

event data. In addition, we will examine whether the distribution of outcomes differ across treatment and control using distribution tests.

In addition, if we observe imbalances in outcomes prior to the start of training, we will use techniques such as controlling and matching to address this. With regard to sample, we plan to analyze administrative outcomes for the complete randomization sample as well as the subset of officers who completed endline assessments, to compare administrative outcomes and assessment outcomes over a common set of officers.

## 5 Outcomes

Uri Simonsohn added numbers next to each outcome to facilitate counting them.

In this section we list the primary outcomes and secondary outcomes we plan to examine. We divide these into two broad sets of outcomes – those that reflect actions by officers and those that reflect other individuals' responses to officers' actions.

### 1. Primary Outcomes

#### (a) Officer's Actions

- From TRR data:
  - Uses of force: measured by the number and occurrence of any TRRs. This will include TRRs of all levels,<sup>1</sup> as well as TRRs of lower levels<sup>2</sup> (1 and 2) and TRRs<sup>3</sup> of higher levels (2 and 3).
  - Uses of force (TRRs) that resulted in subjects being injured, based on subject alleging injury: measured by the occurrence<sup>4</sup> of any such TRR, the number<sup>5</sup> of such TRRs as well as fraction of TRRs that resulted in injury.
- Arrests pre-categorized as "discretionary arrests involving interactions with subjects"<sup>6</sup> which includes categories such as disorderly conduct and resisting or obstructing an officer.

#### (b) Responses to Officer's Actions (from complaint data)

- Number of accusations for excessive force<sup>7</sup>
- Number of accusations for civil rights violations<sup>8</sup>
- Cluster of accusations for verbally abusive behavior, excessive force and civil rights violations for improper search and improper stop / arrest: measured by<sup>9</sup> both the number of accusations that contain at least one allegation of these types<sup>10</sup>; and the number of distinct types of allegations across these accusations.

### 2. Secondary Outcomes

#### (a) Officer's Actions

- Tactics used in TRRs, including (1) movement to avoid attack; zone of safety; tactical positioning; and verbal direction; and (2) open-hand strike, take down, elbow strike, closed-hand strike / punch, knee strike, kicks and push /

for example, here we see outcomes 1-3

11-19:

ambiguous to parse, reads like 3 possible categories, measured as 0/1, number, or fraction. So 9 variables. They could be collapsed, and divided further, this is the judgment call I was less confident of.

Could be just 3 could be way more than 9.

physical re-direction, taser, and (3) any reportable force used against a subject who is handcuffed or otherwise restrained. Measured as the **occurrence** of any such TRR, the **number** of such TRRs as well as **fraction** of TRRs where these tactics were used.

- Other measures of subject injury under TRRs:

20-23

- Uses of force (TRRs) that resulted in subjects being injured, based on **subject being injured according to the officer**: measured by the **occurrence** of any such TRR, the **number** of such TRRs as well as **fraction** of TRRs that resulted in injury.

24-26

- Uses of force (TRRs) that resulted in subjects being hospitalized: measured by the **occurrence** of any such TRR, the **number** of such TRRs as well as **fraction** of TRRs that resulted in injury.

not clear how diff from 24-26, so not counted

- Uses of force (TRRs) that resulted in subjects being hospitalized, in which either the subject alleged injury or based on the subject being injured according to the officer: measured by the occurrence of any such TRR, the number of such TRRs as well as fraction of TRRs that resulted in injury.

27 ambiguity whether to be analyzed aggregated or individually, count just 1 conservatively

- **Aggregate measures of officer activity**, measured by the number of following activities: Warrants; Recovered vehicles; Recovered guns; Traffic stops; Driver stops; ISRs / contact cards; ANOVs; Citations; Curfew violations; CTA checks; Parking citations; and all arrests except those pre-categorized as “discretionary arrests involving interactions with subjects”.

28-31

- **Arrests by type**: total, felonies, misdemeanors, and other / not categorized.
- Commendations and awards: measured as number as well as receipt of any honorable mentions, department commendations, as well as commendations and other high-level awards allocated to individual active duty officers based on performance.<sup>1</sup>

32 (assuming this commits to only all awards)

- If we are able to access data from CPD’s **Force Review Unit (FRU)**, we will examine additional outcomes from this data source. Because we do not know the structure, completeness or quality of this data, or precisely which variables these data contain, we cannot project which **variables** we will be positioned to analyze. Data issues permitting, we hope to examine, for each officer, the number of TRRs that the FRU recommended for advisement; for training or a sit-down with the supervisor; or to COPA. We will also aim to conduct exploratory analysis with **additional relevant outcomes** from this data.

33

this is very broad but counting just 1 variable

(b) Responses to Officer’s Actions

34-35

- **Number** of total accusations (regardless of specific allegation type) and number of **distinct types** of allegations of various types across all accusations.

<sup>1</sup>These high-level awards include: Hundred Club of Cook County Medal of Valor, Superintendent’s Award of Valor, Carter Harrison Award, Life Saving Award, Police Officer of the Month, Problem Solving Award, Thomas Wortham IV Military and Community Service Award, Superintendent’s Honorable Mention, Traffic stop of the month award, Top Gun Arrest Award, Superintendent’s Award of Tactical Excellence.

- 36-37 • Cluster of accusations for arrest / lock-up, conduct unbecoming and traffic, not bribery or excessive force: measured by both the **number of accusations that contain at least one allegation of these types**; and the **number of distinct types of allegations across these accusations**.
- 38-39 (assuming authors committed to including all) • Cluster of accusations that go to COPA, which includes bias, coercion, death in custody, domestic violence, excessive force, improper search / seizure, firearms, taser that results in death / injury, pattern or practice of misconduct, and denial of access to counsel: measured by both the **number** of accusations that contain at least one allegation of these types; and the number of **distinct types** of allegations across these accusations.
- 40-41 • Cluster of accusations for domestic violence, conduct unbecoming for altercations or disturbances, and conduct unbecoming for harassment: measured by both the **number of accusations** that contain at least one allegation of these types; and the **number of distinct types** of allegations across these accusations.
- 42-43 • Cluster of accusations for Operation Personnel Violations for sub-categories inadequate service, neglect of duty, slow response / no response, failure to identify, traffic pursuit violations of procedure (GO G0303-01), and conduct unbecoming measured by both the **number of accusations** that contain at least one allegation of these types; and the number of **distinct types** of allegations across these accusations.

## 6 Dimensions of Heterogeneity

We will examine heterogeneity in the effects based on the following characteristics:

- **District conditions.** Treatment effects may vary based on the crime and demographic composition of the locations in which officers are working. We will therefore examine heterogeneity on the basis of violent crimes, rates of citizen complaints and racial composition in the districts and areas of assignment.
- **Assignments.** Treatment effects may also differ for officers who work in more active assignments, through which they are more likely to interact with citizens. We will therefore look at heterogeneity based on assignment. Our focus will be on the configuration used at the time of randomization, under which officers on tactical teams, mission teams, area saturation teams, gang teams and gun teams were considered to be more active assignments. In 2020, after the Sit-D training was underway, role designations also changed and several of the previous active categories were eliminated while new roles intended to take on active assignments were created (e.g. city-wide units). Contingent on attaining a complete mapping of new and old role designations for the officers in our sample, we will also gauge heterogeneous effects based on active roles under the new configuration. In addition, we will gauge effects specifically for officers considered to have regular policing roles (which encompass designations such as beat officer, beat relief officer, rapid response officer, and police officer).