NBER WORKING PAPER SERIES

PROFESSIONAL MOTIVATIONS IN THE PUBLIC SECTOR: EVIDENCE FROM POLICE OFFICERS

Aaron Chalfin
Felipe M. Gonçalves

Working Paper 31985
http://www.nber.org/papers/w31985

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2023

We are grateful to Bocar Ba, Jordi Blanes i Vidal, Jesse Bruhn, Amanda Geller, Kareem Haggag, Sierra Kuzava, Ilyana Kuziemko, Adriana Lleras-Muney, Bentley MacLeod, Michael Makowsky, Conrad Miller, Michael Mueller-Smith, Emily Owens, Nicola Persico, Till von Wachter, Emily Weisburst, and seminar participants at APPAM, Arnold Ventures Police Accountability Roundtable, Duke, FRB Philadelphia, NBER Summer Institute, Ottawa/Carleton Applied Micro Lab, Penn State, Princeton, SEA, SOLE, Stanford Law School, and UCLA for helpful comments. We thank the Dallas Police Department and the Dallas County Attorney General's office for sharing their records and Lieutenant Stephen Bishopp and Corporal Julie Martin for help with institutional details. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2023 by Aaron Chalfin and Felipe M. Gonçalves. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Professional Motivations in the Public Sector: Evidence from Police Officers
Aaron Chalfin and Felipe M. Gonçalves
NBER Working Paper No. 31985
December 2023
JEL No. J33,J45,K42

ABSTRACT

We study how public sector workers balance their professional motivations with private economic concerns, focusing on police arrests. Arrests made near the end of an officer's shift typically require overtime work, and officers respond by reducing arrest frequency but increasing arrest quality. Days in which an officer works a second job after their police shift have higher opportunity cost, also reducing late-shift arrests. Combining our estimates in a dynamic model identifies officer preferences over workplace activity and overtime work. Our results indicate that officers' private costs of arrests have a first-order impact on the quantity and quality of enforcement.

Aaron Chalfin
Department of Criminology
University of Pennsylvania
558 McNeil Building
3718 Locust Walk
Philadelphia, PA 19104
and NBER
achalfin@sas.upenn.edu

Felipe M. Gonçalves
Department of Economics
University of California, Los Angeles
8283 Bunche Hall
Los Angeles, CA 90095
and NBER
fgoncalves@econ.ucla.edu

1 Introduction

A key feature of the public sector is the delegation of decision-making to individual bureaucrats (Niskanen, 1971; Wilson, 1989). Government agencies are limited in their ability to explicitly dictate how workers behave on the job or to use high-powered incentives such as performance pay. These limits derive, in part, from concerns that public sector work is uniquely susceptible to distortion from extrinsic motives (Holmstrom and Milgrom, 1991; Dixit, 2002; Finan et al., 2017). Instead, governments seek to attract employees who are intrinsically motivated towards their work and who are afforded broad discretion over how to perform their jobs (Dal Bó et al., 2013; Ashraf et al., 2020). The preferences of bureaucrats, and how they are shaped by the balance of intrinsic and extrinsic motives, are thus key determinants of policy outcomes.

This paper studies the professional motivations of police officers, a group of bureaucrats engaged in particularly high-stakes work. We focus on the arrest decisions of patrol officers, who perform the majority of a department's enforcement activity and civilian engagement. After an officer makes an arrest, he typically must spend several hours processing the arrest, a requirement that can lead to working past his scheduled end of shift and the receipt of overtime pay. While an officer who is motivated to do his job well will value apprehending guilty suspects and not arresting innocent suspects, arrest decisions are also affected by the personal costs and benefits of overtime work. We examine the quantity and quality of arrests throughout an officer's shift, how they vary with the value of working overtime, and how they are shaped by officer preferences.

As in much of the public sector, performance incentives are rare in policing. Departments are limited in their ability to pay, promote, or fire officers on the basis of workplace activity. At the same time, scholars have documented instances in which extrinsic economic incentives have distorted police enforcement, such as revenue-motivated traffic enforcement (Makowsky and Stratmann, 2009, 2011) and increased drug arrests after the passage of asset forfeiture laws (Baicker and Jacobson, 2007). Some observers have also raised the particular concern that overtime pay might incentivize officers to make low-quality late-shift arrests in order to secure access to overtime pay, a practice known as as "collars for dollars" (Moskos, 2011). Our study is the first empirical assessment of the collars for dollars hypothesis, which we examine within a larger evaluation of officers' on-the-job preferences.

We leverage a uniquely expansive set of administrative data from Dallas, TX, linking information on Dallas Police Department officer shifts, overtime hours and off-duty work to 911 calls for service, arrests and county court records. We construct a dataset at the officer-by-date-by-hour level, where the main outcome is whether an arrest is made in a given hour. A crucial feature of our setting is that, within a given police district, officer shifts overlap throughout the course of the day, allowing us to empirically distinguish between time-of-day and time-of-shift patterns. Consistent with the institutional details of our setting, the data show that arrests made near shift-end are considerably more likely to lead to overtime pay and work. The average late-shift arrest is worth an additional \$180 in salary, or approximately two-thirds of an officer's gross daily earnings.

Our approach to measuring officer motivation builds on a key insight from Ash and MacLeod (2015), who study state appeals court judges. They find that judges facing fewer time constraints make higher-quality rulings, as measured by future citations. They leverage this shift in the cost of work to infer that judges are intrinsically motivated to write high-quality cases. We apply the same logic to our setting, where the personal cost to officers of making an arrest changes as they reach the end of their shift and on days with an off-duty job post-shift.

Our analysis proceeds in three parts. First, we document how enforcement activity changes throughout the course of a police officer's shift. Despite the presence of strong monetary incentives to make late-shift arrests, we find empirically that the frequency of arrests *decreases* by approximately 30 percent towards the end of an officer's shift. Conversely, the quality of arrests, as measured through both the likelihood of a criminal court conviction and the probability of an incarceration sentence, *increases* throughout the shift. We provide evidence ruling out the importance of formal or informal department policies discouraging officers from taking overtime, the incapacitative effect of early-shift arrests, officer fatigue, and several other alternative explanations for these patterns in the data. Taken together, the results suggest that officers have an aversion to overtime work, expressed through an increase in the threshold used to make an arrest. This finding is consistent with work by Chan (2018) showing a "slacking off" effect for emergency-room doctors. The increase in court convictions from late-shift arrests suggests that officers value arrest quality at the margin, disproportionately abating lower rather than higher-quality arrests.

Second, we evaluate how responsive officers are to *changes* in their extrinsic incentives by examining how they respond to variation in the relative value of overtime work. We exploit the fact that officers face an increased opportunity cost of overtime work on days when they have a scheduled off-duty job shift after their police shift. Off-duty work is a common practice among officers and most often consists of private security work at locations like commercial establishments or residential communities. We find a significant though quantitatively small reduction in the propensity to make late-shift arrests on days when officers have a post-shift off-duty spell. Officers are more responsive to longer off-duty shifts, consistent with these shifts providing a higher opportunity cost for late-shift arrests.

In the final part of the paper, we combine our empirical estimates with a novel dynamic model to formally estimate officers' preferences over different types of arrests, overtime work, and pay from overtime and off-duty work. We build upon the theoretical work of Prendergast (2003, 2007) and model officers as having one margin of decision-making in each hour of their shift, which is the guilt threshold above which they arrest a suspect. Officers face a probability of needing to process the arrest, which may both lead to overtime work and generate different choices of guilt thresholds throughout the shift. We identify the model by matching the regression-based time paths of arrest frequency, court conviction conditional on arrest, overtime pay for arrests in each hour of the day, and off-duty impacts. This final set of moments allows us to identify officers' dollar value for their arrest activity, since it shifts the opportunity cost of arrests by a specific dollar amount.

Our estimates suggest that while officers appear to highly value their leisure time, they also value making accurate arrest decisions. Specifically, their behavior implies a value of several hundred dollars of correctly not arresting an innocent individual and a value of roughly four thousand dollars of correctly arresting a guilty individual. These estimates are large relative to the magnitude of income derived from overtime and off-duty work, indicating that police behavior is unlikely to be significantly distorted by monetary considerations. An implication of this analysis is that the decline in arrests towards the end of an officer's shift is explained by a substantial *non-monetary* cost to working additional police hours.

We use the model estimates to conduct a number of counterfactual exercises. We find that even large increases in overtime pay lead to only modest changes in behavior — the elasticity of arrests with respect to expected overtime pay is approximately 0.04. To consider the overall impact of overtime aversion, we compare the observed arrest patterns to a counterfactual in which officers are indifferent between working late and going home. Absent overtime aversion, patrol officers would increase arrest frequency by 70%, and the conviction rate of arrestees would decline by 7%, indicating that overtime considerations play an important role in shaping the overall quantity and quality of enforcement. Further, when officers are overtime indifferent (i.e. purely "altruistic"), their arrest propensity would, in fact, *increase* towards the end of a work shift, accompanying a decline in the quality of arrests. These effects are driven by the fact that early arrests carry the opportunity cost of precluding officers from making later arrests. This striking result — that altruistic officers would appear to be engaging in "collars for dollars" — highlights the importance of formally modeling the dynamic decision that officers face in order to infer their preferences.

A final insight from the model concerns the implied equity and efficiency impacts of officer overtime preferences. Aversion to overtime work shifts officers along a hypothetical production possibility frontier away from arresting innocent individuals (Type I error) and towards making fewer arrests of guilty individuals (Type II error), generating a significant *equity* impact. However, officers' observed behavior is very close to the frontier, suggesting that there is minimal *efficiency* loss as a result of officers' aversion to overtime work. This surprising result derives from the fact that, while a within-shift decline in officer arrest rates changes the probability of guilt of the marginal arrestee (as indicated by our first analysis), these magnitudes are not large enough to generate an efficiency cost from variation in arrest frequency over the shift.

Several features of our setting are ideal for measuring workplace motivation and the preferences of public sector agents more broadly. First, our data are exceptionally detailed, allowing us to observe actual employee behavior at an hourly level and avoiding the reporting concerns involved with subjective measures of performance such as promotions or evaluations. Second, we use a shift in the cost to a workplace activity that can be denominated in dollar terms, allowing us to estimate the dollar value placed on an employee's professional motivations – in our context, making an arrest of a guilty individual and avoiding an arrest of an innocent individual. Third, officers face relatively low-powered incentives. Dallas officers are promoted based solely on a civil service examination, and their pay is based on rank and experience, neither of which are tied to arrests. We therefore argue that the value officers place on arrest activity reflects a combination of intrinsic pro-social motivations (Bénabou and Tirole, 2003, 2006) and workplace social norms (Bandiera et al., 2009, 2010), which we collectively refer to as officers' professional motivation.

Contribution to Literature — Within the economics of crime, there is rich evidence on the impacts of hiring more police officers (Evans and Owens 2007; Chalfin and McCrary 2018; Mello 2019; Weisburst 2019; Chalfin et al. 2022) and changing the presence of officers (Draca et al., 2011; Blanes i Vidal and Kirchmaier, 2018; Weisburd, 2021). Much less is known about the determinants of enforcement and police behavior for a given set of officers, despite a growing appreciation of the importance of discretion in policing (Ba et al., 2021; Weisburst, 2022; Gonçalves and Mello, 2021, 2023; Abrams et al., 2023). We contribute to this emerging literature, which has studied how enforcement responds to factors such as public scrutiny (Ba and Rivera, 2019; Premkumar, 2020), managerial directives (Bacher-Hicks and de la Campa, 2020), department financial incentives (Makowsky and Stratmann, 2009, 2011; Makowsky et al., 2019), workplace fear (Cho et al., 2023), and officer neighborhood preferences (Ba et al., 2021).

More broadly, our work relates to the literature exploring the balance between the competing motivations of public sector workers. Previous research has examined the interplay between bureaucrats' intrinsic motivation towards their work and factors such as the risk of professional sanction (Prendergast, 2001; Leaver, 2009), promotion concerns (Bertrand et al., 2020), and political and electoral considerations (Berdejo and Yuchtman, 2013; Lim, 2013; Ash and MacLeod, 2015; Cohen and Yang, 2019; Spenkuch et al., 2023). Our key contribution is to show how a ubiquitous labor consideration – working past the end of a shift – can impact the quantity and quality of government work in settings where bureaucrats have significant discretion.

We also build on the labor supply literature on workers with shift-length discretion (Oettinger, 1999; Farber, 2015; Cook et al., 2021). Most closely related is Chan (2018), who finds that emergency department shift schedules induce physicians to "slack off" at the end of their workday by accepting fewer patients near end of shift and spending less time with patients that they do meet with. We show that labor supply considerations are also important in the public sector, even in a setting where workers are nominally unable to set their hours.

Methodologically, our work contributes to a small literature on the structural modeling of criminal justice settings (İmrohoroğlu et al., 2004; Conley and Wang, 2006; Fu and Wolpin, 2018). Most closely to us, Adda et al. (2014) study a cannabis depenalization policy implemented in one London neighborhood, and they combine empirical estimates of the policy's impact with a structurally-estimated model to consider a counterfactual that expands the policy citywide. Further afield, our work also relates to the structural behavioral literature that uses quasi-experimental evidence to estimate behavioral preferences in various settings (DellaVigna et al., 2012, 2016; DellaVigna and Pope, 2018).

2 Institutional Background

Police officers typically have broad discretion over whether or not to make arrests (Owens, 2020), which stems primarily from two sources. First, officers must inevitably make subjective judgments of whether they have probable cause to arrest a suspect. Second, they must decide how proactive to be in searching for and identifying criminal activity. These factors create numerous opportunities for officers to engage in selective enforcement and to potentially exploit the availability of overtime pay or to avoid overtime work.

Processing an arrest requires significant effort. In all cases, officers must prepare a detailed arrest report. The suspect must typically be taken into custody, and, in Dallas, officers are required to remain with the suspect during the entire jail intake process. Even arrests for low-level charges typically require a minimum of several hours of work.[1] If arrest processing requires that an officer work past the end of her regular work shift, she will receive overtime compensation at a rate of 150% her base hourly wage. Officers can choose whether to receive direct pay or compensatory time, in which case they receive time and a half that can be used at their discretion. While in some cities, arrests may not lead to overtime if certain processing steps are not immediately available — for example, sending paperwork to the district attorney's office, which is closed at night — in Dallas, all steps must be completed immediately after the arrest has been made.[2]

Do officers seek to receive overtime pay by making late-shift arrests? While we might

[1] Arrests may also lead to court time on another date. While court overtime is potentially a motivating factor in making arrests, our identification strategy nets out this motivation, focusing solely on changes in the incentive to arrest late versus early in an officer's shift.

[2] Conversations with personnel at the Dallas Police Department suggest that there are no formal or informal limitations on the amount of overtime work allowed. This is consistent with our data, which show a great deal of heterogeneity across officers in the number of overtime spells worked.

expect increased enforcement at the period when overtime payment is most likely, overtime work is not costless for an officer. Ethnographic work has noted that officers' private motivations may not be to maximize overtime pay and that officers may have a variety of personal reasons to avoid making late-shift arrests including a "hot date," a sick baby, a college class or, critically, an off-duty job, all of which would be considerably disrupted by a late-shift arrest (Moskos, 2008; Linn, 2009). Likewise, officers, like other public sector workers, may be intrinsically motivated and find it cognitively or morally costly to make an unjust arrest. The extent to which officers make late-shift arrests will ultimately depend on the relative values they place on their workplace activity (i.e., their professional motivation), overtime pay, the ardor of overtime work, and their planned activities after their shift.

3 Data

Our data come from multiple criminal justice agencies in Dallas, Texas, and cover information for the period spanning January 2015 through March 2019. Arrest records come from the Dallas Police Department (DPD). These data provide information on all arrests made by DPD officers and include the date, time, location, and all of the arrest charges. The arrestee is identified by his or her full name and age at the date of the offense, and the arresting officers are identified by their badge numbers.

To identify whether an arrest leads to a criminal conviction or a sentence, we use criminal court data from the Dallas County Attorney General's office. Each case reports defendant's full name, date of offense, criminal charges, and final case disposition. We link these records to our arrest data using a fuzzy match on first and last name and offense date (Lahiri and Larsen, 2005; Tahamont et al., 2020).[3]

We consider an individual to be convicted if their case is not dismissed and they are not found innocent by judge or jury. Cases that end in a "non-adjudication of guilt" plea, where an individual does not contest the charge but does not formally admit to guilt are considered to be convictions. Our measure of a criminal sentence corresponds to whether

[3]We link records by first and last initial, remove links where the first and last names deviate by more than two characters or the offense dates are more than two days apart, and then keep the strongest link for each arrest record. Our matching aligns closely with an exact matching algorithm: for the set of all first and last initial matches, there is a large spike of cases where the offense dates coincide. When the offense dates and initials match, the majority of names have no disagreement between the two data sets.

any prison or jail sentence is reported in a case, including in cases where the individual is assigned a sentence probation in lieu of serving prison or jail time. In cases in which the arrest does not link to any court records, we conclude that the suspect was not found guilty and was therefore not sentenced.[4] Our measures of arrest quality are court conviction and sentencing, a common approach in both the criminology and economics literature (Forst, 1982; Ater et al., 2014; Weisburst, 2022). While convictions and sentencing do not perfectly reflect an arrestee's guilt or the seriousness of the criminal charges, previous studies suggest that wrongful imprisonment, while highly costly, is relatively rare (Cassell, 2018; Loeffler et al., 2019).

To link every arrest to the relative time during an officer's shift when the arrest was made, we use a record of officer overtime payments from the police department. These data record the date, number of hours, and total payment of the overtime spell. Each entry includes an officer's regular assignment location, regular shift hours, and days of week on which they work, which we use to calculate the number and range of hours in which an officer is regularly working.

To identify arrests that were precipitated by a citizen call for service, we also use data on 911 calls for service from the Dallas Police Department. For each call, these data record the date, time, location, call description, responding officers, and incident numbers for any reports written from the call. Through the incident number, we link these data to arrest records to identify whether an arrest is made. We use these data to identify officers who are working patrol on a given day and are used in our robustness section to restrict attention to arrest propensity from calls.To measure secondary work activity–work that is performed for private employers outside of an officer's official police duties–we use a database of all days when an officer has an off-duty job. These records include the employer, date, shift start and end time, and officer name and badge. Henceforth, we refer to these assignments interchangeably as off-duty or secondary jobs.

[4]The court records comprise all cases seen in the county criminal court system. This system oversees all cases with charges of a Class C Misdemeanor or higher. However, it is not uncommon for arrests with a lower charge in our data to appear in the county court data. Among Violations (below misdemeanor), 10 percent of arrests are linked to a case in the county court. We therefore construct the indicators for guilt and sentenced for all of our arrests and note that the outcome measures guilt but also to some extent reflects the severity of the arrest.

3.1 Sample Construction

The unit of observation for our baseline sample is a date and hour in which an officer is working his or her regular shift. From the overtime data, we observe an officer's regularly scheduled work hours and days of work for each DPD officer. We restrict attention to officers whose listed assignment is one of the seven patrol divisions or the central business district, excluding officers working in the traffic unit and other specialized units such as narcotics, violent crimes, and tactical response and support. One limitation of our data is that we only know an officer's regular shift from their overtime payments. If an officer has two consecutive overtime payments that list different regular shift assignments, hours, or days off, we do not know on which day they changed their assignment. We thus exclude all days-hours in between overtime payments that list different regular shifts. To avoid days where an officer may have called in sick or is otherwise not working patrol, we keep only shifts where an officer appears in the 911 data as taking at least one call. We also restrict our baseline sample to shifts that are eight hours in length, removing the small share of shifts that are nine or ten hours. In Section 5, we expand the sample to include these longer shifts to separately identify coefficients for time from end of shift end time from start of shift. In Appendix A, we provide more detail on the data construction and how the sample size changes with each restriction to create our analysis sample.

3.2 Summary Statistics

In Table 1, we present descriptive data on the 8,614 DPD officer-years in our data. The average DPD officer has served approximately 11 years on the job. The department is fairly diverse — 26 percent of officers are Black and 22 percent are Hispanic. The average officer in our sample earns approximately $81,000 per year in base salary and earns an additional $3,493 in overtime pay. While overtime pay accounts for approximately four percent of an average officer's compensation, overtime pay varies considerably among officers, ranging from zero to more than $84,183 (SD = $11,796). The data are thus consistent with the department's official policy, which does not have a formal cap on overtime pay. To the extent that informal department policies might discourage officers from working overtime spells, these constraints appear to be mostly non-binding.

On average, patrol officers respond to 624 emergency calls per year (minimum = 1, maximum = 2,511). Officers make, on average 26.7 arrests per year, 34 percent of which resulted in a misdemeanor or a felony conviction. 65 percent of arrests by the average officer are initiated by a civilian call for service; the remainder are initiated by the police officer. Table A-1 provides descriptive statistics on the types of arrests that DPD officers make. The most common type of arrests are for warrants (28 percent), assault (15 percent), disorderly conduct (14 percent), narcotics and drugs (13 percent), and public intoxication (3 percent). Overall, 21 percent of arrests were for felony offenses. An average arrest has a 45 percent chance of leading to an overtime spell.

3.3 Descriptive Analysis

Before introducing our empirical strategy, we present descriptive evidence from the raw data. We begin by establishing that the probability of receiving overtime pay varies by the shift-hour in which an arrest is made. In Figure 1, we plot the probability of receiving overtime pay (Panel a) and the number of overtime hours conditional on overtime receipt (Panel b) against the shift-hour of arrest. The -8 hour corresponds with the first hour of an officer's shift; the -1 hour corresponds with the final hour of the officer's shift.

In Panel (a), the blue line indicates that the probability of receiving overtime pay when no arrests are made is approximately 29 percent which reflects the fact that arrests are not the only reason why officers work overtime hours. Nevertheless, the probability of working an overtime spell is considerably higher on days in which an officer makes an arrest. The probability of receiving overtime pay having made an early-shift arrest is approximately 39 percent, and there is a near monotonic increase in the probability of overtime receipt throughout an officer's shift. During the final hour of an officer's shift, the probability of receiving overtime pay is approximately 70 percent.[5]

Panel (b) plots the expected number of overtime hours received conditional on working any overtime, separately by hour of arrest, and shows a small increase for arrests made closer to the end of shift. Given that the average officer's annual salary exclusive of overtime is approximately \$81,000, mean hourly pay is about \$40. The overtime payment rate is 1.5 times an officer's wage, so a typical arrest that receives overtime is worth approximately

[5] In Figure A-2, we present the analysis having conditioned on officer, division × day-of-week × hour, division × day-of-week × shift and division × month-year fixed effects.

$164.

We next explore how the quantity and quality of arrests vary throughout the shift. Figure 2 plots the probability of an arrest (Panel a) and the probability of a misdemeanor or felony court conviction conditional on an arrest (Panel b) against the hour relative to the end of an officer's shift. The probability of an arrest in a given shift-hour is low, peaking at approximately 2.2 percent in the second hour. Interestingly, after initially rising, the probability of an arrest falls considerably — from approximately 2.2 percent to approximately 0.9 percent in the final hours of an officer's shift, a decline of 60 percent. To address the possibility that the decline in arrests is mechanically due to officers processing prior arrests, we also plot the probability of an arrest in a given hour *conditional on not having made an arrest previously*. Given that the shapes of the two curves in Panel (a) are strikingly similar, it is unlikely that the large decline in arrest activity throughout the shift is an artifact of incapacitation from earlier arrests.

If the number of arrests falls later in the shift, it is natural to consider whether the quality of arrests is also changing. A decline in the quality of arrests at the end of the shift would be consistent with the idea that officers reduce their evidentiary threshold for making late-shift arrests, perhaps in order to secure access to overtime pay. An analysis of the arrest quality-shift hour gradient can be thought of as a conceptual analog to the oft-cited "hit rate" test for the presence of racial bias in police stops and searches (Knowles et al., 2001; Feigenberg and Miller, 2022). The presumption is that, in the absence of bias, the perceived guilt threshold for treatment should be the same across racial groups. In our setting, the presumption is that an officer who maintains the same standard for making an arrest throughout their shift should have the same success rate of an arrest in each shift-hour, where success is measured using court convictions.

Panel (b) of Figure 2 shows that the probability that an arrest is sustained by a conviction actually *increases* throughout the work shift. Early shift arrests have a conviction rate of approximately 31 percent, which increases to approximately 35 percent among arrests made at the end of an officer's shift. In Section 5.1 we subject this analysis to further scrutiny using a series of regressions which we describe in the following section.

4 Empirical Methods

4.1 Main Analysis

Our empirical strategy is designed to identify how an officer's arrest propensity changes with his shift-hour–the number of remaining hours in the officer's shift. In order to control for differences throughout the day in criminal activity, we compare officers who work in the same community during the same hours but are at different points in their shift. This approach is possible because DPD officers work overlapping shifts within each sector to avoid significant disruptions during shift changes.[6] We provide a sense for the overlap in officer shifts in Figure 3. For each of the most common work shifts, the figure plots the average number of arrests for each hour in the day. While we plot only the most common shifts, we can see both that there are declines in arrest propensity within each shift and that there is substantial overlap in shifts throughout the day.

Our unit of observation is the officer-by-date-and-hour, which we denote by the double it. As discussed earlier, we include all days during an officer's regular shift where they are assigned to a patrol unit. The primary outcome of interest Y_{it} is an indicator variable for whether an arrest is made, and we run the following regression:

$$Y_{it} = \sum_{k=-7}^{-1} \alpha_k \text{ShiftHour}_{it}^{k} \mu_i + \phi_{dwh} + \theta_{dws} + \xi_{dm} + \epsilon_{it} \tag{1}$$

The coefficients of interest are α_k in (1) which tell us the relative probability of an arrest at each hour of a shift. For instance, α_{-7}, the coefficient on the hour that is 7 hours from the end of an officer's shift measures average arrest incidence relative to the first hour of an officer's shift. Likewise, α_{-1} measures arrest incidence during the final hour of an officer's shift.

In order to identify shift-hour effects, we control extensively for potential confounders using a set of granular fixed effects. We include interacted division $\times$ day-of-week $\times$ hour-of-day fixed effects, ϕ_{dwh}, in order to control for division-specific differences in arrest frequencies across all hours of the week. To account for secular trends in the crime environment in each division, we include interacted division $\times$ year-month fixed effects, ξ_{dm}. To account for

[6] Service disruptions can have large effects on clearance rates. See e.g., Mastrobuoni (2013).

persistent differences across shifts in arrest activity, we include division × shift × day-of-week fixed effects, θ_{dws}. These fixed effects are important insofar as officers in each shift may be routed to different sorts of service calls. Finally, all models condition on officer fixed effects, μ_i, to allow for time-invariant differences in arrest incidence across officers. In Section 5.3, we consider various alternative choices for fixed effects, including a fully-saturated specification that controls for division-by-date-by-hour. We cluster standard errors at the division-by-month level to account for arbitrary serial correlation in outcomes among arrests in the same location and time period.

Two additional analyses merit description. First, we re-estimate Equation (1) focusing on the probability that a 911 dispatch call leads to an arrest in a given shift-hour. These models are estimated at the call-by-officer level. We thus additionally control for 911 call type fixed effects to account for differences in arrest probabilities for different types of emergency calls. We refer to our baseline sample as our "hourly sample," and we refer to this sample of 911 calls as our "dispatch sample." Second, in order to test whether the probability of a criminal conviction and a sentence — both measures of arrest quality — differ throughout the work day, we let our unit of observation be an arrest-by-officer and regress an outcome of the arrest on the vector of shift-hour indicators, conditional on the same fixed effects described in Equation (1).[7]

4.2 Effect of Off-Duty Work

Equation (1) allows us to evaluate how the quantity and quality of arrests change throughout an officer's shift. While these estimates provide a measure of officers' relative preference or aversion to working overtime, they do not allow us to separately identify the importance of the monetary and non-monetary components of overtime work. Decomposing preferences into these two elements is the key to understanding how officer behavior responds to changes in economic incentives.

We therefore seek to identify the impact of *changes* in the relative value of overtime work. To do so, we exploit the fact that officers regularly work second "off-duty" non-police jobs. Off-duty work changes the opportunity cost of an overtime spell, and we test whether hourly

[7] In our primary models, we do not condition on the arrest charge which is potentially endogenous, though we later estimate the model separately by arrest severity (violations, misdemeanors, and felonies).

arrest propensity varies according to whether an officer is scheduled to perform off-duty work on a given day.

Our off-duty data include work that is both regularly occurring and idiosyncratic. Officers are required to notify their supervisor of their off-duty shift and receive approval beforehand. However, a 2018 audit of the department's off-duty activity found that, in practice, officers regularly documented their shift and received approval after the occurrence of the shift (Smith, 2018). A resulting statistical concern is that, if processing an arrest requires that an officer work overtime and cancel a planned off-duty shift, that shift may never be officially reported. Such an occurrence may lead to a mechanical relationship between arrest activity and realized off-duty shifts even in the absence of a behavioral response.

To avoid such a concern, we focus on an officer's *regular* off-duty schedule. Specifically, we split an officer's calendar into days of the week and quarters, and we say that an officer has a regularly planned off-duty shift on that day and quarter if more than 25% of those days has an off-duty spell. We then link these off-duty assignments to the officer's police shifts to measure whether they have a regularly scheduled off-duty shift on that day. We construct indicators for officers working off-duty before their shift, OD^b_{it}, and after, OD^a_{it}.

As in Equation 1, our unit of observation is an officer-date-hour. We regress our key outcomes on the occurrence of off-duty work before or after an officer's police shift. Because we expect impacts to differ by time into shift, we interact the off-duty work indicators by whether an officer is in their first four hours, Early_{it}, or last four hours, Late_{it}:

$$\begin{aligned} Y_{ith} \quad = \quad & \sum_{q\in\{b,a\}} \gamma^q_E OD^q_{it} \times \text{Early}_{ith} + \sum_{q\in\{b,a\}} \gamma^q_L OD^q_{it} \times \text{Late}_{ith} \\ & + \mu_i + \phi_{dwh} + \theta_{dws} + \xi_{dm(t)} + \epsilon_{ith} \end{aligned} \tag{2}$$

In (2), OD^q_{it} is equal to 1 if the officer has a regularly-occurring off-duty shift of type q on that day, where the superscript $q \in \{b, a\}$ refers to either pre-shift or post-shift off-duty work. We additionally control for officer, division-by-day of week-by-hour, division-by-shift and division-by-year-month fixed effects.

Note that γ^q_E and γ^q_L are reduced form coefficients for the relationship between arrests and the presence of scheduled off-duty work and, as such, are lower-bound estimates of the effect of *realized* off-duty work on arrest activity and overtime spells. In Table A-2, we report first stage coefficients from a series of regressions of actual off-duty employment on regular off-duty employment, predicting pre- and post-shift off-duty work using separate models. The entries along the diagonals report the relevant first stage coefficients while the off-diagonal elements establish that predicted pre-shift off-duty work is far less correlated with actual post-shift off-duty work and vice versa. For both before and after off-duty work, the estimated coefficients are approximately 0.5, indicating a strong correspondence between predicted and actual off-duty work.

5 Results

5.1 Main Results

We begin our discussion of the results by formally testing whether the frequency and quality of arrests vary over the course of an officer's shift. Regression coefficients from Equation (1) are plotted in Figure 4. Panel (a) shows that the arrest probability initially rises by approximately 0.6 percentage points relative to the first hour arrest frequency of 1.3 percent, eventually rising to over 0.8 percentage points during the fourth hour of the shift before declining considerably. Consistent with the descriptive findings presented in Figure 2, arrests are approximately 35% less frequent in the final hour of the shift than at mid-shift.

Panels (b) and (c) presents estimates of Equation (1) where the sample is arrests that are made, and the outcomes are a guilty conviction and a criminal sentence in court, respectively. We observe that the probability of a conviction climbs throughout the day, rising by approximately 8 percentage points relative to the first shift hour. The probability of a criminal sentence rises by approximately 5 percentage points throughout the work day. These increases correspond with a 26% increase in the conviction rate and a 28% increase in the probability of a sentence throughout the officer's shift. For both measures, an F-test confirms that the probability in the final three hours of the shift is significantly higher than the probability in the first hour of an officer's shift ($p < 0.05$).

Next, in Panel (d) we examine the shift-time patterns for whether an arrest is for a felony versus a less serious non-felony offense. If, consistent with the arrest frequency and court outcome patterns, officers are increasing their severity threshold for making an arrest as they get near shift-end, we would expect to see an increase in the felony share of arrests. Indeed, we observe an *increase* in the share of arrests for a felony, rising from 17% in the first hour of the shift to 23% in the final hour.

5.2 Heterogeneity

We next turn to exploring heterogeneity in these results across various dimensions. We begin by considering whether the incidence and quality of arrests varies according to the criminal seriousness of the arrest charge. Figure A-4 plots arrest incidence and the conviction rate by shift-hour separately for violations (low-level crimes that are considered by less criminally serious than misdemeanors) as well as misdemeanor and felony arrests. Estimates for all arrest types follow a similar pattern — there is an initial increase in the incidence of arrests early in an officer's work shift followed by a decline during the final hours of the shift. However, the decline in arrest incidence is slightly larger for the least serious crimes — violations and misdemeanors — than for felonies.[8] This result is consistent with the idea that police officers have greater discretion over arrests for violations and misdemeanors than they do for felonies (Smith and Visher, 1981; Linn, 2009). Notably the results are inconsistent with the idea that police officers search for low-level arrests to make at the end of their shift.

Next, recognizing that recent literature finds evidence of racial disparities in policing (Gonçalves and Mello, 2021; Feigenberg and Miller, 2022), we consider whether officers make a larger number of late-shift arrests or lower quality late-shift arrests of minority citizens. We address this possibility in Figure A-5. The left panels on top and bottom show the shift-hour effect on share Black and Hispanic arrestees, respectively. We find no evidence that the composition of arrestees becomes more Black or Hispanic near the end of the shift, though there is weak evidence that officers arrest relatively *fewer* Hispanic individuals as their shifts progress. The right-hand panels present the regression coefficients for court conviction separately by Black and Hispanic arrestees to test for whether the quality of arrests declines for either group. We find no evidence of a decline in court convictions for Black or Hispanic arrestees throughout officers' work shifts.

While our principal findings run contrary to the hypothesis that officers make additional low-quality arrests at the end of the shift in order to receive overtime pay, it is still possible that a fraction of officers engage in this practice even if the behavior of these officers is not

[8]The felony results are nearly identical when we exclude felony drug charges.

detectable in the aggregate data. In Appendix B, we investigate heterogeneity across officers in their late-shift arrest behavior, focusing on the officers with a disproportionately high share of arrests at the end of their shift. While a small minority of officers increase their arrest activity as the shift-end nears, their late-shift arrests are of similar composition and quality to their early-shift arrests.

5.3 Robustness

We argue that the observed patterns of reduced arrest frequency and increased severity of arrests made near shift-end reflect officer preferences and, in particular, officers' aversion to overtime work. In this section we test the validity of our empirical design and present a series of tests for whether our results are driven by officer preferences rather than an artifact of either the data or institutional practices.

Validity of Empirical Design — The key threat to identification in estimating Equations (1) and (2) is unobserved differences in the criminal environment that may be correlated with an officer's shift-hour or off-duty work schedule. With respect to Equation (1) a potential concern is that supervisors may be more likely to have officers work on patrol rather than doing some other form of work in periods in which crime is high. Therefore, certain hours may have a large share of officers working at the beginning of their shift and simultaneously experience high crime that generates higher arrest propensities. A related concern for our off-duty analysis is the possibility that officers who plan to work a second job post-shift only agree to take 911 calls and work on patrol if they expect to not see many serious calls.

Our set of fixed effects partially address these concerns by flexibly controlling for differences in arrest propensities across division $\times$ day-of-week $\times$ hour and division $\times$ day-of-week $\times$ shift, addressing all persistent differences across divisions in their criminal environments and shift-specific practices and their related variation in shift assignments and off-duty schedules. Likewise, our division $\times$ month fixed effects address changes in the environment over time.

However, in order to further probe the validity of our design, we conduct an imperfect but informative balance test for our main variables of interest. Utilizing the same regressions described in Equations (1) and (2), we place the logarithm of the total number of 911 calls and high priority 911 calls in the officer's division-hour on the left-hand side of the equation and test for whether officer shift-time predicts call volume. We present these results in Table A-3. After accounting for the set of fixed effects, the coefficient on a particular shift-hour tells us whether a division-hour with a higher than average share of officers in that shift-hour also has a higher than average number of calls. If idiosyncratic variation in the composition of officer shift-times is uncorrelated with calls, the shift-hour coefficients should be jointly insignificant, which we test using an F-test. With respect to overall call volume, the F-test rejects the null hypothesis that call volume is unrelated to shift-time. However, the effects are very small in magnitude, as our shift-hour coefficients are never larger than 1 percent per hour. When we focus on high-priority calls, the p-value on the F-statistic is insignificant, indicating that there is little evidence that high-priority calls for service vary with the composition of officer shifts.

The last two columns of Table A-3 test for a relationship between call volume and officer off-duty obligations. Similar to our first two columns, the coefficient on an off-duty indicator tells us whether a division-hour with a higher than average share of officers working off-duty also has a higher than average call volume. The joint F-test for column (3) also rejects the null hypothesis of insignificant coefficients, though with similarly small estimated quantities. With respect to high priority service calls, the p-value on the F-test is 0.88, indicating little evidence against balance.

Since call volume is an aggregate variable that does not vary across officers in the same division and hour, these balance tests are an imperfect check on the validity of our design. However, the weak relationships reported in the table are consistent with the idea that the criminal environment faced by officer working in different shifts is close to being equivalent.

We conduct two more analyses to further probe the concern that criminal activity varies throughout officers' shifts. In Figure A-6, we re-estimate our baseline analyses with the inclusion of the log volume of all calls and serious calls. The point estimates are identical to those from our main analyses. Further, as we describe later in this Section, we estimate a version of Equation (1) where the unit of observation is calls taken by officers and where we directly control for type of call, and we find a similar decline in arrest frequency near shift-end.

Functional Form — We next test the robustness of our results to choice of fixed effects. We re-estimate alternative versions of Equation (1), all of which include officer fixed effects but vary the ways in which we account for the importance of place, work duties and time. These estimates are also presented in Figure A-6. In all specifications, we find a similar pattern of arrest frequency reduction. We also observe a broadly consistent pattern for court convictions. The only exception is with the inclusion of the most granular controls that condition on division-by-date-by-hour. Including these effects means the shift-time court conviction patterns are only identified from arrests made in the exact same division, date, and hour, which are a small subset of observations.[9] The court conviction rate in later shift-hours is no longer statistically different from the first hour of the shift, but the effects are still statistically indistinguishable from our baseline patterns.

Alternative Explanations — Next, we consider whether there are alternative explanations for the decline in arrest activity that we observe towards the end of an officer's shift. One possibility is that rather than changing behavior due to an aversion to overtime work, officers reduce arrest frequency because they are fatigued from having already worked for several hours. A fatigue effect could lead to a decline in arrest propensity and an increase in the quality of realized arrests, which we would erroneously ascribe to an aversion to overtime.

To differentiate between changes in arrests due to distance from end of shift and distance from start of shift, we estimate a version of Equation 1 on an expanded sample that also includes officers who work nine-hour and ten-hour shifts. We add to our specification a variable for number of hours into the shift, which captures drivers of time-into-shift effects like fatigue. To increase precision, we include separate indicator variables for only the last four hours of a shift, so that the remaining hours are the leave-out group and contribute to identifying the time-into-shift effect.The results of this analysis are presented in Figure A-3. We plot coefficients from regressions with and without the time-into-shift variable. In both specifications, there is a clear decline in arrest propensity as officers reach the end of their shift. The robustness to including a time-into-shift effect suggests that our results are not due to a fatigue effect. In Section 5.4 we further show that when officers perform off-duty work before their police shift, their arrest activity shows at most a small decline, further indicating that fatigue is not an important driver of arrest activity.

Another potential concern is that a decline in late-shift arrests could be an artifact of either a formal or informal departmental policy that routes officers to fewer calls for service

[9] Of the 90,016 observations in our baseline set of arrests, only 12,482 occur in a division-date-hour where two officers make arrests and are at different points in their shift.

towards the end of their shift. In order to address this possibility, we ask whether an officer's arrest propensity declines *conditional on taking a call for service*. We explore this question in Figure 5 which plots the regression-adjusted probability of an arrest throughout an officer's shift, where the unit of observation is a citizen 911 call for service. As in Figure 4, the relevant shift-hour coefficient is plotted on the y-axis with the hour relative to the end of the shift plotted on the x-axis. Models continue to condition on all of the standard fixed effects in Equation (1), though, since the model is estimated at the service call level, we additionally condition on call type fixed effects. Consistent with the aggregate results, arrests decline significantly — by approximately 20 percent — throughout the officer's shift, thus indicating that our main results are not an artifact of a change in the volume of service calls at the end of the workday.

Focusing on the conditional probability of an arrest given a call for service allows us to address several additional threats to identification. First, we might be concerned that the decline in late-shift arrest activity could be the result of the incapacitative effect of early-shift arrests. That is, arrests might decline mechanically throughout the workday as officers are removed from circulation after having made an arrest in the early in their shift. By estimating the probability of an arrest conditional on a call for service, we remove the source of this concern since all of the officers taking service calls are, by definition, available to make arrests.

A second concern is that there may be measurement error in arrest timestamps. To the extent that there are systematic differences between the timing of an officer's decision to take a suspect into custody and the timestamp of the arrest in our data, we might be concerned that some arrests made later in an officer's shift might be misclassified as having occurred either earlier in the workday or after the officer's official workday has ended. Since the models in which we condition on a call for service obtain a timestamp from the service call (which is documented in the city's 911 system) rather than the arrest (which is documented by the officer), this analysis is robust to the problem of errors in the arrest data.[10]

A third concern is that some officers may engage in "arrest trading," the practice of avoid-

[10]We can also investigate the quality of the timestamps directly. Among dispatch calls in which an arrest was made on the same day, the modal time between the dispatch call and the arrest is 0 hours, and the mean is 0.8 hours. Over 95 percent of arrests occur within 2 hours of the officers being dispatched.

ing overtime work by passing along a late-shift arrest to another officer who co-responded to a particular service call (Linn, 2009). To the extent that this practice occurs, it is possible that the decline in arrest activity that we observe at the end of an officer's shift represents a reallocation of administrative work rather than a true decline in the number of arrests that an officer makes. Figure 5 also directly addresses this issue. Because the model is estimated at the service call level, and the outcome is whether any arrest resulted, these estimates show that when an officer is dispatched to a service call at the end of his shift, an arrest is less likely — regardless of whether the arrest was made by the officer himself. This analysis rules out the possibility that arrest trading is an important contributor to our main estimates.

Finally, we have interpreted our court conviction regressions as evidence that officers' guilt threshold increases near the end of their workday. We consider here two alternative explanations. One possibility is that, for a given underlying offense, officers might increase the number of charges that they issue for a late-shift arrest, which, in the course of the plea bargaining process, could increase the probability of a conviction (Rehavi and Starr, 2014). In the left-hand panel of Figure A-7, we show how the number of arrest charges varies based on the shift-hour of the arrest. While there appears to be a small increase in the number of charges, the magnitude is quite small and unlikely to explain our court conviction finding.

Another explanation that could potentially rationalize our findings is that officers spend more time processing late-shift arrests than they spend processing arrests that are made earlier in their shift. In that case, our conviction results could reflect changes in officer effort rather than changes to the officer's guilt threshold. To test this explanation, we consider cases when the opportunity cost of arrest processing is high, which would reduce the court conviction rate under this alternative story. First, we examine court convictions for shifts that end on Friday and Saturday between 4pm and 10pm, when leisure time may be of greatest value. In the right-hand panel of Figure A-7, we show that the time path of court convictions for these shifts, though more imprecise, are statistically indistinguishable from the full-sample time path. Second, as we discuss in Section 5.4, the presence of an off-duty job after an officer's shift increases the opportunity cost of arrest processing. In Table A-7, we show that the court conviction rate for late-shift arrests is statistically identical on days in which an officer has an off-duty job, running counter to the alternative theory that officers take their time in processing late shift arrests in order to maximize overtime pay.

5.4 Effects of Off-Duty Work

Why don't police officers exploit their broad discretion to make arrests in order to take advantage of overtime pay? We consider two possible drivers of these findings. First, officers may place a high marginal value on their non-work time, such that they prefer to leave work rather than receive the additional pay of overtime (Mas and Pallais, 2019). Second, as an extension of the first case, officers may place a high marginal value on their non-work time, but their propensity to make late-shift arrests may be responsive to changes in overtime pay or other shocks to the marginal value of non-work time.

To distinguish between these possible cases, we exploit information on whether officers are working an off-duty shift on days in which they are also working a police shift. To identify whether officers are responsive to the relative value of overtime pay, we exploit variation in post-shift off-duty work. When officers work an off-duty job after their scheduled work shift, a late-shift arrest compromises their ability to arrive on time for their off-duty job. We therefore interpret variation in post-shift off-duty work as leading to changes in the opportunity cost of making a late-shift arrest, which allows us to differentiate between the two stories above.

In Table 2, we document the frequency with which officers work in a secondary job. The average officer in our sample reports 47 off-duty shifts in a year, and the average length of a shift is 5.6 hours.[11] On 9.5 percent of days with a police shift, officers also have an off-duty shift later in the day; this figure is 6.3 percent for off-duty shifts earlier in the day. As noted in Section 4.2, our empirical analysis will rely on indicators for an officer having a *regularly-scheduled* off-duty shift before or after their police work, which we define as days of the week where an officer reports an off-duty shift for at least 25 percent of dates in that quarter. On 7.6 percent of work days, officers have a regularly-scheduled off-duty shift *after* their work shift. Officers have a regularly scheduled off-duty shift *before* their work shift on 5.4 percent of work days. Consistent with Table A-2, we observe a reported pre-shift (post-shift) off-duty spell on 63 percent (66 percent) of days where we estimate an officer to have a regular off-duty spell; this number is 1.9 percent (3.4 percent) on non-regular days. These values indicate that our measure of regular off-duty work captures much of the variation in off-duty work.

[11] A list of the most common off-duty jobs held by officers can be found in Table A-4.

Our estimates of Equation (2) are presented in Table 3. We present results for both the hourly sample in which the unit of observation is an officer-shift-hour, and the dispatch sample, where the unit of observation is each call taken by an officer. The second and fourth columns reproduce the first and third columns with the inclusion of controls for volume of 911 calls.

The first two rows provide an estimate of the impact of working an off-duty job prior to an officer's police shift. We find only limited evidence that police officers make fewer late-shift arrests on days with pre-shift off-duty work. In the hourly sample, we see no evidence that arrests decline later in the officer's shift; in the dispatch sample we estimate that late-shift arrests may have declined by 7 percent on days in which an officer is predicted to work an off-duty shift prior to his work shift. Given that arrests decline by approximately 28 percent in the second half of the workday, the pre-shift off-duty effects indicate that ability loss explains at most 25 percent of the reduction in arrests that we observe at the end of the shift.

The second pair of rows in Table 3 document the effect of off-duty work *after* an officer's police shift. We see little evidence that early-shift arrests decline on days in which an officer is predicted to moonlight after his on-duty shift. However, late-shift arrests decline on these days by between 4 percent (in the overall sample) and 9 percent (in the dispatch sample). Notably, since these are reduced form estimates, they represent a lower bound on the impact of off-duty work — dividing these estimates by the first stage effect implies an estimate of between 7-15 percent. The estimates are unchanged with the inclusion of controls for 911 call volume, suggesting that our estimates are not reflecting unobserved differences in criminal activity on days when officers have off-duty work. Overall, the evidence suggests that when the opportunity cost of overtime work rises, officers strategically reduce their arrest activity.

We explore the robustness of our findings in Table A-5, which presents estimates with the same set of fixed effects used to test the robustness of our first findings. The effect of post-shift off-duty work is robust to specification and there is little evidence for an effect of pre-shift off-duty work in any of the models. In Table A-6 we present separate estimates for short ($\leq$ 4 hours) versus long ($>$ 4 hours) pre-shift off-duty work spells and find that the late-shift arrest reduction is specific to shifts with a long post-shift job.

While the results above indicate that officers reduce their late-shift arrest activity on days with post-shift off-duty work, a natural question is what happens to the quality of arrests that are made. Since our baseline findings suggest that arrest declines are accompanied by increases in arrest quality, we should expect that arrest quality is higher on days with off-duty work. The results of such an analysis are presented in Table A-7. We use a similar regression of Equation (2), where the unit of observation is an arrest, and each column considers a different outcome. We do not find evidence that arrests made on days when officers have off-duty work are more likely to lead to a guilty conviction or a prison sentence, nor are they significantly more likely to be a felony or a misdemeanor/felony. While these results may be surprising given our previous estimates, our standard errors are likely too wide to make any definitive conclusions. In Section 6, we show that the expected increase in guilty convictions for arrests on post-shift off-duty days, though positive, is small and within the confidence intervals we present in Table A-7.

Our estimates of the off-duty impacts suggest that the decline in officer arrest activity throughout their shift cannot be explained by declining ability, since we find small and inconsistently significant estimates for the impact of pre-shift off-duty work. Instead, we argue that the decline is more consistent with an aversion to working overtime. Our finding that post-shift off-duty work also leads to a decline suggests that officers are responsive to the opportunity cost of arrests. The following section synthesizes these findings with our baseline results in a simple dynamic model. By doing so, we can identify the relative weight officers place on their arrest activity and their personal overtime considerations.

6 Model of Officer Arrest Decisions

We have shown that officers reduce their arrest activity near the end of their shifts while the quality of the arrests made — as proxied by court convictions and sentencing — increases. We then showed that officers who work an off-duty shift after their police shift exhibit a further decline in late-shift arrests. In this section, we present a simple dynamic model of officer arrest decisions. We match this model to our empirical results, allowing us to estimate officer preferences over arrests of guilty individuals and avoidance of arrests of innocent individuals (i.e., their professional motivations), their relative value for working overtime, and the value they place on off-duty and overtime pay. We then use the model to consider the impact of changes to economic incentives on the quantity and quality of arrests.

6.1 Model Setup

An officer works a shift with T periods. In each period $t = 1, ..., T$, the officer encounters a potential arrestee, who is guilty with probability π. We designate whether an individual is guilty with the indicator $g \in \{0, 1\}$. The officer does not observe the individual's guilt but instead observes a noisy signal S that is correlated with guilt: $S|g \sim N(g\mu, 1)$. The mean of the signal distribution for guilty individuals, μ, dictates how well an officer is able to differentiate signals between guilty and innocent individuals. Using this signal and his or her prior on the probability of guilt, the officer generates a Bayes-updated probability of guilt for the individual, $\tilde{p}(s) = \frac{\phi(s-\mu)\rho}{\phi(s)(1-\rho)+\phi(s-\mu)\rho}$.

Officer Per-Period Objective Function: In each period, the officer has the choice of arresting the individual he encounters, and the contemporaneous value that the officer derives from his decision is a weighted sum of the value of correctly arresting a guilty individual and correctly not arresting an innocent individual:

$$v(a, \tilde{p}) = a(1 - \lambda)\tilde{p} + (1 - a)\lambda(1 - \tilde{p})$$

The weight parameter $\lambda \in [0, 1]$ dictates the relative importance that an officer places on avoiding arrests of innocent individuals, $(1 - a)\lambda(1 - \tilde{p})$, compared to arresting guilty individuals, $a(1 - \lambda)\tilde{p}$. Henceforth we refer to this parameter as the officer's risk preference.

If the officer were deciding his arrest activity in a static setting, where the above equation is his sole objective function, he would arrest an individual if their posterior guilt probability satisfied $\tilde{p}(s) \geq \lambda$. In other words, an arrest is made if the probability of guilt exceeds the officer's risk preference. This standard decision rule has been used by Prendergast (2003, 2007) to model public sector agents and Alesina and La Ferrara (2014), Arnold et al. (2018), and Arnold et al. (2020) for modeling judge behavior in similar settings where the agents face a tradeoff between punishing guilty individuals and not punishing innocent individuals. Our model uses this standard objective function and adds a dynamic component.

If the officer makes an arrest, there is a probability ϕ they will need to process the arrest, during which time the officer is not patrolling and is unable to arrest any additional suspects. If the officer does process the arrest, then every period thereafter there is a probability p that the officer must continue to process, until the officer is no longer processing. The model structure is depicted visually in Figure A-10. When processing the arrest, the officer cannot observe individuals on patrol. Therefore, the expected value they receive is the baseline expected value from an innocent individual not being arrested: $v^{np} = \lambda(1-\pi)$.

End of Shift: After the final period of the shift, if the officer is still processing an arrest, he receives a value of being at work and receiving overtime, $V_{ot} = c_{ot} + b \cdot \text{Pay}_{OT}$, where Pay_{OT} denotes the expected overtime pay received by the officer. If the officer is not processing an arrest at the end of the shift, he receives value, $V_0 = [c_{od} + b \cdot \text{Pay}_{OD}] \cdot \text{OD}$, where OD is an indicator for whether the officer is working off-duty after their police shift, and Pay_{OD} is their pay for that off-duty shift. The coefficients in front of Pay_{OT} and Pay_{OD} are the same, so that we assume the officer values a dollar of pay equally from both sources.

Note that we have normalized the value of ending the day without working overtime or working off-duty to 0, so the parameters c_{ot} and c_{od} indicate the costs or benefits to working relative to going home that cannot be explained by pay. For example, one potential cost to missing an off-duty shift is that the officer may be fired from their secondary job, leading to a greater income loss than the pay of the single missed shift. Our specification assumes that this kind of loss is adequately captured by c_{od}, so that the potential future earnings loss is the same regardless of the pay of a single shift.[12]

Value Functions: We now present value functions that denote an officer's utility for the current period and the expectation over utility in later periods. The value function of an officer who is patrolling in period t, conditional on their arrest choice and guilt signal, is denoted by $V_t^p(a,s)$. We represent the value of patrolling prior to conditioning on arrest choice and signal with $V_t^p \equiv E_s\big[\max_a V_t^p(a,s)\big]$. An officer's value function in each period is

$$V_t^p(a,s) = v^p(a,s) + a\big[\phi \cdot V_{t+1}^{np} + (1-\phi) \cdot V_{t+1}^p\big] + (1-a) \cdot V_{t+1}^p, \quad t \leq T$$

For $t \leq T$, V_t^{np} indicates the value of not patrolling in period t,

$$V_t^{np} = v^{np} + p \cdot V_{t+1}^{np} + (1-p) \cdot V_{t+1}^p.$$

Period $T+1$ reflects the time after the end of the shift. If the officer is still processing an arrest, they must work overtime, and they receive value $V_{T+1}^{np} = V_{ot}$. If they are patrolling,

[12]From speaking with officers in the department, it appears to be relatively common to find a fellow officer to take over a missed shift, so the likelihood of being fired from a singled missed shift seems small.

they receive the value from leaving their police shift, $V^p_{T+1} = V_0$.

We add two additional features to the model to better fit the data. First, we impose that officers begin on patrol with probability p_{start}. If an officer begins off of patrol, they have probability $1 - p$ to enter patrol each period. We add this probability to match that officers have a lower arrest rate earlier in the shift than in the middle. Second, we allow an exogenous probability p_{ot} that an officer receives overtime regardless of processing an arrest at the end of the shift. We do so to match that officers sometimes receive overtime on days where they make no arrests.

Note that arrests only enter the officer's utility function through the value of arresting a guilty individual and the value of avoiding arresting an innocent individual. We interpret these expressions as reflecting a composite "professional motivation," which comprises some mix of intrinsic pro-social motivations (Bénabou and Tirole, 2006; Besley and Ghatak, 2018) and workplace social incentives or peer effects (Bandiera et al., 2009, 2010; Mas and Moretti, 2009). One concern with this interpretation is the possibility that officers may receive some workplace benefit from achieving these outcomes. If that is the case, officers may appear to intrinsically value arrests but actually value another outcome that they achieve through arrests, such as career advancement or higher pay. One strength of our setting is that officer incentives are quite low-powered. Promotions from officer to corporal, then sergeant and lieutenant, are allocated through promotion exams. Individuals are ranked by their exam performance and promoted as positions open, with no reference to an officer's number and quality of arrests.[13] So, career advancement for most officers is not a function of arrest activity. However, some workplace benefits may be derived from arrests. For example, supervisors have discretion in providing non-arrest overtime to officers and may favor officers who make many arrests that lead to conviction. We therefore interpret the officers' utility for arrests as reflecting non-pecuniary professional motivation while also noting the important caveat that our estimates may partly capture low-powered workplace incentives.

[13]For higher ranks, which are not in our sample of patrol officers, promotion is discretionary and can be dependent on any measure of job performance.

6.2 Solution

The optimal solution for the officer is characterized by a series of threshold values, s_t^*, for each period, which indicate the guilt signal above which a suspect is arrested. The threshold values have the following solution:

$$\tilde{p}(s_t^*) = \lambda + \phi\big[V_{t+1}^p - V_{t+1}^{np}\big]$$

This equation offers a simple interpretation for the officer's rule. When the value of staying on patrol is higher than the value of being off of patrol ($V_{t+1}^p > V_{t+1}^{np}$) in the following period, the officer raises his guilt threshold. Note that if an officer were simply maximizing the within-hour value, $v(a, s)$, his threshold rule would be $\tilde{p}(s_t^*) = \lambda$ for every period. And if it were the case in the dynamic setting that $\phi = 0$, the threshold rule would also be the same in all periods and equal to λ. Therefore, the distortion in officer behavior relative to the static case comes entirely from the fact that the arrest decision affects whether he or she is on patrol in the following period.

6.3 Model Estimation

We estimate the model by matching four sets of moments from the data and empirical analysis. First, we use the estimated probability of making an arrest in each hour t. These are matched in the model by the probability that an officer is working on patrol in hour t times the probability that an individual's guilt signal, s, is greater than s_t^*. Second, we use the probability of overtime conditional on an arrest as well as the probability of overtime when making no arrest, which we match by calculating that an officer is processing an arrest at the end of the shift conditional on making an arrest in each hour.[14] Third, we match the probability of conviction conditional on arrest. For the purposes of estimating the model, we treat a court conviction as "ground-truth" guilt; the absence of a conviction indicates that the individual is innocent. Another interpretation of this assumption is that we are assuming that the officer's notion of guilt is based on whether the individual can be convicted in court. Therefore, the probability of conviction conditional on arrest is equal to the probability of an individual's guilt conditional on having $s > s_t^*$.

[14]Note that an officer can make an arrest in a certain hour and be processing a *later* arrest when he receives overtime, which is accounted for in constructing these moments from the model.

To estimate the parameters of the officers' value of overtime and off-duty, we use the reduced form arrest impacts of having an off-duty shift after their police shift. Note that the pay levels Pay_{OT} and Pay_{OD} but also the level preferences c_{ot} and c_{od} differ between overtime and off-duty work. We therefore need some variation in payment within off-duty spells to separately identify both parameters. To do so, we run a version of Table 3 where we split off-duty shifts into four hours or shorter and longer than four hours, while maintaining the choice from our main specification of splitting the effects into the first and last four hours of an officer's shift. The results of this regression are presented in Table A-6, which provides our fourth set of moments. We match the effects of short and long after-shift off-duty spells on the first and last four hours of an officer's shift, giving us four moments.

The model also requires that we choose values for Pay_{OT} and Pay_{OD}. For overtime pay, we take the average officer's salary in our sample, \$81,406, and the average number of overtime hours worked, 2.7, multiplied by 1.5, and get an average payment of \$164.8. For off-duty work, we observe that short spells ($\leq$ four hours) have an average duration of 3.4 hours, and long spells ($>$ four hours) have an average duration of 6.4 hours. While we do not observe the hourly pay of officers for their off-duty jobs, the available figures online indicate an hourly rate of \$30-40.[15] We choose a uniform value of \$35 and impute average payments of \$119.3 and \$224.3 for short and long off-duty shifts, respectively.

For the first three sets of empirical moments, we use estimates based on the sample of shifts where an officer is not working off-duty, and the moments are matched to the model with OD = 0. We then estimate Equation (1) with our baseline set of fixed effects. Our moment for each hour and outcome is $\hat{\mu}_8 + \hat{\theta}_k$, where $\hat{\mu}_8$ is the average for the first hour of the shift and $\hat{\theta}_k$ is the estimated coefficient for each hour.

In total, we have ten parameters to be estimated and 29 moments (three moments times eight shift-hours plus overtime probability when not arresting and four off-duty moments) with which to estimate the model. Despite the over-identification of the model, we show that the fit of all the moments is quite good.

[15] https://smallbusiness.costhelper.com/security-guard.html

6.4 Model Estimates and Counterfactuals

Model estimates are presented in Table 4. We estimate that in any given hour, the probability that an officer encounters a guilty individual is 0.033. The officer's average ability μ is 1.585, indicating that the distribution of signals for innocent and guilty individuals is quite overlapping. This low value explains why the true guilt rate is estimated to be significantly higher than any arrest rates we observe throughout the shift. We find that officer's trade-off between releasing innocent individuals and arresting guilty individuals, λ, is 0.13, meaning that, in a static setting, officers would arrest individuals with a posterior probability of guilt higher than 13 percent. When an arrest is made, the probability of the officer processing the arrest is 0.74, and conditional on processing, the probability of continued processing in each successive period is 0.77.

The value of \$1 of either overtime or off-duty income is $\hat{b} = 0.00021$. We can use this estimate to provide a dollar value for an officer's professional motivations. Dividing $\hat{\lambda}$ by $\hat{b}$ yields an estimate of the value, in dollars, of correctly choosing to *not* arrest an innocent individual. This estimate is \$624. Likewise, deflating 1-$\hat{\lambda}$ by this number yields an estimate of the value that officers placed on correctly arresting a guilty individual — \$4,063. These numbers are large, especially in comparison to the daily salary rate of approximately \$320. As such, the model suggests that officers place great value on the avoidance of Type I and especially Type II errors.

One valuable way to interpret our estimates for officer professional motivation is to consider the degree of economic incentives needed to distort their activity in un-ambiguously harmful ways. First, our estimates imply that an officer in the last hour of her shift would need to expect at least \$1,903 in overtime pay to make an arrest of an individual she knows with certainty is innocent.[16] Conversely, an officer in her last hour of the shift would need to have at least \$4,665 of income waiting at an off-duty job to forgo arresting an individual she knows with certainty is guilty.[17,18]

[16] We solve for the value of Pay_{OT} such that $\phi[c_{OT} + b \cdot \text{Pay}_{OT}] = \lambda$.

[17] We solve for the value of Pay_{OD} such that $\phi[c_{OD} + b \cdot \text{Pay}_{OD}] = (1 - \lambda) + \phi[c_{OT} + b \cdot \text{Pay}_{OT}]$, where we use our baseline value of Pay_{OT}.

[18] Note that our estimates indicate that officers care more about Type II errors than Type I errors. These estimates follow from the fact that $\hat{\lambda}$ is lower than 0.5. However, this preference is tempered by the fact that officers' guilt threshold is adjusted *upwards* due to an aversion to working overtime, as can be seen in Figure 6.

Since we do not observe overtime or off-duty pay of this magnitude, these figures necessarily require extrapolation from the smaller degree of variation in off-duty pay we used to estimate the model. Consequently, we think of these estimates as lower bounds for the true dollar values required to make these incorrect arrest decisions. For example, if an officer is willing to reduce their guilt threshold by 5% when they receive a higher value of overtime pay, they will likely need to receive more than ten times that increase to reduce their threshold by 50%, though our estimates assume a linear extrapolation.

In Figures 6, we document the fit of the model to our moments. The estimates fit the empirical moments well, especially when noting the relative parsimony of the parameters relative to the number of moments. We are able to capture the increase and then decline in arrest activity throughout the shift, the increase in court convictions throughout the shift, and the increase in overtime throughout the shift. The model estimate for the time path of overtime rates is positive and convex, driven by our assumption that the processing rate is governed by an hourly rate of continued processing. In contrast, the empirical time path is quite linear, which leads to a slight deviation of model fit from the matched moments.

The bottom right panel of 6 documents the fit of the model to the off-duty effects. All of our model-simulated moments are within the empirical confidence intervals, suggesting a good fit. We also capture the fact that the point estimate are more negative for late-shift arrest effects than early-shift. However, the model appears to simulate slightly more negative effects for early-shift arrests than our empirical moments. This discrepancy is due to the fact that declines in arrests must be driven solely by an increased aversion to overtime work, and the high rate of processing and continued processing leads the model to conclude that even *early-shift* arrests must be reduced. While this finding highlights a slight deviation of the model from the empirical moments, it also documents how the model's predictions and its fit to the data are non-trivial.

As we note in Section 5.5, we do not find an increase in convictions for arrests made on off-duty days, despite the intuition that declining arrests should lead to higher-quality arrests. While we do not include these estimated effects as moments in our model estimation, we simulate what our model would predict for the off-duty conviction effects, and we compare them to our empirical findings in Figure A-11. While the simulated model effects are positive (and the empirical effects are negative and insignificant), the estimates are comfortably within the confidence intervals of the empirical estimates. This finding indicates to us that, though the model and empirical estimates are roughly similar, our sample size limits our ability to precisely estimate these impacts.

A crucial question we can address with this model is how an officer's arrest behavior would change with an adjustment to the value of working overtime, which we present in Figure 7. The red line presents the model fit to our empirical moments. We first consider a tripling of the pay for overtime work, corresponding to the blue line. As expected, we observe an increase in arrest propensity and decline in guilt convictions throughout the shift.[19] Interestingly, these impacts are relatively small. The elasticity of arrest probability to overtime pay is 0.04, indicating that a 100% increase in overtime pay would lead to only a 4% increase in arrest propensity.

A similar counterfactual we consider is how an officer's arrest behavior would appear if he did not exhibit any aversion to working overtime. We simulate this possibility by changing the value of V_{ot} to 0, which we plot in the green line. Surprisingly, this officer exhibits an *increase* in arrest activity throughout the shift and a corresponding decline in court convictions. Intuition might lead us to believe that the overtime-indifferent officer would have a constant arrest rate and guilt threshold throughout the shift. However, arrests in the earlier hours impose the possibility of processing during the shift and precluding future arrests. What is especially striking about this finding is that, without estimation of the dynamic model above, an observed arrest profile that increases throughout the shift would likely be taken as evidence in favor of the "collars for dollars" story, namely that officers reduce their arrest quality in order to receive overtime pay. Our model indicates that, on the contrary, a completely altruistic officer may exhibit an increasing arrest rate and decreasing guilt threshold across their shift because of the higher opportunity cost earlier in the shift.

This counterfactual also tells us how much enforcement overall is affected by overtime aversion. Relative to observed officer behavior, an overtime-indifferent officer has a 70% higher hourly arrest rate (3.1% v. 1.8%), and the court conviction rate of their arrests made is 7% lower (28.1% v. 35.6%). These comparisons provide a stark illustration of

[19] The overall guilt conviction impact is in principle ambiguous: while convictions in all hours are now less likely, the arrest composition more heavily weights late-shift arrests. The overall court conviction rate is lower with triple overtime pay, indicating that the former effect outweighs the latter.

the first-order importance of overtime considerations for the quantity and quality of arrest activity.

A final insight we glean from the model is about the equity and efficiency impacts of officers' overtime aversion. Are arrest decisions unambiguously worse relative to a counterfactual where officers are indifferent to overtime work? To answer this question, we simulate officer choices while setting overtime aversion to 0. We calculate, for different risk preferences λ, the probability of arresting a truly guilty suspect and the probability of not arresting a truly innocent suspect. These outcomes are the converse of Type-I and Type-II errors.

These values are plotted in Figure A-12, and they trace out a "production possibilities" curve of possible outcomes for overtime-indifferent officers. The black triangle depicts the location of officers' current behavior, and the black circle depicts the behavior when overtime aversion is set to zero and all other parameters are kept the same.

This figure clearly indicates that, relative to indifference, overtime aversion reduces the probability of arrest for guilty suspects (i.e. increases Type-II error) and raises the probability of not arresting innocent suspects (i.e. decreases Type-I error), generating an *equity* impact. Surprisingly, however, current officer behavior is very close to the frontier, so there is essentially no *efficiency* impact. Stated differently, officers behave similarly to a counterfactual where they are overtime indifferent but have a higher λ, i.e. place a greater relative weight on not arresting innocent individuals. This result derives from the fact that, while a shift in officer arrest rates changes the probability of guilt of the marginal arrestee (as indicated by our first analysis), these magnitudes are not large enough to generate an efficiency cost from variation in arrest frequency over the shift. These equity-efficiency calculations further highlight the value of using a formal model to consider how overtime preferences impact officer behavior.

6.5 Alternative Modeling Choices

We discuss here two deviations from our baseline model and estimation, and we present updated parameter estimates in Table A-8. The first column presents our baseline model parameter estimates, and subsequent columns present alternative specifications. Our goal here is to consider modifications of the model that might attenuate our estimates of professional motivation in order to infer an approximate "lower bound" for these numbers.

First, our baseline model assumes that an individual is convicted in court if and only if he is truly guilty. In reality, at least some innocent individuals are convicted and at least some guilty individuals are not convicted. Letting true guilt be denoted by g and conviction be denoted by c, we estimate a model variant that supposes that errors occur in 10 percent of cases and are symmetric: $Pr(c = 1|g = 0) = Pr(c = 0|g = 1) = 0.1$. With this additional feature, the court conviction rate for an hour, c_h, and true guilt rate g_h, are related by $c_h = 0.9 * g_h + 0.1 * (1 - g_h)$.

The result of this alternative specification is presented in the second column of Table A-8. The most notable change relative to the baseline estimates is that λ, indicating the relative importance of accurate arrests and non-arrests, declines from 0.133 to 0.064, suggesting that officers place an even higher importance on correct arrests relative to non-arrests. We observe a small increase in the value of an arrest of a guilty individual from \$4,063 in the baseline model to \$4,182, while the value a non-arrest of an innocent individual *declines* from \$626 to \$284.

Second, Our baseline model uses the point estimates from our empirical results, and an important question is how estimates would differ with a larger but still plausible off-duty arrest decline. As we discussed above, a larger decline in arrests on off-duty days corresponds in the model to a *lower* implied valuation for workplace activity. In column 3, we re-estimate our model where we use the largest (most-negative) off-duty effects in our 95% confidence intervals.

Unsurprisingly, the implied dollar values of professional motivation are reduced. Officers' value for not arresting an innocent person declines from \$626 to \$226, and the value for arresting a guilty person declines from \$4063 to \$1515. In column 4, we combine both model modifications, and we find that the non-arrest of innocent individuals declines to \$98, and the arrest of guilty individuals is a nearly identical \$1519. While these numbers are substantially smaller than our baseline estimates, they should be treated as lower bounds for the range of values that are consistent with observed officer behavior.

7 Conclusion

In this paper we evaluate how officers trade off their professional motivation over workplace outcomes with their private economic concerns. In particular, we ask how the quantity and quality of arrests made are affected by the relationship between arrest activity and overtime work. Using unique and highly-detailed administrative data which link records on 911 calls, police officer shift assignments, off-duty work, arrests, and associated court outcomes in Dallas, Texas, we find that officers significantly reduce their frequency of arrests as their shift-end nears. This finding is not explained by officers being taken out of circulation having made earlier shift arrests, nor is it explained by "arrest trading" between officers or a formal or informal policy to route officers to fewer calls at the end of their shift. We next find that the conviction and sentencing rates for arrests increase at the end of the shift. Like Chan's (2018) work on emergency department physicians, our findings are consistent with an aversion to working overtime. Leveraging variation in off-duty work, which causes a shift in the opportunity cost of overtime work, we also find that officers reduce their last-shift arrests when the cost of making an arrest is highest. Feeding our estimates into a dynamic model of officer arrest decisions, we estimate very high valuations by officers on arresting guilty individuals and not arresting innocent individuals, suggesting that, despite an aversion to working overtime, the average officer exhibits a non-trivial degree of professional motivation in their job.

Our study suggests several avenues for future research. First, our estimates of the professional motivations of officers are estimated from a relatively fixed group of individuals who face a constant pay schedule and employment contract. How would the preferences of the police force respond to a change to the employment contract? A large share of the literature on public sector workers focuses on the labor supply of agents and how professional motivation is affected by changes to job characteristics (Dal Bó et al., 2013; Fisman et al., 2015; Ashraf et al., 2020). Connecting our approach to an investigation of police labor supply would be valuable for evaluating any policy changes related to workplace characteristics.

Second, how surprising are our estimates of officers' valuations for arrests and non-arrests? What value would the average civilian place on these outcomes, and what is the socially optimal degree of professional motivation? Our study is among the first to provide evidence of professional motivations identified directly from changes to the opportunity cost of a workplace activity, so there is for now an absence of other estimates to consult as comparison. While there is a large literature on the social value of crime prevention (Cohen and Piquero, 2009; Chalfin, 2015), hardly any research measures the value individuals place on the arrest of a criminal suspect after a crime has been committed. Future research on estimates of professional motivations of public sector workers and on estimates of civilians' valuations of arrests will provide valuable guidance in interpreting our finding that officers appear to place high valuations on their workplace outcomes.

References

Abrams, D., H. Fang, and P. Goonetilleke (2023). Do cops know who to stop? assessing optimizing models of police behavior with a natural experiment. Technical report, National Bureau of Economic Research.

Adda, J., B. McConnell, and I. Rasul (2014). Crime and the depenalization of cannabis possession: Evidence from a policing experiment. *Journal of Political Economy 122*(5), 1130–1202.

Alesina, A. and E. La Ferrara (2014). A test of racial bias in capital sentencing. *American Economic Review 104*(11), 3397–3433.

Arnold, D., W. Dobbie, and C. S. Yang (2018). Racial bias in bail decisions. *The Quarterly Journal of Economics 133*(4), 1885–1932.

Arnold, D., W. S. Dobbie, and P. Hull (2020). Measuring racial discrimination in bail decisions. Technical report, National Bureau of Economic Research.

Ash, E. and W. B. MacLeod (2015). Intrinsic motivation in public service: Theory and evidence from state supreme courts. *The Journal of Law and Economics 58*(4), 863–913.

Ashraf, N., O. Bandiera, E. Davenport, and S. S. Lee (2020). Losing prosociality in the quest for talent? sorting, selection, and productivity in the delivery of public services. *The American Economic Review 110*(5), 1355–94.

Ater, I., Y. Givati, and O. Rigbi (2014). Organizational structure, police activity and crime. *Journal of Public Economics 115*, 62–71.

Ba, B., P. Bayer, N. Rim, R. Rivera, and M. Sidibé (2021). Police officer assignment and neighborhood crime. Technical report, National Bureau of Economic Research.

Ba, B. A., D. Knox, J. Mummolo, and R. Rivera (2021). The role of officer race and gender in police-civilian interactions in chicago. *Science 371*(6530), 696–702.

Ba, B. A. and R. Rivera (2019). The effect of police oversight on crime and allegations of misconduct: Evidence from chicago. *U of Penn, Inst for Law & Econ Research Paper* (19-42).

Bacher-Hicks, A. and E. de la Campa (2020). The impact of new york citys stop and frisk program on crime: The case of police commanders.

Baicker, K. and M. Jacobson (2007). Finders keepers: Forfeiture laws, policing incentives, and local budgets. *Journal of Public Economics 91*(11-12), 2113–2136.

Bandiera, O., I. Barankay, and I. Rasul (2009). Social connections and incentives in the workplace: Evidence from personnel data. *Econometrica 77*(4), 1047–1094.

Bandiera, O., I. Barankay, and I. Rasul (2010). Social incentives in the workplace. *The Review of Economic Studies 77*(2), 417–458.

Bénabou, R. and J. Tirole (2003). Intrinsic and extrinsic motivation. *The review of economic studies 70*(3), 489–520.

Bénabou, R. and J. Tirole (2006). Incentives and prosocial behavior. *The American Economic Review 96*(5), 1652–1678.

Berdejo, C. and N. Yuchtman (2013). Crime, punishment, and politics: an analysis of political cycles in criminal sentencing. *Review of Economics and Statistics 95*(3), 741–756.

Bertrand, M., R. Burgess, A. Chawla, and G. Xu (2020). The glittering prizes: Career incentives and bureaucrat performance. *The Review of Economic Studies 87*(2), 626–655.

Besley, T. and M. Ghatak (2018). Prosocial motivation and incentives. *Annual Review of Economics 10*, 411–438.

Blanes i Vidal, J. and T. Kirchmaier (2018). The effect of police response time on crime clearance rates. *The Review of Economic Studies 85*(2), 855–891.

Cassell, P. G. (2018). Overstating america's wrongful conviction rate: Reassessing the conventional wisdom about the prevalence of wrongful convictions. *Arizona Law Review 60*, 815.

Chalfin, A. (2015). Economic costs of crime. *The Encyclopedia of Crime and Punishment*, 1–12.

Chalfin, A., B. Hansen, E. K. Weisburst, and M. C. Williams Jr (2022). Police force size and civilian race. *American Economic Review: Insights 4*(2), 139–158.

Chalfin, A. and J. McCrary (2018). Are us cities underpoliced? theory and evidence. *Review of Economics and Statistics 100*(1), 167–186.

Chan, D. C. (2018). The efficiency of slacking off: Evidence from the emergency department. *Econometrica 86*(3), 997–1030.

Cho, S., F. Gonçalves, and E. Weisburst (2023). The impact of fear on police behavior and public safety. Technical report, National Bureau of Economic Research.

Cohen, A. and C. S. Yang (2019). Judicial politics and sentencing decisions. *American Economic Journal: Economic Policy 11*(1), 160–191.

Cohen, M. A. and A. R. Piquero (2009). New evidence on the monetary value of saving a high risk youth. *Journal of Quantitative Criminology 25*(1), 25–49.

Conley, J. P. and P. Wang (2006). Crime and ethics. *Journal of Urban Economics 60*(1), 107–123.

Cook, C., R. Diamond, J. V. Hall, J. A. List, and P. Oyer (2021). The gender earnings gap in the gig economy: Evidence from over a million rideshare drivers. *The Review of Economic Studies 88*(5), 2210–2238.

Dal Bó, E., F. Finan, and M. A. Rossi (2013). Strengthening state capabilities: The role of financial incentives in the call to public service. *Quarterly Journal of Economics 128*(3), 1169–1218.

DellaVigna, S., J. A. List, and U. Malmendier (2012). Testing for altruism and social pressure in charitable giving. *The quarterly journal of economics 127*(1), 1–56.

DellaVigna, S., J. A. List, U. Malmendier, and G. Rao (2016). Voting to tell others. *The Review of Economic Studies 84*(1), 143–181.

DellaVigna, S. and D. Pope (2018). What motivates effort? evidence and expert forecasts. *The Review of Economic Studies 85*(2), 1029–1069.

Dixit, A. (2002). Incentives and organizations in the public sector: An interpretative review. *Journal of Human Resources*, 696–727.

Draca, M., S. Machin, and R. Witt (2011). Panic on the streets of london: Police, crime, and the july 2005 terror attacks. *The American Economic Review 101*(5), 2157–81.

Evans, W. N. and E. G. Owens (2007). Cops and crime. *Journal of Public Economics 91*(1-2), 181–201.

Farber, H. S. (2015). Why you cant find a taxi in the rain and other labor supply lessons from cab drivers. *The Quarterly Journal of Economics 130*(4), 1975–2026.

Feigenberg, B. and C. Miller (2022). Would eliminating racial disparities in motor vehicle searches have efficiency costs? *The Quarterly Journal of Economics 137*(1), 49–113.

Finan, F., B. A. Olken, and R. Pande (2017). The personnel economics of the developing state. In *Handbook of Economic Field Experiments*, Volume 2, pp. 467–514. Elsevier.

Fisman, R., N. A. Harmon, E. Kamenica, and I. Munk (2015). Labor supply of politicians. *Journal of the European Economic Association 13*(5), 871–905.

Forst, B. (1982). *Arrest convictability as a measure of police performance*, Volume 73. US Department of Justice, National Institute of Justice.

Fu, C. and K. I. Wolpin (2018). Structural estimation of a becker-ehrlich equilibrium model of crime: Allocating police across cities to reduce crime. *The Review of Economic Studies 85*(4), 2097–2138.

Gonçalves, F. and S. Mello (2021). A few bad apples? racial bias in policing. *American Economic Review 111*(5), 1406–1441.

Gonçalves, F. M. and S. Mello (2023). Police discretion and public safety. Technical report, National Bureau of Economic Research.

Holmstrom, B. and P. Milgrom (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics & Organization 7*, 24.

İmrohoroğlu, A., A. Merlo, and P. Rupert (2004). What accounts for the decline in crime? *International Economic Review 45*(3), 707–729.

Kaplan, J. (2019). Uniform crime reporting (ucr) program data: County-level detailed arrest and offense data.

Knowles, J., N. Persico, and P. Todd (2001). Racial bias in motor vehicle searches: Theory and evidence. *Journal of Political Economy 109*(1), 203–229.

Lahiri, P. and M. D. Larsen (2005). Regression analysis with linked data. *Journal of the American Statistical Association 100*(469), 222–230.

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

Lim, C. S. H. (2013). Preferences and incentives of appointed and elected public officials: Evidence from state trial court judges. *American Economic Review 103*(4), 1360–1397.

Linn, E. (2009). *Arrest Decisions: What Works for the Officer?* Number 5. Peter Lang.

Loeffler, C. E., J. Hyatt, and G. Ridgeway (2019). Measuring self-reported wrongful convictions among prisoners. *Journal of Quantitative Criminology 35*(2), 259–286.

Makowsky, M. D. and T. Stratmann (2009). Political economy at any speed: what determines traffic citations? *The American Economic Review 99*(1), 509–27.

Makowsky, M. D. and T. Stratmann (2011). More tickets, fewer accidents: How cash-strapped towns make for safer roads. *The Journal of Law and Economics 54*(4), 863–888.

Makowsky, M. D., T. Stratmann, and A. Tabarrok (2019). To serve and collect: the fiscal and racial determinants of law enforcement. *The Journal of Legal Studies 48*(1), 189–216.

Mas, A. and E. Moretti (2009). Peers at work. *American Economic Review 99*(1), 112–45.

Mas, A. and A. Pallais (2019). Labor supply and the value of non-work time: Experimental estimates from the field. *American Economic Review: Insights 1*(1), 111–26.

Mastrobuoni, G. (2013). Police and clearance rates: Evidence from recurrent redeployments within a city. *Unpublished Working Paper*.

Mello, S. (2019). More cops, less crime. *Journal of Public Economics 172*, 174–200.

Morris, C. N. (1983). Parametric empirical bayes inference: theory and applications. *Journal of the American Statistical Association 78*(381), 47–55.

Moskos, P. (2008). *Cop in the hood: My year policing Baltimore's eastern district*. Princeton University Press.

Moskos, P. (2011). Collars for dollars: How the drug war sacrifices real policing for easy arrests. *Reason*.

Niskanen, W. A. (1971). *Bureaucracy and Representative Government*. Aldine-Atherton.

Oettinger, G. S. (1999). An empirical analysis of the daily labor supply of stadium venors. *Journal of Political Economy 107*(2), 360–392.

Owens, E. (2020). The economics of policing. *Handbook of Labor, Human Resources and Population Economics*, 1–30.

Premkumar, D. (2020). Intensified scrutiny and bureaucratic effort: Evidence from policing and crime after high-profile, officer-involved fatalities. *SSRN Electron. J. 2020*, 1–132.

Prendergast, C. (2001). Selection and oversight in the public sector, with the los angeles police department as an example. Technical report, National Bureau of Economic Research.

Prendergast, C. (2003). The limits of bureaucratic efficiency. *Journal of Political Economy 111*(5), 929–958.

Prendergast, C. (2007). The motivation and bias of bureaucrats. *The American Economic Review 97*(1), 180–196.

Rehavi, M. M. and S. B. Starr (2014). Racial disparity in federal criminal sentences. *Journal of Political Economy 122*(6), 1320–1354.

Smith, C. A. (2018). *Audit of the Dallas Police Department's Off-Duty Employment Program.*

Smith, D. A. and C. A. Visher (1981). Street-level justice: Situational determinants of police arrest decisions. *Social Problems 29*(2), 167–177.

Spenkuch, J. L., E. Teso, and G. Xu (2023). Ideology and performance in public organizations. *Econometrica 91*(4), 1171–1203.

Tahamont, S., Z. Jelveh, A. Chalfin, S. Yan, and B. Hansen (2020). Dude, where's my treatment effect? errors in administrative data linking and the destruction of statistical power in randomized experiments. *Journal of Quantitative Criminology*.

Weisburd, S. (2021). Police presence, rapid response rates, and crime prevention. *Review of Economics and Statistics 103*(2), 280–293.

Weisburst, E. K. (2019). Safety in police numbers: Evidence of police effectiveness from federal cops grant applications. *American Law and Economics Review*.

Weisburst, E. K. (2022). whose help is on the way? the importance of individual police officers in law enforcement outcomes. *Journal of Human Resources*, 0720–11019R2.

Wilson, J. (1989). Bureaucracy: What government agencies do and why they do it.

Table 1: Summary Statistics for Officer-Years

	(1) Mean	(2) SD	(3) Min	(4) Max	(5) N
Number 911 Calls	624.58	403.76	1	2511	8614
Number Arrests	26.70	29.80	0	440	8614
Share Guilty Conviction	0.34	0.20	0	1	8171
Share Felony Arrests	0.21	0.18	0	1	8171
Share Civilian-Initiated Arrests	0.65	0.26	0	1	8171
Annual Salary	81605	11796	42941	135957	8611
Overtime Pay	3493	5688	0	84183	8614
Number of Off-Duty Shifts	47.53	65.02	0	540	8614
Average Off-Duty Shift Length	5.57	1.79	0	19	6192
Officer Female	0.16	0.36	0	1	8611
Officer Black	0.26	0.44	0	1	8611
Officer Hispanic	0.22	0.42	0	1	8611
Bachelor's Degree +	0.51	0.50	0	1	8611
Officer Tenure	10.87	8.75	-4	48	8611

Note: Table presents summary statistics at the officer-year level.

Table 2: Summary Statistics, Off-Duty Employment

A. Officer-Year Level	(1) Mean	(2) SD	(3) Min	(4) Max	(5) N
Number of Off-Duty Shifts	46.65	64.12	0.00	540.000	9493
Average Off-Duty Shift Length	5.59	1.83	0.00	20.000	6831

B. Daily Police-Shift Level	(1) Mean	(2) SD	(3) Min	(4) Max	(5) N
After-Shift	0.095	0.29	0.00	1.00	614952
Before-Shift	0.063	0.24	0.00	1.00	614952
Regular After-Shift	0.076	0.27	0.00	1.00	614952
Regular Before-Shift	0.054	0.23	0.00	1.00	614952

Note: Table presents summary statistics for data on off-duty employment. Panel A presents information tabulated at the officer-year level — the number of off-duty work shifts and the length (in hours) of an off-duty work shift. Panel B presents information tabulated at the day-by-shift level. Here, we report the share of shifts that are either preceded or proceeded by any off-duty work shift or a regularly-scheduled off-duty work shift.

Table 3: Effect of Off-Duty Employment on Late vs. Early Shift Arrests

	Hourly Sample		Dispatch Sample	
	(1)	(2)	(3)	(4)
	Arrest	Arrest	Arrest	Arrest
Regular Off-Duty Before × Early	0.000120	0.000117	0.000471	0.000518
	(0.000348)	(0.000348)	(0.000740)	(0.000738)
Regular Off-Duty Before × Late	-0.000158	-0.000150	-0.00198*	-0.00202*
	(0.000344)	(0.000344)	(0.000892)	(0.000891)
Regular Off-Duty After × Early	-0.000132	-0.000126	-0.000880	-0.000850
	(0.000321)	(0.000321)	(0.000662)	(0.000664)
Regular Off-Duty After × Late	-0.000521*	-0.000515	-0.00237**	-0.00236**
	(0.000262)	(0.000262)	(0.000761)	(0.000759)
Log Calls		-0.00199***		-0.0131***
		(0.000232)		(0.000564)
Log Serious Calls		0.00289***		-0.00272***
		(0.000181)		(0.000454)
Mean	0.017	0.017	0.032	0.032
Observations	4919616	4919616	1754401	1754401

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Table presents estimates from a series of regressions of an indicator variable for whether an officer made an arrest during a given shift on an indicator for whether the officer was predicted to work a regular off-duty shift either before or after his or her police shift. These regressions correspond with Equation (2). Results are presented separately for early-shift arrests and late-shift arrests. The first two columns pertain to our hourly sample. In columns (3)-(4), we report estimates for the sample of calls which were initiated by a citizen call for service. Standard errors are clustered at the division-by-month level.

Table 4: Model Parameter Estimates

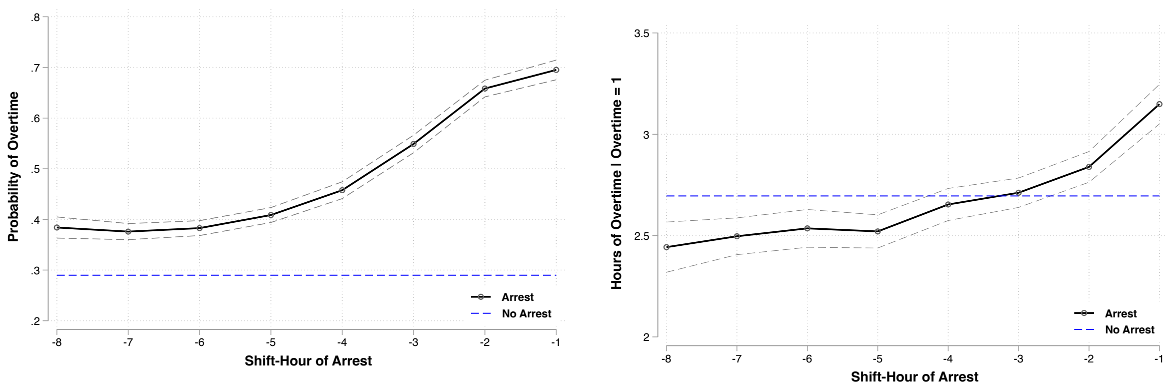
Parameters	Estimates	Description
ρ	0.033	Probability of guilt
μ	1.587	Mean of signal for guilty individuals (officer ability)
λ	0.133	Tradeoff for arresting guilty and not arresting innocent individuals
c_{ot}	-0.227	Intercept value/cost of working overtime
c_{od}	-0.018	Intercept value/cost of working off-duty
b	0.00021	Value of \$1 of overtime/off-duty pay
ϕ	0.741	Probability or processing an arrest
p	0.766	Per-period probability of continued arrest processing
p_{start}	0.349	Probability of beginning work on patrol
p_{ot}	0.290	Probability of receiving overtime when not making an arrest
Parameters	**Estimates**	**Description**
λ/b	624.1	Value of non-arrest of innocent person
$(1-\lambda)/b$	4062.7	Value of arrest of guilty person
$E(v(a,s)\|a=0)$	609.8	Average value of non-arrest
$E(v(a,s)\|a=1)$	1415.2	Average value of arrest
$E(\max_{a_t} v(a,s)) - v_{np}$	29.7	Average value of hour on patrol (relative to hour not on patrol)
$(\lambda/\phi - c_{ot})/b$	1903.4	Overtime pay needed for arrest of certainly innocent suspect
$(1-\lambda)/\phi b + (c_{ot} - c_{od})/b$ $+Pay_{OT}$	4665.0	Off-duty pay needed for non-arrest of certainly guilty suspect

Note: Table presents parameter estimates from the model in Section 6.

Figure 1: The Probability and Amount of Overtime Pay by Shift-Hour of Arrest

A. Probability of Overtime

B. Average Hours of Overtime, Conditional on Arrest

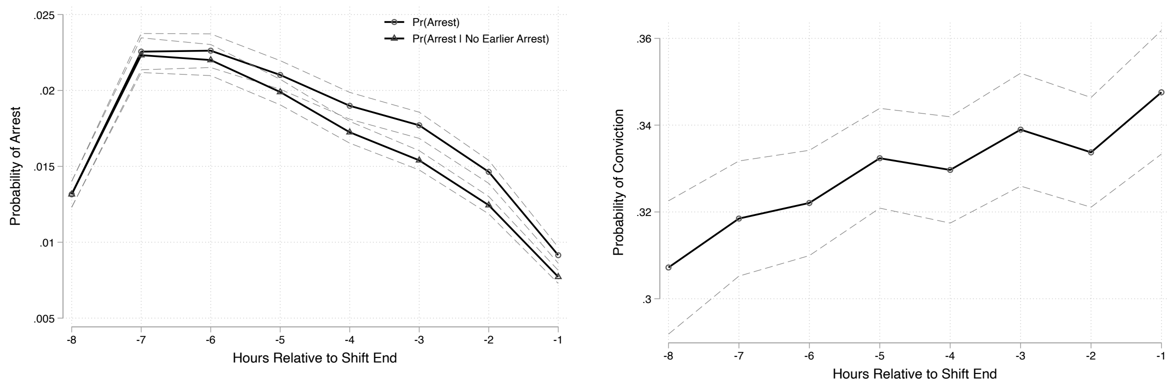


Note: The left-hand panel plots the probability that an arrest made in a given hour of an officer's shift leads to overtime pay. The -8 hour corresponds with the first hour of the shift; the -1 hour corresponds with the final hour of the officer's shift. The dotted blue line represents the probability that overtime hours are worked for a shift in which no arrest is made. The right-hand panel plots the mean number of overtime hours worked for an arrest that is made in a given hour of an officer's shift. Again, the dotted blue line represents the mean number of overtime hours worked when a shift in which no arrest was made leads to an overtime spell. 95 percent confidence intervals are provided for each statistic.

Figure 2: Arrest Frequency and Court Conviction by Shift-Hour

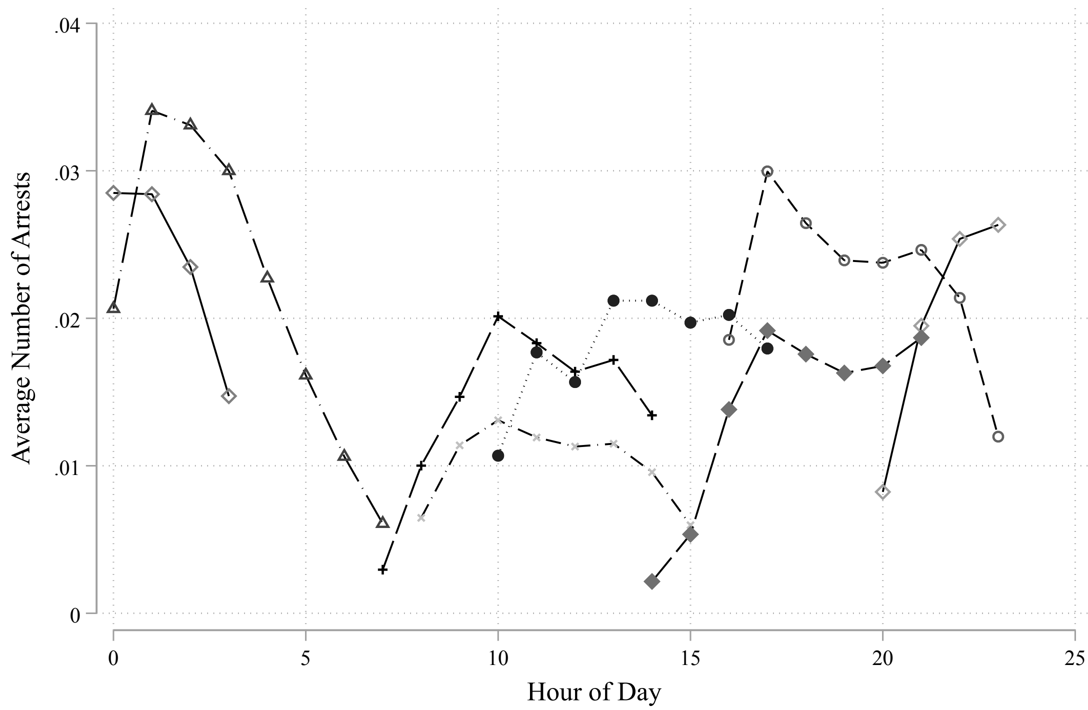
A. Probability of Making Arrest

B. Court Conviction Rate Conditional on Arrest



Note: The left-hand panel plots the probability that an arrest is made in a given hour of an officer's shift. The -8 hour corresponds with the first hour of the shift; the -1 hour corresponds with the final hour of the officer's shift. We also plot the conditional probability that an arrest was made given that an arrest was made earlier in the officer's shift. The right-hand panel plots the probability that an arrest results in a criminal conviction for either a misdemeanor or a felony offense for an arrest made in each hour of an officer's shift. 95 percent confidence intervals are provided for each statistic.

Figure 3: Arrest Propensities By Hour and Shift

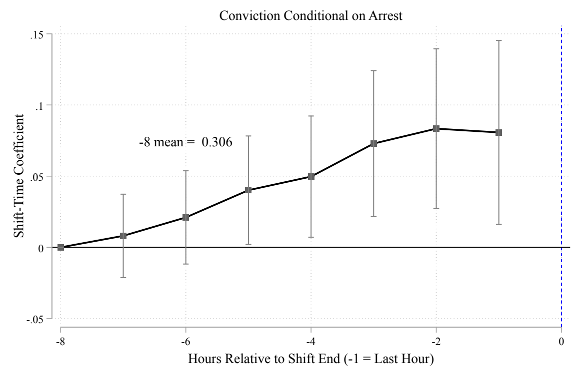


Note: Figure plots the average number of arrests per officer by hour of day, separately for officers in each of the seven most common shifts.

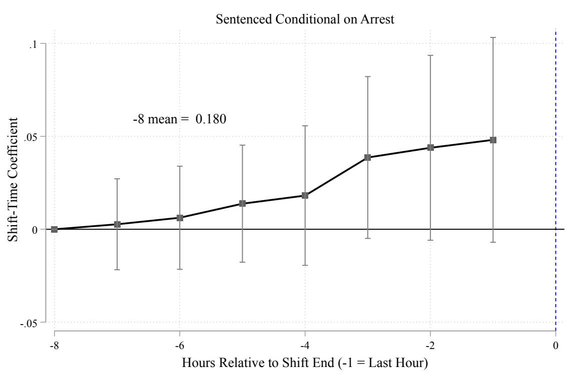
Figure 4: Arrest Frequency and Court Conviction Regressions

A. Arrest

B. Conviction Conditional on Arrest



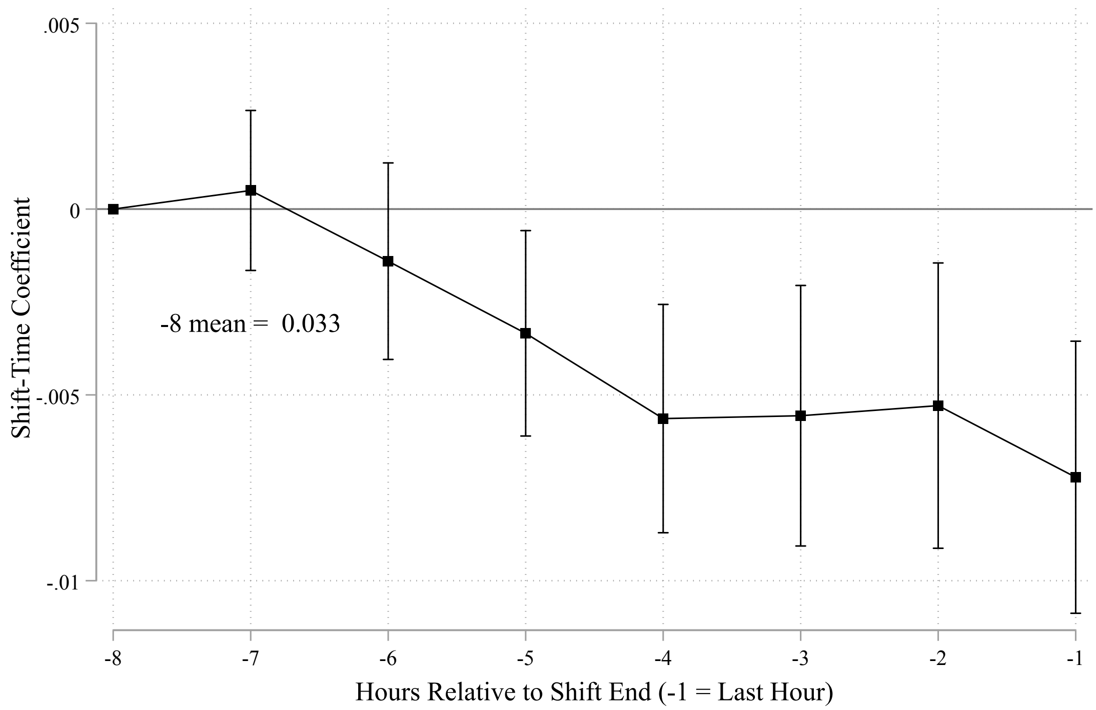
C. Sentenced Conditional on Arrest



D. Felony Arrest

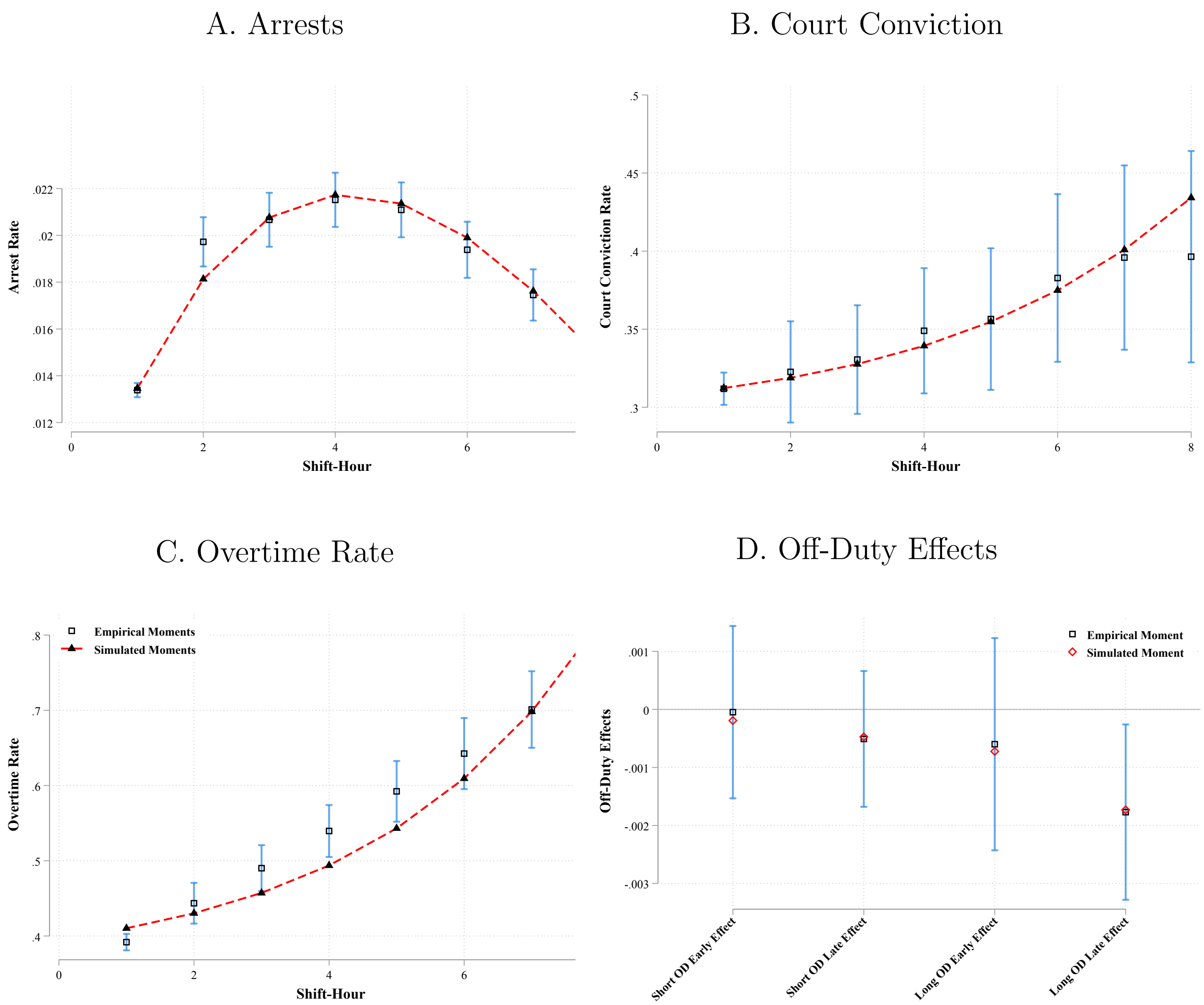
Notes: Figure plots each shift-hour coefficient from a regression of a given outcome variable on a vector of shift-hour indicator variables, conditional on the fixed effects described in Equation (1). 95 percent confidence intervals are presented, computed using standard errors clustered at the division-by-month level. All coefficients are relative the first hour of an officer's shift.

Figure 5: Arrest Propensity From a 911 Call



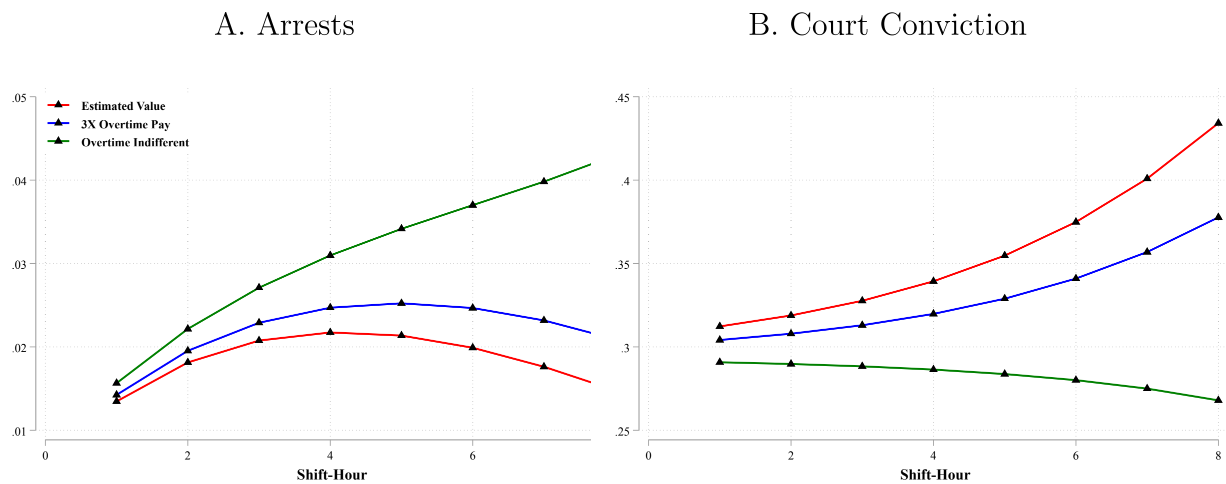
Note: Figure plots each shift-hour coefficient from a regression of a binary arrest indicator on a vector of shift-hour indicator variables, conditional on the fixed effects described in Equation (1). The unit of observation is a 911 call. 95 percent confidence intervals are presented, computed using standard errors clustered at the division-by-month level. All coefficients are relative to the first hour of an officer's shift.

Figure 6: Model Fit, Primary Estimates



Note: Figure plots empirical versus simulated moments as a test of model fit. Estimates are presented for arrest incidence, the court conviction rate, the overtime rate from an arrest, and off-duty effects on arrests.

Figure 7: Model Counterfactual, Changes in Overtime Value



Note: Figure plots the time-path of the arrest rate (left-hand panel) and the court conviction rate of arrests (right-hand panel) across shift-hours under three scenarios. The red line is the estimated value from our model. The blue line refers to a scenario in which overtime pay is increased 3x relative to the baseline. The green line refers to a scenario in which officers are indifferent between working overtime and leaving work.

FOR ONLINE PUBLICATION: APPENDICES

A Data Construction

Section 3 describes the various data sources we use and how we link them and construct our analysis sample. We provide more information here on the sample construction and specifically how many observations are dropped when restricting our raw data to our final analysis sample.

There are 5,469 officers who work in the Dallas Police Department during our sample period. 685 (12.5%) of these officers are dropped because they never appear in the overtime data. An additional 2,354 (43.0%) officers are dropped because they never work a patrol division assignment, never work two adjacent patrol division shifts, or because they are never observed taking a 911 call during their regular shift. These restrictions leave us with 2,430 officers in our analysis sample.

Our raw 911 data consists of 2,144,496 calls made during our sample period. 171,927 (8.0%) of these calls are dropped because the listed responding officers never appear in the overtime data working a patrol shift. 680,485 (31.7%) of calls have a responding officer who does work patrol shifts but the call is outside the hours or dates we know the officer is working a regular shift. Note that these drops can occur for several reasons. These calls may be occurring during an additional shift taken by an officer, in which case we do not know the hours of the shift. They may also be taken during a regular shift for that officer but occur between adjacent overtime payments that list different regular shifts, and we are unable to infer the shift times for that day. These sample restrictions leave us with 1,394,377 calls in our analysis sample.

Our raw arrest data consists of 149,339 arrests made during our sample period. 39,034 (26.1%) of these arrests are dropped because the listed officers never appear in the overtime data working a patrol shift. 69,083 (46.3%) of arrests have an arresting officer who does work patrol shifts but the arrest is outside the hours or dates we know the officer is working a regular shift. Similar to our dropped calls, these arrests may be occurring during an additional shift taken by an officer, in which case we do not know the hours of the shift. They may also be taken during a regular shift for that officer but occur between adjacent overtime payments that list different regular shifts, and we are unable to infer the shift times for that day. These restrictions leave us with 68,569 arrests in our analysis sample.

B Do Some Officers Engage in "Collars for Dollars"?

While our principal findings run contrary to the narrative that officers make additional low-quality arrests at the end of the shift in order to receive overtime pay, it is still possible that a fraction of officers engage in this practice even if the behavior of these officers is not detectable in the aggregate data. In this section, we investigate heterogeneity across officers in their late-shift arrest behavior, focusing on the officers who are especially likely to concentrate their arrest activity at the end of their shift. We assess whether the late-shift arrests made by these "late-arresters" are of lower quality, are more likely to be officer-initiated or disproportionately target minority citizens.

We begin by investigating whether officers, in fact, differ systematically in their propensity to make late-shift arrests. Given that there is evidence that police officers do differ systematically in their overall propensity to make arrests (Weisburst, 2022), any analysis that evaluates differences in end-of-shift arrests must account for overall differences in arrest propensity. Accordingly, we run a regression in which the unit of analysis is a given arrest and calculate whether the arrest i occurs in the last two hours of the arresting officer's shift:

$$Late_i = \alpha_{o(i)} + \phi_{dwh} + \theta_{dws} + \xi_{dm} + \epsilon_i \tag{A-1}$$

In (4), we are evaluating officer differences in the likelihood of making a late arrest conditional on overall arrest activity. As with our baseline analysis in Equations (1) and (2), our fixed effects include division $\times$ day-of-week $\times$ hour, division $\times$ day-of-week $\times$ shift, and division $\times$ year-month. The objects of interest are the set of officer-level fixed effects, $\alpha_{o(i)}$, which document systematic differences in the share of arrests that are made at the end of the day. Notably, because each officer makes a finite number of arrests, each fixed effect will be estimated with error and naturally some fixed effects will be estimated greater precision than others. To adjust the distribution for estimation error, we use a Bayes shrinkage approach similar to that employed in the teacher value added literature (Morris, 1983).

The results of this regression are presented in Figure A-8 which plots both the unadjusted and shrunken distributions of the officer fixed effects.[20] We next use the shrunken fixed effects

[20]Because of the presence of other fixed effects in the regression, the officer fixed effects are approximately centered at zero. The fixed effects are not exactly centered at zero because officers have different numbers of arrests, and the fixed effects distribution is at the level of the officer. The

to explore whether the officer fixed effects are correlated with several signature behaviors of the collars for dollars story. In particular, we ask whether, among officers who are "late arresters," early versus late-shift arrests differ with respect to 1) the probability of conviction, 2) the share of arrests that are officer-initiated, 3) the share of arrests that are of African-American suspects and 4) the types of arrests that are made. To the extent that late arresters are differentially likely to make low quality arrests, officer initiated arrests, arrests of African-Americans or arrests that are more likely to lead to overtime pay, this potentially forms the basis for a claim that such behavior is motivated by the desire to secure access to overtime pay.

We explore these relationships in Figure A-9 which, for each outcome, plots the mean of the dependent variable separately for early shift arrests (the dashed line) and late-shift arrests (the solid line) for officers above a given percentile of the distribution of the shrunken fixed effects. We fail to see evidence that, among the late arresters, early and late-shift arrests differ with respect to the share of arrests they make that are officer-initiated and the share of arrestees who are non-white. With respect to conviction rates, if anything arrests during the final four hours of the shift have a higher conviction rate. We also generate a predicted overtime variable by multiplying each arrest by its expected number of overtime hours using statistics presented in Table 2. For each officer we compute predicted overtime hours on the basis of the distribution of that officer's arrest charges. There is little evidence that late arresters are differentially likely to make the types of arrests that are more likely to lead to overtime hours. in their shift. Had this been the case, we would have expected the dashed and solid lines to have switched positions at the top of the distribution.

Taken as a whole, the evidence suggests that while some officers consistently make late-shift arrests, this behavior cannot be explained by the type of strategic behaviors that have been suggested as pillars of the collars for dollars story. Instead, it is possible that officers who tend to make late-shift arrests simply have greater ability or skill to make late-shift arrests, either because their skills erode over the course of a shift at a lower rate or for some other reason.

average of the fixed effects weighted by the number of arrests is equal exactly to 0.

Table A-1: Summary Statistics for Ten Most Common Arrest Types

UCR Offense	(1) Share of Arrests	(2) Felony	(3) Convicted	(4) Hour of Shift	(5) Overtime Paid
All Arrests	1.00	0.21	0.34	4.22	0.45
Warrant	0.28	0.00	0.06	4.24	0.39
Assault	0.15	0.27	0.29	4.18	0.49
Disorderly Conduct	0.14	0.00	0.08	4.18	0.34
Narcotics & Drugs	0.13	0.51	0.71	4.30	0.49
Public Intoxication	0.03	0.01	0.07	4.16	0.37
Dwi	0.03	0.11	0.88	4.26	0.61
Trespass	0.02	0.02	0.81	3.99	0.44
Theft-Retail	0.02	0.32	0.74	4.13	0.45
Other-Misdemeanor	0.02	0.00	0.12	3.67	0.35
Theft-Other	0.02	0.70	0.73	4.16	0.56

Note: Table presents summary statistics for the ten most common types of arrest charges in the data. For each arrest charge, we note the share of arrests, the share that are felonies, the share resulting in a conviction, the mean shift-hour and the probability that the arrest charge leads to an overtime spell.

Table A-2: "First Stage" Regression: Relationship Between Predicted and Actual Off-Duty Work

	(1) Before OD	(2) After OD
Regular Off-Duty Before Shift	0.509***	0.0292***
	(0.00586)	(0.00291)
Regular Off-Duty After Shift	-0.000737	0.516***
	(0.00150)	(0.00556)
Mean	0.063	0.095
Observations	614921	614921

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Table presents estimates from a regression of an indicator for off-duty work on an indicator for predicted off-duty work, where the predicted work variable is equal to 1 if the officer worked an off-duty shift on a given day of the week at least 25 percent of the time in a given quarter. Models condition on the fixed effects described in Equation (1). Standard errors are clustered at the division-by-month level.

Table A-3: Balance Table: 911 Calls for Service by Shift-Hour

	(1) Log Calls	(2) Log Serious Calls	(3) Log Calls	(4) Log Serious Calls
1hr From End	-0.00417	-0.00241		
	(0.00231)	(0.00226)		
2hr From End	-0.00412	-0.00250		
	(0.00253)	(0.00214)		
3hr From End	-0.00282	0.00191		
	(0.00243)	(0.00214)		
4hr From End	-0.000946	0.00442*		
	(0.00245)	(0.00209)		
5hr From End	-0.000449	0.00417		
	(0.00241)	(0.00217)		
6hr From End	-0.000277	0.0000215		
	(0.00188)	(0.00174)		
7hr From End	0.000254	-0.0000129		
	(0.00145)	(0.00158)		
Before-Shift			-0.00122	-0.00163
Off-Duty			(0.00101)	(0.00107)
After-Shift Off-Duty			0.00362**	0.000495
			(0.00115)	(0.00103)
Mean	2.040	1.452	2.040	1.452
Observations	4919616	4919616	4919616	4919616
F-value	0.772	3.461	5.796	1.280
F-test	0.611	0.001	0.003	0.279

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Table presents estimates from a regression of the natural logarithm of the total number of 911 calls and 911 calls for high-priority calls on a vector of shift-hour dummies, conditional on the fixed effects in Equation (1). Results are presented separately for the hourly sample and the dispatch sample. Below the estimated coefficients and standard errors we report the F-statistic along with the its associated p-value on the joint significance of the shift-hour terms in predicting the number of service calls.

Table A-4: Most Common Off-Duty Jobs

American Airlines Center (AAC)
Greenway Parks Home Owners Association
Kalua Discoteque
Green Oaks Hospital
Cowboys Red River
Prestonwood PID ENP
Crescent Hotel
Hunt Oil
Southwest Center Mall
Preston Hollow North Inc ENP
North Bluffview ENP
Texas Scottish Rite Hospital for Children
Children's Medical Center
Bank of America Plaza
Inwood National Bank
Pegasus Link Constructors, LLC
Royalwood ENP
Watermark Church
Meadows Foundation Inc
Medical City Dallas Hospital ER

Note: Table presents a list of the most common off-duty jobs worked by Dallas police officers during the study period.

Table A-5: Robustness of Off-Duty Estimates to Alternative Models

	(1) Arrest	(2) Arrest	(3) Arrest	(4) Arrest
Regular Off-Duty Before × Early	0.000120 (0.000348)	0.000150 (0.000346)	0.000221 (0.000342)	0.000147 (0.000345)
Regular Off-Duty Before × Late	-0.000158 (0.000344)	-0.000107 (0.000349)	-0.000150 (0.000335)	-0.000128 (0.000343)
Regular Off-Duty After × Early	-0.000132 (0.000321)	-0.000224 (0.000325)	-0.000219 (0.000311)	-0.000187 (0.000320)
Regular Off-Duty After × Late	-0.000521** (0.000262)	-0.000607** (0.000265)	-0.000549** (0.000255)	-0.000497* (0.000263)
Mean	0.017	0.017	0.017	0.017
Observations	4919616	4919616	4919603	4917547
Baseline FE	X			
Baseline + Division-Date FE		X		
Division-Month-Weekend-Hour FE			X	
Division-Date-Hour FE				X

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Table presents estimates of variants on Equation (2) with different sets of fixed effects. Standard errors are clustered at the division-by-month level.

Table A-6: Effect of Off-Duty Employment, Short versus Long Pre-Shift Work

	Hourly Sample		Dispatch Sample	
	(1)	(2)	(3)	(4)
	Arrest	Arrest	Arrest	Arrest
Regular Off-Duty Before × Early × Short	0.000238	0.000237	-0.000127	-0.0000552
	(0.000441)	(0.000442)	(0.000908)	(0.000908)
Regular Off-Duty Before × Late × Short	-0.000220	-0.000213	-0.00270**	-0.00281***
	(0.000423)	(0.000423)	(0.00105)	(0.00105)
Regular Off-Duty Before × Early × Long	-0.0000821	-0.0000894	0.00154	0.00155
	(0.000508)	(0.000508)	(0.00113)	(0.00113)
Regular Off-Duty Before × Late × Long	-0.0000644	-0.0000518	-0.000716	-0.000646
	(0.000501)	(0.000501)	(0.00138)	(0.00138)
Regular Off-Duty After × Early × Short	-0.0000243	-0.0000215	-0.000968	-0.000941
	(0.000391)	(0.000391)	(0.000849)	(0.000849)
Regular Off-Duty After × Late × Short	-0.000262	-0.000259	-0.00234***	-0.00233***
	(0.000308)	(0.000308)	(0.000857)	(0.000855)
Regular Off-Duty After × Early × Long	-0.000309	-0.000298	-0.000747	-0.000712
	(0.000481)	(0.000481)	(0.000892)	(0.000893)
Regular Off-Duty After × Late × Long	-0.000913**	-0.000904**	-0.00241**	-0.00238**
	(0.000398)	(0.000398)	(0.00112)	(0.00112)
Log Calls		-0.00199***		-0.0131***
		(0.000232)		(0.000564)
Log Serious Calls		0.00289***		-0.00272***
		(0.000181)		(0.000454)
Mean	0.017	0.017	1.000	1.000
Observations	4919616	4919616	1754401	1754401

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Table presents estimates from a series of regressions of an arrest indicator on the interaction between either predicted before or after-shift off-duty work and indicators for 1) whether a shift-hour is early (first four hours) or late (second four hours) in an officer's shift and 2) whether the off-duty is long (more than 4 hours) or short (less than four hours). Standard errors are clustered at the division-by-month level.

Table A-7: Effect of Off-Duty Employment on Arrest Outcomes

	(1) Guilty	(2) AnySentence	(3) Felony	(4) Mis/Fel
Regular Before Shift	0.00505	0.00249	0.00757	-0.00460
	(0.0106)	(0.00927)	(0.00787)	(0.00900)
Regular After Shift	-0.00896	-0.000893	0.00345	-0.00593
	(0.00975)	(0.00871)	(0.00763)	(0.00851)
Mean	0.355	0.247	0.187	0.724
Observations	62042	62042	62042	62042

Note: Table presents estimates from a series of regressions of different outcomes of an arrest on an indicator for whether the officer has a regular off-duty shift either before or after his or her police shift. Models condition on the fixed effects described in Equation (2). Standard errors are clustered at the division-by-month level.

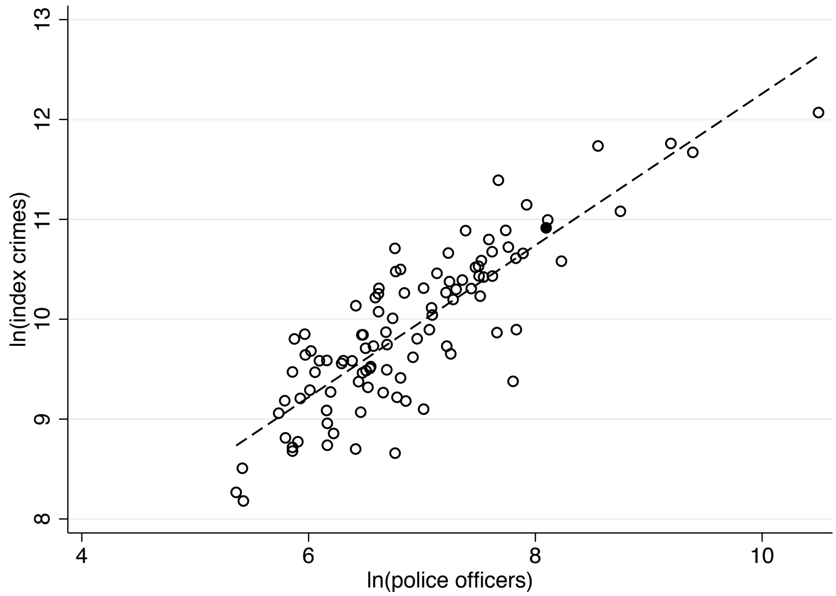
Table A-8: Model Parameter Estimates, Alternative Specifications

Parameters	Baseline	Court Errors	Biggest OD Decline	Both	Description
ρ	0.033	0.019	0.032	0.019	Probability of guilt
μ	1.585	2.023	1.610	2.053	Mean of signal for guilty individuals (officer ability)
λ	0.133	0.064	0.130	0.061	Tradeoff for arresting guilty and not arresting innocent individuals
c_{ot}	-0.219	-0.229	-0.270	-0.283	Intercept value/cost of working overtime
c_{od}	-0.018	-0.019	-0.060	-0.066	Intercept value/cost of working off-duty
b	0.00021	0.00022	0.00057	0.00062	Value of $1 of overtime/off-duty pay
ϕ	0.741	0.742	0.740	0.742	Probability or processing an arrest
p	0.766	0.765	0.766	0.765	Per-period probability of continued arrest processing
p_{start}	0.349	0.336	0.348	0.335	Probability of beginning work on patrol
p_{ot}	0.290	0.290	0.290	0.290	Probability of receiving overtime when not making an arrest
Parameters					Description
λ/b	626.0	284.6	226.9	98.1	Value of non-arrest of innocent person
$(1-\lambda)/b$	4063.2	4182.0	1515.4	1519.2	Value of arrest of guilty person
$E(v(a,s)\|a=0)$	611.5	281.7	221.9	97.2	Average value of non-arrest
$E(v(a,s)\|a=1)$	1415.1	1300.3	528.6	473.4	Average value of arrest
$E(\max_{a_t} v(a,s)) - v_{np}$	29.7	33.0	11.2	12.1	Average value of hour on patrol (relative to hour not on patrol)

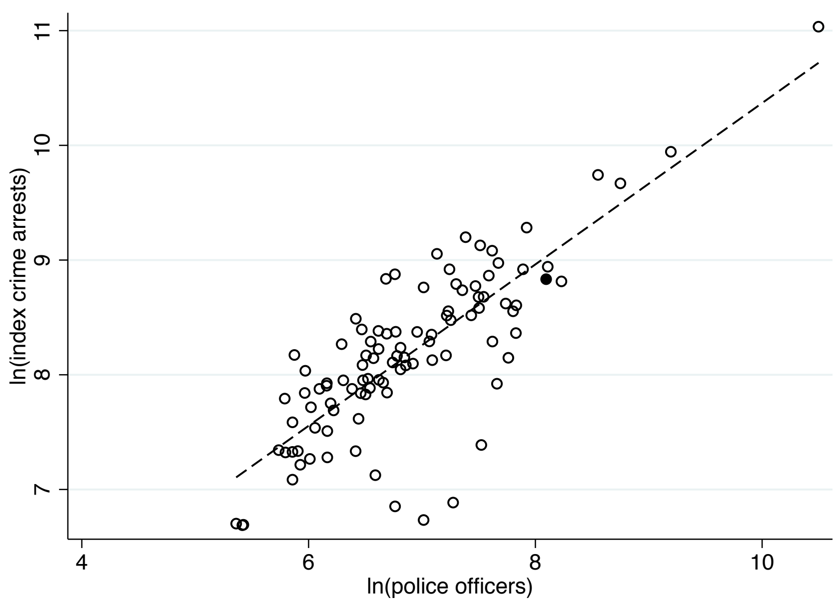
Note: Table presents parameter estimates from the model in Section 6 under alternative modeling choices described in Section 6.5.

Figure A-1: Police, Arrests and Index Crimes: U.S. Cities, 2016

(A) Police Officers and Index Crimes

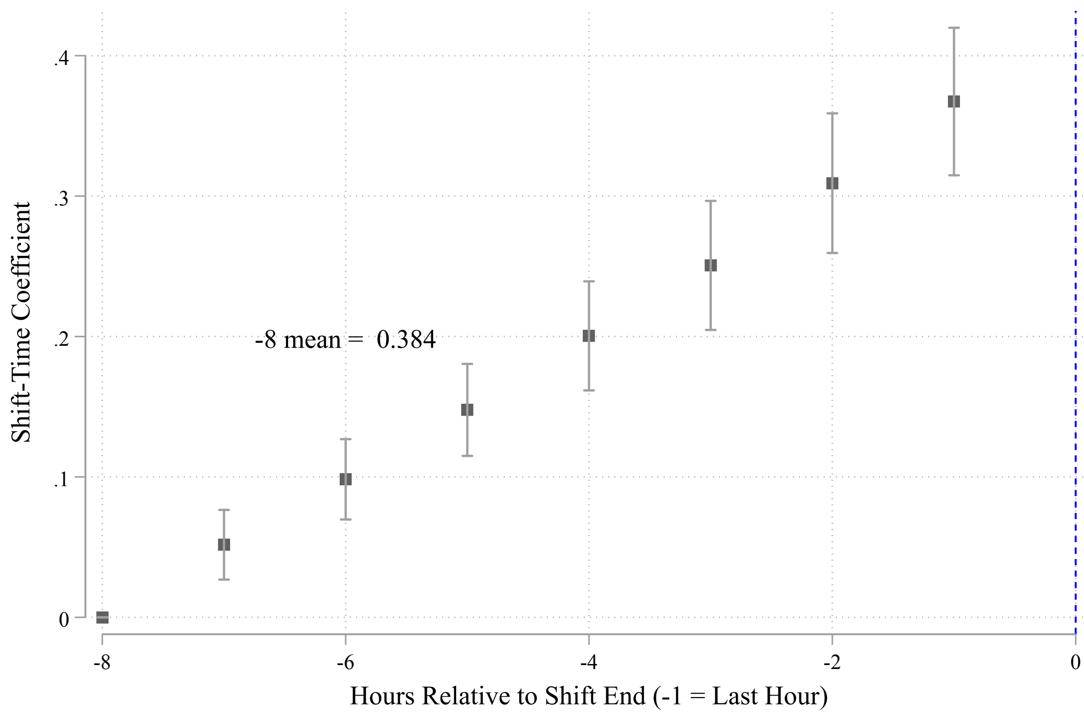


(B) Police Officers and Index Crime Arrests



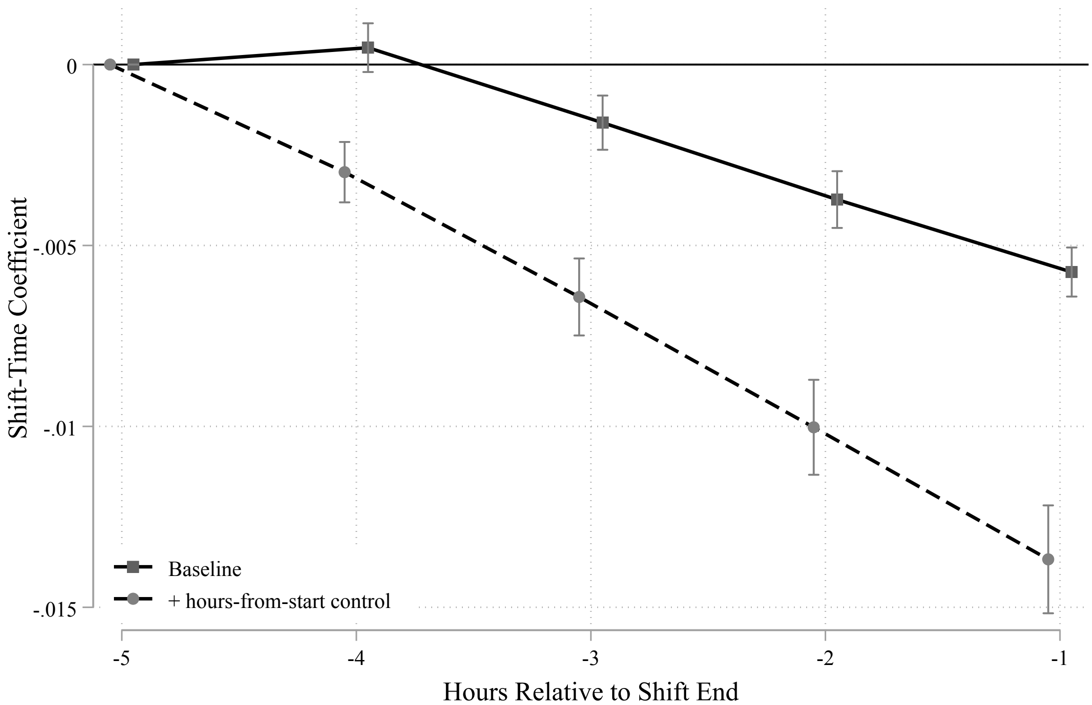
Note: Figure plots the natural logarithm of index crimes (Panel A) and the natural logarithm of index crime arrests (Panel B) against the natural log of sworn police officers for cities with populations over 250,000 residents in 2016. Data on crimes, arrests and police manpower come from the Federal Bureau of Investigation's Uniform Crime Reports accessed from Kaplan (2019).

Figure A-2: Probability of Overtime Pay by Shift-Hour, Regression Adjusted



Note: Figure plots the probability that an arrest made in a given hour of an officer's shift leads to overtime pay, conditional upon the fixed effects in Equation (1). The -8 hour corresponds with the first hour of the shift; the -1 hour corresponds with the final hour of the officer's shift. 95 percent confidence intervals, clustered at the division-by-month level, are provided for each statistic.

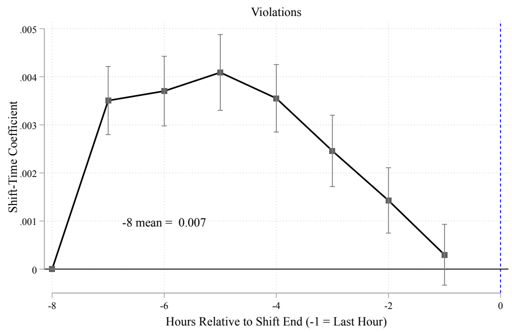
Figure A-3: Accounting for Time-Into-Shift Effects



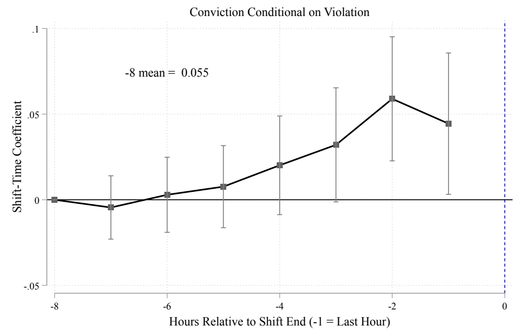
Note: Figure plots robustness analyses described in Section 5.3. The baseline sample is expanded to include nine-hour and ten-hour shifts. The "Baseline" coefficients are from estimation of Equation (1) with only coefficients for the last four hours, and the second set of coefficients are from the same equation with the inclusion of a variable for hours into the shift.

Figure A-4: Arrest Frequency and Court Conviction Regressions

A. Violation Arrest

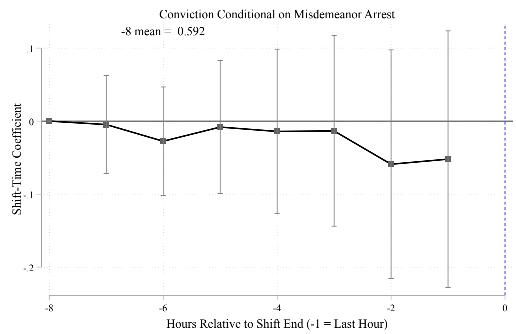


B. Conviction Conditional on Violation Arrest

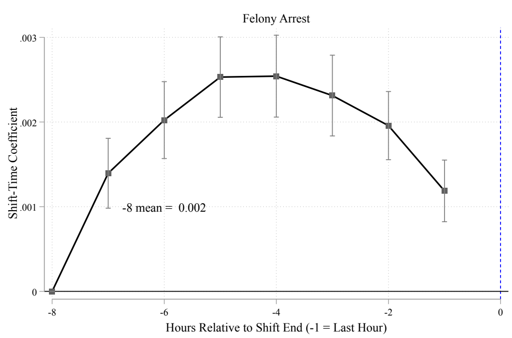


C. Misdemeanor Arrest

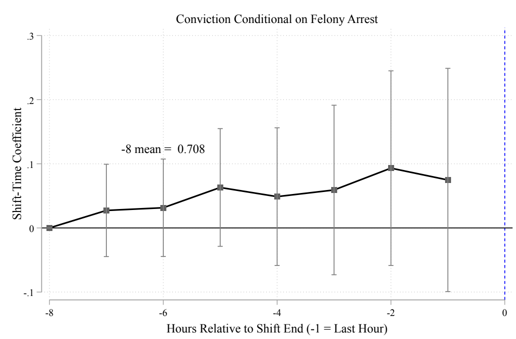
D. Conviction Conditional on Misdemeanor Arrest



E. Felony Arrest



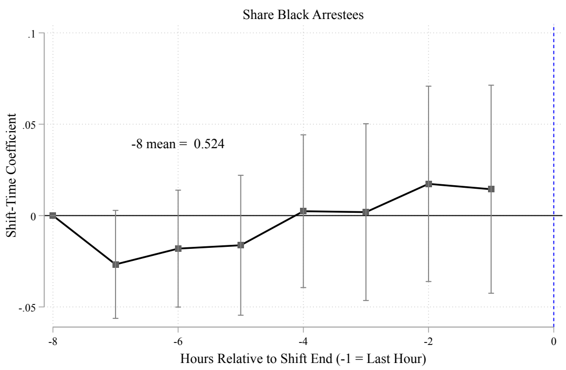
F. Conviction Conditional on Felony Arrest



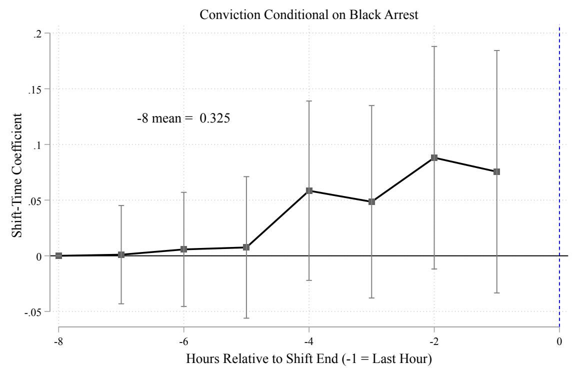
Notes: Figure plots each shift-hour coefficient from a regression of a given outcome variable on a vector of shift-hour indicator variables, conditional on the fixed effects described in Equation (1). We present estimates separately for arrests for violations, misdemeanor arrests, felony arrests and felony arrests excluding drug crimes. For each crime type, we also plot the probability of a conviction given an arrest. 95 percent confidence intervals, computed using standard errors that are clustered at the division-by-month level, provide a boundary around the point estimates. All coefficients are relative to the arrest incidence during the first hour of an officer's shift.

Figure A-5: Arrest Frequency and Court Conviction Regressions

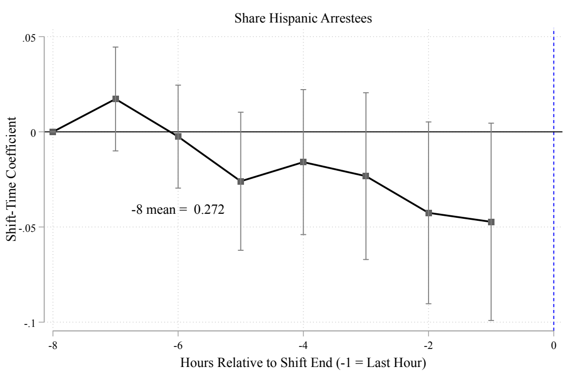
A. Share Black Arrestee



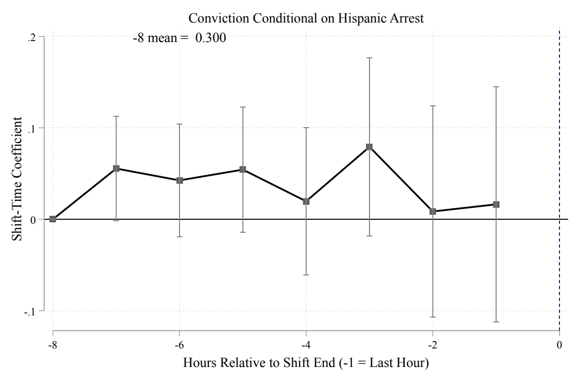
B. Conviction Conditional on Black Arrest



C. Share Hispanic Arrestee



D. Conviction Conditional on Hispanic Arrestee

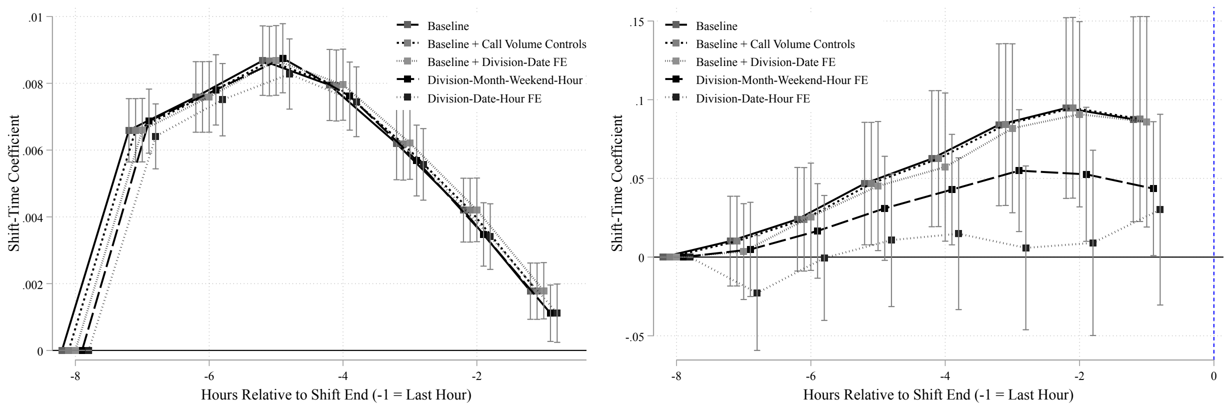


Note: Figure plots each shift-hour coefficient from a regression of a given outcome variable on a vector of shift-hour indicator variables, conditional on the fixed effects described in Equation (1). We present estimates for the share of arrestees who are Black and Hispanic as well as the conviction rate for Black and Hispanic arrestees. 95 percent confidence intervals are presented, computed using standard errors that are clustered at the division-by-month level. All coefficients are relative to the arrest incidence during the first hour of an officer's shift.

Figure A-6: Robustness to Alternative Specifications

A. Arrest Regression

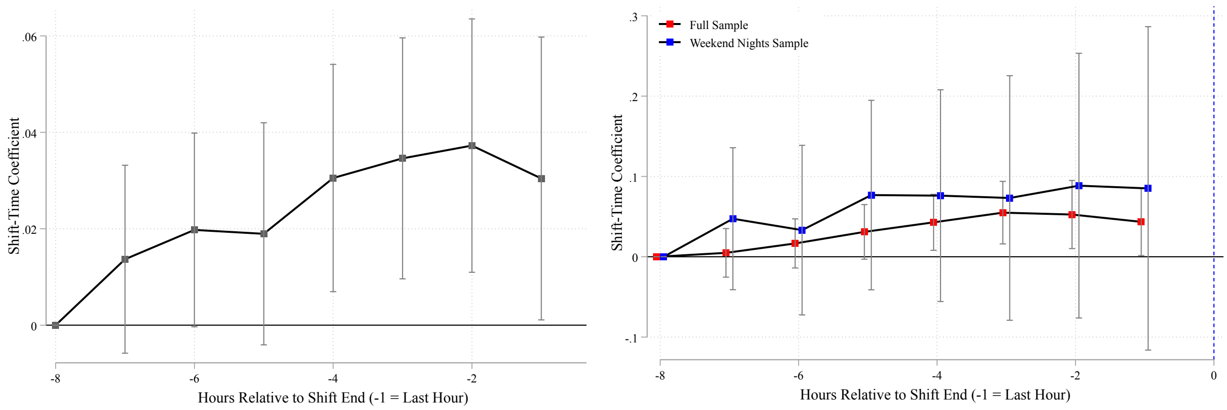
B. Court Conviction Regression



Note: Figures present versions of the Arrest and Court Conviction Regressions from Figure 4 with alternative sets of fixed effects.

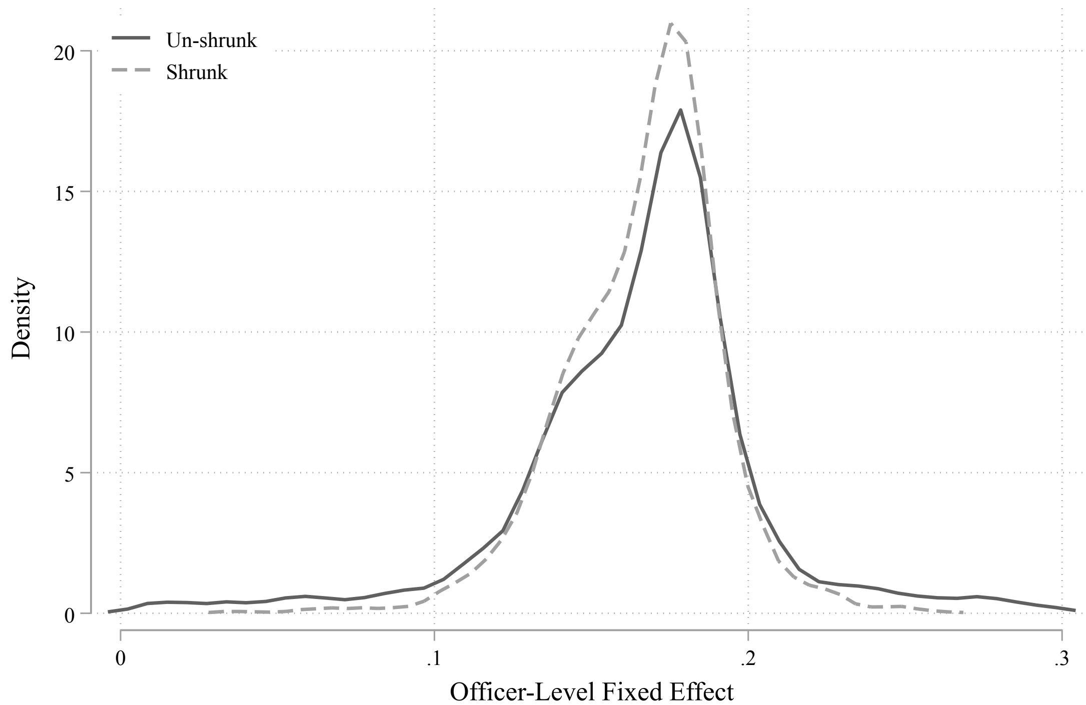
Figure A-7: Ruling Out Alternative Hypotheses

A. Number of Charges   B. Court Conviction Effect



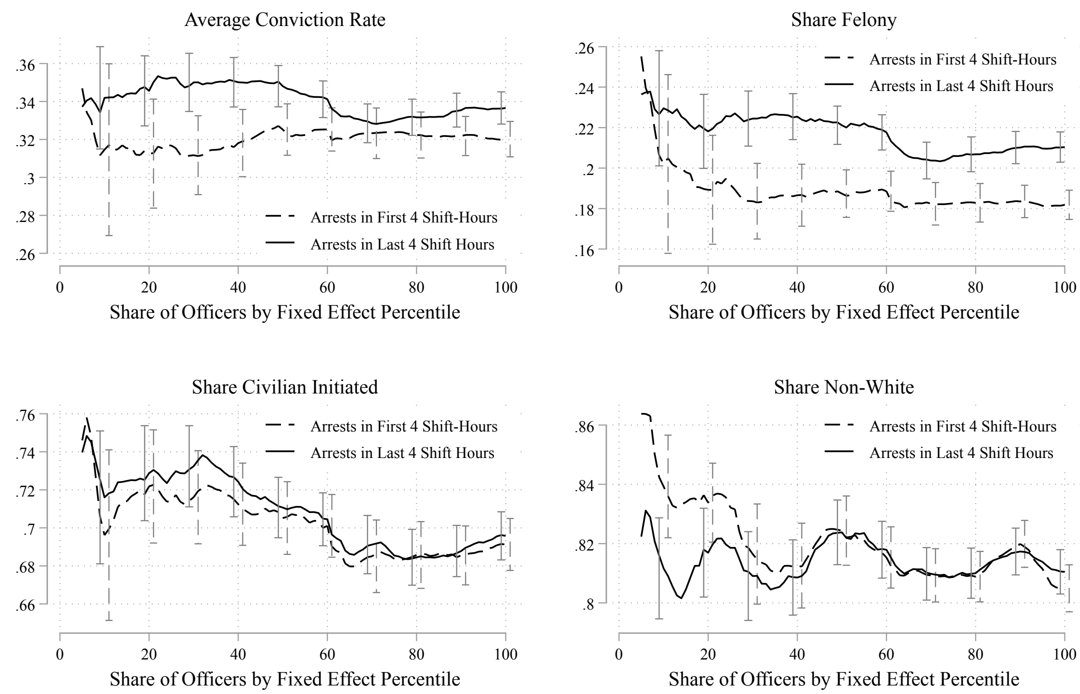
Note: The left-hand figure plots the number of arrest charges as a function of given hour of an officer's shift. The right-hand figure plots the probability of a court conviction in each hour of an officer's shift, separately for the full sample and for shift ending between 4:00pm and 10:00pm on Friday and Saturdays. All estimates condition upon the fixed effects in Equation (1). The -8 hour corresponds with the first hour of the shift; the -1 hour corresponds with the final hour of the officer's shift. 95 percent confidence intervals, clustered at the division-by-month level, are provided for each statistic.

Figure A-8: Distribution of Officer Fixed Effects



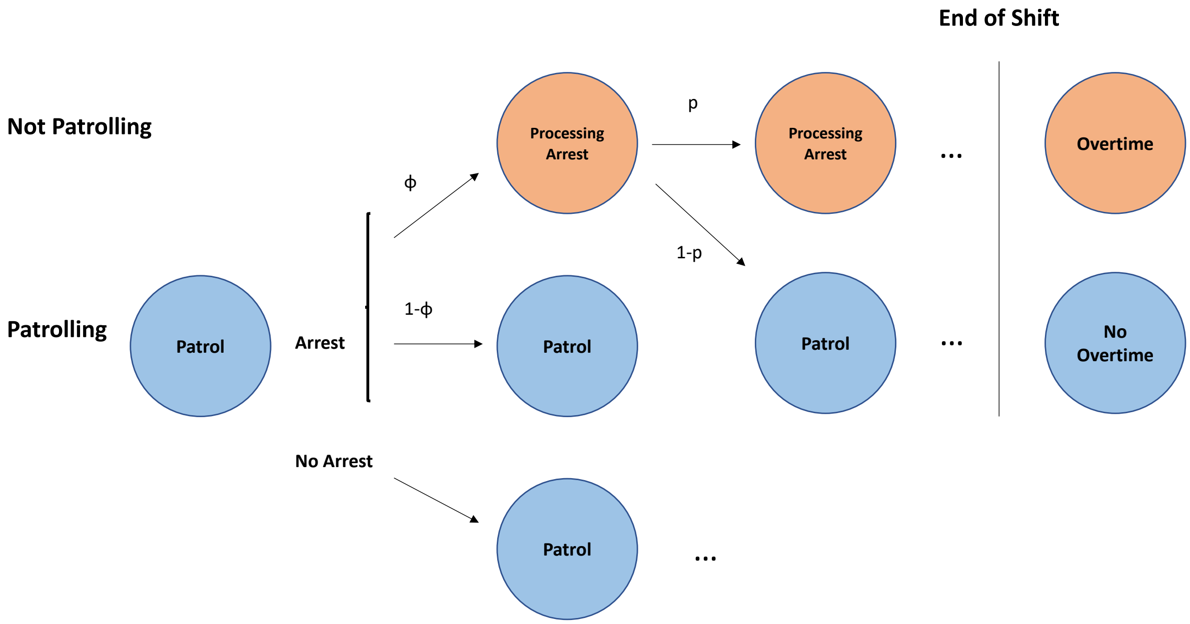
Note: Figure plots the the unadjusted distribution (the solid line) and the shrunken distribution (the dotted line) of officer fixed effects with respect to late-shift arrest activity. The fixed effects are estimated using a regression of whether an arrest occurs in the final two hours of an officer's shift on officer fixed effects. Because each officer makes a finite number of arrests, each fixed effect will be estimated with error and naturally some fixed effects will be estimated greater precision than others. To adjust the distribution for estimation error, we use a Bayes shrinkage approach similar to that employed in the teacher value added literature (Morris, 1983).

Figure A-9: Arrest Outcomes Across Officer Fixed Effect Percentile Cutoffs



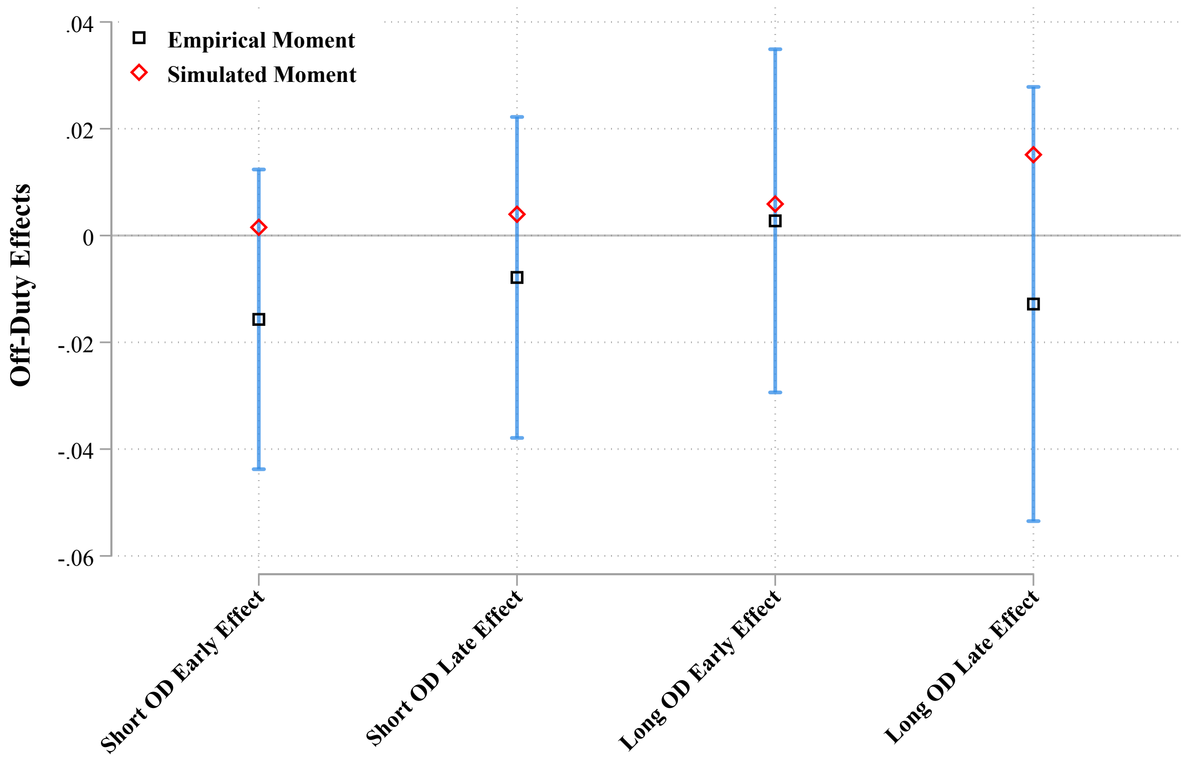
Note: Figure plots the mean of a given outcome variable separately for early shift arrests (the dashed line) and late-shift arrests (the solid line) for officers above a cutoff percentile of the distribution of the shrunken fixed effects. We present means for the conviction rate, the share of arrests that are officer-initiated, the share of arrests of non-white suspects and predicted overtime by crime type.

Figure A-10: Visualization of Model



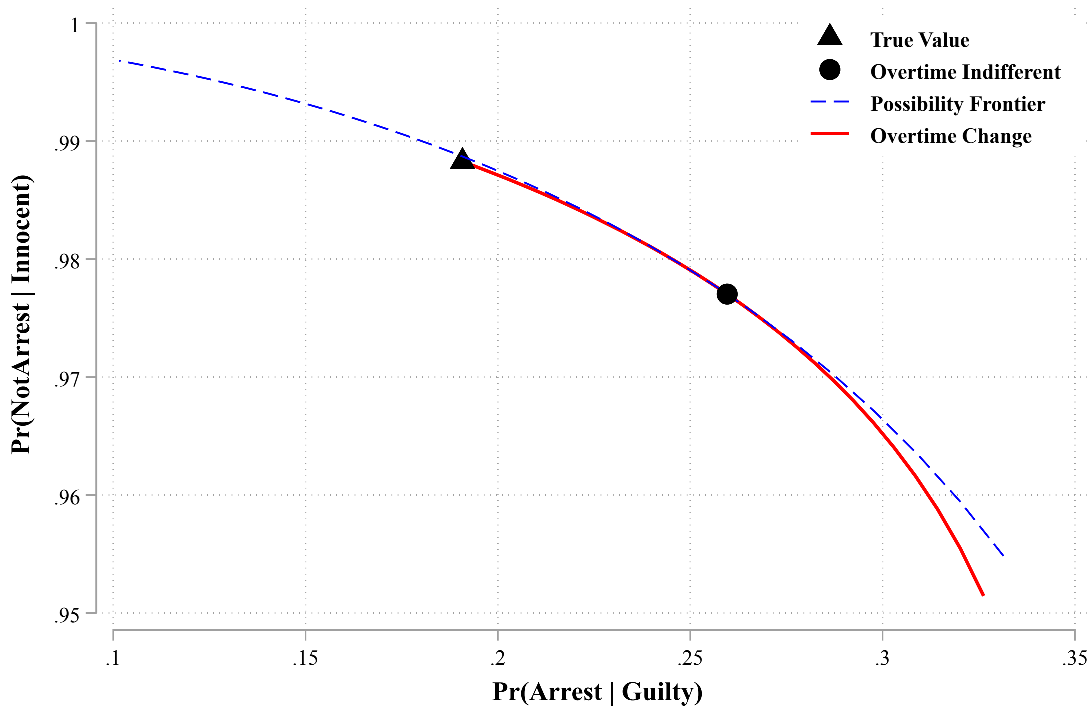
Note: Figure presents a schematic of the model presented in Section 6.

Figure A-11: Model Fit, Off-Duty Guilt Effects



Note: Figure plots empirical versus simulated moments as a test of model fit for regular off-duty arrest impacts.

Figure A-12: Production Possibilities Frontier



Note: This figure uses the model estimates to calculate the probability of not arresting a suspect who is innocent and the probability of arresting a suspect who is guilty for various parameter values, as described in Section 6. The dotted line is constructed by setting $V_{ot} = 0$ (i.e. overtime indifference) and spanning different values of λ, the relative weights on arresting guilty suspects and not arresting innocent suspects. The solid triangle is the value for the baseline estimated parameters, and the solid circle is the value when V_{ot} is then set to 0. The red line traces the outcome for the baseline estimated parameters with progressive increases in overtime pay.