-stakes setting of the criminal justice system. The agency issues I study in this paper are ubiquitous throughout the economy and have previously been studied in the context of microfinance (Kim, Sudhir, and Uetake 2022), consumer lending (Dobbie et al. 2021), and emergency departments (KC 2014), among many others.[9] I study trooper responses to misaligned incentives in the consequential setting of policing and public safety, where police officers are responsive to differences in pay (Mas 2006), overtime opportunities (Chalfin and Goncalves 2023), and external oversight (Campbell 2023). A novel general feature of my setting is that I am able to precisely measure performance on both the contracted (enforcement) and non-contracted (non-enforcement) actions, rather than being limited to just the former. In doing so, I examine the degree to which salient incentives distort the allocation of effort across these two actions, as well as the downstream and distributional consequences of such distortions. Thus, my results highlight both the efficacy and efficiency implications of misaligned incentive schemes in this setting and their contribution to the production of negative externalities.

This paper also builds on a rich literature documenting sources of racial disparities and suboptimal outcomes in the criminal justice system. Racial disparities are pervasive and have previously been documented in traffic searches (Knowles, Persico, and Todd 2001; Antonovics and Knight 2009; Feigenberg and Miller 2022), ticketing (Goncalves and Mello 2021), pretrial release (Arnold,

[9] More generally, this paper builds on the literature studying the labor supply responses of workers with fixed shift lengths (e.g., Chan 2018).

Dobbie, and Yang 2018; Arnold, Dobbie, and Hull 2022), prosecution (Harrington and Shaffer 2022), and incarceration decisions (Rehavi and Starr 2014). While much of the previous literature documents and develops tests for racial animus, I build on a nascent literature studying the equity consequences of the decision-making constraints judicial agents face (Jordan 2024; Rao and Reeves 2024). Rather than strictly animus, I provide evidence that the incentives and constraints agents face can give rise to racial and ethnic disparities due to the distortions they impose on decision makers, especially in the presence of multitasking responsibilities.[10] I build on this previous work by more tightly aligning the competing demands of judicial agents to changes in both contracted and non-contracted behavior, as well as further examining the efficiency implications of such incentive schemes, rather than identifying a purely relative equity effect. Moreover, I illustrate how adverse reactions to incentives can not only harm equity but other downstream race-neutral outcomes as well.

Finally, this paper highlights how better aligning the abilities of the workforce to location- and task-specific demands can reduce the welfare consequences of incomplete contracts in multitask settings. I show how this matching objective can be empirically cast in an optimal transport framework, maintaining a fundamental connection to the social planner's welfare maximization problem while accommodating variation in both the trooper and location characteristic distributions. Combining the underlying principal-agent model with the optimal assignment matrix, I illustrate how to construct counterfactual trooper responses under alternative assignments using the estimated empirical moments. In doing so, I provide an alternative approach for empirical economists to estimate partial equilibrium outcomes under reassignment counterfactuals, a common exercise across many strands of applied economics (e.g., Fenizia 2022).

The remainder of the paper is structured as follows. Section 2 provides an overview of the Washington State Patrol, the data infrastructure, and presents a parsimonious principal-agent model of multitasking, from which I derive a core proposition on the effort substitution patterns for troopers along with empirical tests for the model-predicted behavior. Section 3 estimates trooper responses to incentives, Section 4 examines the degree of task substitution across contracted and non-contracted actions, and Section 5 quantifies trooper-specific estimates of task substitution and enforcement patterns. Section 6 explores the feasibility of alternative policies using a novel matching approach as a benchmark and Section 7 concludes.

2 Institutional Setting and Data

In this section, I discuss the history of the Washington State Patrol and draw contrasts to the Patrol's responsibilities, relative to typical municipal police departments. To develop intuition for the trooper and Patrol's desired effort allocation, I develop a simple multitask principal-agent model

[10] This phenomenon has previously been identified in the context of policing, where revenue shortfalls and multitasking responsibilities generate disparities in the spatial distribution of parking fine enforcement in Chicago (Luh, Pyle, and Reeves 2024) and hypothesized as a mechanism for racial disparities in emergency medicine (Cone et al. 2003).

and derive sharp predictions which help guide and formalize the subsequent empirical analysis. Finally, I discuss the unique data infrastructure I develop to examine incentives in policing.

2.1 The Washington State Patrol

Founded in 1921, the Washington State Patrol (WSP or Patrol) is Washington State's largest law enforcement agency (Washington State Patrol 2019).[11] As a statewide highway patrol agency, the Patrol's primary function is to conduct traffic enforcement on the state's highways and interstates while having concurrent jurisdiction on all other roads in the state. In addition, troopers working in the Patrol also conduct non-enforcement actions, such as responding to calls for service, investigating collisions, and assisting stranded motorists.

The Patrol currently employs around 2,000 individuals, about half of whom are commissioned troopers. Geographically, the Patrol is organizes its highway patrol troopers into eight mutually exclusive geographic districts (see Appendix Figure A1), with each district comprising of several autonomous patrol areas which are further subdivided into detachments. Throughout the rest of the paper, I refer to these detachments as "locations" or "beats" to align the exposition with more canonical language around organizing law enforcement workforces. Across these beats, troopers work either eight or ten hour shifts, with shift times staggered throughout the day to provide maximal visibility during peak traffic periods. Furthermore, troopers generally rotate through day- and night-shift assignments on either four or eight week rotations, depending on the beat.

Comparison to Typical Police Officer: While many of their tasks may overlap, the daily operations of highway patrol troopers are generally distinct from those of a police officer in a typical metropolitan area. Specifically, highway patrol troopers often have a primary emphasis on conducting traffic enforcement and upholding roadway safety, which also includes responding to collisions and assisting motorists. Additionally, troopers, at least in Washington State, are not generally the primary investigating agent for typical daily crimes. Instead, the Patrol will assist local agencies with more serious criminal investigations on an as-needed basis. This increased focus on traffic enforcement and away from crime responses is distinct from typical police officers, who face exactly the opposite responsibility emphasis.[12]

Moreover, troopers have a substantial degree of latitude in both the geographic patrol areas and actions they can take during a shift. While troopers are assigned to specific patrol areas, the Patrol has concurrent jurisdiction on all Washington State roads and troopers operate "independently and with minimal supervision."[13] In contrast, metropolitan police officers may patrol narrowly defined beats and often engage with the public in a reactive manner (Ba et al. 2022). Given their multidimensional responsibility set and wide latitude, troopers have substantial discretion in how to balance effort across these competing tasks, both within and across their shifts.

[11] See `https://www.wsp.wa.gov/about-us/history/` for additional details on the agency's history.
[12] Police officers will also respond to collisions and assist civilians in their daily operations.
[13] This description comes from the position description for a Washington State Patrol trooper (`https://www.wsp.wa.gov/wp-content/uploads/2022/03/trooper-position-description.pdf`).

2.2 The Role of Enforcement Targets in Personnel Evaluations

The substantial latitude and minimal supervision for troopers lends itself to a classic hidden action problem that is ubiquitous in employer-employee relationships. Specifically, because troopers operate with such minimal oversight, the Patrol is unable to ensure that troopers are consistently taking actions in the interests of the Patrol, which is to "enhanc[e] the safety and security of all people and communities."[14] A standard prediction of agency theory is that in the absence of oversight or sufficiently strong incentives, agents will choose to shirk, as effort is costly to them.

Like many public sector agencies, the Patrol is limited in the politically feasible tools it can use to tackle the underlying agency issue. For example, contract parameters are constrained by union negotiations and legislative restrictions on bonus payments. To circumvent this imperfect monitoring problem, the Patrol therefore develops specific traffic enforcement targets for their troopers. Throughout the rest of the paper, I refer to this production of traffic enforcement toward the target as violations. These enforcement targets are designed to encourage troopers to undertake costly (in an effort sense) enforcement actions and improve roadway safety through finding, sanctioning, and educating law-violating motorists. Targets vary across locations to account for spatial variation in traffic volumes but are generally time-invariant within the year, leading to some periods where it's more difficult for troopers to meet their goals. The per-shift target generally ranges from eight to ten violations, with performance averaged across shifts within months, affording troopers some potential to intertemporally smooth their effort across high- and low-productivity days. Across some beats, targets may be higher or lower, depending on traffic volume and the exact responsibilities of the position. Moreover, while targets for troopers are broadly centered around aggregate enforcement volume, the Patrol also places increased emphasis on a subset of offenses which they deem to be high-value, including driving under the influence, speeding, cell phone use, seatbelt violations, and finding drivers with suspended licenses. These latter offenses are highly valued because in the Patrol's view they are tightly connected to vehicular safety and accident-causing behaviors, in contrast to minor offenses such as expired registration.

While these enforcement metrics are part of a broader set of personnel evaluation criteria, including report timeliness, misconduct complaints, and other criteria, they are among the most salient features of a trooper's job description. In internal surveys, troopers report being "stressed out" that they won't "get in the right number" and express a general sentiment that the Patrol "focuses on data rather than actual results" (Branson et al. 2016). The purportedly strict emphasis on a numbers-based personnel evaluation system has led to general trooper dissatisfaction and high turnover rates.[15] In legislative hearings concerning whether quota-based systems should be outlawed in the state, Patrol leadership has argued that some type of numbers-based performance metric is crucial to providing trooper-level feedback and ensuring they are taking the right actions

[14] See `https://www.wsp.wa.gov/wp-content/uploads/2021/07/370008_Mission_Statement_letter_6-21_Color.pdf`.

[15] Historically, WSP also places a greater emphasis on performance and activity measures compared to other highway patrol agencies (Ewing & Associates 2004).

(SB 6316).[16]

The overarching goal of incentivizing troopers to conduct traffic enforcement relies on the assumption that traffic enforcement is negatively correlated with traffic collisions. Empirically, this assumption is corroborated among a number of settings (Makowsky and Stratmann 2011; Luca 2015; Goncalves and Mello 2023). Thus, while the Patrol may ideally prefer to deploy an evaluation scheme that conditions on outcomes (e.g., collisions), such a contract is infeasible since collisions are idiosyncratic events.[17] Therefore, the Patrol emphasizes a performance metric that is tightly correlated with their true outcome of interest, while also being more representative of actual individual trooper effort. However, placing such emphasis on a single performance metric, even if unintended, may distort trooper behavior and lead to troopers prioritizing enforcement effort over other tasks, even though the Patrol values performance on all of them.[18] This heightened saliency is partly due to the frequent measuring and tracking of traffic enforcement effort compared to other metrics which track non-enforcement effort such as collision report completion time.

2.3 A Principal-Agent Model with Task Substitution

I theoretically ground this tension between tasks and effort allocation with a parsimonious multitask principal-agent model. The goal of this simple framework is to highlight how differences in the relative cost of tasks can shift the allocation of effort. This model follows and builds on MacDonald and Marx (2001) and illustrates how the choice of enforcement effort depends on the relative cost of enforcement versus non-enforcement actions. A key feature of the model is that tasks will be substitutes from the standpoint of the trooper and complements for the Patrol. For expositional simplicity, I focus on the agent's problem in the main text (Lazear 2000; De Philippis 2021) and defer an explicit characterization of the Patrol's problem and additional model extensions to Appendix B.[19]

Preliminaries: The agent is endowed with one unit of effort which she divides between traffic enforcement t_e, non-enforcement activities t_n, and all other uses of effort t_o, which has marginal cost of zero. To map into the contracting literature, traffic enforcement is the "contracted" action and non-enforcement is the "non-contracted" action. The relative cost of enforcement $\omega(a) \in (0, 1)$ such that total effort is $t_e\omega(a) + t_n$. Let a be an exogenous factor that shifts the relative cost of enforcement, with $\omega'(a) > 0$. For example, poor recent performance ($\downarrow a$) on enforcement tasks (such as missing the target) lowers the cost of specialization in enforcement.

The Patrol does not observe the agent's effort allocation between enforcement and non-

[16] Even if troopers are pro-social (Chalfin and Goncalves 2023), appropriate incentives may still be required to ensure troopers are taking the "right" actions.

[17] More generally, developing a contract over a highly stochastic outcome will be infeasible as the correlation between effort and the observed output will be close to zero.

[18] Baker (1992) shows that performance measures not based on the principal's true objective will in general not lead to first-best effort provision. Baker, Gibbons, and Murphy (1994) study conditions under which a combination of subjective and objective criteria can be optimal.

[19] Appendix B also presents a version of the model where troopers choose between high- and low-quality enforcement, rather than between tasks, as I present here.

enforcement tasks, but does observe whether the agent succeeds on these tasks. Linking to the institutional details in a stylized manner, let "success" on traffic enforcement be defined as meeting the enforcement target and let "success" on non-enforcement activities be fully completing them in a timely and thorough manner. The Patrol receives a payoff only if both tasks are completed successfully and receives zero payoff otherwise. Outcomes are determined independently. Allocating effort t to a task successfully completes that task with probability equal to the allocated effort. For example, if the agent exerts effort $t_e = 1$, then she meets the enforcement goal with probability one and completes the other task with probability zero.

The agent can receive one of three distinct payoffs as a function of the tasks they complete:

$$\text{Agent's payoffs} \begin{cases} b & \text{if succeed on both } e, n \text{ tasks} \\ o & \text{if succeed on one of } e, n \text{ tasks} \\ n & \text{if succeed on neither } e, n \text{ task} \end{cases}$$

Following MacDonald and Marx (2001), define the following differences in payoffs for the agent

$$\Lambda^b = u(b) - u(o)$$
$$\Lambda^o = u(o) - u(n)$$

Agent's Problem: Let $u(\cdot)$ be the agent's vNM utility function associated with a given payoff and which satisfies the usual conditions of strict concavity and is twice continuously differentiable. Given a contract $\mathbf{c} \equiv (n, o, b)$ which specifies the payoffs above, the agent's problem is to choose effort allocations to solve the following maximization problem:

$$\max_{t_e, t_n} U(t_e, t_n, \mathbf{c}) = u(n) + t_e(\Lambda^o - \omega(a)) + t_n(\Lambda^o - 1) + t_e t_n(\Lambda^b - \Lambda^o) \tag{1}$$

Taking as given an arbitrary contract which generates an interior solution (see MacDonald and Marx (2001) and Appendix B for additional details), the trooper's first-order condition is:

$$t_e^* = 1 - t_n = \frac{1}{2}\left(1 + \frac{1 - \omega(a)}{\Lambda^b - \Lambda^o}\right) \tag{2}$$

It is straightforward to see that the only incentive schemes which will generate an interior solution are those with $\Lambda^b - \Lambda^o > 1 - \omega(a)$. For example, a monotone contract $\mathbf{c}$ with $b > o > n$ may generate an interior solution. Intuitively, lower relative costs of enforcement will shift or ultimately break this relationship, such that $\Lambda^b - \Lambda^o \leq 1 - \omega(a)$ and lead to troopers specializing in enforcement. The following proposition formalizes this idea and derives two empirical implications.

Proposition 1: Constrained troopers ($\omega(\underline{a})$) will increase enforcement effort (t_e) relative to unconstrained troopers ($\omega(\overline{a})$), where $(\underline{a}, \overline{a})$ are arbitrary values in the support of a, with $\underline{a} < \overline{a}$.

This proposition follows from a simple comparative static on the first-order condition, namely:

$$\frac{\partial t_e^*}{\partial a} = -\frac{\omega'(a)}{2(\Lambda^b - \Lambda^o)} \tag{3}$$

Since $\omega'(a) > 0$, greater values of a will lower t_e^* and vice versa.[20] Therefore, holding fixed $\Lambda^b - \Lambda^o$, troopers with lower relative costs of enforcement effort will choose to increase time spent on traffic enforcement, relative to troopers with relatively higher costs of enforcement effort.

Proposition 1 has two testable implications in the data, which neatly map into the identification strategy I leverage below in my investigation of trooper responses to misaligned incentives. I present the formal tests here and then discuss identification in greater detail in each of their respective sections.

To set some notation, suppose that $a \in A$ has two values, $(\underline{a}, \overline{a})$ with $\underline{a} < \overline{a}$. Respectively, these values of a correspond to constrained and unconstrained troopers. Let $p \in P$ be the value of some performance measure, such that if $p < 0$, troopers have $\underline{a}$ and if $p > 0$, troopers have $\overline{a}$. That is, the arbitrary cutoff of $p = 0$ induces a sharp change in the relative cost of enforcement.

Empirical Test: Marginally constrained troopers should have greater enforcement volume than marginally unconstrained troopers. Put differently, troopers with poor performance in one period should compensate in the following period since repeated poor performances can lead to negative personnel evaluations, performance improvement plans, or a lower chance of being tasked with desirable assignments. Intuitively, this test gives rise to a regression discontinuity design comparing the enforcement patterns of troopers with marginally different values of p, which generate sharp changes in a and subsequently, their relative cost of enforcement effort. Formally, letting Y_t denote a measure of enforcement volume (e.g., violations), the implication is:

$$\lim_{p \to 0^-} \underbrace{\mathbb{E}[Y_t | P_{t-1} = p]}_{\text{Constrained}} > \lim_{p \to 0^+} \underbrace{\mathbb{E}[Y_t | P_{t-1} = p]}_{\text{Unconstrained}}$$

Similarly, the increased enforcement volume implies that troopers substitute away from non-enforcement effort and towards enforcement effort, such that

$$\lim_{p \to 0^-} \underbrace{\mathbb{E}[t_e | P_{t-1} = p]}_{\text{Constrained}} > \lim_{p \to 0^+} \underbrace{\mathbb{E}[t_e | P_{t-1} = p]}_{\text{Unconstrained}}, \Big\} \text{ enforcement effort}$$

$$\lim_{p \to 0^-} \underbrace{\mathbb{E}[t_n | P_{t-1} = p]}_{\text{Constrained}} < \lim_{p \to 0^+} \underbrace{\mathbb{E}[t_n | P_{t-1} = p]}_{\text{Unconstrained}} \Big\} \text{ non-enforcement effort}$$

Together, each of these empirical tests of Proposition 1 indicate that raising the tightness or saliency of the enforcement target (through narrowly missing it) will lead to increased emphasis on enforcement effort since it has lowered the cost of specializing in the lower-cost task. Alternative models highlighting the decision between enforcement quality and enforcement equity in Appendix

[20] See Appendix B for additional details.

B predict that raising the constraint saliency will also decrease enforcement quality and disparately affect minority motorists.[21]

2.4 Data Infrastructure

In this paper, I assemble a rich data infrastructure to study the public safety consequences of misaligned incentives and to empirically test the predictions of the simple theory. I link nearly nine years (2010-September 2018) of incident-level activity records from the Washington State Patrol, known as Time and Activity Reports.[22] Crucially, these records include the date, time, location, trooper name, action taken, and associated violations and enforcement actions. These data also include basic motorist demographics including race, age, and gender.

I merge these activity records to administrative schedules which detail when and where troopers are assigned to be working. For each entry in the scheduling data, I observe trooper name, date, unit assignment (i.e., beat/location), and shift beginning and end hours. Combining these data together, I construct work histories at the shift-hour level for troopers over the sample period. To allow for a small degree of measurement error in the recorded time of the activity, I group any activity that occurs in the hour before or after the shift ends into the focal shift.[23] The typical scheduled shift for a trooper is ten hours in length. Less than one-fifth of scheduled employees work eight-hour shifts, but these generally reflect administrative positions or non-standard assignments and I exclude these shifts from my sample, along with a small number of highly specialized assignments.

I merge this data infrastructure with detailed information on personnel evaluation schemes describing both the construction and role of the enforcement target in the broader constellation of trooper performance evaluation. I leverage a combination of internal administrative documents including location-year-specific targets from monthly worksheets, year-end reviews, personnel evaluation forms, senior Patrol management-level feedback, and instructions for setting monthly targets to formulate and assign location-by-year specific targets.[24]

Table 1 presents descriptive statistics for troopers in my sample. Unsurprisingly, given the

[21] An alternative theoretical modeling approach is to specify a fully-dynamic model of trooper behavior. I choose to focus on the static model for two reasons. First, my focus is on the effect of period-over-period responses in incentives, rather than within-period responses, though below I empirically examine trooper behavior within-period. Second, under certain model parameterizations (e.g., Holmström and Milgrom 1987), the dynamic principal-agent model effectively reduces down to a sequence of static problems, suggesting that many of the core insights can be gleaned from a tractable static framework.

[22] In general, my data infrastructure leverages a subset of the raw underlying data from Pierson et al. (2020) along with independently sourced records.

[23] For example, if the shift ends at 4 p.m., but I observe a stop as 5 p.m. on the same day, I group this activity into the 4 p.m. hour of the preceding shift, and similarly for activity preceding the focal shift. Such a pattern may occur if the activity occurred at the very end of the regularly scheduled shift hour or was a lengthy stop and the time was only recorded ex-post. In practice, this procedure affects only a small handful of observations and makes little difference to the empirical results described below.

[24] I supplement these records and follow the institutional instructional documents on target setting and use the within location-year per-shift average as the relevant metric when it is otherwise missing. Benchmarking these estimated targets against administratively recorded ones yields an average deviation of less than one-tenth. I also show in falsification tests below that alternative "placebo" targets do not yield any discrete changes in trooper behavior. Moreover, any measurement error in the assigning of targets or the construction of the trooper performance measure should attenuate any findings.

population demographics of the state, troopers are predominantly white and male. The average trooper records around nine violations per shift, about two-thirds of which are moving violations (e.g., speeding, following too close, improper lane change, etc.) and the other one-third of which are non-moving violations (e.g., licensing, paperwork, etc.). While vehicle searches are relatively rare, the typical trooper conducts between two and three total assists per shift, reflecting the varied responsibility set of troopers. Finally, just under thirty percent of shifts contain a collision response, implying that the average trooper receives between four and five collision responses per month.

3 Enforcement Targets as Effort Incentives

In this section, I provide empirical evidence that enforcement targets have a causal impact on troopers' traffic enforcement levels. I also illustrate how these salient goals affect trooper behavior on an evaluation period-by-evaluation period basis, affecting both trooper traffic enforcement volume and the distribution of enforcement.

3.1 Trooper Responses to Varying Traffic Enforcement Targets

I first provide empirical evidence that trooper enforcement volume is a function of the enforcement targets that they face. To do so, I leverage trooper transfers across locations with varying enforcement objectives. This identifying variation is common in the wage premium (Abowd, Kramarz, and Margolis 1999; Card, Heining, and Kline 2016; Song et al. 2019; Bonhomme et al. 2023) literature and in other settings such as healthcare (Finkelstein, Gentzkow, and Williams 2016) and education (Jackson 2013).

I start by examining the dynamic impacts of a change in enforcement target around the time of the transfer across locations using an event-study design. Conditional on a move, this research strategy identifies the dynamic path of troopers' average change in enforcement volume as a linear function of the change in the target. Formally, I estimate

$$Y_{is(j,j')t} = \sum_{\substack{k=-4 \\ k\neq -1}}^{4} \underbrace{\tau_k \times \Delta Target_{s(j,j')}}_{\text{Event-time coefficients}} + \zeta_1 \Delta Target_{s(j,j')} + \underbrace{\mathbf{X}'_{is(j,j')t}\delta}_{\substack{\text{Fixed} \\ \text{effects}}} + \varepsilon_{is(j,j')t} \tag{4}$$

for trooper i, transfer s from location j to j', and time t (month and year). The coefficients of interest are τ_k which trace out the dynamic response of trooper per-shift average enforcement, interacted the change in enforcement target moving from j to j'. $\mathbf{X}_{is(j,j')t}$ includes trooper, month, year, and "origin" (j) and "destination" (j') location fixed effects. For simplicity, I align trooper movements so that troopers are always transferring to locations with greater expected enforcement volume.

One challenge with this empirical design is that trooper movements may be endogenously determined. For example, underperforming troopers may be systematically transferred to beats with less demanding responsibilities or trooper preferences over location may generate predictable sorting across beats. While the Patrol retains ultimate control over the assignment of troopers to beats, such endogenous behaviors may appear as violations of the parallel trends assumption.

Figure 1 presents the results of this exercise. Reassuringly, I find little evidence of any systematic parallel trend violations, as all pre-period coefficient estimates are closely centered around zero and statistically insignificant. The path of the pre-period estimates is also remarkably flat, suggesting little evidence in favor of the systematic sorting described above. In the post-transfer period, there is an immediate jump in enforcement activity. The magnitude of the period zero estimate is slightly attenuated as some transfers may occur within a given month, which will attenuate the change in per-shift enforcement volume. After this initial period however, enforcement volume sharply increases and levels off in the second month after the transfer, indicating minimal adjustment frictions such as trooper learning. Taking the average of the post-period coefficients as a benchmark, the estimates imply that increasing the per-shift enforcement target by one increases average per-shift violations by 0.52, or more than fifty percent of the change.[25]

While I have oriented trooper transfers to be in the positive direction for simplicity, the effects on trooper behavior need not be symmetric with respect to the change in target. I test whether this asymmetry is present in Appendix Figure A3 and estimate Equation (4) separately for troopers who transfer to locations with higher and lower targets. Appendix Figure A3 reveals some degree of asymmetry between the type of trooper transfer. Specifically, troopers who transfer to locations with higher targets increase their per-shift enforcement by around 0.8. In contrast, troopers who move to less demanding locations reduce their per-shift enforcement by only around two-fourths, though the long-run absolute magnitude of both directions is similar.

These mild asymmetric patterns are consistent with an increase in the target being more salient to transferring troopers than a decrease. Holding traffic patterns fixed, upward transferring troopers must exert additional effort to meet the more demanding goals, where as downward transferring troopers can hold their effort fixed and outperform the new, lower target.[26]

One limitation of the event-study approach is that the estimated responses are inherently local to the time of the transfer and may mask long-run learning behavior, where troopers initially only partially adjust their behavior before fully accommodating to the new target. Indeed, there is some evidence of these short-run dynamics where trooper enforcement volume grows over the first two months after the transfer before leveling off in the second post-transfer month. I explore this possibility in Panel B of Figure 1, where I plot changes in average per-shift violations against changes in the enforcement target in the raw data. Consistent with troopers adjusting more completely in the long-run, the magnitude of the slope coefficient indicates that trooper per-shift violation production increases by approximately two-thirds, an estimate that is larger in magnitude than the implied average event-study effect, although closely aligned with the estimated medium-run effects

[25]For completeness, Appendix Figure A2 presents robustness checks for a number of alternative event-study specifications. All of the estimates reveal sharp and similar magnitude changes to trooper per-shift enforcement in response to target changes.

[26]I also present Abowd, Kramarz, and Margolis (1999) (AKM) style transitions in Appendix Figure A4 for selected moves, grouping locations into bins which roughly correspond to quartiles of their target.

beginning in the second post-transfer month and onward. However, trooper responses still do not lie perfectly on the 45 degree line, indicating some degree of adjustment frictions.

Trooper traffic enforcement volume may not adjust on a one-to-one basis with the target because not all troopers are marginal to the target. Put differently, there are troopers who will consistently underperform and troopers who will consistently over-perform the target in their current beat. Transferring underperforming troopers to a more demanding beat means that these troopers may not be able to fully adjust because they were also unable to consistently meet their prior, lower target. Similarly, transferring over-performing troopers means that they need not adjust their behavior on a one-to-one basis to still meet the target since they were already exceeding the prior level. Either of these phenomena will lead to an attenuated response.[27] Nonetheless, both the event-study and cross-sectional approaches yield similar magnitude responses and empirically confirm that trooper traffic enforcement volume is strongly correlated to the location-specific demands they face.

3.2 Trooper Responses to Missing Enforcement Target

The above results highlight how trooper enforcement effort is malleable to location-specific demands. However, these findings may simply reflect supply-side changes in the prevalence of law-violating motorists across areas, rather than responses to the enforcement target *per se*. To rule out this supply-side channel as a sole explanation for these patterns, I next examine how missing the enforcement target influences trooper behavior in the following evaluation period among troopers who remain in the same location for consecutive months. Trooper performance on traffic enforcement is among the most salient features of their job design, with performance on this metric being closely tracked by Patrol leadership, and based on the qualitative evidence, by troopers themselves (Branson et al. 2016). Troopers who miss may have conversations with their sergeants about poor performance, be placed on a performance improvement plan, or anecdotally, be denied leave (Jones 2007). Consequently, troopers have strong incentives to perform consistently well on this performance metric.

I examine how troopers respond to narrowly missing their enforcement targets using a regression discontinuity design (RD). Intuitively, this research design compares troopers who just miss versus just make the stated enforcement goal. In effect, the target increases the saliency of the importance of conducting traffic enforcement, relative to other tasks. Formally, I estimate regressions of the following form:

$$\begin{aligned} Y_{ijt} = \alpha_0 + \alpha_1 \mathbf{1}[\text{Miss Target}_{ij,t-1}] + \alpha_2 \underbrace{\text{Relative Performance}_{ij,t-1}}_{\text{Enforcement - target}} \\ + \alpha_3 \mathbf{1}[\text{Miss Target}_{ij,t-1}] \times \text{Relative Performance}_{ij,t-1} + \underbrace{\mathbf{X}'_{ijt}\delta}_{\substack{\text{Fixed} \\ \text{effects}}} + \varepsilon_{ijt} \end{aligned} \tag{5}$$

[27] An alternative explanation is measurement error.

for trooper i, location j and time t. Y_{ijt} is a measure of per-shift enforcement activity. The coefficient of interest is α_1 which measures the change in per-shift enforcement activity for narrowly target-missing troopers, relative to target-achieving troopers. If a trooper's enforcement production is a key feature of their job then we would expect narrowly missing (constrained) troopers to increase their enforcement effort the following month to compensate. I also include linear measures of relative performance and a series of fixed effects, including location-by-year and location-by-month fixed effects which are designed to capture common shocks across troopers. Throughout, I weight the regressions using a triangular kernel to place greater emphasis on troopers near the cutoff, though I show robustness to alternative weighting schemes below.

The key identifying assumption underlying this research design is that troopers are unable to systematically manipulate their enforcement volume to perfectly reach the target. Unlike many RD settings, there is greater scope for potential manipulation since troopers in this context individually determine their enforcement volume through the actions that they take. The key detail is that troopers are unable to *perfectly* target their enforcement volume to match the target expectation or sort just above it (Lee and Lemieux 2010). Such imperfect control can arise from idiosyncratic shocks, such as higher or lower volumes of law-breaking motorists, weather patterns, traffic flows, or variation in non-enforcement responsibilities. The implication, therefore, is that the density of troopers should be smooth through the cutoff. I formally test whether this is the case in Panel A of Appendix Figure A5 using the nonparametric estimator of Cattaneo, Jansson, and Ma (2018; 2020) and find no evidence, either visually or statistically, of any systematic sorting ($p = 0.92$).[28] Panel B of Appendix Figure A5 also illustrates that trooper characteristics are smooth through the cutoff, providing additional evidence for the internal validity of the exercise.[29]

With the validity of the natural experiment established, I begin by estimating the probability that a trooper meets the enforcement target in time t as a function of their performance in the previous month $t-1$. The blue dots in Panel A of Figure 2 plot residualized means of the probability that a trooper meets the target in time t. Marginally missing troopers are 3.2 percentage points more likely to meet their enforcement goal in the following month compared to marginally meeting (unconstrained) troopers.[30] This response gap across troopers is larger (4.8 percentage points) for troopers who are on the margin of a second consecutive miss (hollow orange diamonds), consistent with troopers responding to signals about the potential negative career consequences associated with consistently poor performance.[31]

I further confirm these findings with two additional exercises in Appendix Figure A7. First, I include trooper fixed effects, effectively limiting the variation to comparing the same trooper on either side of the discontinuity. If anything the behavioral response appears to be even stronger

[28]The $[-1.5, 1.5]$ bandwidth corresponds approximately to plus or minus ten percentile ranks. I show below that results are robust, if not even stronger, when using narrower bandwidths.

[29]Column 2 of Table 1 displays summary statistics for this subsample.

[30]Mean reversion alone is unlikely to explain these patterns as troopers on both the left and right hand sides of the cutoff would experience this phenomenon, leading to no visual or statistical discontinuity.

[31]Appendix Figure A6 conducts placebo tests using alternative lower and higher "cutoffs," and finds precisely estimated null responses, confirming that the focal cutoff is capturing a key point in the trooper evaluation scheme.

and more precisely estimated within-trooper. Second, I examine whether troopers are similarly responsive in periods where it is easier or more difficult to reach the target due to variation in traffic volume since the target time-invariant within-year. For parsimony, I define the easy periods as the six months with the highest daily average volume of enforcement and the more difficult periods as the other six months.[32] While splitting the sample in this manner limits statistical power, I find slightly larger responses for troopers in easier months. Constrained troopers in more difficult months are as differentially likely to subsequently meet the target as in the pooled sample, although this estimate is less precise, even though there is a visual change at the discontinuity. Together, I take these results as evidence that the incentive scheme is a salient feature of troopers' daily activities throughout the course of the year.

Having established that troopers are responsive to enforcement targets and respond by changing their enforcement efforts, I now turn to examining the activity which generates such a response. In Panel B of Figure 2, I report RD coefficients along with associated confidence intervals.[33] Consistent with marginally-missing troopers increasing their enforcement effort, I find increases in per-shift violations, concentrated among non-moving violations, and stops, though the latter is less precisely estimated.[34] Moreover, that violations increase by a greater margin than stops also suggests that troopers are differentially more stringent and cite more violations per stop. I empirically test this prediction in the next series of coefficients, showing that marginally missing troopers cite more violations per stop, which is spread almost equally between moving and non-moving violations - an overrepresentation of non-moving violations relative to the equilibrium per-stop share.

Given this differential stringency, a natural question is whether troopers are equally differentially stringent across the motorist distribution. To examine these implications of trooper responses I compare the stringency gaps between white and minority (all other) motorists in Panel C.[35] I find suggestive evidence evidence of differential stringency across motorists, with minority motorists receiving weakly more violations per stop than white motorists. This disparity is concentrated among non-moving violations, a class of infractions where troopers have substantial discretion and are lower priority given their looser association with traffic safety.[36] In Appendix Table A1, I

[32]Easier months are January, February, March, April, May, and July.

[33]For completeness, Appendix Table A1 provides formal point estimates and standard errors.

[34]Non-moving violations are generally more minor infractions that are often related to equipment, licensing, and paperwork violations rather than inappropriate driving behavior such as speeding, improper lane changes, or cell phone use, among other infractions. I classify moving violations following the relevant statues and for expositional simplicity, classify other violations as non-moving.

[35]I group all non-white motorists together to reflect that the population demographics of the state vary by region (e.g., higher rates of Black individuals in metropolitan areas and higher rates of Hispanic individuals in rural areas) as well as to capture that there will be substantial traffic flows across the state throughout the day. Moreover, recent work has highlighted measurement error in how race and ethnicity are recorded in the criminal justice system (Finlay, Luh, and Mueller-Smith 2024) - aggregating motorists together in this binary fashion should minimize the degree of measurement error, but attenuate the magnitude of any disparities.

[36]In Appendix Figure A8, I undertake a battery of robustness checks to ensure the stability of the above empirical patterns. These findings are robust to alternative bandwidths, weighting schemes, and using within-, rather than across-trooper variation. In fact, narrowing the bandwidth or using within-trooper variation generally increases the magnitude and precision of the RD estimate. The choice of weighting scheme also does little to impact the magnitude of the regression discontinuity estimates.

further show that these disparities and increased stringency (Panels B and C) are larger during periods of the year where it is more difficult to meet the enforcement target, although splitting the sample limits statistical power, indicating that greater degrees of target saliency increase the prevalence of deleterious outcomes at the margin.

Effectiveness vs Efficiency: The available evidence indicates that enforcement targets are effective at inducing additional enforcement effort from troopers. Thus, the incentive scheme performs as intended. However, much of this effectiveness response comes from generally minor non-moving violations and differentially greater stringency for minority motorists. If the goal is to minimize collisions, then the deterrence effects from minor non-moving violations are likely less impactful than those for more severe offenses like speeding or reckless driving. Moreover, this latter disparity also has the potential to harm the financial stability of low-income motorists (e.g., Mello 2023; Luh, Pyle, and Reeves 2024), as race and ethnicity are correlated with income in Washington State, as in the United States more broadly (Appendix Figure A9). Such an enforcement disparity may be socially efficient however, if there are differences in the driver risk distribution across race and ethnicity. Put differently, the increased enforcement may be socially valuable if it results in increased deterrence effects for risky drivers.

While I am unable to directly test for individual deterrence effects given data limitations, measured either by receiving future driving sanctions or collisions, I partially test whether minority drivers are relatively riskier than white drivers by examining collision rates over the age profile. Specifically, I calculate the average number of yearly collisions by motorist age and racial or ethnic group and scale by the corresponding population.[37] While this exercise is an imperfect measure of driver risk, it offers a first-order approximation of driving behavior in the population using a metric that is objective and not confounded by potential differences in enforcement patterns.

Panel A of Appendix Figure A10 reports local polynomial smoothed collision risk distributions separately for white and minority drivers.[38] While the two distributions share the same shape, minority drivers exhibit lower collision rates uniformly across the age distribution, suggesting that the stringency documented above is likely unwarranted on the basis of risk alone. Panel B disaggregates minority motorists into each group separately. Black motorists have higher collision risk than white motorists across the age profile. Younger Hispanic motorists have slightly lower collision rates than white motorists while the opposite is true for older Hispanic motorists. Other racial minorities, such as Asian and Pacific Islander and Native American/Alaskan Native have sharply lower collision rates compared to white motorists.[39]

[37]I measure population using the 2010-2018 waves of the American Community Survey.

[38]While WSP has concurrent jurisdiction over all roads in the state, the set of collisions, and therefore estimated risk distributions, they respond to may not be broadly representative. To test this representativeness, I generate similar risk distributions by motorist gender in Appendix Figure A11 and compare them to the corresponding distribution from the Washington State Department of Transportation database. Relative to the universe of collisions in Washington State, WSP-responded collisions are slightly more likely to involve younger drivers, perhaps a sign that the collisions themselves are more severe, though the shapes of the distributions themselves are highly similar. Taken together, I conclude that the estimated risk distributions by race/ethnicity are a reasonable approximation to the statewide equivalent.

[39]This relative ranking is also broadly representative of fatal collision rates in public data

I further extend this investigation into differences by motorist risk and estimate a "risk score" using race-by-gender fixed effects and a quadratic in motorist age. I define "high-risk" drivers as the top quintile and "low-risk" drivers as the bottom four quintiles of this measure. Appendix Figure A12 examines racial differences in per-stop stringency among these two risk groups. If anything, I find that the majority of the stringency disparity is concentrated among non-moving violations for high-risk motorists with few differences in the rest of the distribution. Together, these results illustrate that minority motorists receive differentially greater per-stop stringency for relatively minor infractions from constrained troopers even when compared to white motorists in the same part of the risk distribution.[40]

Aside from differences in driver risk, there may be underlying differences in vehicle quality (e.g., expired registration) between motorist race/ethnicity that lead to pre-textual or "mixed-motive" stops.[41] If these differences are effective tags (Akerlof 1978) for driver risk, either with respect to driving behavior or the likelihood of carrying contraband, then the efficiency gain may outweigh the disparity cost from a public safety standpoint. I test this mechanism in the context of vehicle searches - if the non-moving violations are effective tags, then the search outcomes for minority motorists should also be differentially successful. One drawback to this exercise is that I do not observe the reason for the search. However, conditional on a search, around half of stops also have a non-moving violation associated with them, suggesting that there is a possibility for non-moving violations either being used in a mixed-motive manner or the search uncovers additional non-moving violations.

Appendix Figure A13 reports regression discontinuity estimates using various measures of trooper search behavior as the outcome. I first document that there is an aggregate search response, as marginally missing troopers weakly conduct more searches, particularly for minority motorists. These searches are not differentially more successful however as "hit rates" weakly fall.[42] Finally, I test whether troopers cite more violations to these additional searches as suggestive evidence of non-moving violations being used as the basis for a search. I find that violations during unsuccessful searches, particularly for minority motorists, sharply increase, with no detectable changes for white

(https://wtsc.wa.gov/dashboards/traffic-fatality-rates-dashboard/). One exception is for Native American/Alaskan Native drivers who have elevated fatal collision rates compared to white drivers in contrast to the overall collision rates illustrated in Appendix Figure A10.

[40] These patterns are not just driven by Black motorists who have the highest collision risk, as dropping areas with a high prevalence of Black motorists only exacerbates these patterns. This pattern of differential stringency could also be consistent with officer preferences for allocating more sanctions to high-risk motorists (Goncalves and Mello 2023) interacting with imperfect priors over motorist risk by subpopulation. Moreover, from an efficiency standpoint, high-risk drivers are the least likely to be deterred from sanctions (Goncalves and Mello 2023).

[41] Non-moving violations may be used as the basis for conducting stops and searches in the absence of any specific moving violation. While strictly speaking, pre-textual stops were outlawed in the state in 1999, "mixed-motive" stops were ruled constitutional in 2012. Mixed motive stops are those which the officer has an independent reason for stopping and searching a motorist, such as an illegally altered muffler as in *State v. Arreola 2012*, which may lead to further search and investigation.

[42] These estimates on search quality represent the average outcome among searches conducted, not the marginal. The marginal search outcome can be recovered by scaling by the probability of a search being conducted and under an additional assumption that the enforcement target only influences the probability of a search occurring. Since the probability of a search increases for minority motorists, the marginal search outcome is even less likely to reveal contraband.

motorists or during successful searches. Taken together, these results provide little evidence to support the fact that either searches are differentially warranted on the basis of efficiency grounds or that non-moving violations are sufficiently sharp tags for contraband yield. These results also provide an alternative mechanism for, and corroborate, historical reporting documenting that white drivers are in fact more likely to carry contraband in the state, even though they are searched at lower rates than minority drivers (Buch and Borkholder 2020).

In the absence of measurable risk differences or effective targeting and search due to non-moving violations, there are two alternative mechanisms which may explain these enforcement patterns. The first is differential "hassle costs" across motorist groups. When a motorist receives a traffic citation, they have the opportunity to contest the citation in court and offer evidence in their favor. If the infraction is contested, the citing trooper is required to appear in court and testify that the infraction occurred. Given differential opportunity costs and access to political capital, white and higher-income motorists face lower barriers to contesting marginal infractions. Thus, all else equal, minority or lower-income motorists should have lower contesting rates. Prior work has found evidence in favor of hassle costs driving differential search patterns in Texas (Feigenberg and Miller 2023) and differential ticketing patterns in Massachusetts (Makowsky and Stratmann 2009).[43]

While I do not have information on individual downstream violation outcomes in my data, I source caseload reports from the Washington State Administrative Office of the Courts. These caseload reports include court-level outcomes on the number of traffic infractions that have hearings or are contested. Although it is an imperfect test for the presence of hassle costs, I test whether population characteristics of the county are predictive of caseload outcomes. Following prior work, I focus on the relationship between caseload outcomes and poverty rates, with the implication that poverty and racial demographics are also positively correlated (Appendix Figure A9). Appendix Figure A14 reports these cross-sectional correlations. I find suggestive evidence that counties with higher poverty rates have lower rates of traffic cases with any hearings and unconditional on hearing, are less likely to be contested. Together, these correlations are consistent with troopers facing lower hassle costs for sanctioning lower-income and minority motorists. Consistent with the version of the principal-agent model in Appendix B, this should lead to greater enforcement for minority motorists.

A final potential mechanism is racial animus across troopers. Under this mechanism, troopers have differential preferences for enforcement that co-vary with motorist race and ethnicity. I conduct the Antonovics and Knight (2009) test for racial bias in the context of vehicle searches. While this test relies on potentially strong econometric assumptions and there is an imperfect mapping from search behavior to general traffic enforcement, it still provides some useful information on the potential for differential treatment of motorists of different races and ethnicities. I find little evidence of racial animus in the contexts of vehicle searches ($p = 0.54$), suggesting that animus

[43] Police officers more generally also alter their behavior when they face overtime opportunities, which such court appearances may fall under (Chalfin and Goncalves 2023).

alone is unlikely to fully explain the differential enforcement patterns I document above.

Taken together, the preceding exercises demonstrate that, consistent with prior qualitative evidence, troopers are highly responsive to their enforcement targets. However, while this personnel management tool is effective in the sense of increasing enforcement effort, it also induces differential changes to trooper behavior, including increased production of minor violations and per-stop stringency, which differentially affects minority motorists. This change in enforcement distribution is unlikely to be fully attributable to differences in motorist risk, tagging, or racial animus, but is instead consistent with differential costs of enforcement interacting with trooper performance. These results offer a novel, systemwide explanation for previously documented differential stringency (Lovrich et al. 2003) that does not hinge on an interpretation of racial animus, which I find no evidence of, at least in the context of searches. Instead, these empirical patterns are the product of troopers adversely reacting to their incentive scheme, producing negative externalities. Moreover, from a deterrence standpoint, the production of such minor violations is unlikely to be as impactful as a focus on higher-risk offenses such as speeding or reckless driving. Below, I examine how the presence of this enforcement target also influences the substitution of trooper effort across enforcement and non-enforcement tasks.

4 Multitasking and Competing Objectives

In addition to altering the quality of enforcement actions, enforcement targets may also divert trooper away from less salient responsibilities. In this section, I examine how troopers trade off effort across enforcement (contracted) and non-enforcement (non-contracted) responsibilities. I provide evidence that the design of the existing incentive scheme induces constrained troopers to focus more on short-run enforcement and delay completion of other non-contracted responsibilities. In additional exercises, I also study the composition of this short-run enforcement emphasis and conduct a variance decomposition exercise to disentangle the role of troopers and locations in generating these adverse outcomes.

4.1 Tradeoff Between Contracted and Non-Contracted Actions

During their shifts, Patrol troopers are responsible for completing a variety of tasks, including both traffic enforcement and non-enforcement actions, such as responding to calls for service and collisions. Using a similar regression discontinuity design as in Section 3.2, I examine how marginally missing and marginally meeting troopers allocate their effort on these non-contracted responsibilities as measured using collision responses.[44]

[44] I modify the previous research design slightly and estimate Equation (5) at the trooper-month-year-collision event level. This modification allows me to include hour-of-shift, hour-of-day, and district-by-date fixed effects which help control for variation in traffic flows and influence how troopers allocate their effort across competing tasks in addition to including trooper fixed effects. In the RD exercises below, I also examine events which occur in the interior hours $h \in [3, 8]$ of the shift to allow estimation of pre-event balance, observation of any follow-up activity, and to maintain consistency with exercises in Section 4.2 and the remainder of the paper.

Figure 3 presents estimates of missing the target on trooper non-contracted effort. I first show in Panel A that marginally missing troopers are no more or less likely to be assigned collision responses. However, marginally missing troopers are more likely to have observed non-collision activity immediately following a collision response, as measured either in number of hours until or the probability of having any activity in the following hour.[45] Thus, marginally missing troopers return to active patrol and traffic enforcement duties more quickly than marginally meeting troopers. I further investigate the composition of this activity in Appendix Figure A16 and Appendix Table A2, where I confirm that marginally missing troopers are weakly more likely to conduct only enforcement activity in the following hours and demonstrate that the burden of their enforcement efforts predominately falls on minority motorists, consistent with the conclusions of previous exercises.[46]

One interpretation of these results is that the incentive scheme induces constrained (marginally missing) troopers to efficiently complete collision and non-contracted responsibilities before returning to their usual activities. However, I show in Panel B that this is not the case. Obligations from collision responses include both the immediate response of clearing the road and controlling the scene along with the downstream completion of lengthy narrative reports detailing the incident. Using a subsample of collision reports (2014-2018) probabilistically linked to events, I find that constrained troopers are 1.4 percentage points (around 4%) more likely to file collision reports more than one week after the event, despite having no difference in observed response times.[47] This delayed reporting is exacerbated, although imprecise, among the sparse subsample of events which involve driving under the influence or end in felony charges, where constrained troopers are also more likely to file reports after five business days, an internal reporting deadline for felony events.[48] Together, these results suggest that constrained troopers ex-post prioritize short-run quantifiable activity related to job performance at the expense of the public through lower longer-run performance on less salient responsibilities, despite behaving ex-ante similarly to unconstrained troopers.

However, the short-run emphasis on enforcement is not without cost. Internal documents highlight how the Patrol views the obligations of troopers as to the victims with respect to collision responsibilities and how the delayed reporting implicitly hampers the expediency of the legal

[45]In this exercise, activity includes both enforcement and motorist assists that the trooper initiated. Both of these actions require active driving and patrol efforts, capturing troopers conducting non-collision response obligations. For troopers who have no observed activity over the remainder of the focal shift, I assign them the number of hours remaining in the shift plus one to differentiate them from troopers whose activity occurred at the end of the shift. This is equivalent to these troopers conducting their first observed activity in the first hour of the following shift. All analysis controls for the within-shift timing of the event. Appendix Figure A15 presents a number of robustness checks.

[46]Appendix Figure A17 shows that pre-collision activity remains balanced across the discontinuity and Appendix Figure A18 displays robustness checks.

[47]Given the subsample and probabilistic nature of the linkage as well as the paucity of well-identified events with felony charges, I focus on the mean difference between constrained and unconstrained troopers, rather than the effect at the discontinuity itself.

[48]Driving under the influence may be charged as a felony or a misdemeanor, depending on the driver's prior history. Given data limitations, I am unable to precisely distinguish between these two charge grades in the data and so present results with DUIs included to maximize sample.

process and postpones the payment of victim compensation. I estimate the impacts of this report completion lag on social welfare in Equation (6), comparing the difference in net present value from the victim's settlement due to the delayed report with the implied government revenue and deterrence effects from additional traffic enforcement.

$$\underbrace{\Delta\text{NPV Settlement}}_{\text{1 week delay}} = \left[\underbrace{\alpha_1(\text{Moving Viols.})}_{\text{RD Estimate}}\right] \times \left(\underbrace{\text{Rev. from enforce.}}_{\substack{\text{Ticket value } \times \\ \text{Pr(payment)}}} + \underbrace{\text{Deterrence effect}}_{\substack{\text{Deter treat. eff. } \times \\ \text{settlement cost}}}\right) \tag{6}$$

On the left-hand side of Equation (6) is the change in the net present value (NPV) of the collision victim's settlement payment to reflect the delayed report. On the right-hand side is the benefit from additional enforcement specialization, combining the additional government revenue from enforcement and the deterrence effect of additional collisions due to increased enforcement - see Appendix D for additional details and the sources for all key parameters.

Figure 4 presents bounds on the net welfare impact of this effort tradeoff under a range of assumptions. In Panel A, I assume that the enforcement activity deters a minor collision. For each shaded region I place bounds on the welfare change using a range of settlement values for the collision whose report is delayed. I then compute the social value of additional enforcement against this change in settlement NPV over a range of trooper activity hours. If the additional enforcement deters a minor collision, then the breakeven point is between three and seven hours of additional enforcement. Put differently the social cost of the delayed report is offset by between three and seven hours of additional traffic enforcement. This breakeven point shrinks if the deterred collisions are of moderate severity (Panel B).[49] Together, these calculations suggest that the welfare impact of enforcement emphasis on this margin is approximately welfare-neutral or slightly welfare-negative with respect to the settlement payment. Over a range of plausible parameters, the increased enforcement emphasis does not offset the reduced net present value of the settlement payment to the victim. The exceptions are when troopers devote at least half a shift to enforcement and the additional enforcement deters collisions of moderate severity. One caveat to these interpretations is that the exercise abstracts from distributional implications and the potential harm on cited motorists if the marginal sanction harms household balance sheets (Mello 2023; Luh, Pyle, and Reeves 2024). If any of the enforcement benefits instead disproportionately harm individuals with high social welfare weights or reduce the social benefit through poor downstream financial outcomes, then the conclusions of this exercise favor timely report completion even more.

4.2 Examining Non-Local Responses to Multitasking Obligations

The results in the previous subsection reveal that marginally missing troopers prioritize short-run enforcement effort at the expense of longer-run obligations. However, a drawback with the

[49]For expositional purposes I display the same calculus for "severe" collisions, which have settlement bounds between \$489,000 and \$2.7 million in Appendix Figure A19. For the range of these outcomes, the benefit of additional enforcement does not offset the difference in NPV for the settlement payment even when devoting an entire shift to enforcement.

regression discontinuity approach is that the effect is only identified for troopers directly at the cutoff. In practice many troopers may be located away from the cutoff. Moreover, if the ultimate goal is to examine the potential for alternative personnel policies to mitigate the distortion of the existing incentive scheme, then estimates across the trooper distribution will be needed.

In Appendix C, I develop a complementary dynamic difference-in-differences approach which compares collision responding and non-responding troopers who work the same location and shift, recovering the difference in hourly activity across these groups of troopers. Intuitively, this approach identifies heterogeneity in how troopers trade off contracted and non-contracted actions when facing the same working conditions. Specifically, this exercise can be thought of as examining the composition of the enforcement troopers prioritized when abstracting from longer-run non-contracted action obligations and to what extent troopers make these cross-task tradeoffs. The key advantage of this approach is that effects are identified for troopers away from the cutoff, at the cost of an auxiliary parallel trends assumption - see Appendix C for full details. Moreover, all troopers are "treated" multiple times throughout the course of the sample period, suggesting that trooper-specific responses are recoverable from this identification strategy.

I use the dynamic difference-in-differences approach to further investigate how the incentive scheme encourages the production of socially suboptimal outcomes in two ways. First, I explore how troopers respond over the course of the month as the saliency of the target also increases near the end of the month. The intuition is that near the end of the month, the disruption cost of collision responses is greater for marginal troopers, since they are afforded less time to smooth over these shocks. Mapping into the two groups of troopers, constrained troopers should maintain elevated rates of enforcement emphasis throughout the course of the month and unconstrained troopers should exhibit an upward sloping gradient, since the constraint is only salient for them during the end of the period.

Figure 5 presents kernel-weighted estimates of trooper responses as a function of the time left in the month for various outcomes, with Appendix Figure A20 showcasing additional outcomes for completeness. Consistent with the predictions, constrained troopers exhibit uniformly larger degrees of enforcement specialization, whether measured as violations (Panel A) or on the extensive margin of effort substitution (Panels B and C). In general, marginally unconstrained troopers also exhibit a gradient of increasing specialization over the course of the month, consistent with their behavior becoming more similar to constrained troopers when the enforcement target is more salient. Moreover, constrained troopers produce enforcement disparities across the entire month, with the behavior of unconstrained troopers becoming increasingly disparate near the end of the month. Finally, for all of these outcomes, I additionally plot the behavior of "highly unconstrained" troopers, a subset of troopers far to the right-hand side of the RD. These troopers represent an empirical "placebo" group, whose behavior should be relatively unaffected over the course of the month since the enforcement target is non-binding to them. While substantially noisier due to smaller samples, the within-month behavior of this group is generally less sensitive than that of even the unconstrained troopers, as would be predicted. If anything, these placebo troopers exhibit degrees of specialization towards non-enforcement activities and generate enforcement disparities that move in the opposite direction. Thus, these results further confirm the distortion of trooper effort and behavior is greatest when the enforcement target is most binding.

Second, I investigate the broader consequences of these compensating enforcement responses on traffic court caseload outcomes using the court data from Section 3.2. If some of the compensating enforcement activity is of "marginal" quality, then the outcomes for these sanctions should be worse and manifest themselves in the cross-sectional correlation between the court caseload composition and the degree of enforcement emphasis. While I am unable to directly link the outcomes of individual sanctions to their case outcomes given data limitations, I construct county-level measures of the enforcement responses and correlate these with the mean caseload outcomes from 2010-2018 in Appendix Table A3. Perhaps strikingly, I find evidence that counties where troopers are more enforcement-focused have measurably worse caseload outcomes, including higher rates of dismissal and higher rates of the adjudicating body determining the infraction was not committed. These patterns are consistent across multiple approaches to measuring the degree of enforcement focus. One exception is correlating the race-neutral outcomes against the enforcement disparity, which if anything, reveals correlations in the opposite direction. However, this correlation is rationalized by previous results where minority and lower-income motorists have greater opportunity costs of contesting infractions in court. Together, these simple correlations highlight a significant downstream consequence of trooper enforcement emphasis and provide suggestive evidence that at least some of this compensating enforcement response results in lower quality citations.[50]

4.3 Trooper vs Location Heterogeneity

The documented production of negative externalities could reflect either heterogeneous responses to the salient target and how troopers uniquely react to multitasking obligations (i.e., some troopers are more sensitive than others) or it may simply reflect fixed differences across locations. Thus, understanding whether trooper or location heterogeneity is first-order is crucial for designing potential alternative policies which help mitigate these adverse responses.

To disentangle whether troopers or locations contribute more to response heterogeneity, I begin by collecting the full set of collision responses with well-defined control groups from the difference-in-differences design.[51] Define the response $R_{ije(h,t)}(y)$ as the mean hourly collision response (event-specific estimate), for trooper i, in location j, for event e as a function of collision shift-hour h and date-time t, for a given outcome y. I decompose the variation in $R_{ije(h,t)}(y)$ through a two-way fixed effects framework which relates the magnitude of the response to vectors of trooper and location

[50] For example, troopers could reduce the threshold at which they contact motorists for speeding. Unfortunately, I do not have information on charged and posted speeds and cannot examine this mechanism directly.

[51] An aggregate difference-in-differences estimate is the average across each of these event-specific estimates. The intuition is similar to a "stacked" event study (Cengiz et al. 2019) where "event-specific" estimates are recoverable.

fixed effects:

$$R_{ije(h,t)}(y) = \underbrace{\alpha_i}_{\substack{\text{Trooper}\\\text{fixed}\\\text{effects}}} + \underbrace{\psi_j}_{\substack{\text{Location}\\\text{fixed}\\\text{effects}}} + \underbrace{\mathbf{W}'_{e(h)}\delta}_{\substack{\text{Auxiliary}\\\text{controls}}} + \varepsilon_{ije(h,t)} \tag{7}$$

where $\mathbf{W}_{e(h)}$ includes indicators for the shift-hour of the collision response h, which I partial out in a first-step. Equation (7) is similar in spirit to an Abowd, Kramarz, and Margolis (1999) type regression, except troopers may experience multiple collision response events $e(h,t)$ within a given location j, rather than working at the same firm for multiple years.[52] The goal is to estimate the relative contributions of the variances of the trooper and location effects in explaining the variance of $R(y)$. Define the variance of trooper effects as $\mathbb{V}_{\alpha_i}$ and the variance of location effects as $\mathbb{V}_{\psi_j}$.[53]

Simple "plug-in" approaches to estimating $\mathbb{V}(\cdot)$ will be upwardly biased due to estimation error in the trooper and location fixed effects. I follow the leave-out estimation procedure from Kline, Saggio, and Sølvsten (2020) and obtain unbiased estimates of $\mathbb{V}_{\alpha_i}$ and $\mathbb{V}_{\psi_j}$. I then take the ratio of these two components to compactly summarize the relationship between the variance estimates and a given $R(y)$.[54] This magnitude of this ratio can be interpreted as the following:

$$\frac{\mathbb{V}_{\alpha_i}(y)}{\mathbb{V}_{\psi_j}(y)} \begin{cases} > 1, \text{ trooper effects more important,} \\ < 1, \text{ location effects more important} \end{cases}$$

Figure 6 summarizes this ratio across a series of outcomes measuring trooper enforcement patterns and effort substitution across tasks. Respectively, each set of bars represents the relative contributions of trooper and location variances for three measures of enforcement activity, effort substitution across tasks, and accompanying enforcement disparities. Across all outcomes, I find variance ratios that are consistently above one, ranging from 1.7 to 4.1, indicating that trooper effects are substantially more important for explaining the variance in responses to non-contracted actions than location effects.[55] The importance of trooper effects is also increasing from within each set of activities as the outcome shifts to those where troopers have greater discretion (e.g., moving vs non-moving violations).[56]

Finally, I also confirm that trooper heterogeneity plays a key role in determining not only

[52] Whether $\mathbf{W}_{e(h)}$ is partialled out or omitted entirely makes no difference for the conclusions from this decomposition as I discuss below. Conducting this exercise collapsed down to the trooper-location-year level yields similar results.

[53] To simplify the notation, I write $\mathbb{V}(R(y))$ as $\mathbb{V}(y)$, where $R(\cdot)$ is implicit in the argument.

[54] See Appendix E for additional details on the estimation procedure and a derivation of the ratio measure along with its relation to more traditional decomposition statistics.

[55] I confirm the enforcement target is a strong predictor of these outcomes in Appendix Table A4 where I project estimates of the trooper fixed effects onto the relative performance at the time of the event. Consistent with the previous results, I find that more constrained troopers place greater emphasis on enforcement and have larger enforcement disparities.

[56] For completeness, Appendix Figure A21 presents simple "plug-in" variance estimates. Appendix Figure A22 presents alternative methods of conducting this decomposition, omitting the first-step partialling out $\mathbf{W}_{e(h)}$ and various degrees of winsorization to reduce the influence of potential outlier estimates. These permutations are repeated for both the simple and unbiased approaches. The conclusion that trooper effects contribute substantially more to explaining the variance of $R(y)$ holds across all of these permutations.

trooper collision responses, but per-shift enforcement as well. I conduct the decomposition from Equation (7) at the trooper-by-shift level in Appendix Figure A23. Consistent with trooper heterogeneity playing a key factor in determining enforcement patterns, trooper effects explain up to 9.9 times the variance of location effects, depending on the outcome.

Taken together, these results imply that the observed variation in multitasking responses and per-shift enforcement reflect substantially more than just mean differences across locations. In fact, trooper effects explain many more times the variation in observed outcomes than corresponding location effects. This decomposition further underscores the importance of quantifying and accounting for trooper heterogeneity when developing policies which may mitigate the prevalence of these adverse outcomes.

5 The Distribution of Trooper Responses

The previous variance decomposition highlights how trooper heterogeneity is of first-order importance in generating socially suboptimal outcomes. Therefore, policies which may wish to target this heterogeneity to minimize the prevalence of negative externalities require estimates of trooper-specific behavior. To quantify differences across troopers, I begin by constructing trooper-specific estimates ($\theta_i(y)$) as the weighted average of all trooper-specific difference-in-differences estimates.

While it is attractive to work with the empirical estimates of $\theta_i(y)$ directly, these estimates will reflect a combination of true differences across troopers as well as estimation error. In extreme cases, this estimation error will be substantial enough as to generate observed differences across troopers that are entirely just noise. One way to assess whether the distribution of trooper estimates contains entirely estimation error is to construct a bias-corrected estimate of its variance (Krueger and Summers 1988; Kline, Rose, and Walters 2022):

$$\mathbb{V}(\theta_i(y)) = \left(\frac{N-1}{N}\right)\Bigg[\underbrace{\frac{1}{N-1}\sum_{i=1}^{N}\left(\theta_i(y) - \overbrace{\frac{1}{N}\sum_{l=1}^{N}\theta_l(y)}^{\overline{\theta_i}(y)}\right)^2}_{\text{Sample variance}} - \underbrace{\frac{1}{N}\sum_{i=1}^{N}\sigma_i^2(y)}_{\text{Bias correction}}\Bigg] \tag{8}$$

The bias correction term is the average trooper-specific squared standard error. Intuitively, if the cross-trooper dispersion is predominantly estimation error, then $\mathbb{V}(\cdot)$ should be close to zero. Appendix Table A5 reports estimates of mean cross-trooper responses along with naive plug-in and bias-corrected variances. Consistent with significant variation across troopers that reflects additional signal beyond just noise, I find sizable remaining variation, even after correcting for estimation error.

Documenting significant variation across troopers does not solve the challenge of quantifying individual trooper estimates in the presence of estimation error however. Below, I explore two potential solutions. The first approach groups troopers into discrete bins or “types” which reflect the covariance across measures of trooper activity and performance. While this clustering approach leverages noisy but unbiased estimates, it only permits identification of a finite number of types, rather than trooper-specific estimates. A second approach is to employ Empirical Bayes (EB) tools to shrink noisy trooper-specific estimates towards a common prior, trading off bias for variance. While the individual EB estimates themselves are biased, they allow the researcher to recover trooper-specific estimates instead of group-level aggregates. Below, I employ both techniques to quantify the heterogeneity across troopers.

5.1 Identifying Discrete Trooper Typologies

In this subsection, I use k-means clustering, an unsupervised machine learning algorithm, to classify troopers into a finite number of distinct groups.[57] As inputs, I use the trooper-specific responses for violations, violation disparity, and the probability of having only enforcement activity, along with trooper mean per-shift violations, such that troopers are characterized by this four-dimensional joint distribution. I define per-shift enforcement as the trooper fixed effect from the following auxiliary regression:

$$Y_{ijst} = \sum_{i=1}^{N} \gamma_i + \mathbf{X}'_{jt}\beta + \varepsilon_{ijst} \tag{9}$$

for trooper i, location j, shift s, and time t, where $\mathbf{X}_{jt}$ includes both location-year and date fixed effects. In this context, γ_i recovers an estimate of the trooper-specific average violations per shift, accounting for differences in targets and responsibilities across locations, along with idiosyncratic daily shocks.[58] Similar "benchmarking" regressions have previously been used to study the average behavior of judicial agents in pretrial release (Arnold, Dobbie, and Hull 2022) and policing (Gelman, Fagan, and Kiss 2007; Luh, Pyle, and Reeves 2024). Using these four dimensions as inputs, the clustering algorithm optimally identifies five distinct groups of troopers.

Figure 7 plots the lower triangular matrix of these joint distributions. Each subgraph plots a two-way marginal distribution along with troopers grouped according their resulting classification. Starting with the upper left subgraph, it is unsurprising that troopers who are more likely to have only observed enforcement activity also produce greater volumes of violations, as indicated by the positive correlation between the two measures. However, holding fixed one value, there is still substantial variation in the other dimension. This pattern of heterogeneity persists across other outcomes as well. Moreover, consistent with relative, rather than absolute performance playing a role in determining trooper enforcement responses, there is sizable variation in all three collision response measures, conditional on a given level of per-shift enforcement. Put differently, it is not the case that absolutely low-performing troopers always have positive enforcement responses or

[57] See Appendix E for additional details. Beyond just used for classifying data into groups, k-means clustering has also been used as a dimensionality-reduction device in high-dimensional fixed effect models (Bonhomme, Lamadon, and Manresa 2019). In other applications with agent heterogeneity, Kim, Sudhir, and Uetake (2022) use a structural model to identify two latent types of loan officers who differ in their ability to acquire and collect on microfinance loans.

[58] Following Angelova, Dobbie, and Yang (2023) I truncate a small handful of observations which have implied negative γ_i coefficients, replacing their estimates with the minimum positive value. Specifications with higher dimensional or alternatively specified fixed effects yield similar results.

disparities and vice versa.

I summarize these group-level outcomes in Panel B, where I take the cross-trooper average within each cluster. Across these groups there are clear differences in average trooper behavior. The dark gray diamonds correspond to the set of troopers who exhibit little cross-task effort distortions and produce effectively zero enforcement disparities, in addition to being among the most productive troopers on a per-shift basis. In contrast, troopers belonging to the light gray x's group are very sensitive to balancing effort across responsibilities, and as a result, predominately focus on enforcement, implying they are more likely to delay non-contracted obligations. Finally, troopers in the medium blue triangles generate large enforcement disparities. Jointly, there may therefore be both efficiency (in the sense of increased enforcement to generate deterrence) and equity gains to assigning these high-performing troopers to high-risk, high-collision areas and better optimizing the match of trooper heterogeneity with location demands. I explore the potential of this alternative and other counterfactual policies below.

5.2 Trooper-Specific Posterior Estimates

Rather than reducing variance through dimensionality reduction, an alternative approach is to employ shrinkage estimators to recover biased, but lower variance estimates of trooper-specific effects. Conventional solutions to this noisy estimates problem use Empirical Bayes tools to shrink trooper-specific estimates towards the average under an auxiliary assumption that the prior distribution of trooper effects is normally distributed (e.g., Morris 1983).[59] However, it is easy to imagine scenarios where the prior distribution is in fact not normally distributed, leading to shrunken estimates that no longer correspond to posterior means.[60]

I relax this normality assumption by flexibly estimating the prior distribution using the deconvolution approach of Efron (2016). Following Kline, Rose, and Walters (2022), I assume that the prior distribution belongs to an exponential family and estimate this distribution using a fifth-order spline and penalized maximum likelihood, calibrated so that the mean and variance of the estimated prior distribution match the sample mean and finite-sample bias-corrected variance above. This approach relies on an inherent assumption that the observed trooper effects are independent of their standard errors. Appendix Table A6 tests whether this assumption is plausible and finds little evidence that it is consistently violated.

Appendix Figure A24 presents the empirical distribution, deconvolved prior, and posterior mean estimates of trooper-specific responses across a battery of measures. Across many of the outcomes, the prior distribution takes the form of a slightly skewed normal distribution. More important however, is that there is still substantial heterogeneity across trooper posterior responses,

[59] I shrink the estimates toward the overall cross-trooper mean. An alternative approach is to shrink trooper estimates toward the location-specific mean. This is an alternative modeling assumption, though I note that the overall mean is reasonable here since troopers experience responses across multiple locations and the responses themselves are quasi-experimental estimates which include location fixed effects, which account for time-invariant mean differences between areas.

[60] For prior work in economics which uses an assumption of a normally distributed prior, see Kane and Staiger (2008) and Abaluck et al. (2021).

whether measured by aggregate enforcement, non-enforcement actions, or enforcement disparities, consistent with the bias-corrected variance estimates in Appendix Table A5.

The vast majority of troopers have negative enforcement responses, consistent with multitasking responsibilities, although there are a handful of (perhaps highly risk-averse) troopers who are very highly responsive and even increase enforcement relative to their untreated counterfactual, as also highlighted in the clustering exercise above. Moreover, more than half of troopers have enforcement responses that disparately affect minority motorists. However, this relative symmetry of the trooper distribution stands in stark contrast to the sharp differences documented in prior empirical exercises. Taken together, these results support a mechanism where the incentive structure of missing or being behind target induces or exacerbates disparities in enforcement behavior, rather than solely being explained by a handful of "bad apple" officers (Goncalves and Mello 2021). The residual variation in responsiveness across troopers is then both a function of how salient the target is and how troopers respond to that state.[61]

6 Examining Alternative Personnel Policies

There is substantial heterogeneity across troopers, both in their overall traffic enforcement production and how they trade off effort across competing responsibilities, some of whom generate negative externalities. Alternative personnel policies to mitigate these undesired outcomes could target either this heterogeneity or location-specific demands through manipulating the enforcement target. In this section, I develop a method to optimally match the population of heterogeneous troopers to the varying demands across locations and benchmark the gains from this policy against alternatives which hold the existing match of troopers and locations fixed.

6.1 Matching Trooper Heterogeneity to Location Demands

Eliminating the targets or instituting a "first-best" policy of assigning individually optimal, non-distortionary targets to each trooper is infeasible without placing additional structure on trooper behavior in addition to being practically difficult due to union bargaining agreements. As a "second-best" alternative, a social planner may wish to reassign troopers across locations in a way that minimizes the distortion of effort between enforcement and non-enforcement actions. Alternatively stated, taking the policy environment as given, this solution attempts to harmonize the task-specific performance of the workforce with locations that place different degrees of emphasis on each task. From a theoretical standpoint, this optimal matching approach is a feasible implementation of the optimal task grouping and specialization result from Holmström and Milgrom (1991) and offers an alternative method to mitigating the welfare losses from incomplete contracts in multitask principal-agent settings. For example, assigning troopers with high levels of per-shift enforcement to locations with high enforcement targets may be desirable because it's easier for these troopers to reach the

[61] Appendix Figure A25 presents similar distributions for estimates of regression-adjusted trooper per-shift enforcement activity. I constrain the support of the prior distribution to be non-negative (Kline, Rose, and Walters 2022), which is motivated by natural restrictions on the support of the potential values of the outcome.

enforcement goals. Similarly, troopers who generate particularly large negative externalities may be best placed in areas with low traffic volumes and few collisions or low targets. Such reassignments are also attractive as it would ensure that high enforcement-producing troopers are placed in areas in need of high enforcement levels, maximizing potential deterrence effects.

I first assess whether the observed match of troopers and locations could be improved through reassignment. In Panel A of Figure 8 I plot the observed flow of trooper typologies to location typologies, identified using a similar k-means clustering approach.[62] There are notable areas of potential improvement. Nearly half of "high-disparity" troopers work in either high minority share or high target locations. Similarly, the vast majority of "high-enforcement" troopers work in locations where the demands are relatively balanced or there is a high volume of non-enforcement demands. Thus, there may be gains to improving the matches between typologies, such as assigning high-enforcement troopers to areas with high targets.

The empirical challenge is how to determine the optimal match of troopers to locations, accounting for the rich heterogeneity across troopers and locations, while simultaneously respecting the location-specific capacity constraints. I cast this problem through the lens of optimal transport theory, a mathematical subfield which has a rich history in economics and whose primal problem has a deep connection with a social planner's optimal allocation problem.[63] To simplify the exposition, I present an intuitive overview of the social planner's problem in the main text and provide a discussion of the underlying implementation in Appendix E.

Consider a social planner (the Patrol) who wants to optimally allocate a set of troopers to a set of locations. Let X denote trooper characteristics and Z denote location characteristics. The social planner's problem is to choose an optimal match of troopers and locations π^* which solves the following problem

$$\max_{\pi \in \Pi} \mathbb{E}_{\pi}[\Phi(X, Z)] \tag{10}$$

where $\Phi(X, Z)$ denotes the payoff from matching a trooper of characteristics X with a location of characteristics Z.[64] Solving for π^* yields the optimal match between troopers and locations, given the focal characteristics X and Z, where optimal is in the welfare-maximizing sense. For example, if troopers are solely characterized by their per-shift violation volume and locations are

[62] The location groupings are constructed for illustrative purposes. In the counterfactual exercises below I use the underlying location means as these are directly observed in the data and not subject to the same degree of estimation error as the quasi-experimental trooper estimates. Using the means has the added benefit of harmonizing the mechanics of the counterfactual exercises across permutations below. Appendix Figure A26 plots the same lower triangular matrix and mean cluster outcomes as Figure 7 using information for locations. For simplicity, I use location characteristics in 2016 and drop a small handful of specialized assignments.

[63] In fact, many classic results in economics are disguised optimal transport problems or implicitly use results from this literature, including optimal matching in marriage markets with transferable utility (Becker 1973), identification in models with nonadditive functions (Matzkin 2003), and the changes-in-changes estimator (Athey and Imbens 2006; Gunsilius, Rigollet, and Torous 2021). See Galichon (2016) for an overview of optimal transport applications in economics and Villani (2009) for a comprehensive overview.

[64] In computer science, this problem is often written as a cost-minimization problem. That is, minimize the cost of transforming the source distribution into the destination distribution. This alternative method of setting up the problem has a natural interpretation from the lens of the principal-agent model in Section 2.3 - minimizing the cost disparity of enforcement specialization.

solely characterized by their targets, an optimal matching matrix would assign the most productive troopers to locations with the highest targets.[65] After identifying the optimal matching matrix given the focal characteristics, the remaining empirical challenge is to predict the counterfactual change in outcomes of moving trooper i from assignment j to j^* - that is, quantifying the change of moving from $\hat{\pi}_{ij}$ (the observed empirical match) to π^*_{ij} (the optimal match).

I employ a data-driven approach to quantifying such changes, using information from other troopers in the counterfactual match to predict responses of the reassigned trooper. Consider a response $\theta_{ij^*}(y)$ of trooper i reassigned to j^*. Moving trooper i from j to j^* has three effects. First, holding fixed the target E_j and trooper violation production Y_i, there is a level difference in the responses across locations. This level difference may reflect differences in the motorist populations, enforcement opportunities and road patterns, or relationships with local Patrol leadership. Second, moving a trooper across locations induces a change in the relative saliency of the enforcement target $(E_{j^*} - E_j)$, holding fixed their enforcement volume. Put differently, if $E_{j^*} > E_j$, then the target saliency is heightened for the reassigned trooper. This component is a movement along the $Y_i - E_j$ gradient of trooper responses within the new location, j^*. The gradient and level difference above are identified from the observed behavior of the existing troopers in location j^*. Finally, given the transfer event-study results from Section 3, there are first-order asymmetric effects from increasing or decreasing the target on trooper violation volume, which influence the value of Y_i under the new assignment.

Formally, I model the within-location gradient across troopers using linear regression.[66] Specifically, I regress the observed outcomes of troopers in location j^* on their relative performance $Y_i - E_j$ and use both the constant and slope coefficient from this regression to predict the response of a trooper who is reassigned from j to j^*. Intuitively this method has a connection to a classic Oaxaca-Blinder decomposition, where there are both level and slope differences in the responses across locations. Appendix Figure E2 provides a visual illustration of this intuition and I provide additional details in Appendix E.

Table 2 presents the results of this counterfactual reassignment exercise using modally observed assignments from 2016. I characterize troopers using their regression-adjusted per-shift enforcement volume from Equation (9), their θ_i estimates for violations and the probability of having only observed enforcement activity, which capture both an intensive and extensive margin response, and the θ_i estimate of enforcement disparities. For locations, I similarly characterize them using their enforcement targets and per-shift demands of collisions, calls for service, and local population racial composition. I use these sets of inputs to define the payoff from matching trooper i with location j, respectively, the performance of the trooper with the demands of the location.

I begin by studying the potential gains from optimally matching troopers to locations when

[65] More generally, when the source and destination distributions are unidimensional, as is common in applied work, π^* can be solved in closed form as positive assortative matching (i.e., the sorting algorithm). In richer settings, solving for the optimal sparse matrix is possible using standard linear programming techniques.

[66] While more flexible functional forms such as local linear regression are possible, extrapolating outside of the observed support can quickly lead to unstable predictions.

quantifying troopers based on their clustered typologies, illustrating the resulting optimal match in Panel B of Figure 8. The figure highlights a number of changes to the match of troopers and locations, with high-enforcement troopers being assigned to high target locations and low-disparity troopers replacing high-disparity troopers in high minority share locations, among other such potential improvements.

I quantify the gains from this reassignment in Column 2. Consistent with the visual improvements in the sankey diagram, there are noticeable gains in observed outcomes, including reduced response disparities and greater ex-ante enforcement, which lowers collision volume, in expectation. This policy does come with a drawback of a mild increase in the probability of only conducting enforcement activity, which may further delay collision report filing, an undesired outcome. While using the clustered typologies avoids introducing bias, it may reduce the efficacy of the reassignment as troopers cannot be precisely matched due to the discretization. Accepting a small degree of bias in exchange for quantifying trooper-specific estimates with EB tools may be useful for improving the match between troopers and locations. Indeed, in Column 3 I find that there are measurable gains to permitting this bias, including a reduced emphasis on only conducting enforcement, and further improving on the outcomes over the cluster-based match.

The previous two exercises consider "global" reassignments, where the social planner is unrestricted in their ability to match troopers and locations across the state. While the Patrol ultimately reserves the ability to move troopers around as needed, such counterfactual policies place little weight on trooper location preferences, as troopers are required to live near their beats. Ignoring trooper location preferences implies that some troopers could be counterfactually required to move across the entire state, which may be undesirable from a trooper retention standpoint. In Column 4, I instead consider reassignments within each of the eight districts, implicitly incorporating trooper location preferences through their observed postings. While some of the gains are mitigated relative to the unconstrained reassignment, I continue to find benefits across the entire suite of outcomes.

I benchmark these results against two alternative personnel policies which instead alter location-specific targets, holding the observed match of troopers and locations fixed. Specifically, I examine the effect of uniformly raising or lowering the enforcement target. Consistent with increased saliency, uniformly raising the target heightens the degree to which troopers focus on solely enforcement activity, in addition to exacerbating observed response disparities (Column 5). However, this policy also increases ex-ante enforcement and subsequently lowers collision volume, highlighting the tension between these competing objectives. In contrast, lowering the target has exactly the opposite effects, reducing enforcement disparities and the cross-task distortion of effort, but raising collision volume through lower ex-ante deterrence (Column 6). However, neither of these policies achieves the full suite of outcome improvements as do the optimal reassignment approaches. Taken together, these results provide evidence that improving the match between trooper performance and location-specific task demands can be a fruitful approach to mitigating the welfare losses associated with incomplete contracts in multitask environments, a policy solution which may be applicable in many settings beyond the criminal justice system.

7 Discussion and Conclusion

Firms and organizations across the economy face fundamental agency issues throughout their day-to-day operations. When agents are multitaskers, designing an incentive scheme which appropriately encourages monitored effort while not undermining incentives for non-monitored effort on competing responsibilities is practically challenging. In this paper, I study the efficacy and efficiency implications of misaligned incentives in the high-stakes setting of policing and public safety. Even though their performance is evaluated across a broad range of objective and subjective criteria, I illustrate how troopers in my setting are strongly differentially responsive to salient, quantifiable expectations over a subset of their responsibilities, indicating a degree of misalignment in the existing incentive scheme.

Guided by a parsimonious principal-agent model, I derive sharp predictions about how the tightness and saliency of a traffic enforcement target distorts trooper behavior and devise empirical tests concerning the nature of effort and the substitution across competing tasks. These predictions are consistently borne out in the data. I show that target-constrained troopers devote greater effort to traffic enforcement responsibilities and substitute away from non-enforcement effort. However, this efficacy gain with respect to enforcement volume does not come for free. Constrained troopers also enforce minor violations, generate racial inequities in their compensating enforcement patterns, and delay completion of non-enforcement responsibilities. I also find suggestive evidence that this compensating enforcement is of lower quality and more likely to lead to charges which are dismissed or were not committed.

Using a variance decomposition exercise, I highlight the critical role that trooper heterogeneity plays in governing these observed response patterns, with some troopers generating larger socially negative outcomes even though other similarly constrained troopers do not respond so adversely. I leverage this rich heterogeneity to develop a tractable approach for conducting partial equilibrium counterfactual simulations based on optimally matching troopers and locations in a manner that closely aligns with a social planner's welfare maximization objective function. Compared to alternative policies which simply manipulate the saliency of the enforcement target, optimally matching troopers across locations can both reduce the degree of cross-task effort distortion and production of racial enforcement disparities in addition to reducing equilibrium collision volume, the primary goal of the Patrol. These exercises illustrate how better matching the abilities of the workforce to heterogeneous task demands is a feasible solution to reducing the welfare losses associated with incomplete contracts in multitask environments.

My findings highlight the real-world challenges in optimally designing incentives schemes and job responsibilities for multitasking agents, along with the consequences for social welfare when the incentive schemes are suboptimal. In the context of the criminal justice system, my results suggest that one approach to improving the function of the justice system may be through targeting the incentives of judicial agents. Increasingly, firms and agencies across the economy are deploying granular data analytics to optimize employee performance through targeted incentive schemes. However, the use of these tools invokes a delicate tradeoff between leveraging narrow, high-powered incentives which maximize productivity against the potential for reduced effort on less-emphasized tasks. Moving forward, correctly balancing these competing objectives will play a crucial role in maximizing social welfare across many consequential settings beyond criminal justice, including medicine, education, and government.

References

Abaluck, Jason, Mauricio Caceres Bravo, Peter Hull, and Amanda Starc. 2021. "Mortality Effects and Choice Across Private Health Insurance Plans." *Quarterly Journal of Economics*, 136(3): 1557–1610.

Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica*, 67(2): 251–333.

Akerlof, George A. 1978. "The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning." *American Economic Review*, 68(1): 8–19.

Alexander, Diane. 2020. "How Do Doctors Respond to Incentives? Unintended Consequences of Paying Doctors to Reduce Costs." *Journal of Political Economy*, 128(11): 4046–4096.

Angelova, Victoria, Will S. Dobbie, and Crystal Yang. 2023. "Algorithmic Recommendations and Human Discretion." *NBER Working Paper No. 31747.*

Antonovics, Kate and Brian G. Knight. 2009. "A New Look at Racial Profiling: Evidence from the Boston Police Department." *Review of Economics and Statistics*, 91(1): 163–177.

Applebaum, Steven H., Adam Marchionni, and Arturo Fernandez. 2008. "The multitasking paradox: perceptions, problems, and strategies." *Management Decision*, 46(9): 1313–1325.

Arnold, David, Will Dobbie, and Crystal Yang. 2018. "Racial Bias in Bail Decisions." *Quarterly Journal of Economics*, 133(4): 1885–1932.

Arnold, David, Will Dobbie, and Peter Hull. 2022. "Measuring Racial Discrimination in Bail Decisions." *American Economic Review*, 112(9): 2992–3038.

Asch, Beth J. 1990. "Do Incentives Matter? The Case of Navy Recruiters." *Industrial and Labor Relations Review*, 43(3): 89S–106S.

Athey, Susan and Guido W. Imbens. 2006. "Identification and Inference in Nonlinear Difference-in-Differences Models." *Econometrica*, 74(2): 431–497.

Ba, Bocar, Patrick Bayer, Nayoung Rim, Roman Rivera, and Modibo Sidibé. 2022. "Police Officer Assignment and Neighborhood Crim." *NBER Working Paper No. 29243.*

Baker, George P. 1992. "Incentive Contracts and Performance Measurement." *Journal of Political Economy*, 100(3): 598–614.

Baker, George, Robert Gibbons, and Kevin J. Murphy. 1994. "Subjective Performance Measures in Optimal Incentive Contracts." *Quarterly Journal of Economics*, 109(4): 1125–1156.

Baron, David P. 1972. "Incentive Contracts and Competitive Bidding." *The American Economic Review*, 62(3): 384–394.

Becker, Gary S. 1973. "A Theory of Marriage: Part I." *Journal of Political Economy*, 81(4): 813–846.

Blinder, Alan S. 1973. "Wage Discrimination: Reduced Form and Structural Estimates." *Journal of Human Resources*, 8(4): 436–455.

Bolton, Patrick and Mathias Dewatripont. 2005. *Contract Theory.* MIT Press.

Bonhomme, Stéphane, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad, and Bradley Stezler. 2023. "How Much Should We Trust Estimates of Firm Effects and Worker Sorting?" *Journal of Labor Economics*, 41(2): 291–322.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa. 2019. "A Distributional Framework for Matched Employer Employee Data." *Econometrica*, 87(3): 699–738.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. Forthcoming. "Revisiting Event Study Designs: Robust and Efficient Estimation." *Review of Economic Studies.*

Branson, Russ, Rick Braziel, Karen Coffee, Alyson Cummings, and Mary Fleckenstein. 2016. "Washington State Patrol Trooper Recruitment and Retention Study."

Buch, Jason and Joy Borkholder. 2020. "Report: Washington State Patrol Singles out Native american drivers." *The Spokesman-Review.*

Campbell, Romaine. 2023. "What Does Federal Oversight Do to Policing and Public Safety? Evidence from Seattle." *Unpublished Working Paper.*

Card, David, Jöerg Heining, and Patrick Kline. 2016. "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact for Firms on the Relative Pay of Women." *Quarterly Journal of Economics*, 131(2): 633–686.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma. 2018. "Manipulation Testing Based on Density Discontinuity." *The Stata Journal*, 18(1): 234–261.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma. 2020. "Simple Local Polynomial Density Estimators." *Journal of the American Statistical Association*, 115(531): 1449–1455.

Cengiz, Doruk, Arindrajit Dube, Attila Lundner, and Ben Zipperer. 2019. "The Effect of Minimum Wages on Low-Wage Jobs." *Quarterly Journal of Economics*, 134(3): 1405–1454.

Chalfin, Aaron and Felipe Goncalves. 2023. "Professional Motivations in the Public Sector: Evidence from Police Officers." *NBER Working Paper No. 31985.*

Chan, David C. 2018. "The Efficiency of Slacking Off: Evidence from the Emergency Department." *Econometrica*, 86(3): 997–1030.

Chiappori, Pierre-André and Bernard Salanié. 2003. "Testing Contract Theory: A Survey of Some Recent Work." , ed. Mathias Dewatripont, Lars P. Hansen and Stephen T. Turnovsky *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, 115–149. Cambridge University Press.

Chung, Doug J., Das Narayandas, and Dongkyu Chang. 2021. "The Effects of Quota Frequency: Sales Performance and Product Focus." *Management Science*, 67(4): 1993–2656.

Clemens, Jeffrey and Joshua D. Gottlieb. 2014. "Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?" *American Economic Review*, 104(4): 1320–1349.

Cone, David C., Lynne D. Richardson, Knox H. Todd, Joseph R. Betancourt, and Robert A. Lowe. 2003. "Health Care Disparities in Emergency Medicine." *Academic Emergency Medicine*, 10(11): 1176–1183.

Cox, James C., R. Mark Isacc, Paula-Ann Cech, and David Conn. 1996. "Moral Hazard and Adverse Selection in Procurement Contracting." *Games and Economic Behavior*, 17(2): 147–176.

Dee, Thomas, Will Dobbie, Brian Jacob, and Jonah Rockoff. 2019. "The Causes and Consequences of Test Score Manipulation: Evidence from the New York Regents Examinations." *American Economic Journal: Applied Economics*, 11(3): 282–423.

Deming, David and Lisa B. Kahn. 2018. "Skill Requirements across Firms and Labor Markets: Evidence from Job Postings for Professionals." *Journal of Labor Economics*, 36(S1): S337–S369.

De Philippis, Marta. 2021. "Multi-Task Agents and Incentives: The Case of Teaching and Research for University Professors." *Economic Journal*, 131(636): 1643–1681.

Dobbie, Will, Andrew Liberman, Daniel Paravisini, and Vikram Pathania. 2021. "Measuring Bias in Consumer Lending." *Review of Economic Studies*, 88(6): 2799–2832.

Edwards, Griffin and Stephen Rushin. 2023. "The Effect of Police Quota Laws." *Iowa Law Review*, 109(5): 2127–2184.

Efron, Bradley. 2016. "Empirical Bayes deconvolution estimates." *Biometrika*, 103(1): 1–20.

Ewing & Associates, René. 2004. "Review of Performance and Outcome Measures of the Washington State Patrol." *Washington State Legislative Transportaton Committee: Transportation Performance Audit Board.*

Feigenberg, Benjamin and Conrad Miller. 2022. "Would Eliminating Racial Disparities in Motor Vehicle Searches Have Efficiency Costs?" *Quarterly Journal of Economics*, 137(1): 49–113.

Feigenberg, Benjamin and Conrad Miller. 2023. "Class Disparities and Discrimination in Traffic Stops and Searches." *Unpublished Working Paper.*

Fenizia, Alessandra. 2022. "Managers and Productivity in the Public Sector." *Econometrica*, 90(3): 1063–1084.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams. 2016. "Sources of Geographic Variation in Health Care: Evidence from Patient Migration." *Quarterly Journal of Economics*, 131(4): 1681–1726.

Finlay, Keith, Elizabeth Luh, and Michael Mueller-Smith. 2024. "Race and Ethnicity (Mis)measurement in the U.S. Criminal Justice System." *NBER Working Paper No. 32657.*

Galichon, Alfred. 2016. *Optimal Transport Methods in Economics.* Princeton University Press: Princeton, NJ.

Gardner, John. 2021. "Two-Stage Differences in Differences." *Unpublished Working Paper.*

Gelman, Andrew, Jeggry Fagan, and Alex Kiss. 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.

Gibbons, Robert. 1998. "Incentives in Organizations." *Journal of Economic Perspectives*, 12(4): 115–132.

Gittleman, Maury and Brooks Pierce. 2013. "How Prevalent is Performance-Related Pay in the United States? Current Incidence and Recent Trends." *National Institute Economic Review*, 226(1): R4–R16.

Goncalves, Felipe and Steve Mello. 2021. "A Few Bad Apples? Racial Bias in Policing." *American Economic Review*, 111(5): 1406–1441.

Goncalves, Felipe and Steve Mello. 2023. "Police Discretion and Public Safety." *NBER Working Paper No. 31678.*

Gunsilius, Florian, Philippe Rigollet, and William Torous. 2021. "An Optimal Transport Approach to Causal Inference." *Unpublished Working Paper.*

Harrington, Emma and Hannah Shaffer. 2022. "Brokers of Bias in the Criminal Justice System: Do Prosecutors Compound or Attenuate Racial Disparities Inherited from Arrest." *Unpublished Working Paper.*

Hart, Oliver and Bengt Holmström. 1987. "The Theory of Contracts." , ed. Truman Fassett Bewley *Advances in Economic Theory: Fifth World Congress*, 71–156. Cambridge University Press.

Holmström, Bengt. 1979. "Moral Hazard and Observability." *The Bell Journal of Economics*, 10(1): 74–91.

Holmström, Bengt. 2017. "Pay for Performance and Beyond." *American Economic Review*, 107(7): 1753–1777.

Holmström, Bengt and Paul Milgrom. 1987. "Aggregation and Linearity in the Provision of Intertemporal Incentives." *Econometrica*, 55(2): 303–328.

Holmström, Bengt and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, & Organization*, 7: 24–52.

Jackson, C. Kirabo. 2013. "Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence from Teachers." *The Review of Economics and Statistics*, 95(4): 1096–1116.

Jacob, Brian A. and Steven D. Levitt. 2003. "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating." *Quarterly Journal of Economics*, 118(3): 843–877.

Jones, Jesse. 2007. "Investigators: Are WSP Troopers Filling Ticket Quotas?" *King 5 News*.

Jordan, Andrew. 2024. "Racial Patterns in Approval of Felony Charges." *Unpublished Working Paper*.

Kane, Thomas J. and Doublas O. Staiger. 2008. "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation." *NBER Working Paper No. 14607*.

KC, Diwas Singh. 2014. "Does Multitasking Improve Performance? Evidence from the Emergency Department." *Manufacturing & Service Operations Management*, 16(2): 168–183.

Khan, Adnan Q., Asim I. Khwaja, and Benjamin A. Olken. 2014. "Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors." *Quarterly Journal of Economics*, 131(1): 219–271.

Kim, Minkyung, K. Sudhir, and Kosuke Uetake. 2022. "A Structural Model of a Multitasking Salesforce: Incentives, Private Information, and Job Design." *Management Science*, 68(6): 3975–4753.

Kline, Patrick, Evan K. Rose, and Christopher R. Walters. 2022. "Systemic Discrimination Among Large U.S. Employers." *Quarterly Journal of Economics*, 137(4): 1963–2036.

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten. 2020. "Leave-Out Estimation of Variance Components." *Econometrica*, 88(5): 1859–1898.

Knowles, John, Nicola Persico, and Petra Todd. 2001. "Racial Bias in Motor Vehicle Searches: Theory and Evidence." *Journal of Political Economy*, 109(1): 203–229.

Knutsson, Daniel and Björn Tyrefors. 2022. "The Quality and Efficiency of Public and Private Firms: Evidence from Ambulance Services." *Quarterly Journal of Economics*, 137(2): 2213–2262.

Krueger, Alan B. and Lawrence H. Summers. 1988. "Efficiency Wages and the Inter-Industry Wage Structure." *Econometrica*, 56(2): 259–293.

Kuhn, Peter J. and Lizi Yu. 2021. "Kinks as Goals: Accelerating Commissions and the Performance of Sales Teams." *NBER Working Paper No. 28487.*

Lazear, Edward P. 2000. "Performance Pay and Productivity." *American Economic Review*, 90(5): 1346–1361.

Lee, David S. and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*, 48(2): 281–355.

Lewis, Gregory and Patrick Bajari. 2014. "Moral Hazard, Incentive Contracts, and Risk: Evidence from Procurement." *Review of Economic Studies*, 81(3): 1201–1228.

Lovrich, Nicholas, Michael Gaffney, Clay Mosher, Mitchell Pickerill, and Michael R. Smith. 2003. *WSP Traffic Stop Data Analysis Project Report.* Pullman: Washington State University.

Luca, Dara Lee. 2015. "Do Traffic Tickets Reduce Motor Vehicle Accidents? Evidence from a Natural Experiment." *Journal of Policy Analysis and Management*, 34(1): 85–106.

Luh, Elizabeth, Benjamin Pyle, and James Reeves. 2024. "Agency Incentives and Disparate Revenue Collection: Evidence using Chicago Parking Tickets." *Unpublished Working Paper.*

MacDonald, Glenn and Leslie M. Marx. 2001. "Adverse Specialization." *Journal of Political Economy*, 109(4): 864–899.

Makowsky, Michael D. and Thomas Stratmann. 2009. "Political Economy at Any Speed: What Determines Traffic Citations?" *American Economic Review*, 99(1): 509–527.

Makowsky, Michael D. and Thomas Stratmann. 2011. "More Tickets, Fewer Accidents: How Cash-Strapped Towns Make for Safer Roads." *Journal of Law and Economics*, 54(4): 863–888.

Mangasarian, O.L. 1979. "Uniqueness of solution in linear programming." *Linear Algebra and its Applications*, 25: 151–162.

Manson, Steven, Jonathan Schroeder, David Van Riper, Katherine Knowles, Tracy Kugler, Finn Roberts, and Steven Ruggles. 2023. "IPUMS National Historical Geographic Information System: Version 18.0 [dataset]. Minneapolis, MN: IPUMS."

Mas, Alexandre. 2006. "Pay, Reference Points, and Police Performance." *Quarterly Journal of Economics*, 131(3): 783–821.

Matzkin, Rosa. 2003. "Nonparametric Estimation of Nonadditive Random Functions." *Econometrica*, 71(5): 1339–1375.

Mello, Steven. 2023. "Fines and Financial Wellbeing." *Unpublished Working Paper.*

Morris, Carl N. 1983. "Parametric Empirical Bayes Inference: Theory and Applications." *Journal of the American Statistical Association*, 78(381): 47–55.

Oaxaca, Ronald. 1973. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review*, 14(3): 693–709.

Pierson, Emma, Camelia Simoiu, Jan Overgoor, Sam Corbett-Davies, Daniel Jenson, Amy Shoemaker, Vignesh Ramachandran, Phoebe Barghouty, Cheryl Phillips, Ravi Shroff, and Sharad Goel. 2020. "A Large-Scale Analysis of Racial Disparities in Police Stops Across the United States." *Nature Human Behavior*, 4: 736–745.

Prendergast, Canice. 2007. "The Motivation and Bias of Bureaucrats." *American Economic Review*, 97: 180–196.

Rao, Nikhil and James Reeves. 2024. "Measuring Discrimination using Natural Experiments." *Unpublished Working Paper.*

Rehavi, M. Marit and Sonja B. Starr. 2014. "Racial Disparity in Federal Criminal Sentences." *Journal of Political Economy*, 122(6): 1320–1354.

Rubin, Donald B. 1981. "The Bayesian Bootstrap." *The Annals of Statistics*, 9(1): 130–134.

Shearer, Bruce S., Nibene Habib Somé, and Bernard Fortin. 2018. "Measuring Physicians' Response to Incentives: Evidence on Hours Worker and Multitasking." *IZA Discussion Paper No. 11565.*

Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter. 2019. "Firming Up Inequality." *Quarterly Journal of Economics*, 134(1): 1–50.

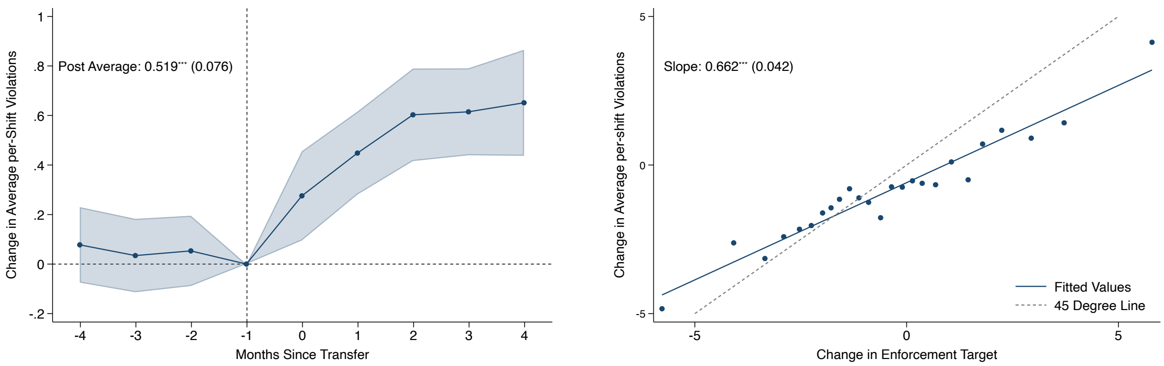
Villani, Cedric. 2009. *Optimal Transport: Old and New.* Springer-Verlag.

Washington State Patrol. 2019. "Washington State Patrol Strategic Plan 2019-2022."

Figure 1: Changes in Enforcement Targets Influence Trooper Enforcement Volume

Panel A: Dynamic Response Around Transfer

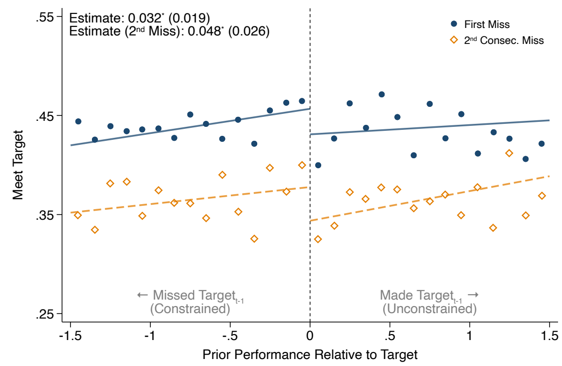
Panel B: Cross-Sectional Correlation



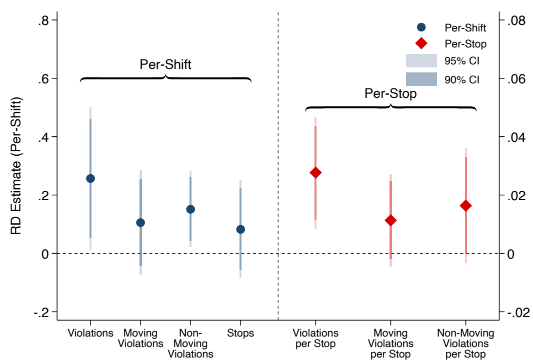
Notes: This figure reports estimates of changes in enforcement targets on trooper per-shift enforcement volume. Panel A reports the event-study estimates around the time of a transfer across locations with varying targets. Each point represents a coefficient of the interaction of event-time and a linear target change measure. Estimates are normalized to the month before the transfer. Regression also controls for trooper, month, year, and location origin and destination fixed effects. The reported average is the average of event-time coefficients 0 to 4. 95 percent confidence interval indicated by the shaded region based on standard errors clustered at the trooper level. Panel B reports the long-run cross-sectional correlation between changes in location targets and changes in per-shift enforcement. The dashed line is the 45 degree line and the solid line is the line of best fit.

Figure 2: Trooper Enforcement Responses to Missing Target

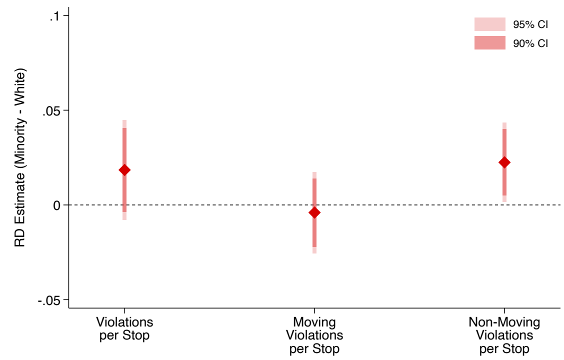
Panel A: Meet Target



Panel B: Changes in Enforcement Activity



Panel C: Changes in Stringency Disparities

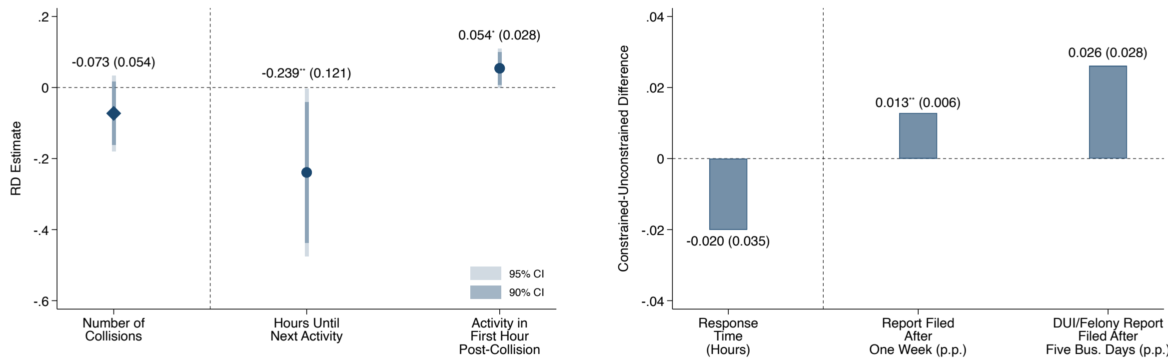


Notes: This figure reports regression discontinuity estimates of trooper enforcement effort. Panel A reports a regression discontinuity figure of the probability of meeting the enforcement target in time t as a function of the relative performance in time $t-1$. Navy circles indicate a miss in $t-1$, hollow orange diamonds indicate troopers on the margin of a potential second consecutive miss. Outcomes are residualized on location-by-month and location-by-year fixed effects. Panels B and C report regression discontinuity point estimates for changes in measured enforcement activity and disparities in per-stop stringencies. Navy circles indicate per-shift averages and red diamonds indicate per-stop averages. Regressions in Panels B and C control for location-by-month and location-by-year fixed effects. Lighter and darker shaded regions represent 95 and 90 percent confidence intervals with standard errors clustered at the trooper level.

Figure 3: Measuring Trooper Completion of Non-Enforcement Responsibilities

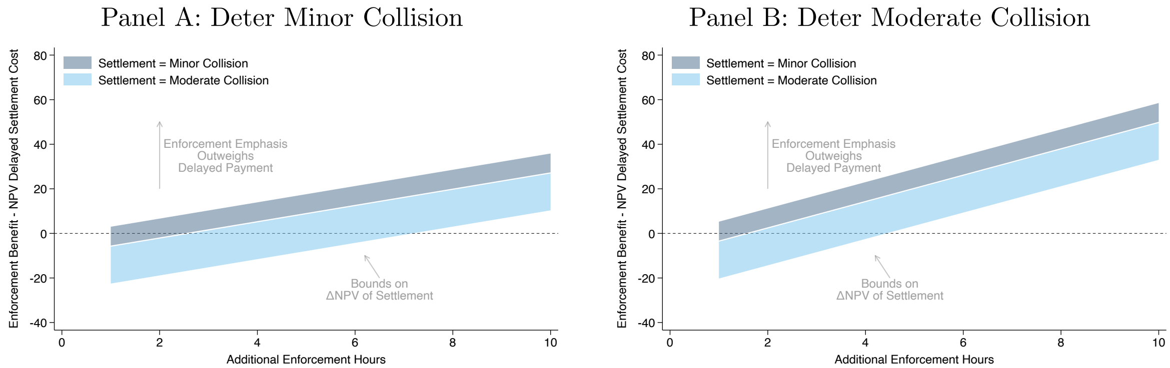
Panel A: Return to Non-Collision Effort

Panel B: Report Completion



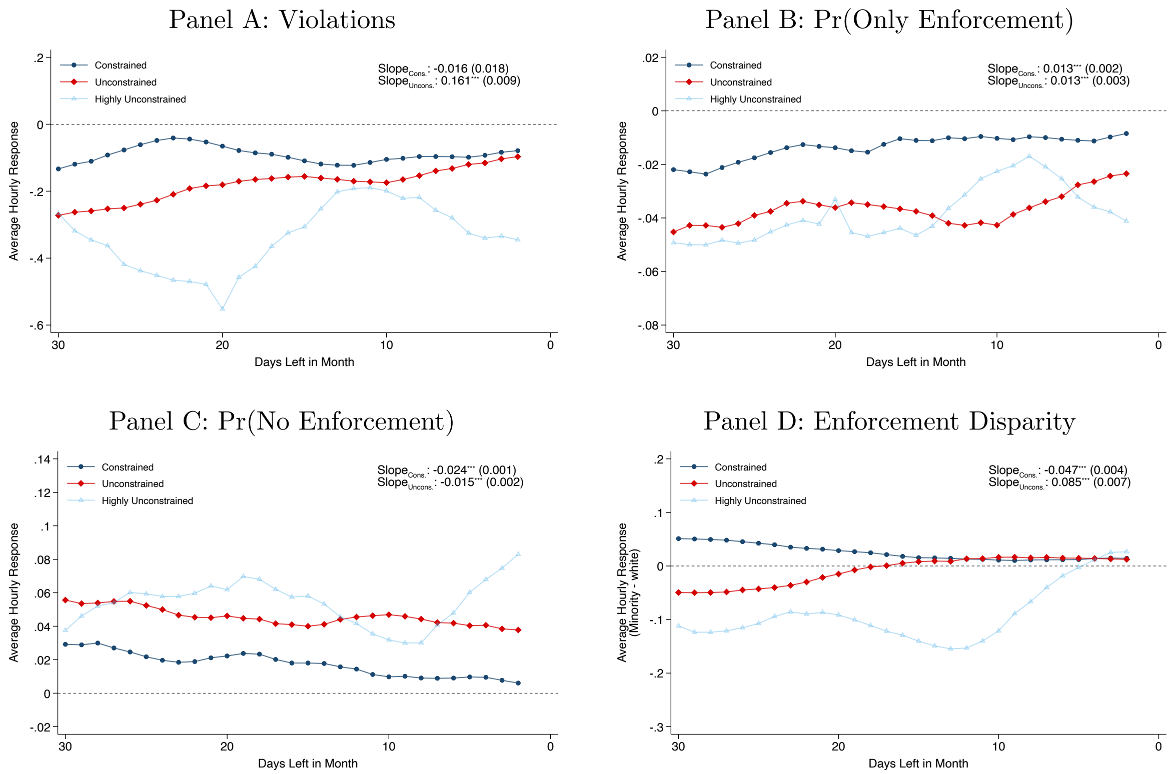
Notes: This figure reports measures of trooper effort substitution across enforcement and non-enforcement responsibilities, as measured using collision responses. Panel A reports regression discontinuity estimates. Each point represents a coefficient from a separate regression, comparing troopers who marginally missed (constrained) versus marginally made (unconstrained) the prior month's target. Regressions are weighted using a triangular kernel and control for location-month, location-year, trooper, hour of day, hour of shift, and district-date fixed effects. Vertical shaded regions denote 95 and 90 percent confidence intervals based on standard errors clustered at the trooper level. Panel B reports mean differences in constrained and unconstrained report-writing behavior using a sample of collision reports. Bars report the constrained-unconstrained difference, controlling for location-month, location-year, trooper, hour of day, hour of shift, and district-date fixed effects. DUI/Felony reports control only for county-month, county-year, and hour of day fixed effects. Standard errors clustered at the trooper level are reported in parentheses. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.

Figure 4: Bounds on Change in Welfare from Delayed Collision Report



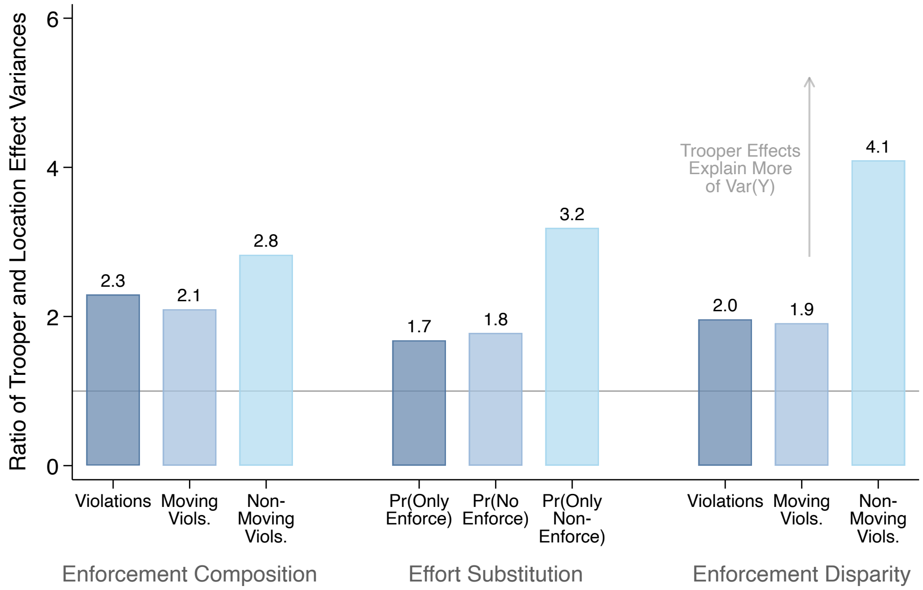
Notes: This figure reports bounds on the net benefit of a delayed collision report under various assumptions. In each panel, the x-axis is the number of hours of additional enforcement and the y-axis is the value of additional enforcement less the change in net present value from the delayed settlement. Darker blue regions assume the settlement is for a "minor" [$1,090, $15,000] collision and lighter blue regions assume the collision is for a "moderate" [$15,850, $43,063] collision. Panel A assumes the value of the deterred collision is "minor" and Panel B assumes the value of the deterred collision is "moderate." See Appendix D for additional details.

Figure 5: The Evolution of Trooper Multitasking Responses Over Course of the Month



Notes: This figure reports day-by-day estimates of trooper multitasking collision responses based on how many days are left in the month. Each point represents the hourly activity difference-in-differences estimate for events which occurred with the number of days left in the month listed on the x-axis, combined with estimates from nearby days using an Epanechnikov kernel. The navy dots represent estimates for constrained troopers, the red diamonds represent estimates for unconstrained troopers, and the light blue hollow triangles represent estimates for highly unconstrained troopers, or troopers who are far above the discontinuity. Panel A reports results for total violations, Panel B reports results for only having enforcement activity, Panel C reports results for having no enforcement activity, and Panel D reports results for the minority-white enforcement disparity. The reported slope coefficients are the cross-day gradient, rescaled so that the coefficient estimate represents the effect of going from 30 days to 0 days remaining. See Appendix C for additional details on the estimation strategy. Heteroskedastic-robust standard errors are reported in parentheses. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.

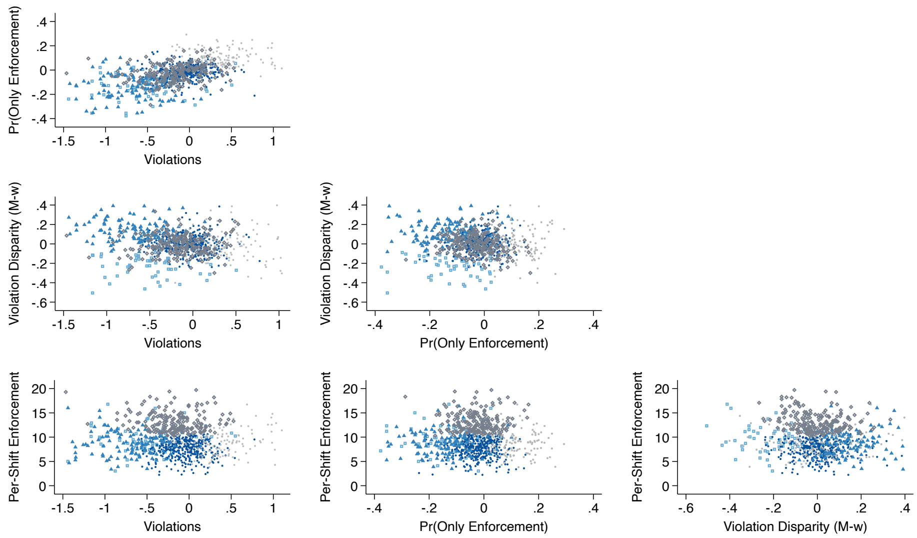
Figure 6: Trooper Effects Explain More Variation in Responses than Location Effects



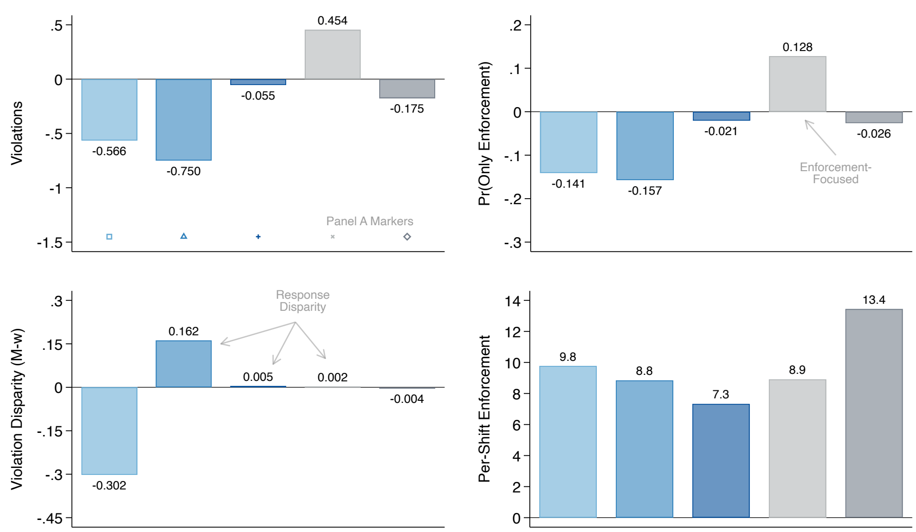
Notes: This figure reports the ratio of the variance of trooper effects and the variance of location effects in determining the distribution of multitasking collision responses. Trooper and location effects are estimated using Equation (7) after partialling out hour of event indicators in a first-step. Unbiased estimates of the effect variances are calculated following the procedure in Kline, Saggio, and Sølvsten (2020). Values greater than one indicate the trooper effects explain a relatively greater share of the variance in the outcome compared to location effects. The outcome and outcome category are listed on the x-axis. The first three bars measure the relative contribution of trooper effects to enforcement activity. The second set of bars measure trooper effort substitution across enforcement and non-enforcement tasks. The third set of bars measure the relative contribution of trooper effects to enforcement disparities.

Figure 7: Identifying Trooper Typologies Using k-means Clustering

Panel A: Matrix of Trooper Means and Identified Clusters



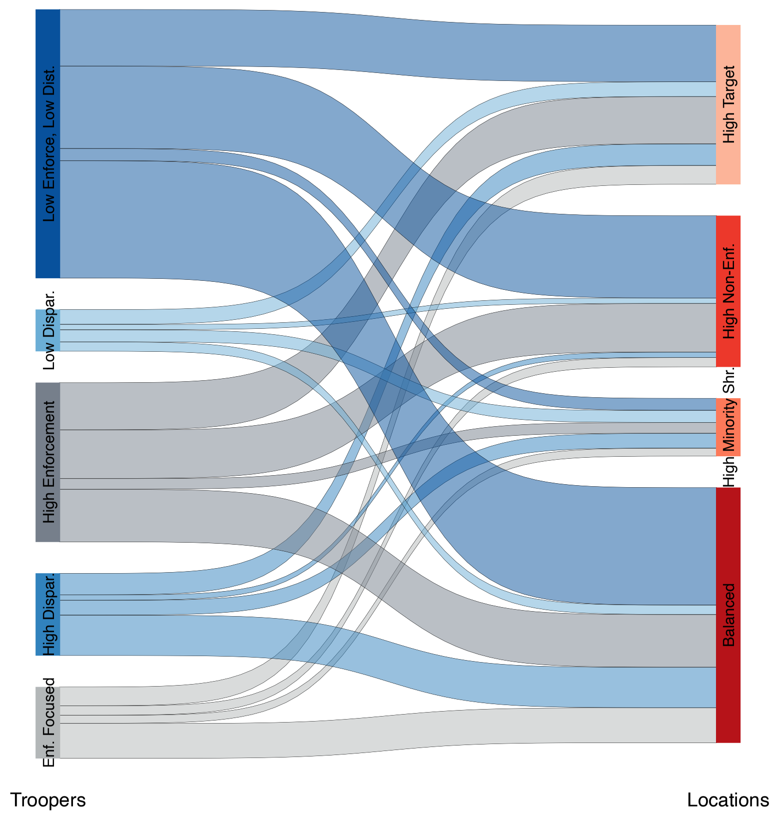
Panel B: Within-Cluster Averages



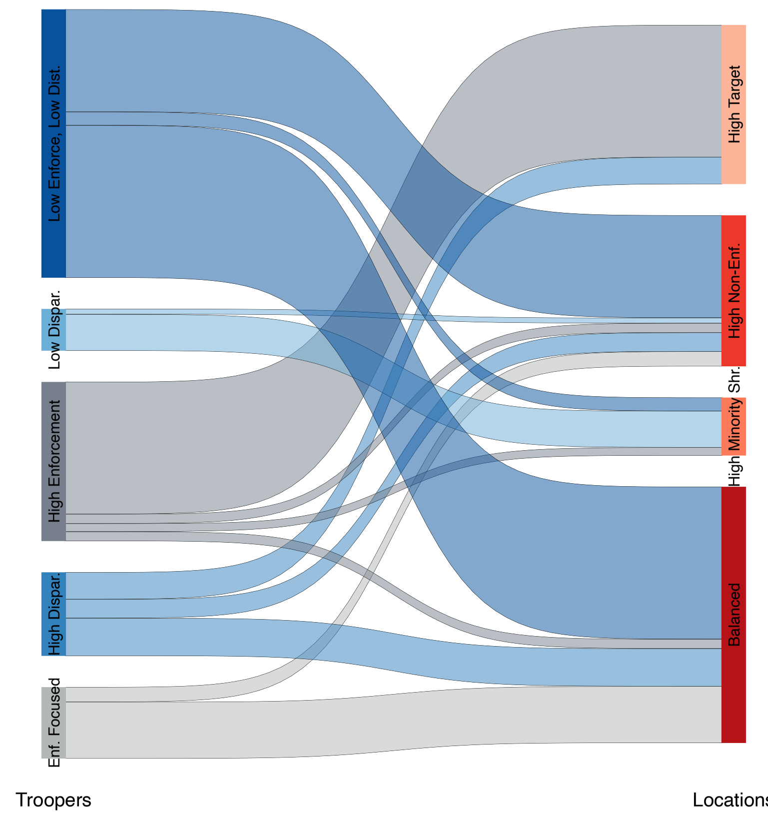
Notes: This figure reports the joint distribution of trooper-specific mean estimates of multitasking collision responses and regression-adjusted per-shift mean enforcement. Multitasking collision responses are measured as the average hourly activity difference, estimated using the approach described in Appendix C. Panel A reports a lower triangular matrix of two-way marginal distributions of the listed outcomes from the broader four-dimensional underlying distribution. Points are classified into specific clusters using a k-means classification algorithm. For exposition, visual outcomes exclude outcomes outside the 1^{th} and 99^{th} percentiles, though all data is used in the clustering stage. See Appendix E for additional details. Panel B reports the within-cluster means across the four input outcomes.

Figure 8: Empirical and Optimal Match of Trooper and Location Typologies

Panel A: Empirical Match



Panel B: Optimal Match



Notes: This figure illustrates the empirical (observed) and optimal match of trooper and location typologies using data from 2016. Trooper and location typologies are classified using k-means clustering using the inputs described in the main text. The optimal match is calculated using the optimal transport approach with troopers classified using the mean value of their typologies.

Table 1: Descriptive Statistics

	All Troopers	RD Sample
Panel A: Trooper Characteristics	(1)	(2)
Male	0.92	0.92
White	0.88	0.88
Black	0.03	0.03
Hispanic	0.05	0.05
Asian/Pacific Islander or Native American/Alaskan Native	0.04	0.04
Panel B: Average Per-Shift Enforcement		
Violations	9.11	8.53
Moving Violations	5.99	5.67
Non-Moving Violations	3.12	2.87
Violations (White Motorists)	6.88	6.46
Violations (Minority Motorists)	2.21	2.05
Searches	0.14	0.14
Calls for Service	1.68	1.72
Self-Initiated Physical Assist	0.87	0.87
Any Collision Response	0.29	0.30
Number of Trooper-Shifts	713,200	202,857

Notes: This table reports descriptive statistics. Panel A reports means at the trooper-level and Panel B reports means at the trooper-shift level. Violations by motorist race/ethnicity exclude a small handful of violations with missing motorist race. Column 1 reports means across all troopers and Column 2 reports means for troopers in the RD sample as described in Sections 3 and 4.

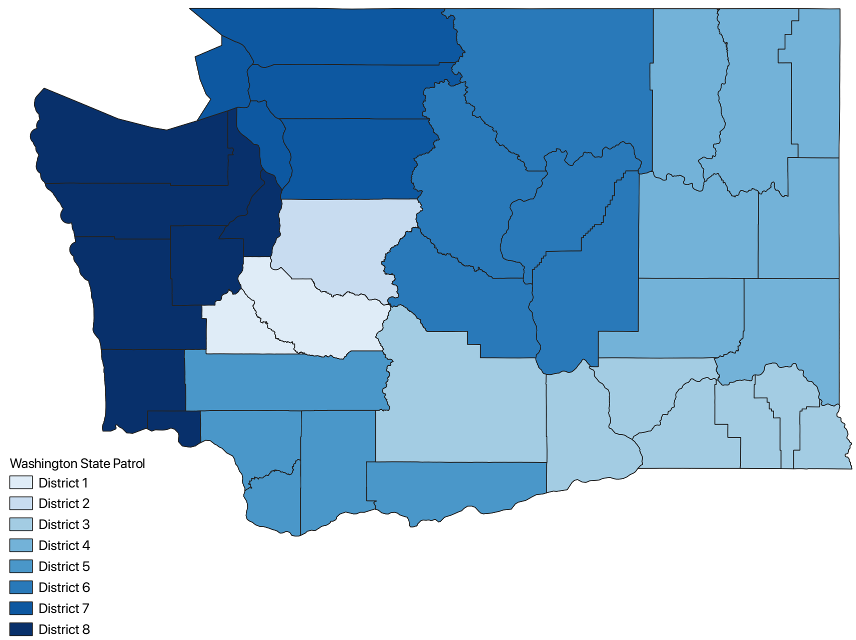
Table 2: Counterfactual Outcomes from Alternative Personnel Policies

		Counterfactual Percent Change in Outcome				
	Average Value	Optimal Match (Cluster)	Optimal Match (EB)	Geographic Constrained (EB)	Raise E_j by 1	Lower E_j by 1
Panel A: Trooper Responses	(1)	(2)	(3)	(4)	(5)	(6)
Violations	−0.142	−1.1%	−12.3%	−7.9%	1.6%	−0.3%
Pr(Only Enforcement)	−0.028	1.2%	−6.2%	−3.2%	1.0%	−1.0%
Violation Disparity	0.002	−51.1%	−217.1%	−290.1%	38.3%	−72.2%
Panel B: Location Outcomes						
Total Violations	7,527	5.2%	6.1%	5.5%	9.1%	−3.7%
Collisions	480	−7.2%	−4.8%	−4.1%	−10.3%	4.2%

Notes: This table reports the counterfactual outcomes associated with alternative enforcement policies using observed assignments in 2016. Panel A reports the average of trooper multitasking collision responses and Panel B reports the average location outcome. Column 1 reports the average observed value, Columns 2-6 report the percent change in the outcome under the observed policy listed in the column title. Column 2 employs the optimal transport approach described in the main text using identified trooper clusters in the matching and prediction stages. Column 3 combines the optimal transport approach with Empirical Bayes posteriors in both stages. Column 4 combines the optimal transport approach with Empirical Bayes posteriors in both stages with an additional district-level geographic restriction. Column 5 uniformly increases the observed target by one in all locations, Column 6 uniformly lowers the observed target by one in all locations. See Appendix E for additional details.

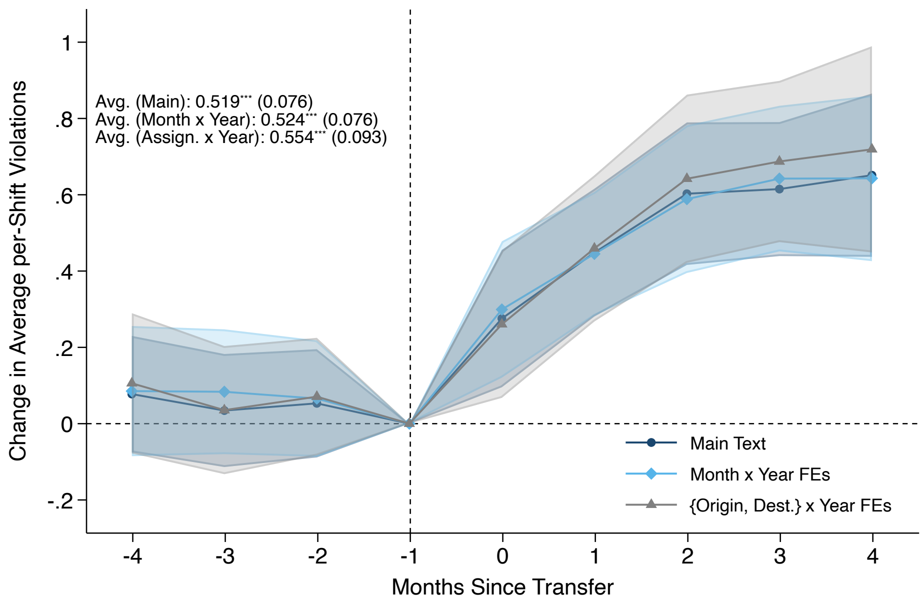
Appendix A: Additional Results

Appendix Figure A1: Map of Washington State Patrol Districts



Notes: This figure reports the spatial distribution of districts in the Washington State Patrol. Black outlines denote counties.

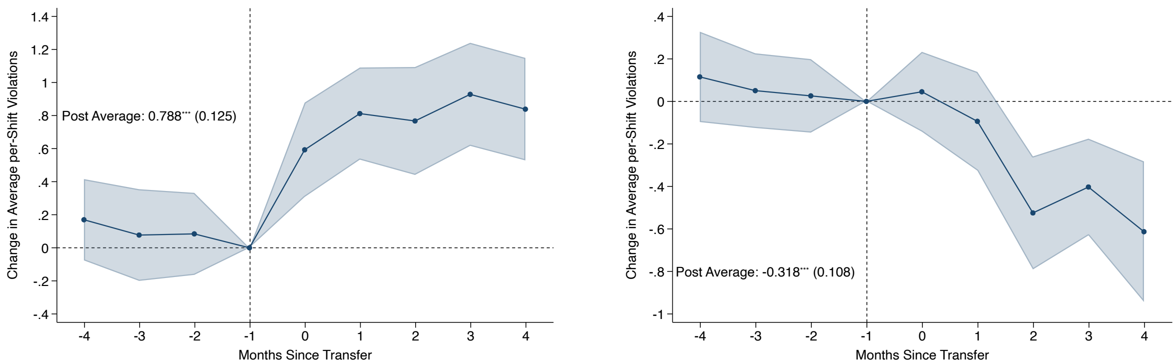
Appendix Figure A2: Alternative Specifications of Transfers Event Study



Notes: This figure reports estimates of changes in enforcement targets on trooper per-shift enforcement volume under alternative specifications. Each point represents a coefficient of the interaction of event-time and a linear target change measure. The navy dots report the estimate from the main text which controls for trooper, origin, destination, month, and year fixed effects. The light blue diamonds replace month and year fixed effects with month-by-year fixed effects and the gray triangles replace the origin, destination, and year fixed effects from the base specification with origin x year and destination x year fixed effects. Estimates are normalized to the month before the transfer. The reported average is the average of event-time coefficients 0 to 4. 95 percent confidence intervals indicated by the shaded regions based on standard errors clustered at the trooper level.

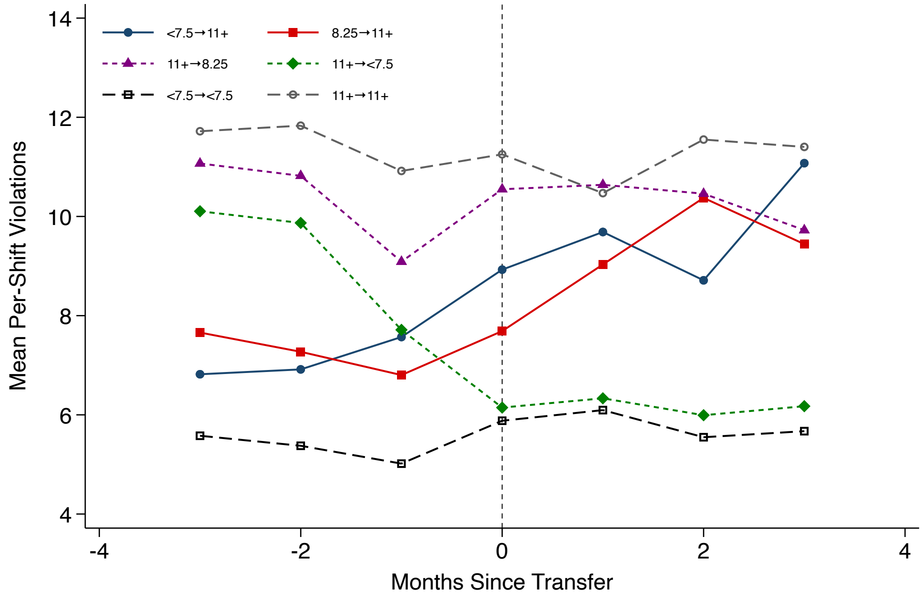
Appendix Figure A3: Asymmetries in Transferring Trooper Enforcement Responses

Panel A: Higher Target | Panel B: Lower Target



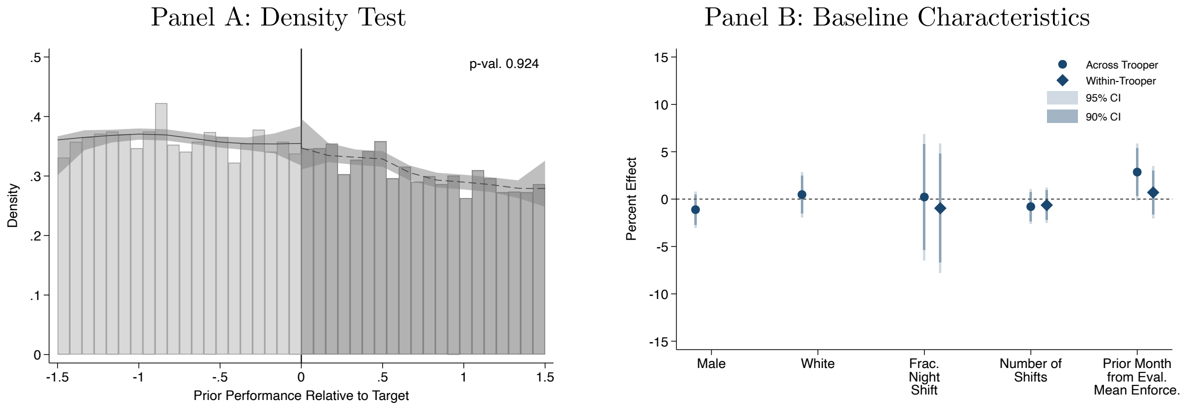
Notes: This figure reports estimates of changes in enforcement targets on trooper per-shift enforcement volume. Panel A reports the event-study estimates around the time of a transfer across locations with transfers to units with higher targets and Panel B reports similar event-study estimates for transfers to units with lower targets. Each point represents a coefficient of the interaction of event-time and a linear target change measure. Estimates are normalized to the month before the transfer. Regression also controls for trooper, month, year, and location origin and destination fixed effects. The reported average is the average of event-time coefficients 0 to 4. 95 percent confidence interval indicated by the shaded region based on standard errors clustered at the trooper level.

Appendix Figure A4: Mean Per-Shift Enforcement for Movers by Origin and Destination Target



Notes: This figure reports average per-shift enforcement around the months of a location transfer. Each point and line represents the average per-shift enforcement of troopers who transferred from (j) and to (j') a location with the listed target. The set of cutoffs $\{7.5, 8.15, 9, 10+\}$ roughly corresponds to quartiles of the enforcement target distribution. Dashed lines denote transfers within the same general bin, solid lines denote transfers where the target increased, and dotted lines denote transfers where an target decreased. Data are residualized on month and year fixed effects.

Appendix Figure A5: Testing Experimental Validity of Performance Regression Discontinuity Design

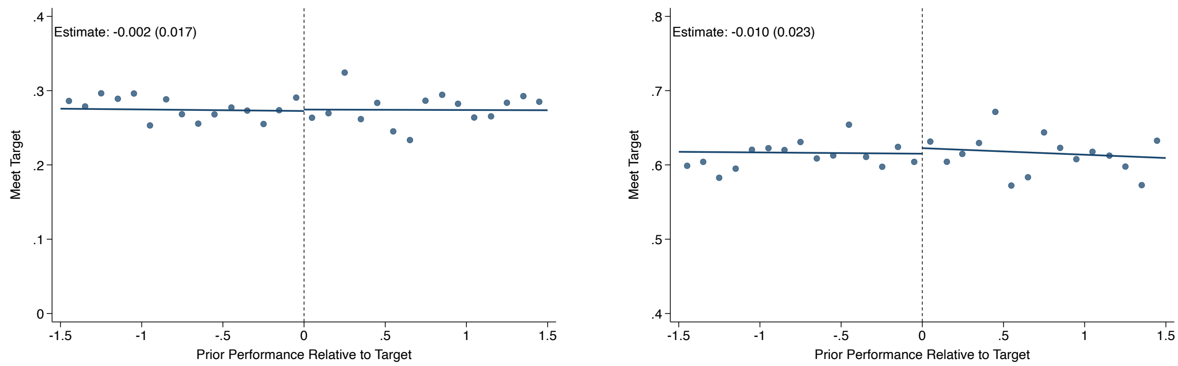


Notes: This figure reports tests of the experimental validity of the regression discontinuity design. Panel A reports the distributional density of the running variable. Local linear estimates, confidence bands, and p-value computed following Cattaneo, Jansson, and Ma (2018; 2020). Panel B reports regression discontinuity estimates using the listed characteristic as the outcome. Dots denote estimates that use variation across troopers and diamonds denote estimates that use within-trooper variation. All estimates are re-scaled to represent percent effects relative to the outcome mean. All regressions control for location-by-month- and location-by-year fixed effects and are weighted using a triangular kernel. Lighter and darker shaded regions represent 95 and 90 percent confidence intervals, with standard errors clustered at the trooper level.

Appendix Figure A6: Placebo Estimates of Trooper Responses to Missing Enforcement Targets

Panel A: Lower Cutoff

Panel B: Higher Cutoff

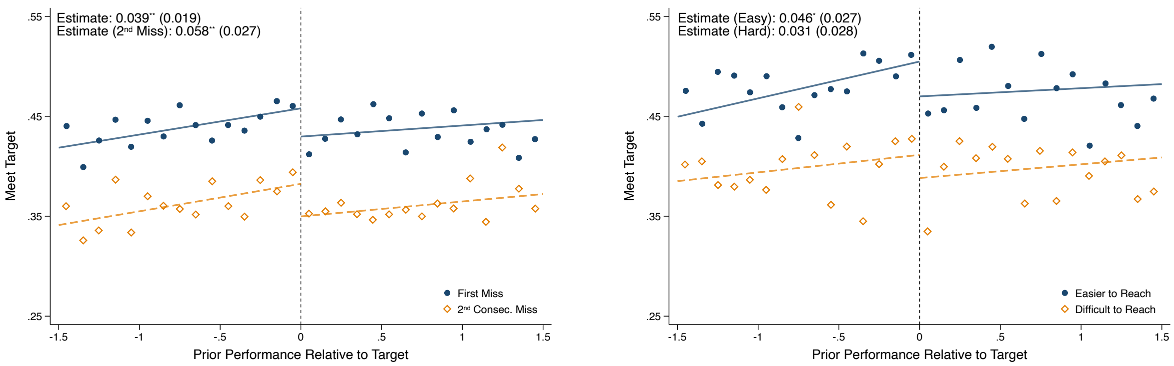


Notes: This figure reports placebo regression discontinuity estimates of the probability of meeting the enforcement target. Each panel uses a different placebo cutoff and estimates the probability of meeting the target using this alternative placebo cutoff. All regressions control for location-by-month and location-by-year fixed effects and are weighted using a triangular kernel. Points represent residualized means that residualize out the fixed effects and an aggregate linear trend. Linear lines of best fit denoted by solid lines. Standard errors clustered at the trooper level are reported in parentheses. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.

Appendix Figure A7: Within-Trooper and Time of Year Enforcement Responses to Missing Target

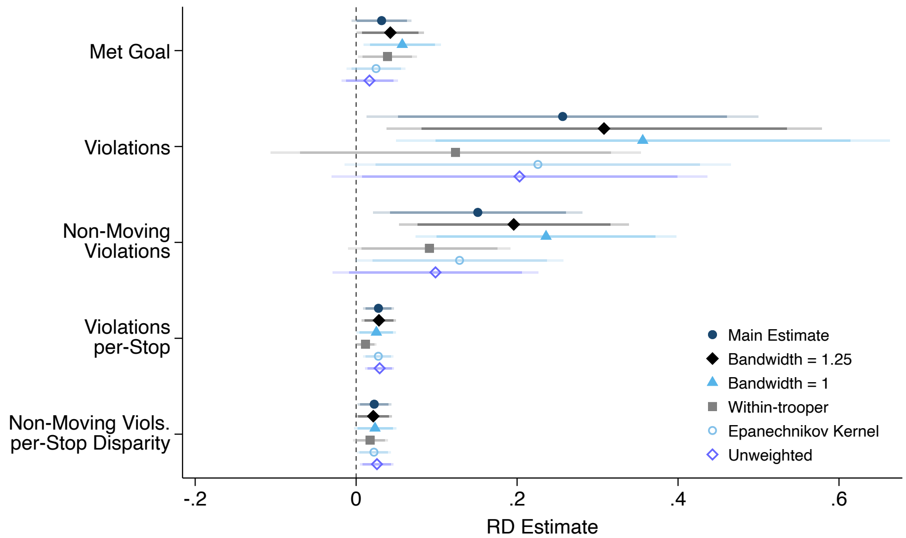
Panel A: Within-Trooper

Panel B: By Time of Year



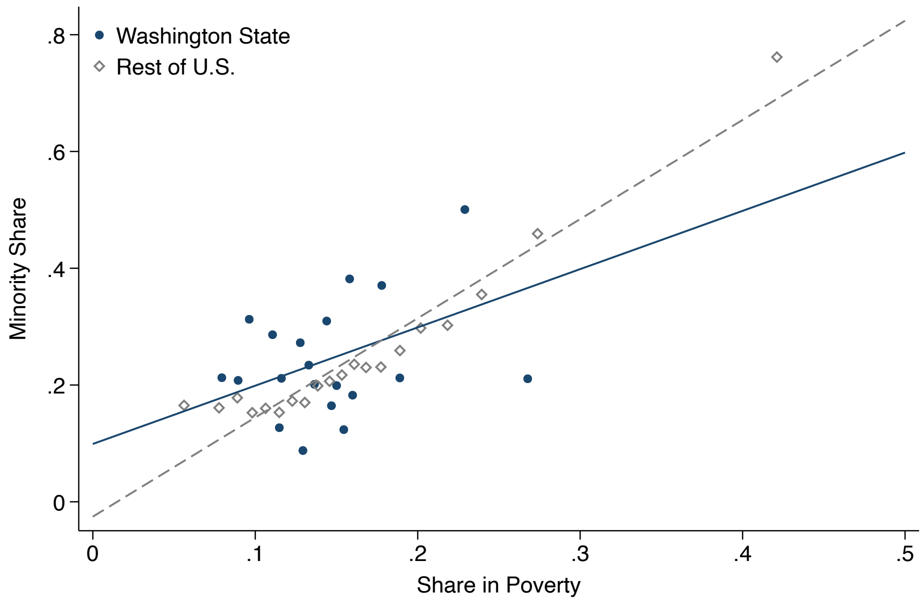
Notes: This figure reports regression discontinuity estimates of the probability of meeting the enforcement target. The x-axis is the prior month's relative performance. Panel A reports estimates which leverage within-trooper variation. Navy dots report within-bin means for a first miss and orange diamonds report within-bin means for troopers on the margin of a potential second consecutive miss. Bin means are residualized on location-month and location-year fixed effects, as well as an aggregate trend. Lines represent lines of best fit. Panel B reports cross-trooper estimates for months with higher and lower daily enforcement volume, on average. Formal point estimates are reported in the upper left part of the figure. Standard errors clustered at the trooper level are reported in parentheses. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.

Appendix Figure A8: Alternative Specifications of Regression Discontinuity Design



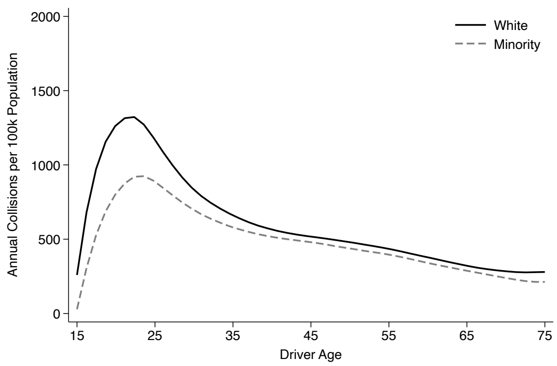
Notes: This figure reports robustness checks of the regression discontinuity design to alternative specifications. Each point represents an alternative specification, indicated in the legend. The outcome is listed on the y-axis. All regressions include location-month and location-year fixed effects. Lighter and darker shaded regions indicate 95 and 90 percent confidence intervals based on standard errors clustered at the trooper level.

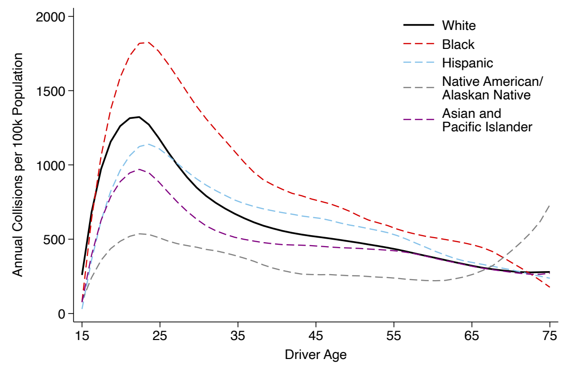
Appendix Figure A9: Correlation Between Minority and Poverty Share in United States and Washington State



Notes: This figure reports a binscatter of the correlation between the minority (all subpopulations but white, non-Hispanic) share of the population and the share of the population in poverty. Navy circles represent means for Washington State and gray diamonds represent means for the rest of the United States. The solid and dashed lines represent linear lines of best fit and are estimated on the underlying, non-binned data.

Appendix Figure A10: Motorist Risk Distributions Across Age Profile by Race/Ethnicity



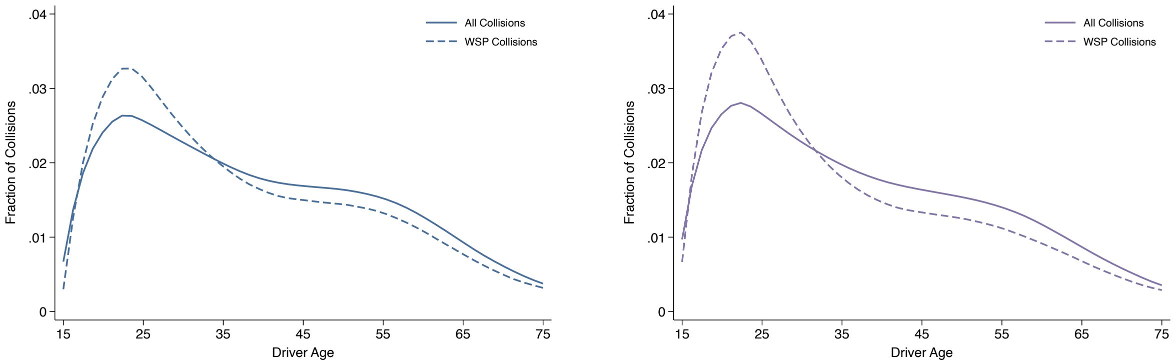


Notes: This figure plots smoothed annual collision rates over the age profile by motorist race and ethnicity. The sample includes collisions which were responded to by a Washington State Patrol trooper. Annual rates are smoothed using a local linear regression. Panel A reports collision rates for white and minority motorists. Panel B decomposes the collision rates for minority motorists into the component parts.

Appendix Figure A11: Comparing Motorist Risk Distributions Across Gender-Age Profile for State Patrol and Statewide Collisions

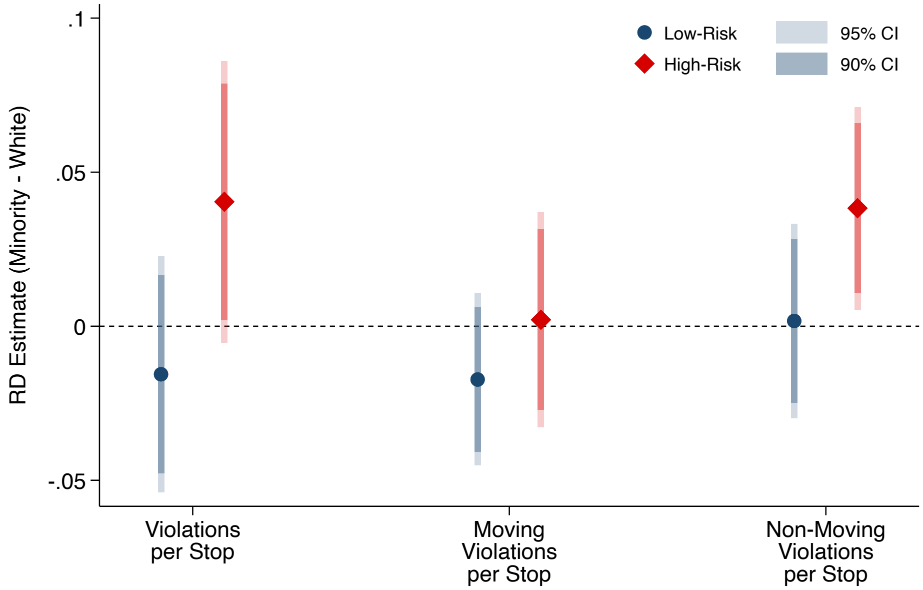
Panel A: Male Motorists

Panel B: Female Motorists



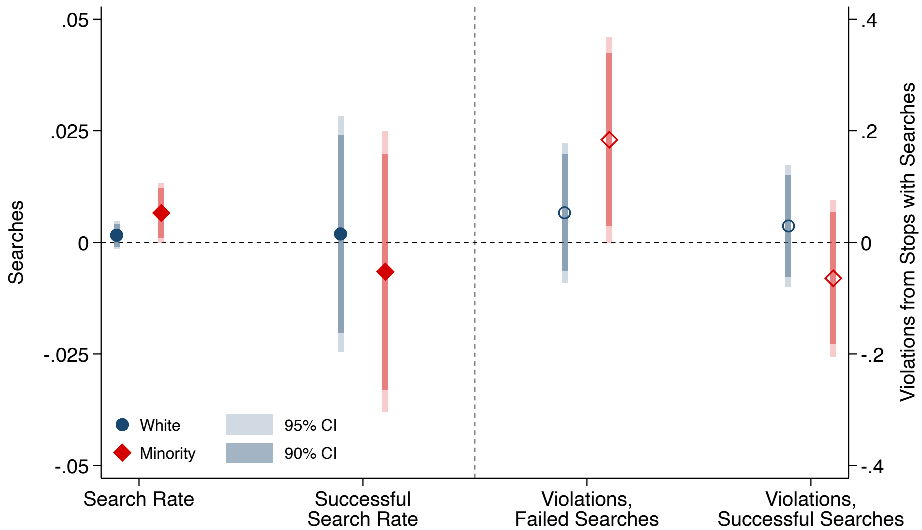
Notes: This figure plots the fraction of annual collisions across the age profile by motorist gender and data source. Collision fractions by age are smoothed using a local linear regression. The solid line represents the annual collision rate using data from the Washington State Department of Transportation, which includes all responding agencies. The dashed line represents the annual collision rate using data from the Washington State Patrol. Panel A reports collision rates for male motorists and Panel B reports collision rates for female motorists.

Appendix Figure A12: Stringency by Segment of Risk Distribution



Notes: This figure reports regression discontinuity estimates of differential stringency for different outcomes and segments of the risk distribution. Risk is calculated by predicting population collision rates using race by gender fixed effects and a quadratic in age. Low-risk motorists are defined as the bottom four quintiles and high-risk motorists are defined as the top quintile. Each coefficient represents the minority-white gap for that particular outcome, among that risk cell. Navy circles report estimates for low-risk motorists and red diamonds report estimates for high-risk motorists. All regressions are weighted using a triangular kernel and include location-month location-year fixed effects. Lighter and darker shaded regions indicate 95 and 90 percent confidence intervals based on standard errors clustered at the trooper level.

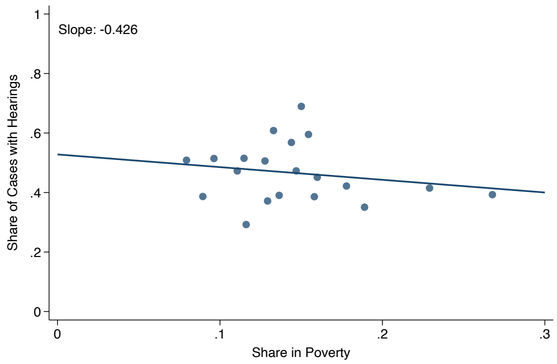
Appendix Figure A13: Trooper Search Responses to Missing Target



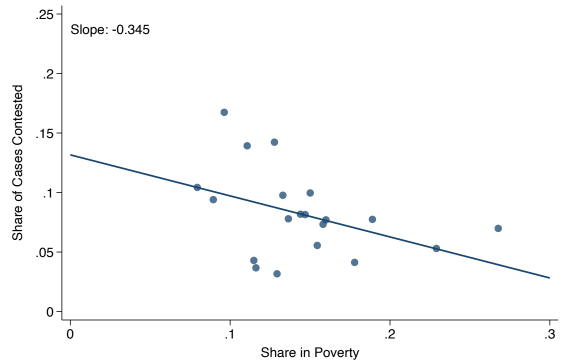
Notes: This figure reports regression discontinuity estimates of trooper average per-shift search behavior. Each coefficient is the effect of missing the enforcement target in the previous evaluation period. Navy dots report estimates for white motorists and red diamonds report outcomes for minority motorists. The first outcome is the search rate on a per-stop basis, the second outcome is the "hit rate", and the last two outcomes are the number of violations during failed or successful searches on a per search basis. Denominators are race-specific. All regressions are weighted using a triangular kernel and include location-month and location-year fixed effects. Lighter and darker shaded regions indicate 95 and 90 percent confidence intervals based on standard errors clustered at the trooper level.

Appendix Figure A14: Correlations of Court Outcomes for Traffic Infractions and County Demographics

Panel A: Share Cases with Hearing

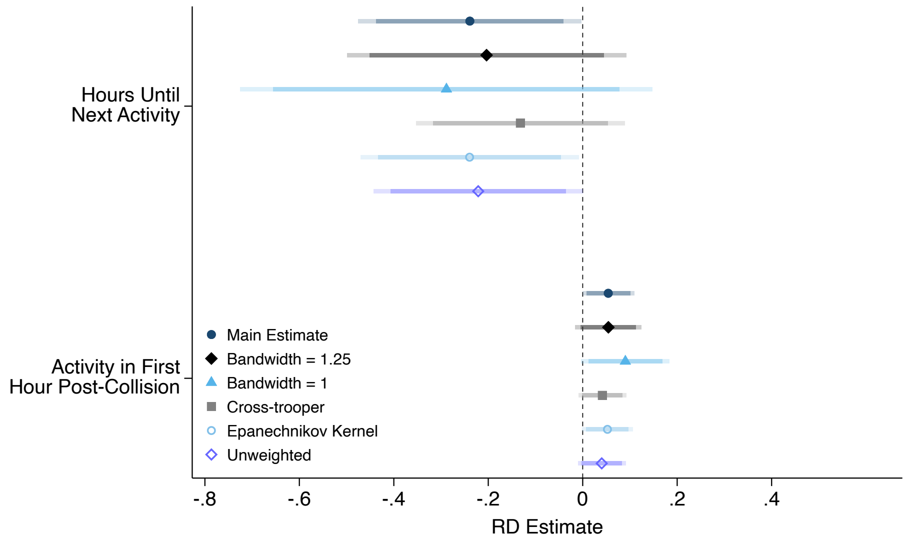


Panel B: Share Cases Contested



Notes: This figure reports the correlation between caseload outcomes for traffic violations and county-level characteristics. Each panel lists the variable on the y-axis and is measured using 2010-2018 caseload reports from the Washington State Administrative Office of the Courts. County share in poverty is measured using the 2014-2018 American Community Survey from Manson et al. (2023).

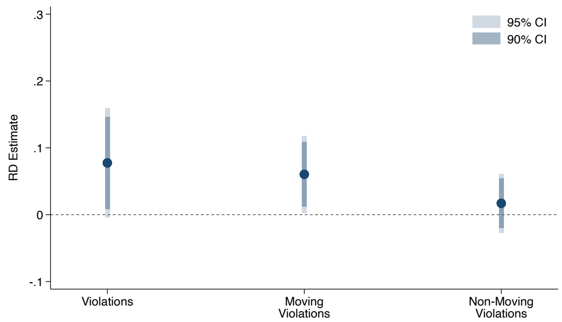
Appendix Figure A15: Robustness Checks of Trooper Completion of Non-Enforcement Responsibilities



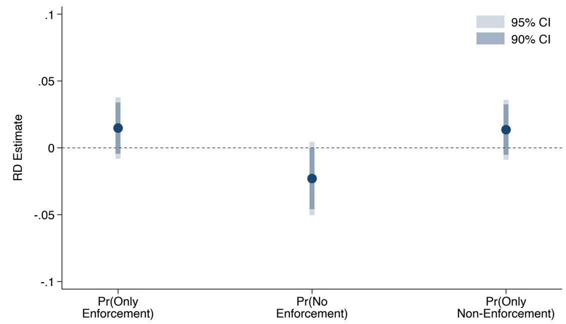
Notes: This figure presents robustness checks for measuring trooper effort substitution across enforcement and non-enforcement responsibilities, as measured using collision responses. Each point estimate represents a separate regression. All regressions include trooper, location-month, location-year, district-date, hour of shift, and hour of day fixed effects. Specification permutations are listed. Lighter and darker shaded regions correspond to 95 and 90 percent confidence intervals based on standard errors clustered at the trooper level.

Appendix Figure A16: Trooper Enforcement Activity Changes Following Collision Response

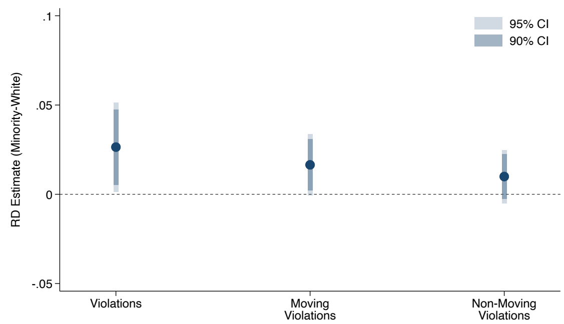
Panel A: Changes in Enforcement Activity



Panel B: Effort Substitution

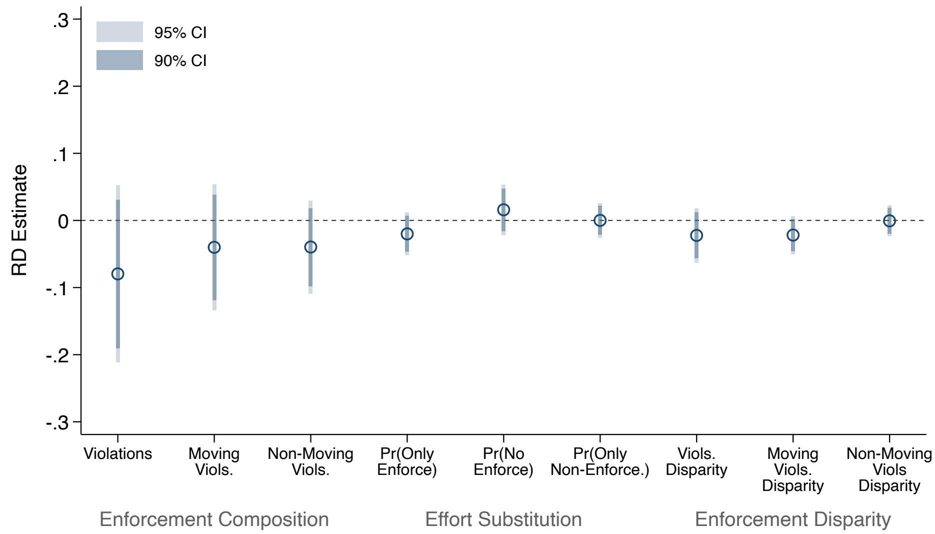


Panel C: Changes in Enforcement Disparities



Notes: This figure reports estimates of changes in enforcement behavior troopers following an assigned collision response. All outcomes have been normalized to the hourly level over the remainder of the focal shift. Regression discontinuity estimates control for location-month, location-year, trooper, district-date, hour of day, and hour of shift fixed effects. Panel A reports changes in enforcement activity, Panel B reports estimates of the substitution between enforcement and non-enforcement responsibilities, and Panel C reports estimates of enforcement disparities after adjusting for differences in population size. Lighter and darker shaded regions represent 95 and 90 percent confidence intervals.

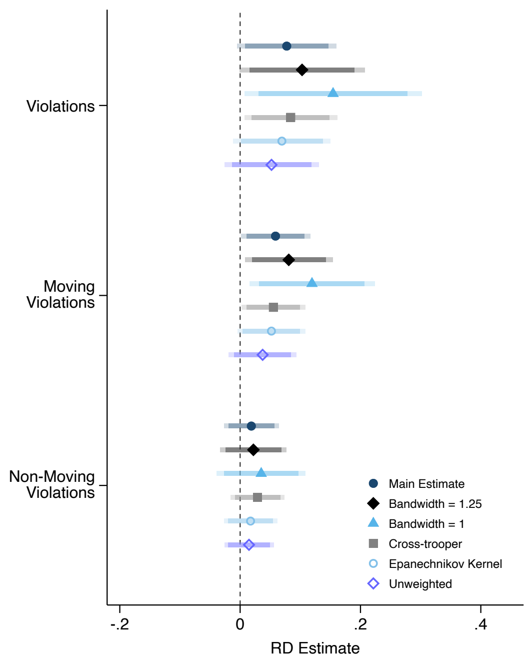
Appendix Figure A17: Pre-Collision Response Balance in Enforcement and Non-Enforcement Activity



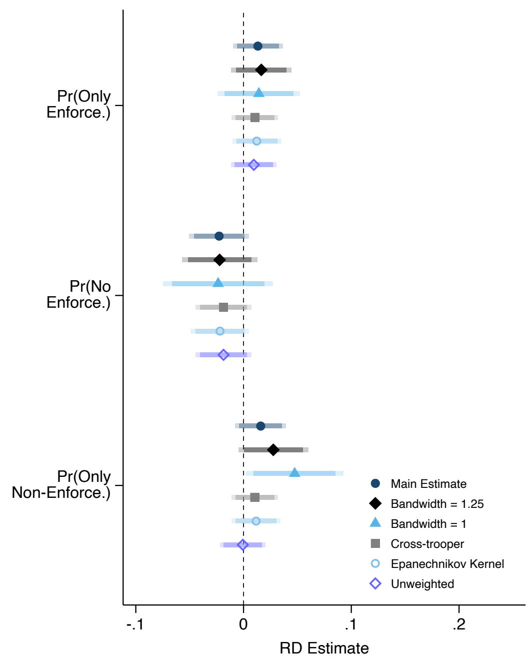
Notes: This figure presents regression discontinuity estimates of constrained and unconstrained troopers on pre-collision response hourly enforcement and non-enforcement activity. Each point estimate represents a separate regression. All regressions include trooper, location-month, location-year, district-date, hour of day, and hour of shift fixed effects and are weighted using a triangular kernel. Lighter and darker shaded regions correspond to 95 and 90 percent confidence intervals based on standard errors clustered at the trooper level.

Appendix Figure A18: Alternative Specifications of Collision Regression Discontinuity Design

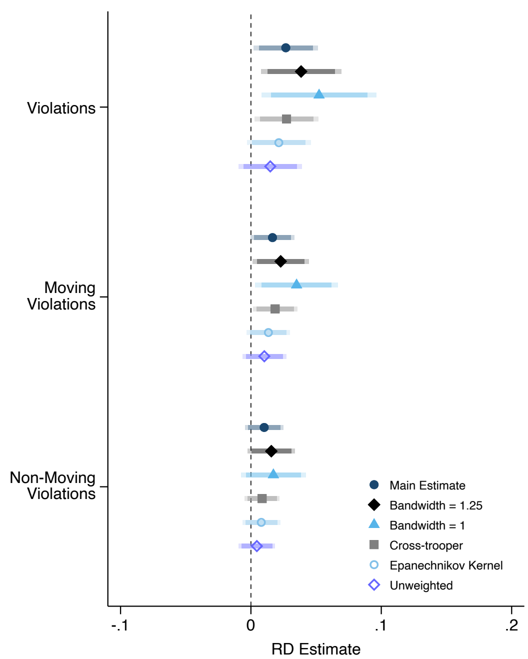
Panel A: Changes in Enforcement Activity



Panel B: Effort Substitution



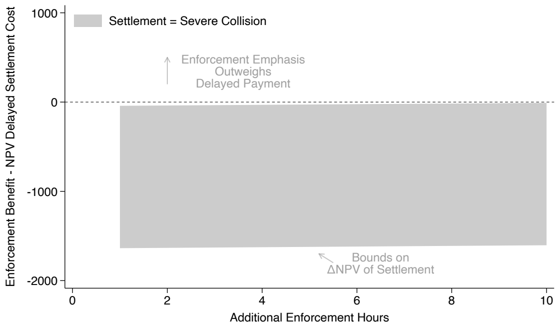
Panel C: Changes in Enforcement Disparities



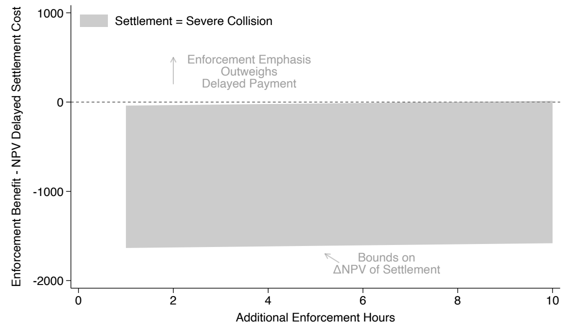
Notes: This figure reports robustness checks of the regression discontinuity design investigating changes to enforcement composition post collision response to alternative specifications. Each point represents an alternative specification, indicated in the legend, with the magnitude of the estimate representing the hourly effect. The outcome is listed on the y-axis. All regressions include trooper, location-month, location-year, district-date, hour of day, and hour of shift fixed effects, unless otherwise stated. Lighter and darker shaded regions indicate 95 and 90 percent confidence intervals based on standard errors clustered at the trooper level.

Appendix Figure A19: Bounds on Change in Welfare from Delayed Collision Report for Severe Collision Settlements

Panel A: Deter Minor Collision



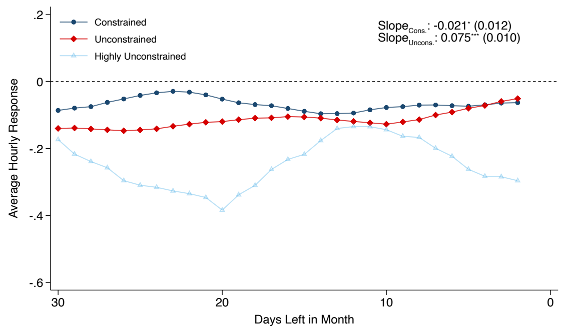
Panel B: Deter Moderate Collision



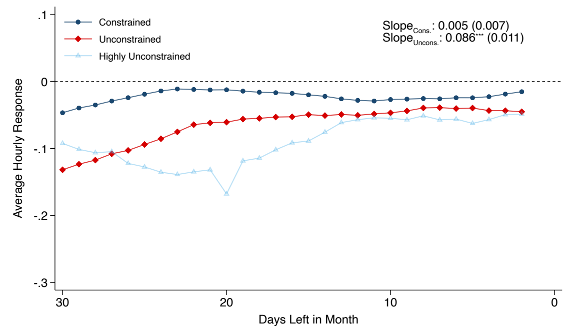
Notes: This figure reports bounds on the net benefit of a delayed collision report under various assumptions. In each panel, the x-axis is the number of hours of additional enforcement and the y-axis is the value of additional enforcement less the change in net present value from the delayed settlement. The gray region assumes the delayed settlement is for a "severe" [$489,000, $2.7 million] collision. Panel A assumes the value of the deterred collision is "minor" and Panel B assumes the value of the deterred collision is "moderate." See Appendix D for additional details.

Appendix Figure A20: Additional Trooper Multitasking Responses Over Course of Month

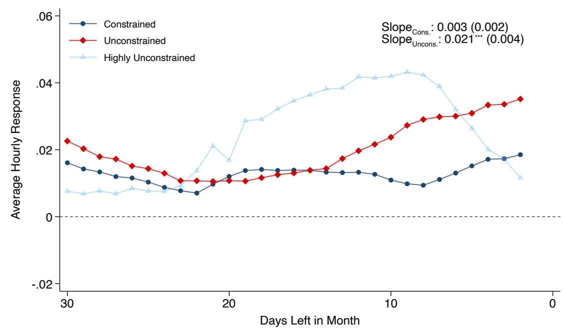
Panel A: Moving Violations



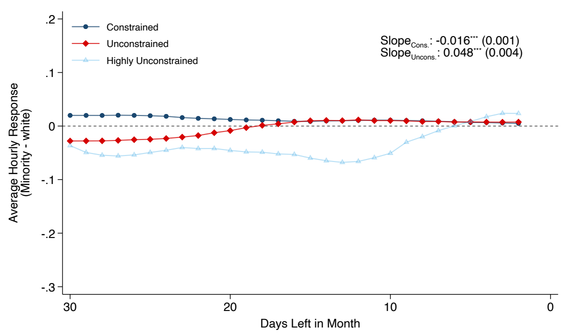
Panel B: Non-Moving Violations



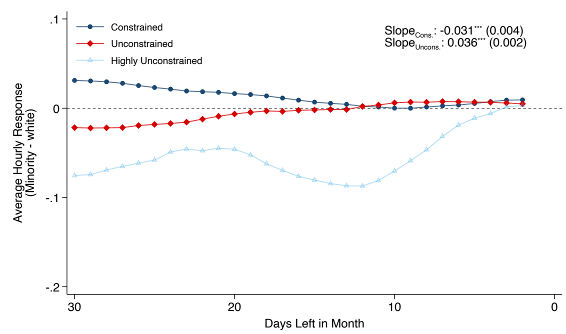
Panel C: Pr(Only Non-Enforcement)



Panel D: Moving Violation Disparity

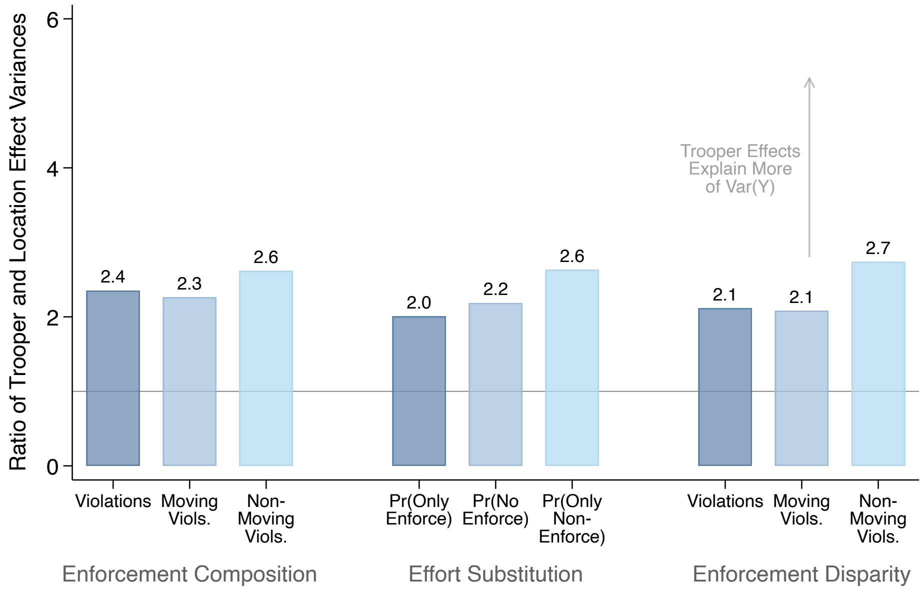


Panel E: Non-Moving Violation Disparity



Notes: This figure reports day-by-day estimates of trooper multitasking collision responses based on how many days are left in the month. Each point represents the hourly activity difference-in-differences estimate for events which occurred with the number of days left in the month listed on the x-axis, combined with estimates from nearby days using an Epanechnikov kernel. The blue dots represent estimates for constrained troopers, the red diamonds represent estimates for unconstrained troopers, and the light blue hollow triangles represent estimates for highly unconstrained troopers, or troopers who are far above the discontinuity. Panels A and B report results for moving and non-moving violations, Panel C reports results for only having non-enforcement activity, and Panels D and E report results for moving and non-moving violation disparities. The reported slope coefficients are the cross-day gradient, rescaled so that the coefficient estimate represents the effect of going from 30 days to 0 days remaining. See Appendix C for additional details on the estimation procedure. Heteroskedastic-robust standard errors are reported in parentheses. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.

Appendix Figure A21: Trooper Effects Explain More Variation in Responses than Location Effects - “Plug-in” Estimates

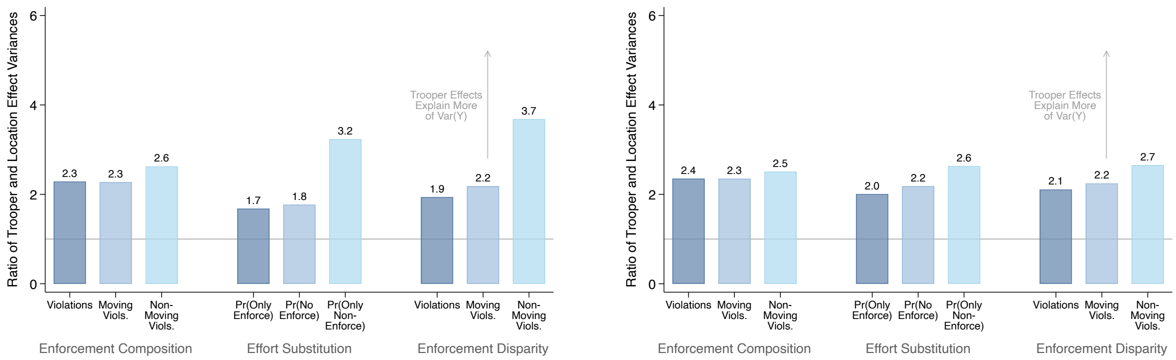


Notes: This figure reports the ratio of the variance of trooper effects and the variance of location effects in determining the distribution of multitasking collision responses. Trooper and location effects are estimated using Equation (7) after partialling out hour of shift indicators in a first-step. Estimates of the effect variances are calculated as standard empirical measures. Values greater than one indicate the trooper effects explain a relatively greater share of the variance in the outcome compared to location effects. The outcome and outcome category are listed on the x-axis. The first three bars measure the relative contribution of trooper effects to enforcement activity. The second set of bars measure trooper effort substitution across enforcement and non-enforcement tasks. The third set of bars measure the relative contribution of trooper effects to enforcement disparities.

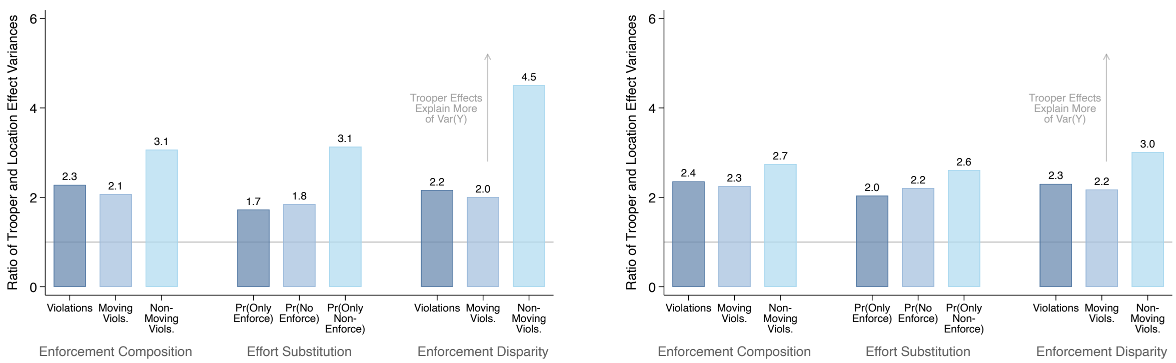
Appendix Figure A22: Alternative Specifications of Ratio Trooper and Location Effect Variances Across Different Outcomes

Bias-Corrected Estimate | "Plug-in" Estimate

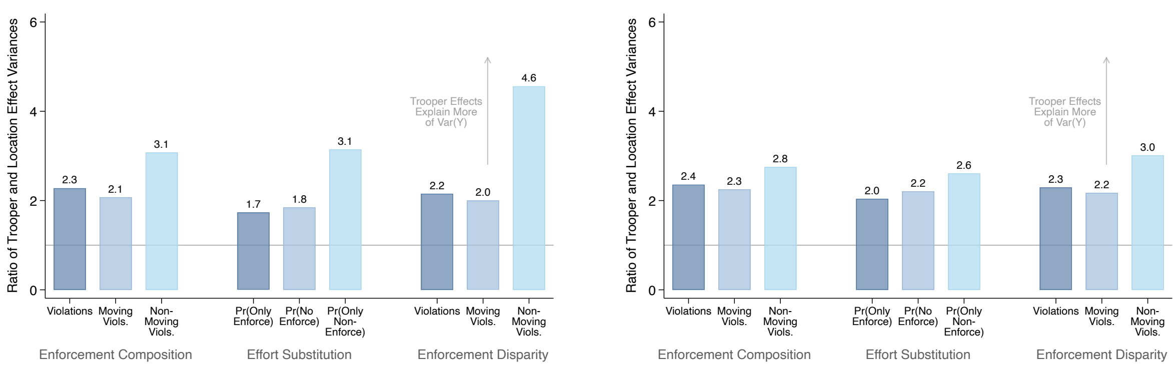
Panel A: No First-Step Adjustment



Panel B: Winsorize Responses

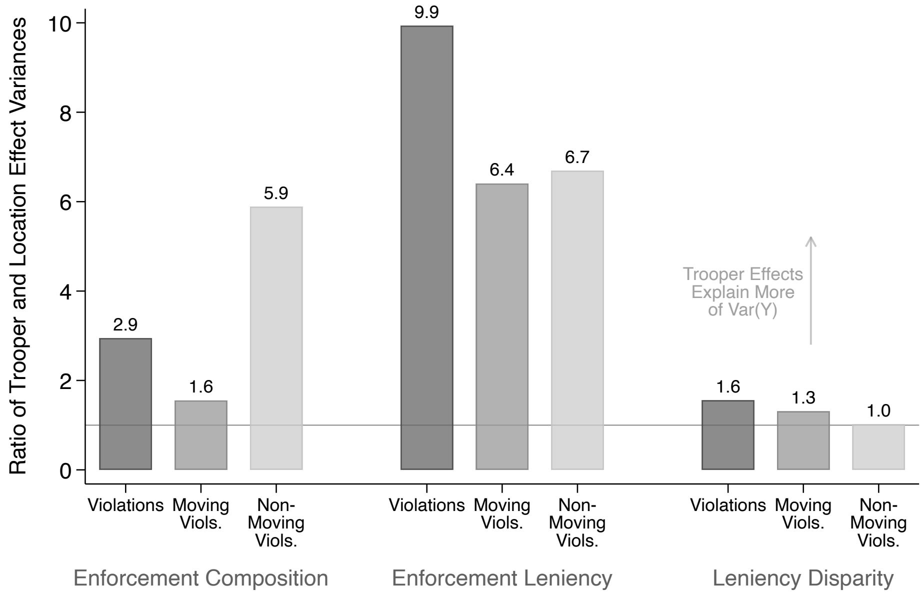


Panel C: Winsorize Responses and No First-Step Adjustment



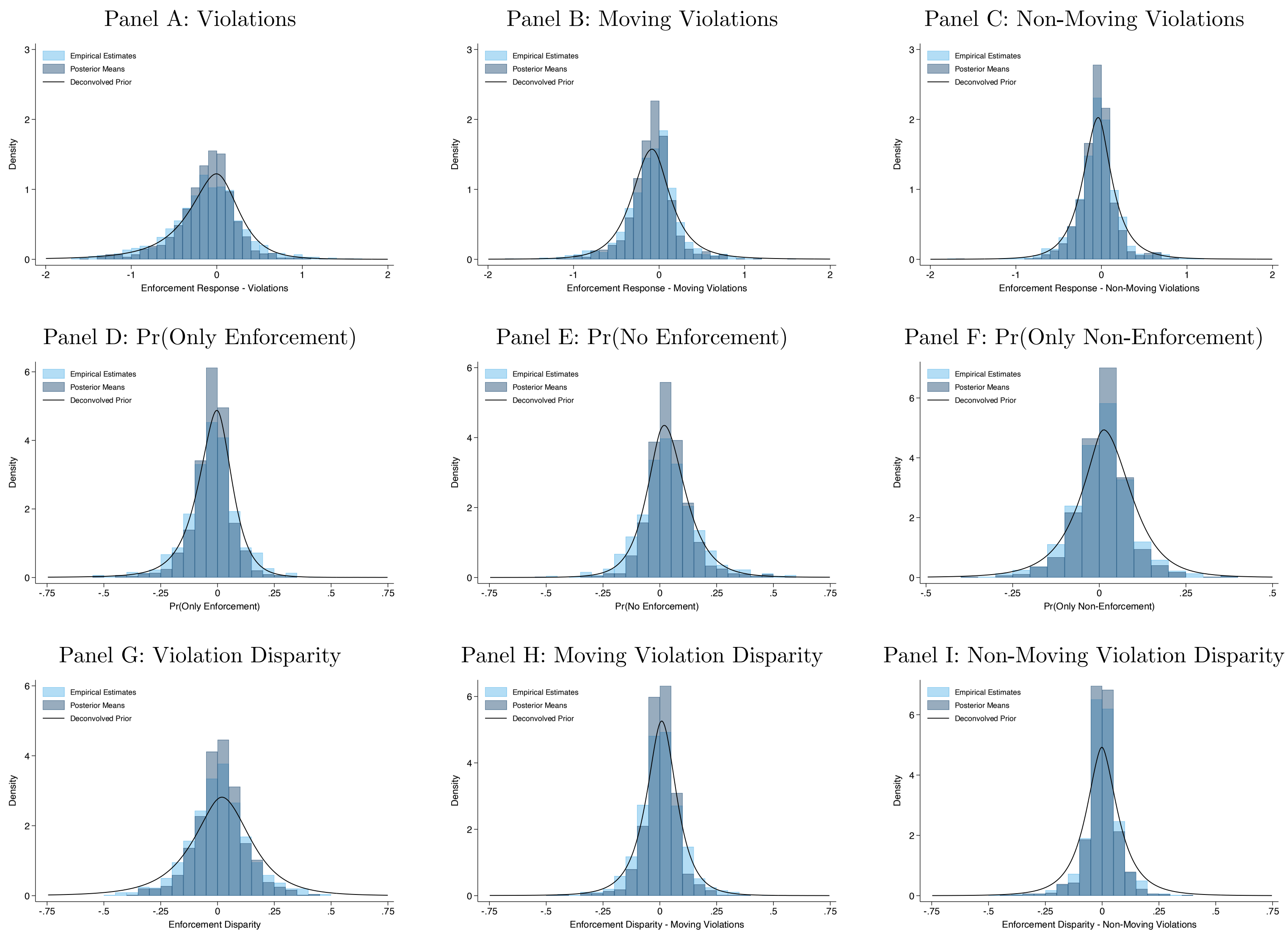
Notes: This figure reports the ratio of the variance of trooper effects and the variance of location effects in determining the distribution of multitasking collision responses using various specification choices. Trooper and location effects are estimated using Equation (7) after partialling out hour of collision response indicators in a first-step. Unbiased estimates of the effect variances in the left-hand column are calculated following the procedure in Kline, Saggio, and Sølvsten (2020). Empirical "plug-in" estimates are reported in the right-hand column. Panel A omits the first-step adjustment of partialling out $\mathbf{W}_{e(h)}$, Panel B winsorizes multitasking collision responses at the 0.5^{th} and 99.5^{th} percentiles while still partialling out $\mathbf{W}_{e(h)}$ in a first-step, and Panel C winsorizes responses without the first-step adjustment procedure. Values greater than one indicate the trooper effects explain a relatively greater share of the variance in the outcome compared to location effects. The outcome and outcome category is listed on the x-axis. The first three bars measure the relative contribution of trooper effects to enforcement activity. The second set of bars measure trooper effort substitution across enforcement and non-enforcement tasks.The third set of bars measure the relative contribution of trooper effects to enforcement disparities.

Appendix Figure A23: Trooper Effects Explain More Variation in Per-Shift Enforcement than Location Effects



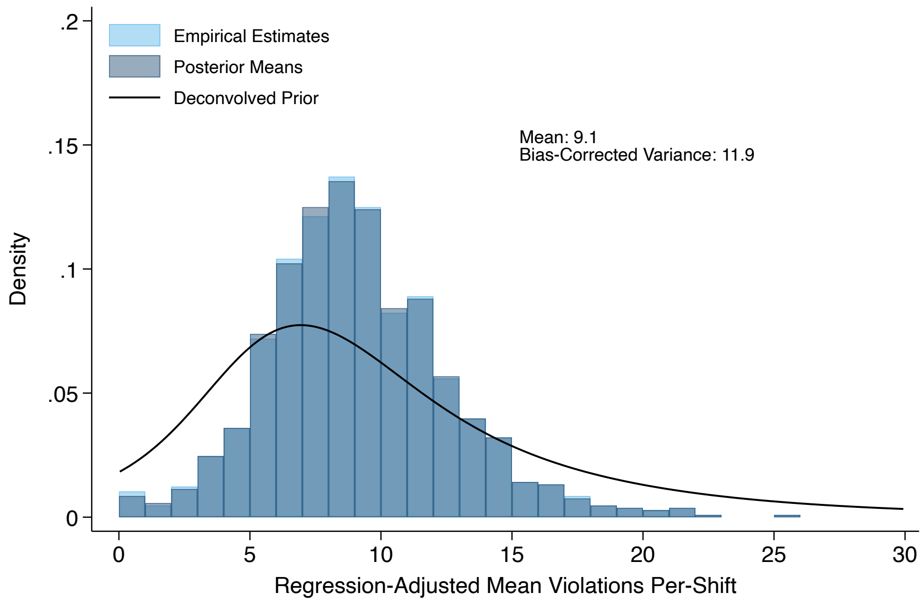
Notes: This figure reports the ratio of the variance of trooper effects and the variance of location effects in determining the distribution of trooper-by-shift enforcement volume. Trooper and location effects are estimated using Equation (7) after partialling out hour of collision response indicators in a first-step. Unbiased estimates of the effect variances are calculated following the procedure in Kline, Saggio, and Sølvsten (2020). Values greater than one indicate the trooper effects explain a relatively greater share of the variance in the outcome compared to location effects. The outcome and outcome category is listed on the x-axis. The first three bars measure the relative contribution of trooper effects to enforcement activity. The second set of bars measure trooper enforcement leniency, measured as the ratio of violations per stop. The third set of bars measure the relative contribution of trooper effects to enforcement leniency disparities, measured as the difference in ratios of violations per stop for minority and white motorists.

Appendix Figure A24: Distributions of Empirical Bayes Posteriors, Empirical Estimates, and Deconvolved Prior Density of Multitasking Collision Responses



Notes: This figure reports empirical estimates, deconvolved prior, and Empirical Bayes posterior mean estimates of trooper multitasking collision responses across a range of outcomes. The outcome is listed in the panel title. The deconvolved prior is estimated using the approach from Efron (2016). Empirical estimates are reported in the light blue histogram, posterior means are reported in the darker blue histogram, and the deconvolved prior is illustrated in the black line. A small handful of empirical outlier estimates are trimmed for exposition.

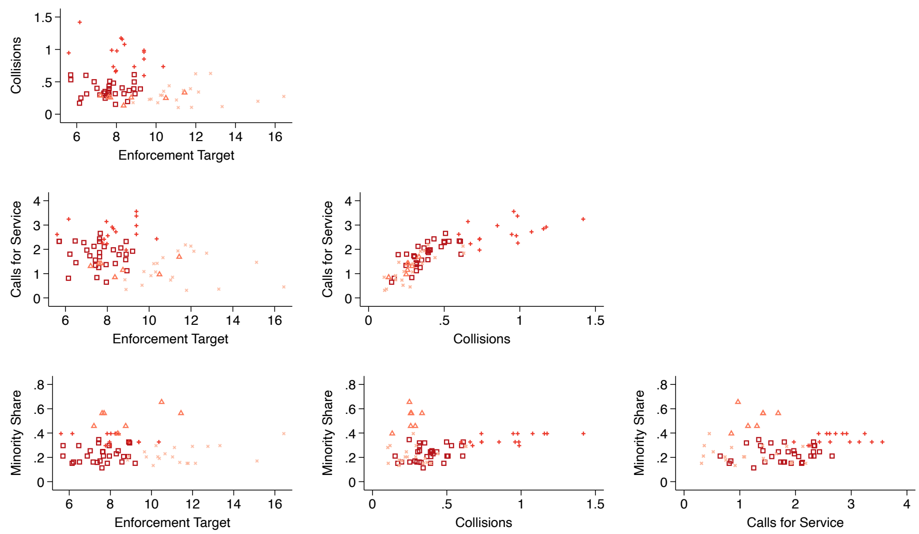
Appendix Figure A25: Distributions of Empirical Bayes Posteriors, Empirical Estimates, and Deconvolved Prior Density of Trooper Regression-Adjusted Per-Shift Enforcement



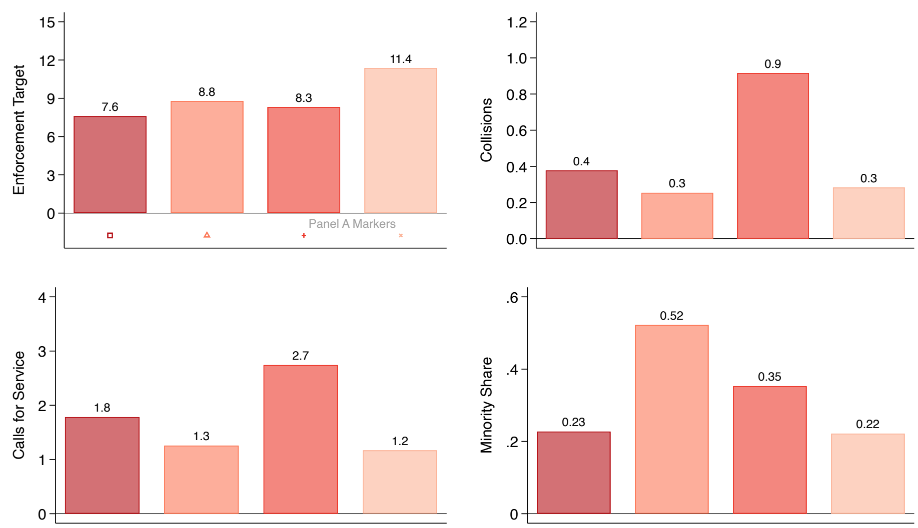
Notes: This figure reports empirical estimates, deconvolved prior, and Empirical Bayes posterior mean estimates of regression-adjusted trooper-specific per-shift enforcement. Trooper fixed effects are estimated using Equation (9). The deconvolved prior is estimated using the approach from Efron (2016). Empirical estimates are reported in the light blue histogram, posterior means are reported in the darker blue histogram, and the deconvolved prior is illustrated in the black line. The support of the prior distribution is restricted to be non-negative.

Appendix Figure A26: Identifying Location Typologies Using k-means Clustering

Panel A: Matrix of Location Means and Identified Clusters



Panel B: Within-Cluster Averages



Notes: This figure reports the joint distribution of location-specific demands at the shift-level. Panel A reports a lower triangular matrix of two-way marginal distributions of the listed outcomes from the broader four-dimensional underlying distribution. Points are classified into specific clusters using a k-means classification algorithm. See Appendix E for additional details. Panel B reports the within-cluster means across the four input outcomes.

Appendix Table A1: Regression Discontinuity Estimates of Missing Enforcement Target on Per-Shift Enforcement Behavior

	Full Sample		Easier Months		More Difficult Months	
	Unconstrained Mean	RD Estimate	Unconstrained Mean	RD Estimate	Unconstrained Mean	RD Estimate
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Enforcement Activity						
Violations	8.940	0.257**	9.213	0.402**	8.668	0.182
		(0.124)		(0.176)		(0.177)
Moving Violations	5.906	0.105	6.074	0.233*	5.739	0.064
		(0.091)		(0.125)		(0.130)
Non-Moving Violations	3.034	0.151**	3.139	0.169*	2.929	0.118
		(0.066)		(0.095)		(0.094)
Panel B: Per-Stop Stringency						
Violations	1.507	0.028***	1.509	0.020	1.506	0.031
		(0.010)		(0.014)		(0.014)
Moving Violations	0.977	0.011	0.975	0.015	0.978	0.009
		(0.008)		(0.011)		(0.012)
Non-Moving Violations	0.531	0.016	0.534	0.005	0.527	0.022
		(0.010)		(0.013)		(0.016)
Panel C: Per-Stop Stringency Disparity						
Violation Disparity	0.071	0.018	0.071	0.008	0.072	0.030
		(0.013)		(0.019)		(0.019)
Moving Violation Disparity	0.073	−0.004	0.076	−0.011	0.071	0.012
		(0.011)		(0.016)		(0.014)
Non-Moving Violation Disparity	−0.002	0.022**	−0.005	0.019	0.002	0.018
		(0.011)		(0.015)		(0.016)
Observations		13,588		6,883		6,705

Notes: This table reports regression discontinuity estimates of missing the enforcement target on average per-shift enforcement behavior in the following month. Columns 1-2 report estimates for the full sample, Columns 3-4 report estimates for easier months, and Columns 5-6 report estimates for more difficult months. See the main text for a description of the subsample specification. The outcome is listed in each row. Panel A reports estimates for measures of violations, Panel B reports estimates for measures of per-stop stringency, and Panel C reports estimates for measures of per-stop stringency disparities, with positive values indicating greater stringency towards minority motorists. Odd-numbered columns report the control (unconstrained trooper) mean. Even-numbered columns report the regression discontinuity estimate, controlling for location by month and location by year fixed effects. Observation counts reflect effective observation counts after dropping singleton observations. Standard errors clustered at the trooper level are reported in parentheses. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.

Appendix Table A2: Examining Composition of Post Collision Response Behavior: Enforcement Activity, Effort Substitution, and Enforcement Disparities

	Unconstrained Mean (1)	RD Estimate (2)
Panel A: Enforcement Activity		
Violations	0.733	0.077*
		(0.042)
Moving Violations	0.464	0.060**
		(0.029)
Non-Moving Violations	0.269	0.017
		(0.023)
Panel B: Effort Substitution		
Pr(Only Enforcement)	0.194	0.015
		(0.012)
Pr(No Enforcement)	0.703	−0.023*
		(0.014)
Pr(Only Non-Enforcement)	0.173	0.014
		(0.011)
Panel C: Enforcement Disparity		
Violations	−0.006	0.026**
		(0.013)
Moving Violations	0.000	0.016*
		(0.009)
Non-Moving Violations	−0.006	0.010
		(0.008)
Number of Events		13,412

Notes: This table reports average hourly multitasking collision responses for the focal shift. Column 1 reports the mean for unconstrained troopers and Column 2 reports point estimates from the regression discontinuity design. Each point estimate comes from a separate regression. Panel A reports results for enforcement activity, Panel B reports results for task substitution and non-enforcement activity, and Panel C reports results for enforcement disparities. Observation counts for the RD sample reflect effective observation counts after dropping singleton observations. Standard errors clustered at the trooper level are reported in parentheses. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.

Appendix Table A3: Correlating County-Specific Multitasking Collision Responses with Traffic Court Caseload Outcomes

	Caseload Outcomes	
	Share Dismissed	Share Not Committed
Panel A: Violations	(1)	(2)
Violations	0.123*	0.038**
	(0.067)	(0.017)
Non-Moving Violations	0.017	0.086**
	(0.212)	(0.039)
Panel B: Activity Composition		
Pr(Only Enforcement)	0.348	0.169**
	(0.222)	(0.064)
Panel C: Enforcement Disparities		
Violation Disparity	0.078	−0.132*
	(0.292)	(0.068)

Notes: This table reports bivariate correlations of county-level traffic court caseload outcomes on county-specific multitasking collision responses, estimated using the difference-in-differences approach in Appendix C. Each cell represents a separate regression with the type of response listed in each row. In Column 1 the dependent variable is the share of charges dismissed and in Column 2 the dependent variable is the share of charges deemed "not committed." Outcomes are calculated as the share of charges disposed. Regressions are weighted using the number of events which contribute to the collision response estimation. Positive values of the row-listed covariate indicate increasing degrees of enforcement emphasis or socially suboptimal outcomes. Heteroskedasticity-robust standard errors are reported in parentheses. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.

Appendix Table A4: Correlating Trooper Fixed Effects with Relative Performance

	Distance from Target
	(1)
Panel A: Enforcement Composition	
Violations (×100)	-0.751^{***}
	(0.094)
Moving Violations (×100)	-0.378^{***}
	(0.063)
Non-Moving Violations (×100)	-0.373^{***}
	(0.046)
Panel B: Effort Substitution	
Pr(Only Enforcement) (×100)	−0.017
	(0.019)
Pr(No Enforcement) (×100)	0.035
	(0.022)
Pr(Only Non-Enforcement) (×100)	0.007
	(0.014)
Panel C: Enforcement Disparities	
Violation Disparity (×100)	-0.065^{**}
	(0.026)
Moving Violation Disparity (×100)	−0.028
	(0.017)
Non-Moving Violation Disparity (×100)	-0.037^{**}
	(0.014)

Notes: This table reports bivariate correlations of trooper fixed effects and relative performance at the time of the collision response. Negative values of distance indicate a trooper is behind their target (constrained) and positive values indicate a trooper is ahead of their target (unconstrained). Outcomes are rescaled by ×100 for readability. Bias-corrected standard errors estimated using the approach from Kline, Saggio, and Sølvsten (2020) are reported in parentheses. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.

Appendix Table A5: Variation Across Trooper-Specific Multitasking Collision Responses

	Mean	Plug-in Variance	Bias-Corrected Variance
Panel A: Enforcement Activity	(1)	(2)	(3)
Violations	−0.133	0.227	0.156
Moving Violations	−0.086	0.114	0.081
Non-Moving Violations	−0.047	0.068	0.047
Panel B: Effort Substitution			
Pr(Only Enforcement)	−0.026	0.013	0.009
Pr(No Enforcement)	0.038	0.018	0.012
Pr(Only Non-Enforcement)	0.007	0.009	0.007
Panel C: Enforcement Disparities			
Violation Disparity	0.002	0.027	0.019
Moving Violation Disparity	0.001	0.012	0.008
Non-Moving Violation Disparity	0.002	0.007	0.005
Number of Troopers	892	892	892

Notes: This table reports mean and variance estimates of the cross-trooper distribution of multitasking collision responses. Panel A reports statistics for enforcement activity, Panel B reports statistics for enforcement disparities, and Panel C reports statistics for non-enforcement activity. This table drops a handful of troopers with only one collision response. Column 1 reports the average cross-trooper estimate, Column 2 reports the empirical "plug-in" variance, and Column 3 reports the bias-corrected estimate.

Appendix Table A6: Testing Independence of Z-scores and Trooper Standard Errors

Panel A: Enforcement Activity	(1)
Violations	-0.331
	(0.776)
Moving Violations	-1.556
	(1.176)
Non-Moving Violations	-1.492
	(1.732)
Panel B: Effort Substitution	
Pr(Only Enforcement)	-2.670
	(3.609)
Pr(No Enforcement)	-6.086^{*}
	(3.583)
Pr(Only Non-Enforcement)	-16.916^{***}
	(5.718)
Panel C: Enforcement Disparities	
Violation Disparity	1.805
	(2.880)
Moving Violation Disparity	1.702
	(3.383)
Non-Moving Violation Disparity	-2.309
	(5.018)
Number of Troopers	892

Notes: This table reports tests of independence of z-scores and trooper standard errors across a range of outcomes. Each estimate comes from a separate regression of the z-scores on the standard errors. Panel A reports results for enforcement activity, Panel B reports results for enforcement disparities, and Panel C reports results for non-enforcement activity. The outcome from which the z-score and standard error originate from are listed in each row. Heteroskedasticity-robust standard errors are reported in parentheses. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.

Appendix B: Model Appendix

In this appendix, I provide additional results and discussion on the model in Section 2.3. I also illustrate how two alternative formulations of the model develop empirical predictions about the type/quality of enforcement and the distribution of enforcement across motorists.

B.1 Derivation of the Agent's Problem

In the main text, I present a compact version of the trooper's optimization problem and associated first-order condition. This subsection provides additional details on the derivation.

Given an arbitrary contract $\mathbf{c} \equiv (n, o, b)$, the trooper's optimization problem is:

$$\begin{aligned}\max_{t_e, t_n} U(t_e, t_n, \mathbf{c}) = \quad & \underbrace{t_e t_n u(b)}_{\text{payoff complete both tasks}} + \underbrace{[t_e(1-t_n) + (1-t_e)t_n]u(o)}_{\text{payoff complete one task}} \\ & + \underbrace{(1-t_e)(1-t_n)u(n)}_{\text{payoff complete neither task}} - \underbrace{(\omega(a)t_e + t_n)}_{\text{effort cost}}\end{aligned} \tag{B1}$$

I define the following two differences in payoffs, $\Lambda^b = u(b) - u(o)$ and $\Lambda^o = u(o) - u(n)$, which when substituted into Equation (B1) yields the optimization problem in the main text:

$$\max_{t_e, t_n} u(n) + t_e(\Lambda^o - \omega(a)) + t_n(\Lambda^o - 1) + t_e t_n(\Lambda^b - \Lambda^o) \tag{B2}$$

Given an interior solution, the trooper will exert $t_e + t_n = 1$, with $t_o = 0$. See MacDonald and Marx (2001) for a formal argument.
Since $t_e = 1 - t_n$, we can substitute this relationship into Equation (B2), yielding

$$\max_{t_e} u(n) + t_e(\Lambda^o - \omega(a)) + (1-t_e)(\Lambda^o - 1) + (t_e - t_e^2)(\Lambda^b - \Lambda^o) \tag{B3}$$

The corresponding first order condition is:

$$\frac{\partial U(\cdot)}{\partial t_e} = (\Lambda^o - \omega(a)) - (\Lambda^o - 1) + (1 - 2t_e)(\Lambda^b - \Lambda^o) = 0 \tag{B4}$$

$$\omega(a) - 1 = (1 - 2t_e)(\Lambda^b - \Lambda^o) \tag{B5}$$

$$t_e^* = \frac{1}{2}\left(1 + \frac{1 - \omega(a)}{\Lambda^b - \Lambda^o}\right) \tag{B6}$$

B.2 The Patrol's Problem

In the main text, I abstract from an explicit characterization of the Patrol's problem for expositional ease. This subsection provides additional on the Patrol's maximization problem.

Recall from the main text that the Patrol receives a payoff, B, only when both enforcement and non-enforcement tasks are completed successfully, and zero payoff otherwise. The Patrol's problem is then the following:

$$V(t_e, t_n, \mathbf{c}) = t_e t_n (B - b) - [t_e(1 - t_n) + t_n(1 - t_e)]o - (1 - t_e)(1 - t_n)n \tag{B7}$$

It's straightforward to see that the Patrol would like to induce the agent to exert effort on both tasks, since they receive a positive payoff in expectation only when the agent does so. In particular, the Patrol would like $t_e = t_n = \frac{1}{2}$, which maximizes $V(\cdot)$. This relationship between tasks, such that they are complements for the principal and substitutes for the agent, is a key ingredient of this model. Thus, the Patrol needs to design a contract or set of payoffs such that it is optimal for the trooper to devote effort towards both enforcement and non-enforcement tasks.

B.3 Multitask Principal-Agent Model with Motorist Group-Specific Enforcement Costs

In the main text, I develop a model which illustrates the degree of substitution across enforcement (contracted) and non-enforcement (non-contracted) tasks. In this subsection, I present an alternative model where the trooper divides enforcement effort between enforcement for minority and white motorists, providing a theoretical rationale for the observed empirical disparity.

Suppose the trooper divides her time between enforcement for minority (m) and white (w) motorists. Minority motorists have a lower cost of enforcement ($\omega(a) \in (0, 1)$) due to downstream hassle costs and probability of court contesting, which I illustrate in the main text. Alternatively, one could imagine there are differences in vehicle quality across motorist groups, such that minority motorists are more likely to have vehicles with hidden non-moving violations. Therefore the marginal cost of enforcement is also lower for these motorists. To formalize this, assume there's an arbitrary contract $\mathbf{c} \equiv (n, o, b)$ which generates an interior solution. The trooper's optimization problem is then to choose t_m, t_w to solve:

$$\max_{t_m, t_w} u(n) + t_m(\Lambda^o - \omega(a)) + t_w(\Lambda^o - 1) + t_m t_w(\Lambda^b - \Lambda^o) \tag{B8}$$

which yields the following first order condition, with $t_m = 1 - t_w$:

$$\frac{\partial U(\cdot)}{\partial t_m} = (\Lambda^o - \omega(a)) - (\Lambda^o - 1) + (1 - 2t_m)(\Lambda^b - \Lambda^o) = 0 \tag{B9}$$

$$\omega(a) - 1 = (1 - 2t_m)(\Lambda^b - \Lambda^o) \tag{B10}$$

$$t_m^* = \frac{1}{2}\left(1 + \frac{1 - \omega(a)}{\Lambda^b - \Lambda^o}\right) \tag{B11}$$

Thus, when $1 - \omega(a)$ is sufficiently large, relative to $\Lambda^b - \Lambda^o$, the distribution of a trooper's enforcement pattern will shift towards that of minority motorists.

B.4 Multitask Principal-Agent Model with Two Types of Enforcement Tasks

In the main text, I develop a model which illustrates the degree of substitution across enforcement (contracted) and non-enforcement (non-contracted) tasks. In this subsection, I present an alternative model where the trooper divides effort between producing low- and high-quality enforcement tasks and show that it leads to similar predictions as before.

Suppose the trooper divides her time between high-quality (h) and low-quality (l) enforcement tasks, respectively devoting t_h and t_l to each of them. Low-quality enforcement tasks have relative cost of effort $\omega(a) \in (0,1)$. This simple division of enforcement tasks can be motivated by the Patrol's emphasis on particular "targeted" enforcement action, such as driving under the influence or seatbelt usage, however these generally require greater effort to seek out and enforce. Assuming the same arbitrary contract which generates an interior solution, the trooper's optimization problem is to choose t_h and t_l to solve:

$$\max_{t_h, t_l} u(n) + t_l(\Lambda^o - \omega(a)) + t_h(\Lambda^o - 1) + t_l t_h(\Lambda^b - \Lambda^o) \tag{B12}$$

which yields the following first order condition, with $t_l = 1 - t_h$:

$$\frac{\partial U(\cdot)}{\partial t_l} = (\Lambda^o - \omega(a)) - (\Lambda^o - 1) + (1 - 2t_l)(\Lambda^b - \Lambda^o) = 0 \tag{B13}$$

$$\omega(a) - 1 = (1 - 2t_l)(\Lambda^b - \Lambda^o) \tag{B14}$$

$$t_l^* = \frac{1}{2}\left(1 + \frac{1 - \omega(a)}{\Lambda^b - \Lambda^o}\right) \tag{B15}$$

Thus, when $1-\omega(a)$ is sufficiently large, relative to $\Lambda^b - \Lambda^o$, the trooper will substitute towards low-quality enforcement tasks that are easier to generate, at the expense of high-quality enforcement.

Appendix C: Complementary Difference-in-Differences Design

In this appendix, I provide additional details on the complementary difference-in-differences strategy I develop to supplement the primary regression discontinuity specification. Relative to the regression discontinuity approach, the key advantage of the difference-in-differences approach is that it allows for the estimation of causal effects away from the discontinuity as well as providing sufficient power and ease of aggregation for subsample estimates, such as the study of within-month behavior. I develop a semi-parametric aggregation scheme that allows for pooling information across many unique events while still ensuring common support of the relative time coefficients to test the identifying assumptions and illustrate dynamic effects. I also extend the structure of the data to accommodate changes in trooper behavior across multiple shifts, allowing for medium-term intertemporal effort substitution.

C.1 Difference-in-Differences Design

I estimate the impact of multitasking via collision responses on the allocation of trooper enforcement activity using a dynamic difference-in-differences design. In the simplest case, this design intuitively compares two troopers who are working the same shift and beat, one who was assigned a collision response and one who was not. The "untreated" trooper(s) provide a counterfactual for the enforcement activity the collision responding trooper would have conducted in the absence of the collision response. Since these troopers are working the same shift and beat, they face the same enforcement expectations and traffic patterns. Formally, I estimate regressions of the following general form:

$$Y_{ijt}^h = \alpha_i^h + \delta_{jt}^h + \sum_{k \in \mathcal{K}} \beta_k^h D_{ik}^h \cdot Collision_i^h + \varepsilon_{ijt}^h \tag{C1}$$

for trooper i, beat j, time t, and the shift-hour of the collision response h. The coefficients of interest are β_k^h which trace out the dynamic differences in the outcome Y_{ijt} between treated and untreated troopers.

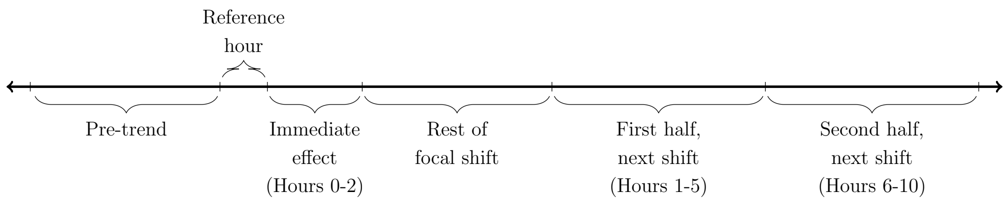
I estimate equation (C1) using a two-step imputation approach akin to Gardner (2021) and Borusyak, Jaravel, and Spiess (Forthcoming).[1] I conduct inference using a Bayesian boostrap (Rubin 1981), accounting for variation in both the first- and second-steps of the estimation strategy. Finally, to capture both the immediate and medium-term responses, I examine the dynamics of trooper responses in both the focal shift and in the following shift I observe the trooper working.

One empirical challenge with this approach is that collision responses may arrive at any point within a shift. Implicitly, the set of event time indicators $\mathcal{K}$ therefore depends on h, such that:

$$\mathcal{K} := \big\{k : k \in \underbrace{[-h+1, -2]}_{\text{Pre-trend}} \cup \underbrace{[0, 10-h]}_{\text{Same shift}} \cup \underbrace{[10, 20-h]}_{\text{Next shift}} \big\}.$$

[1] Specifically, imputation estimators use untreated and not-yet-treated observations to form estimates of the untreated counterfactual ($Y(0)$). The difference between the realized outcome and $Y(0)$ is an estimate of the treatment effect.

Thus, $k = 4$ may represent an hour at the end of one shift and the beginning of another shift, depending on the shift-hour of the collision response, h. Moreover, the fact that collision responses arrive at different hours means that event-time coefficients near the boundaries of $\mathcal{K}$ will have only partial support across values of h. I overcome both of these empirical challenges by estimating Equation (C1) separately for each h and then aggregate event-time coefficients to represent common periods according to the following semi-parametric aggregation scheme:



where the first two post-periods (immediate effect and rest of the focal shift) capture enforcement responses during the focal shift and the last two periods capture spillover effects into the following shift while permitting some dynamics within that following shift. Within each of these groups, I calculate the average hourly response by taking the weighted average across h, using the number of events as weights.

Following the main text, I focus on collision responses which occur in the interior hours $h \in [3, 8]$ of worked shifts in my sample for two reasons, the first of which is specific to this research design. First, collision responses which occur in the first two hours of a shift may not have a well-defined pre-period, which precludes testing the identifying assumptions. Second, collisions which occur in the tails of the shifts may not be quasi-randomly assigned. For example, when shifts overlap, collisions may be more likely to be assigned to troopers who are just beginning their shifts rather than troopers who are ending their shifts. Appendix Figure C1 plots the mean collision assignment rates by hour for common shifts and depicts exactly this type of strategic phenomenon in the shift tails. While the focus on interior shifts potentially limits the generalizability of the estimates to all collisions, it also ensures that the experimental design is internally valid.[2]

There are two identifying assumptions underlying this research design. The first is that enforcement patterns of collision-responding and non-responding troopers should evolve in parallel in the absence of the collision response. This assumption could be violated if collisions are systematically assigned to troopers who have dynamically higher or lower enforcement volume within a shift and would present as non-parallel trends in the pre-period. A second concern is that collision responses are systematically assigned to troopers across shifts, perhaps in an attempt to equalize collision volume across troopers. I empirically test whether this behavior occurs in Appendix Figure C2 and find little evidence a collision response in a given shift is consistently correlated with being assigned a collision response in either the prior or following shift.

[2] Appendix Table C1 shows that the set of shifts with collision responses and untreated control groups is broadly representative both spatially and temporally. I also note here that the set of collision responses is a subset of the events from the regression discontinuity application, as the difference-in-differences approach requires an untreated control group, a condition that may not always be satisfied when using only troopers in the same assignment.

Appendix Figure C3 presents difference-in-differences estimates using the semi-parametric aggregation scheme. Reassuringly there are little consistent parallel trends violations for the vast majority of outcomes. One caveat is that the disparity estimates for constrained troopers are somewhat elevated in the pre-period, although the implied trend would suggest enforcement patterns trending in the opposite direction. Moreover, the RD estimates in Appendix Figure A17 suggest little pre-existing differences for across troopers in addition to returning similar magnitude and precision estimates as the difference-in-differences design. However, the results should be interpreted with these caveats in mind.

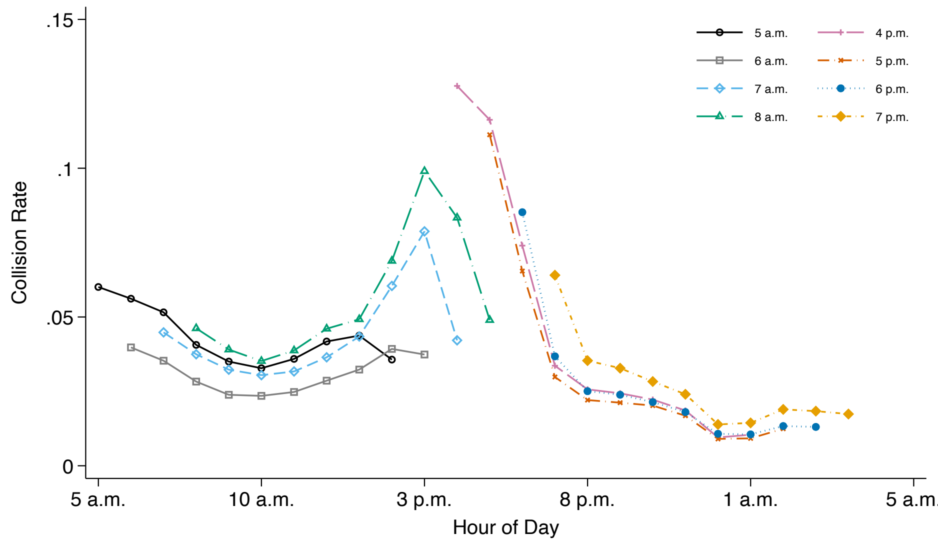
Consistent with collision responses having dynamic impacts on troopers, I find sharp changes in trooper enforcement activity and allocation of effort. Immediately following the collision response, enforcement activity drops sharply, consistent with the collision response triggering additional obligations. During the following hours, activity rises but does not fully recover on average. These patterns are likely due to troopers completing collision reports and otherwise fulfilling additional non-enforcement obligations, given the pattern of effort in Panels D-F.

DiD Empirical Test of Proposition 1: In the main text, I define two empirical tests or implications of Proposition 1. Here I define the analogous test in the difference-in-differences framework. The analogous test compares the difference in mean enforcement activity for constrained and unconstrained troopers. Formally, this is:

$$\underbrace{\mathbb{E}[\beta_k | k \geq 0, \underline{a}]}_{\substack{\text{Average of dynamic} \\ \text{enforcement estimates}}} > \mathbb{E}[\beta_k | k \geq 0, \overline{a}]$$

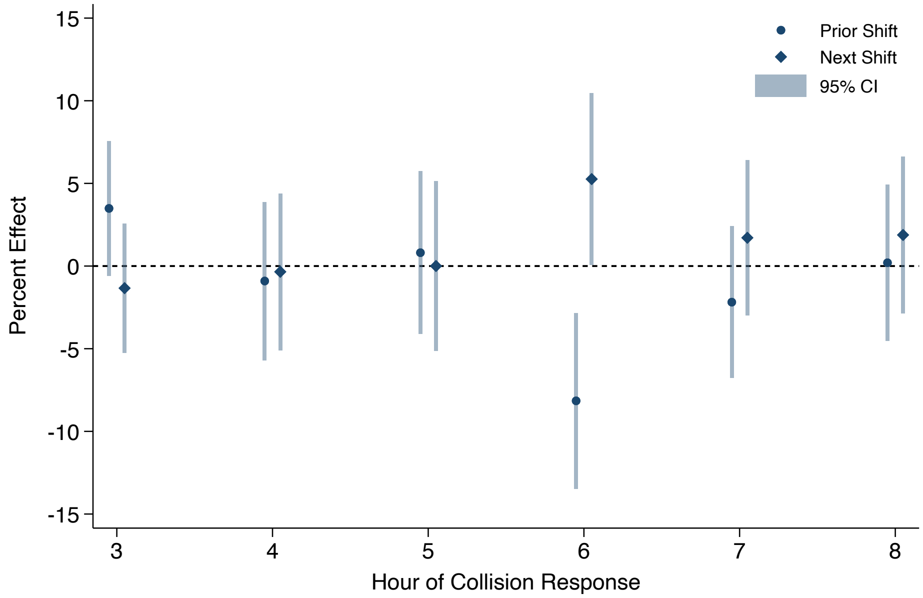
where troopers with $\underline{a}$ are "constrained" and troopers with $\overline{a}$ are "unconstrained." Formal point estimates for the mean hourly differences between these two groups are reported in Appendix Figure C4.

Appendix Figure C1: Collision Rates by Hour of Shift for Common Shifts



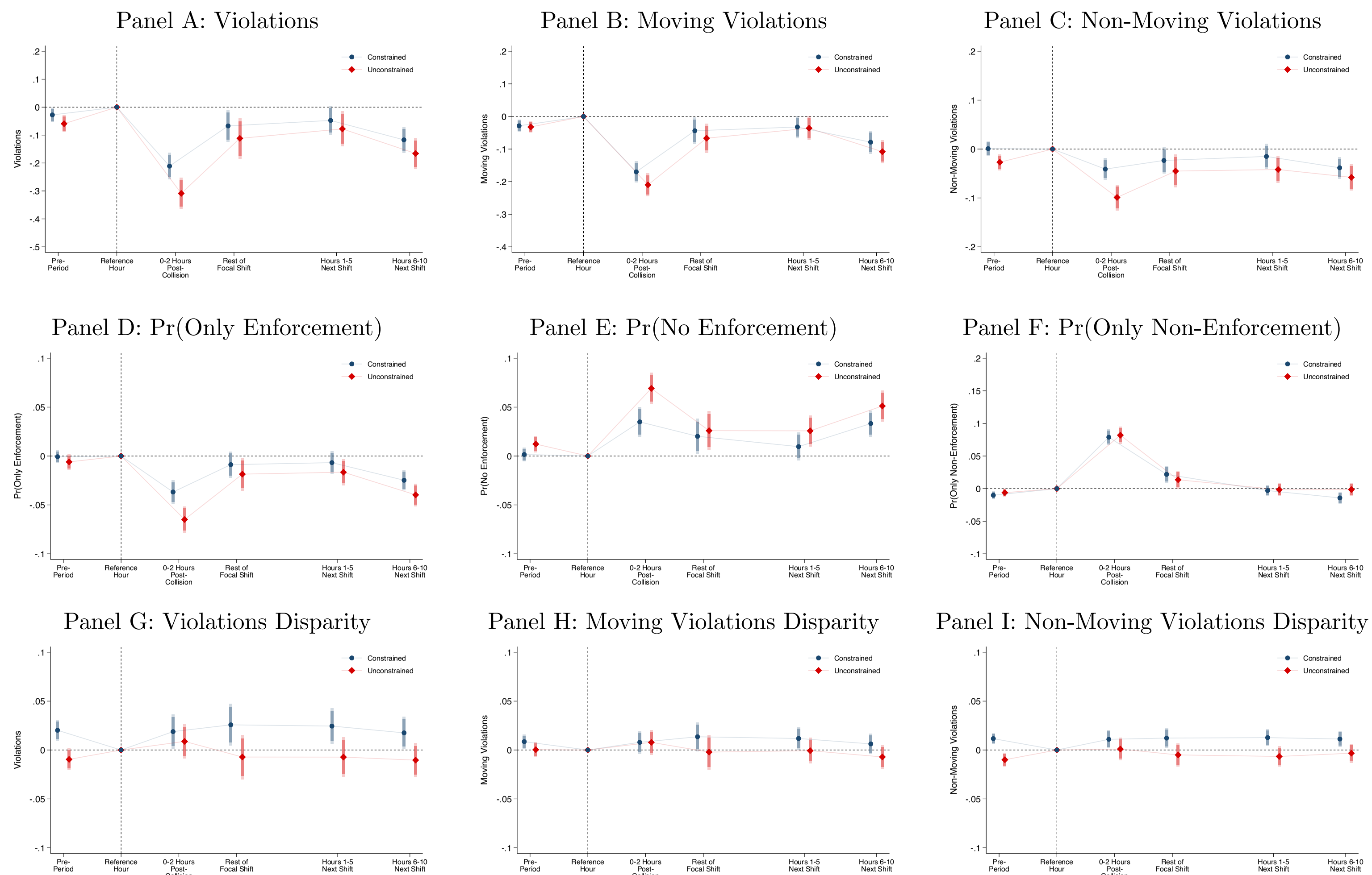
Notes: This figure presents collision assignment rates by hour of day and hour of shift for common ten-hour shifts in the Washington State Patrol. Each point represents the mean probability of receiving a collision response in that hour for that shift, pooling data across all districts for the period 2010-2018. Various color and shape markers denote shifts that begin in the listed hour.

Appendix Figure C2: Testing for Systematic Assignment of Collision Responses Across Shifts



Notes: This figure tests for a systematic assignment of collision responses across adjacent shifts. Specifically, conditional on a collision response in shift t, each coefficient tests whether shift $t-1$ or $t+1$ are also more or less likely to have collision responses assigned in them. Each point represents a separate regression. Solid dots represent prior shifts and solid diamonds represent following shifts. Regressions also control for trooper, location by date, and date by shift start hour fixed effects. Coefficients are converted to percent effects relative to the sample mean. Shaded regions represent 95 percent confidence intervals with standard errors clustered at the trooper level.

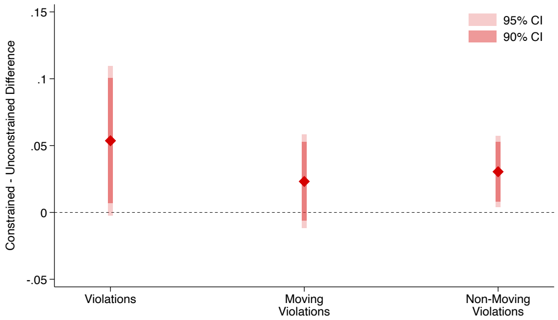
Appendix Figure C3: Difference-in-Differences Estimates of Trooper Activity Post-Collision Response Under Semi-Parametric Aggregation Scheme



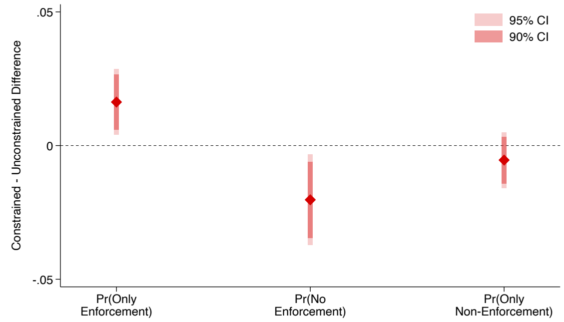
Notes: This figure reports semi-parametric difference-in-differences estimates of trooper responses to collisions. Each point represents the event-weighted average, following the semi-parametric aggregation scheme discussed in Appendix C. Navy dots represent constrained troopers and red diamonds represent unconstrained troopers using the definitions and samples from the main text. Lighter and darker shaded regions correspond to 95 and 90 percent confidence intervals.

Appendix Figure C4: Trooper Enforcement Activity Changes Following Multitasking Collision Response in Difference-in-Differences Framework

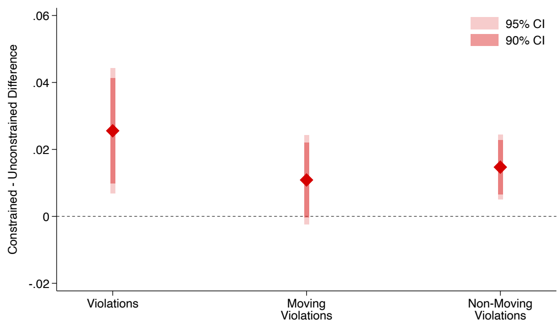
Panel A: Changes in Enforcement Activity



Panel B: Effort Substitution



Panel C: Changes in Enforcement Disparities



Notes: This figure reports difference-in-differences estimates testing for cross-group differences in multitasking collision responses between constrained and unconstrained troopers. Each coefficient represents the difference in the average DiD estimate for constrained and unconstrained troopers. Estimates denote the difference in average hourly activity. Panel A reports results for changes in enforcement activity, Panel B reports results for the composition of trooper activity, and Panel C reports enforcement disparities, adjusting for differences in the population prevalence of motorists. Lighter and darker shaded regions represent 95 and 90 percent confidence intervals.

Appendix Table C1: Representativeness of Sample Shifts with the Broader Set of Collision-Involved Shifts

	Sample Mean	All Collision Shifts Mean
Panel A: Districts	(1)	(2)
District 1	0.176	0.187
District 2	0.356	0.320
District 3	0.035	0.046
District 4	0.079	0.076
District 5	0.045	0.079
District 6	0.050	0.067
District 7	0.218	0.162
District 8	0.040	0.063
Panel B: Years		
2010	0.120	0.103
2011	0.120	0.101
2012	0.108	0.098
2013	0.111	0.099
2014	0.112	0.108
2015	0.097	0.119
2016	0.104	0.134
2017	0.125	0.137
2018	0.103	0.100

Notes: This table compares district and year characteristics of shifts with collision responses that are in the primary estimation sample with shifts that had a collision response in $h \in [3, 8]$ in the same sample period. Column 1 reports means for the primary estimation sample and Column 2 reports means for the broader sample of shifts with collision responses.

Appendix D: Welfare Calculation Appendix

This appendix provides additional details on the parameters and approach I use in Section 4.2 to estimates the welfare impacts of delayed collision reports.

Recall from the main text that I relate the "costs" of delayed reporting with the "benefits" of additional enforcement emphasis through the following framework:

$$\underbrace{\Delta\text{NPV Settlement}}_{\text{1 week delay}} = \left[\underbrace{\alpha_1(\text{Moving Viols.})}_{\text{RD Estimate}}\right] \times \left(\underbrace{\text{Rev. from enforce.}}_{\substack{\text{Ticket value } \times \\ \text{Pr(payment)}}} + \underbrace{\text{Deterrence effect}}_{\substack{\text{Deter treat. eff. } \times \\ \text{settlement cost}}}\right) \quad \text{(D1)}$$

Costs of Delayed Reports: The left-hand side of Equation (D1) is the cost to the victim from receiving their settlement one week later. I take bounds on the value of the settlement and assume an average payment lag of 10.7 months.[3] Minor collision settlement values are between [$1,090, $15,000] and moderate collisions are between [$15,850, $43,063]. I assume a discount rate of 3 percent compounded monthly and calculate the left-hand side as the difference in the net present value of a lump-sum payment under the shorter and longer waiting horizon.

Benefits of Enforcement Emphasis: The right-hand side of Equation (D1) is the social benefit of increased enforcement and abstracting from report-writing. The first term $\Delta_{below-above}$ is the difference in hourly enforcement between constrained and unconstrained troopers from the RD estimate for moving violations in Appendix Figure C4, multiplied by the number of hours. I remain agnostic on when the report-writing takes place and how long the report takes to complete and therefore compute Equation (D1) over a range of values from 1-10 hours, where 10 hours is one shift. For the government revenue component, I take $150 as the average value of the ticket, which is approximately the cost of speeding between 6-10 miles per hour over the limit, and multiply this value by the caseload weighted probability of a traffic violation ending in payment from 2010-2018.[4] For the deterrence component, I use the average values of a minor ($6,827) and moderate ($25,181) collision settlement and the one-year collision deterrence estimate from Goncalves and Mello (2023).

[3] Respectively, these values come from *InjuryClaimCoach.com* (Washington State values) and `https://www.forbes.com/advisor/legal/auto-accident/typical-car-settlement-amounts/`.

[4] Implicitly, I assume that all violations result in a ticket. Multiplying the right-hand side revenue component by mean ticketing rates would further shift the calculus towards welfare-negativity.

Appendix E: Econometric Appendix

In this appendix I provide additional detail and derivations for some of the key exercises I conduct in the main text.

E.1 Variance Decomposition

Recall from Section 4 that I am interested in decomposing the variance across collision responses into a component ascribed to trooper effects and a component ascribed to location effects. This decomposition begins with the simple two-way fixed effects model

$$R_{ije(h)}(y) = \underbrace{\alpha_i}_{\substack{\text{Trooper}\\\text{fixed}\\\text{effects}}} + \underbrace{\psi_j}_{\substack{\text{Location}\\\text{fixed}\\\text{effects}}} + \underbrace{\mathbf{W}'_{e(h)}\beta}_{\substack{\text{Auxiliary}\\\text{controls}}} + \varepsilon_{ije(h)} \tag{E1}$$

Simple estimates of the variance of the trooper ($\hat{\mathbb{V}}_{\alpha_i}$) and location effects ($\hat{\mathbb{V}}_{\psi_j}$) will be upwardly biased due to estimation error in the fixed effects. This upward bias is problematic for correctly quantifying the share of the variance in $R(y)$ that's due to trooper or location effects since the degree of estimation error may differ across each of these groups. Kline, Saggio, and Sølvsten (2020) illustrate a simple leave-out approach to purging the variance estimates of this estimation error. I illustrate my implementation of their procedure below.

I begin by first partialling out the vector $\mathbf{W}_{e(h)}$, which includes shift-hour of collision response indicators for each $h \in [3, 8]$. I then work with the simplified model

$$\tilde{R}_{ije(h)}(y) = \alpha_i + \psi_j + \varepsilon_{ije(h)} \tag{E2}$$

where $\tilde{R}_{ije(h)}(y) = R_{ije(h)}(y) - \mathbf{W}'_{e(h)}\hat{\beta}$.

I then compute leave-out estimates of the standard error of each trooper and location fixed effect, which, when squared, I denote $\hat{s_i}^2$ and $\hat{s_j}^2$. Kline, Saggio, and Sølvsten (2020) show that standard variance component estimates can be corrected for estimation error-induced bias by subtracting the weighted average of these squared standard errors from empirical variance estimates, which I use to form bias-corrected estimates of trooper and location fixed effects.

$$\mathbb{V}_{\alpha_i}(y) = \overbrace{\hat{\mathbb{V}}_{\alpha_i}(y)}^{\substack{\text{“Plug-in”}\\\text{estimate}}} - \overbrace{\sum_{i=1}^{N} \underbrace{\frac{|R_i|}{|R|}}_{\substack{i\text{'s share}\\\text{of}\\\text{events}}} \hat{s_i}^2(y)}^{\text{Bias correction}}, \quad \mathbb{V}_{\psi_j}(y) = \overbrace{\hat{\mathbb{V}}_{\psi_j}(y)}^{\substack{\text{“Plug-in”}\\\text{estimate}}} - \overbrace{\sum_{j=1}^{J} \underbrace{\frac{|R_j|}{|R|}}_{\substack{j\text{'s share}\\\text{of}\\\text{events}}} \hat{s_j}^2(y)}^{\text{Bias correction}} \tag{E3}$$

where $\hat{\mathbb{V}}_{\alpha_i}$ and $\hat{\mathbb{V}}_{\psi_j}$ are the empirical “plug-in” variance estimates.

Derivation of Ratio Measure: Canonical variance decompositions (e.g., Abowd, Kramarz, and Margolis 1999, Card, Heining, and Kline 2016) are often interested in quantifying the contribution of the variance of some set of fixed effects to the variance of the outcome. For example, the contribution of the variance of firm effects to the variance of log earnings. In my setting, I am interested in determining whether trooper or location effects explain a greater share of the variance in collision responses. Rather than calculating the variance shares for each component, I instead simply compute the ratio of the two variances. This ratio compactly summarizes the relative contributions of each set of fixed effects to the overall variance and falls neatly out of working with the standard variance share estimates.

$$\frac{\mathbb{V}_{\alpha_i}(y)/\mathbb{V}(y)}{\mathbb{V}_{\psi_j}(y)/\mathbb{V}(y)} = \frac{\mathbb{V}_{\alpha_i}(y)}{\mathbb{V}_{\psi_j}(y)} \tag{E4}$$

E.2 Discussion of Clustering Algorithm

In Section 5.1, I provide evidence that the variation in multitasking collision responses reflects distinct differences in trooper behavior and performance, which I use to subset troopers into specific typologies. I use these typologies, combined with the simple multitask principal-agent model of differential costs of enforcement as motivation for counterfactual exercises which change this differential cost. In this subsection, I provide additional details on the classification algorithm I employ to identify trooper typologies.

k-Means Clustering: k-means clustering is a parsimonious method to grouping together unlabeled (i.e., without pre-determined group membership) data points. At a high-level, the method chooses a partition (clusters) with centroids for a given number clusters to minimize within-cluster variance. This requires a single input parameter of the number of clusters. I determine the "optimal" number of clusters using established techniques, including examining the relationship between the sum of squared errors and number of clusters for an "elbow" and the relationship between the silhouette score and number of clusters.

Appendix Figure E1 reports the results from these cluster choice exercises when applied to both troopers and locations. The first row presents results for troopers and both methods highlight five as the choice of clusters. The second row presents results for locations. While Panel C highlights four as the optimal number of clusters, the story in Panel D is less clear. Although taken together, the dip when moving from four to five combined with the results in Panel C suggests that using four clusters is a reasonable approach.

With the number of clusters determined, the next objective is to find the centroids of these clusters and assign nearby data points to these clusters as appropriate. Let $(\mathbf{y}_1, \dots \mathbf{y}_N)$ represent a d-vector of trooper-specific or location-specific means. The goal is to find a partition $P = \{p_1, \dots, p_k\}$ with centroids $\boldsymbol{\kappa}$ which minimize the following objective function:

$$\arg\min_{P} \sum_{i=1}^{k} \sum_{\mathbf{y} \in p_i} ||\mathbf{y} - \boldsymbol{\kappa}_i||_2 \tag{E5}$$

where $|| \cdot ||_2$ is the L2-norm. That is, we are looking for P which minimizes the sum of the within-cluster sum of squares (WSS). The centroids of each partition are given by $\boldsymbol{\kappa}$.

There are two drawbacks with leveraging such a parsimonious clustering mechanism. First, since the measure of within-cluster variation is distance-based, dimensions of $\mathbf{y}$ with greater variances will implicitly receive more weight when solving the objective function. Following standard approaches, I first standardize $\mathbf{y}$ so that each dimension of $\mathbf{y}$ has mean zero and unit variance. I perform the full algorithm on this standardized data, before presenting the unstandardized and labeled data in Figure 7 and Appendix Figure A26. Second, k-means in general is sensitive to the initial starting conditions. I therefore randomly choose starting points 1,000 times, perform the full clustering algorithm, and then choosing the final set of cluster centroids from the iteration with the lowest WSS.

E.3 Allocating Troopers to Locations

In this subsection I provide additional details on the approach I use to solve the social planner's optimal allocation problem.

Recall from the main text that the social planner would like to choose a matching between troopers and locations which maximizes some payoff function $\Phi(x_i, z_j)$. Empirically, this problem can be written in a general form of

$$\max_{\pi_{ij}} \sum_{ij} \pi_{ij} \Phi(x_i, z_j) \tag{E6}$$

$$\text{s.t.} \quad \sum_{j=1}^{J} \pi_{ij} = p_i \tag{E7}$$

$$\sum_{i=1}^{N} \pi_{ij} = q_j \tag{E8}$$

$$\pi_{ij} \geq 0 \tag{E9}$$

which is a linear program and can be tackled using standard tools. π_{ij} is a joint distribution or optimal matching matrix which matches troopers i to locations j. I take the payoff function, $\Phi(x_i, z_j)$, to be the scalar product of the input characteristics for troopers and locations as described in the main text.

Existence and Uniqueness of π^*_{ij}: Existence of an optimal matching matrix is trivial since any arbitrary matching of characteristics (joint distribution) of the marginal distributions is a potential solution. See Mangasarian (1979) for a discussion on the uniqueness of linear program solutions. More generally in the optimal transport literature, a unique solution exists when at least one of the source or target distributions is absolutely continuous with respect to the Lebesgue measure. See Galichon (2016) for a high-level discussion of duality in this space.

E.4 Predicting Trooper Behavior Under Counterfactual Policies

Once the optimal matching matrix π^*_{ij} is discovered, the remaining challenge is how to predict trooper responses under the new matching regime. Unlike standard linear programs where the object of interest may be the value of Equation (E6), here I am interested in predicting counterfactual values of $\theta_i(y)$ when moving from the empirical coupling $\hat{\pi}_{ij}$ (observed match) to the optimal one, π^*_{ij}. Returning to the underlying principal agent problem, moving a trooper i from location j to j^* effectively changes the saliency of the enforcement target in the new location, in addition to there being a common level shift in trooper behavior across locations, along with any behavioral responses to the change in target. Thus, any prediction of counterfactual trooper behavior needs to account for all of these components.

In practice, I predict counterfactual trooper responses using linear regression. Specifically, I regress trooper responses $\theta_i(y)$ on the relative performance $Y_i - E_j$ within each location. This gradient illustrates the difference in trooper responses as a function of the constraint saliency, a relationship which has a tight connection to the underlying principal-agent model in Section 2. Formally, this is

$$\theta_i(y) = \varrho_j + \rho_j(Y_i - E_j) + \varepsilon_{ij} \tag{E10}$$

Changing E_j of Assignments: To start, consider a counterfactual policy which raises or lowers the enforcement target E_j by some value Δ_j. Under this alternative policy, the observed and optimal assignments are the same, $j = j^*$. Thus, trooper counterfactual responses are simply given by

$$\mathbb{E}[\theta^*_i(y)|j^*] = \hat{\varrho}_{j^*} + \hat{\rho}_{j^*}(Y_i - E_{j^*} - \Delta_j) \tag{E11}$$

and this policy represents a movement along the $\hat{\rho}_{j^*}$ gradient. Adding trooper responses from the change in target is straightforward by adding $\hat{\rho}_{j^*}(\Delta_j \times \mathbb{E}[\tau_{k;k\geq 0}])$ where the latter expectation is the average of the asymmetric trooper transfer responses as a function of changes in the enforcement target. Note that the effect of changing the target will in general be partially mitigated the behavioral response to the change in the target. For example, increasing the target by one holding fixed relative performance means the target is more salient, but on average troopers adjust their behavior by some amount which will offset some of this increased saliency, meaning the effective change in relative performance $(Y_i - E_j)$ is less than one.

Intuition to Oaxaca-Blinder Decomposition: In this subsection, I illustrate how the prediction of a trooper response with movement across locations has a similar intuition to an Oaxaca-Blinder decomposition. To begin, write the response under the observed assignment as

$$\theta_i(y) = \hat{\varrho}_j + \hat{\rho}_j(Y_i - E_j) + d_{ij} \tag{E12}$$

where d_{ij} is a known residual. For simplicity and to connect with Appendix Figure E2, assume

$d_{ij} = 0$. Let $\Delta_{j,j^*} = E_{j^*} - E_j$ and write the *change* in trooper response moving from j to j^* as

$$\mathbb{E}[\theta_i^*(y)|j^*] - \theta_i(y) = \underbrace{\hat{\varrho}_{j^*} + \hat{\rho}_{j^*}(Y_i - E_{j^*} + \overbrace{\Delta_{j,j^*} \times \mathbb{E}[\tau_{k;k\geq 0}]}^{\substack{\text{Behavioral response} \\ \text{(from transfer)}}})}_{\text{Counterfactual response } (\pi^*_{ij^*})} - \underbrace{\big(\hat{\varrho}_j + \hat{\rho}_j(Y_i - E_j)\big)}_{\text{Observed response } (\hat{\pi}_{ij})} \tag{E13}$$

$$= \hat{\varrho}_{j^*} + \hat{\rho}_{j^*}(Y_i - E_{j^*}) + \hat{\rho}_{j^*}\Delta_{j,j^*} \times \mathbb{E}[\tau_{k;k\geq 0}] - \hat{\varrho}_j - \hat{\rho}_j(Y_i - E_j) \tag{E14}$$

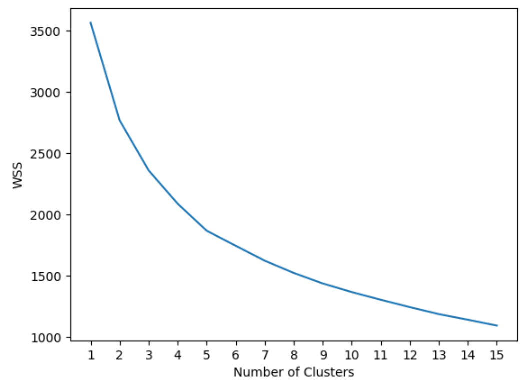
$$= \underbrace{\big(\hat{\varrho}_{j^*} - \hat{\varrho}_j\big)}_{\substack{\text{Level difference} \\ \text{across locations}}} + \underbrace{\big(\hat{\rho}_{j^*}(Y_i - E_{j^*}) - \hat{\rho}_j(Y_i - E_j)\big)}_{\substack{\text{Response to updated target} \\ \text{across locations}}} + \underbrace{\hat{\rho}_{j^*}\Delta_{j,j^*} \times \mathbb{E}[\tau_{k;k\geq 0}]}_{\substack{\text{Behavioral response to} \\ \text{updated target}}} \tag{E15}$$

The first term highlights a level difference between locations, holding fixed both the target and performance. The second term quantifies the change in response due to the change in target, taking into account any potential slope differences in the new location. The third term incorporates trooper behavioral responses due to changes in the target from the event-study results in the main text. In a simple Oaxaca-Blinder decomposition, the first term corresponds to the differences in intercepts. The second term (supposing for simplicity that $E_{j^*} = E_j$) represents the difference in the relationship between the observed response and relative performance (slope) across locations and the third term represents a movement along the relative performance curve in the new location.

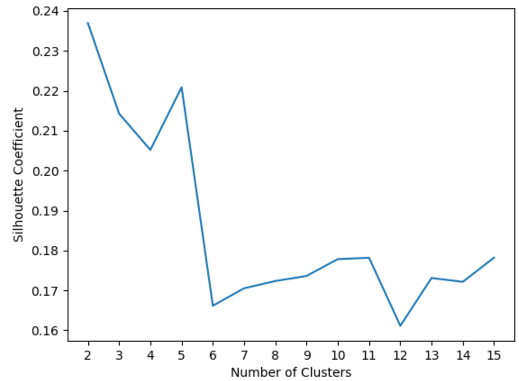
Appendix Figure E1: Testing for Optimal Number of Clusters

Troopers

Panel A: Within-Cluster Sum of Squares

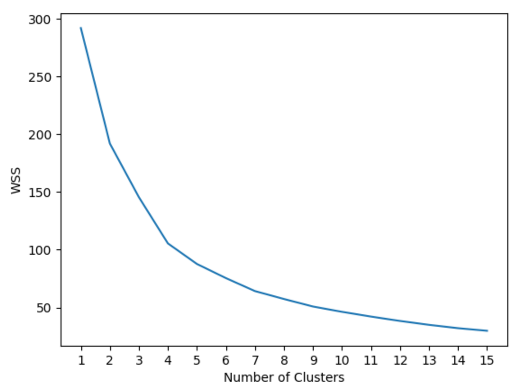


Panel B: Silhouette Coefficient

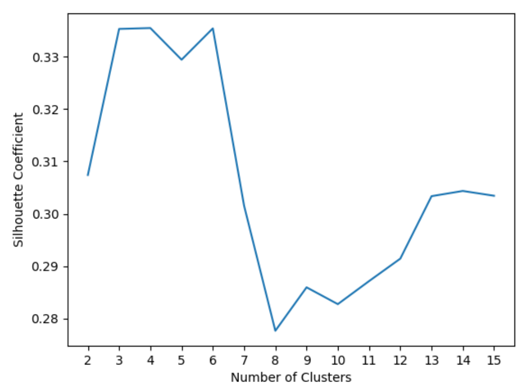


Locations

Panel C: Within-Cluster Sum of Squares

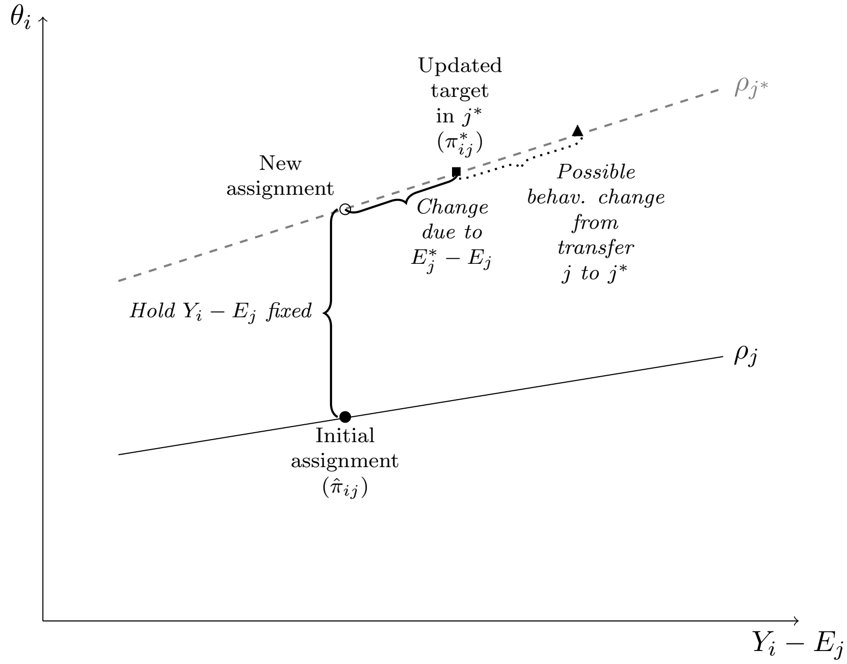


Panel D: Silhouette Coefficient



Notes: This figure reports the results of two exercises used to determine the "optimal" number of clusters. Panels A and C report the relationship between within-cluster sum of squares and number of clusters. Panels B and D report the average silhouette coefficient against the number of clusters.

Appendix Figure E2: Illustration of Counterfactual Posterior Mean Prediction



Notes: This figure illustrates a stylized example providing the graphical intuition behind the construction of predicted counterfactual responses. The x-axis is a measure of relative performance and the y-axis is the trooper posterior mean collision response. For visual clarity, I draw both the movement and behavioral transfer responses as oriented in the same direction.