# Appendix to:
# Detecting Anomalies in Data on Government Violence

# Appendices

## A CDCR Data

Table A1 summarizes the data, reporting the total number of reported incidents, disaggregated by force type.[1] Importantly, however, the institutions overseen by the CDCR differ considerably in their propensity to use force against inmates, as A1 illustrates. To create Figure A1, we generate the sum of all uses of force in prison, $i$, in year $t$, and plot those values for each institution for every year between 2008 and 2017.

**Table A1:** Total Reported Incidents by Type

| **Type of Force** | **Total Incidents** | **"Zero" Incidents** |
|---|---|---|
| Use of Oleoresin Capsicum (Pepper Spray) | 43,569 | 120 |
| Physical Force | 24,841 | 339 |
| Discharge of a 37mm and/or 40mm Launcher | 11,999 | 1,624 |
| Use of a Baton | 6,659 | 1,458 |
| Use of Chloroacetophenone (Tear Gas or Chemical Cace) | 2,079 | 3,469 |
| Other Force Options Not Otherwise Mentioned | 665 | 3,766 |
| Firing of Semi-Automatic Rifle Shots (Warning) | 350 | 4,099 |
| Use of High-Pressure Water Hose System | 105 | 3,573 |
| Use of Non-Conventional Force | 76 | 565 |
| Firing of Semi-Automatic Rifle Shots (Contact Intended) | 71 | 4,229 |

Note: More than one type of force can be used per incident.

The 35 institutions overseen by the CDCR differ considerably in their propensity to use force against inmate. Figure A1 plots the total uses of force over time for each institution. One notices right away that the considerable variation in total across institutions. There is also considerable variation over time within institutions. Some high security prisons like the Salinas Valley State Prison (SVSP), the California State Prison in Los Angeles Country (LAC), and the California State Prison in Sacramento (SAC) all average over 550 uses of force per year while Pelican Bay (PBSP)—another high security institution—averages only 200 incidents. Some general population prisons, like Avenal State Prison (ASP), average only about 100. There is also considerable variation over time within institutions. Consider, for example, the Richard J. Donovan Correctional Facility (RJD), where reported uses of force fell considerably in 2010, and California State Prison, Corcoran (COR), which has seen reports of use of force steadily increase over time.

[1] Institutions are not mandated to report when force was *not* used. The choice to report zeroes, rather than to leave the category blank, may itself reflect strategic considerations.

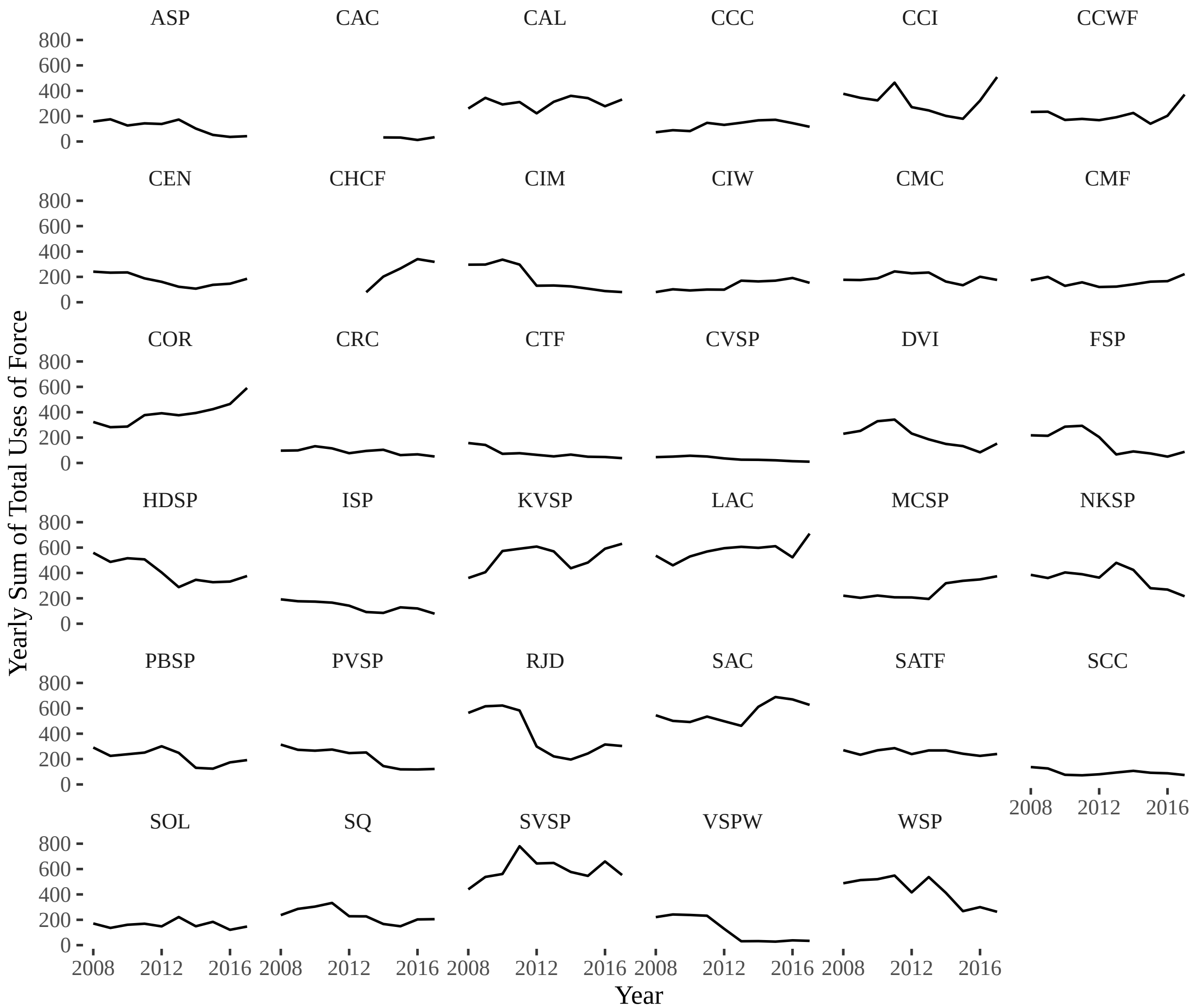


**Figure A1:** Total Uses of Force, by Institution (2008-2017)

# Appendices

## B List of Institutions by Institution Type

**Table B1:** List of Institutions by Institution Type

| Institution Type | Institutions |
|---|---|
| **Gen Pop** | Avenal State Prison; Calipatria State Prison; California State Prison, Centinela; Correctional Training Facility; Chuckawalla Valley State Prison; Ironwood State Prison; Mule Creek State Prison; Pleasant Valley State Prison; California State Prison, Solano; Valley State Prison |
| **High Sec** | California City Correctional Facility; California Correctional Institution; California State Prison, Corcoran; High Desert State Prison; Kern Valley State Prison; California State Prison, Los Angeles County; Pelican Bay State Prison; California State Prison, Sacramento; California Substance Abuse Treatment Facility and State Prison, Corcoran; Salinas Valley State Prison |
| **Reception** | California Correctional Center; California Institution for Men; California Men's Colony; California Rehabilitation Center; Deuel Vocational Institution; North Kern State Prison; Richard J. Donovan Correctional Facility; Sierra Conservation Center; San Quentin State Prison; Wasco State Prison-Reception Center |
| **Female** | Central California Women's Facility; California Health Care Facility, Stockton; California Institution for Women; California Medical Facility; Folsom State Prison;[†] Valley State Prison[‡] |

*Source*: `https://www.cdcr.ca.gov/adult-operations/` (Accessed September 2019).

[†] In 2013 a women's wing was added to the otherwise all male prison.

[‡] Valley State Prison was a female prison from 2008–2012. In 2013, it was converted to a prison for low-risk male inmates.

# Appendices

## C Digit Distribution By Year

**Table C1:** Distribution of First Digits of Total Uses of Force in California Prisons, Compared to Benford's Law

| Digit | Benford | All Years | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.301 | 0.333 | 0.338 | 0.313 | 0.323 | 0.285 | 0.381 | 0.353 | 0.363 | 0.378 | 0.301 | 0.289 |
| 2 | 0.176 | 0.219 | 0.278 | 0.283 | 0.207 | 0.225 | 0.184 | 0.232 | 0.175 | 0.179 | 0.217 | 0.217 |
| 3 | 0.125 | 0.128 | 0.131 | 0.139 | 0.144 | 0.149 | 0.136 | 0.093 | 0.120 | 0.116 | 0.133 | 0.123 |
| 4 | 0.097 | 0.103 | 0.109 | 0.104 | 0.111 | 0.136 | 0.111 | 0.111 | 0.084 | 0.080 | 0.091 | 0.091 |
| 5 | 0.079 | 0.066 | 0.053 | 0.056 | 0.078 | 0.071 | 0.051 | 0.053 | 0.072 | 0.065 | 0.067 | 0.099 |
| 6 | 0.067 | 0.048 | 0.033 | 0.023 | 0.061 | 0.051 | 0.045 | 0.045 | 0.053 | 0.041 | 0.062 | 0.062 |
| 7 | 0.058 | 0.037 | 0.013 | 0.033 | 0.020 | 0.030 | 0.030 | 0.040 | 0.050 | 0.041 | 0.052 | 0.057 |
| 8 | 0.051 | 0.033 | 0.020 | 0.020 | 0.038 | 0.035 | 0.030 | 0.035 | 0.043 | 0.039 | 0.037 | 0.035 |
| 9 | 0.046 | 0.033 | 0.025 | 0.030 | 0.018 | 0.018 | 0.030 | 0.038 | 0.038 | 0.061 | 0.040 | 0.027 |

# Appendices

## D Robustness Checks

The results of Tables 1 and 2 of the main text show evidence that the uses-of-force data often do not conform to Benford's Law. When discussing those results we have given particular attention to the $\chi^2$ test statistic and its associated p-value. However, as Pericchi and Torres (2011) and Sellke, Bayarri, and Berger (2001) remind, frequentist p-values are not the same as the posterior probability that the null hypothesis is true. In studies like ours, the latter quantity is very much of interest however. To that end, Sellke, Bayarri, and Berger (2001) and Perichi and Torres (2011) offer techniques to convert frequentist p-values to such posterior probabilities. We refer the interested reader to those articles for further discussion and mathematical derivations, but the essence of the technique is to apply a variety of Bayesian prior distributions to the problem and to then calculate the absolute minimum posterior probability that the null is true across that large class of prior distributions. That minimum probability is referred to as the Ultimate Lower Bound that $PR(H0|data)$. The conversion can be quite important—Pericchi and Torres (2011, Table 5) show that a p-value of 0.05 still has an Ultimate Lower Bound of 0.29, indicating that there is *at least* a 30% chance of the null being true given the data. We apply that technique to our $\chi^2$ test statistics below. Another approach to the same issue is the use of Bayes Factors to produce a statistic indicating the strength of the evidence in favor of the null hypothesis compared to the alternative (Pericchi and Torres 2011; Sellke, Bayarri, and Berger 2001). We apply this technique as well.[2]

Another potential concern with the results in Tables 1 and 2 is that testing the data in every year and for every institution type produces a multiple comparisons problem (Benjamini and Hochberg 1995). To account for that possibility, we conduct the standard p-value adjustment to correct for the false discovery rate, as presented in Benjamini and Hochberg (1995).[3]

Table D1 below shows the results of these three additional tests. The first two columns are the $\chi^2$ test statistic as presented in Tables 1 and 2 of the main text. The third column generates the frequentist p-value associated with those $\chi^2$ test statistics. Column "ULB" is the Ultimate Lower Bound. Next, is the Bayes Factor. The last column in the p-value adjust to control the rate of false discovery. Read the table across the columns as follows: For all years combined, for example, the p-value of 0.00 obliges one to reject the null hypothesis that the data are distributed according to Benford's Law. The Ultimate Lower Bound correction affirms that result, as it implies that the probability the null is true is *at least* 0.00. The Bayes Factor is also very small (0.00), further obliging us to conclude that data are not distributed according to Benford. This is a conclusion we can maintain even after adjusting for the rate of false discovery. Meanwhile, for the year 2016, the test statistic is small and the p-value is large, indicating, as we say in the manuscript, that we cannot reject the null hypothesis that the data are Benford. The Ultimate Lower Bound, Bayes Factor, and adjusted p-value all concur that we cannot reject the null in this case.

For the most part, none of these tests obliges us to alter our conclusions as discussed in the manuscript. In general, where the $\chi^2$ statistic and associated p-value obliged us to reject the null,

[2]The Ultimate Lower Bound is calculated using the `pcal` function of the `pcal` package (Fonseca and Paulo 2020) for the R environment (R Core Team 2020). The Bayes Factor is calculated using the `bcal` function of the same package.

[3]To do so, we use the `p.adjust` function in R, with the `fdr` argument specified.

**Table D1:** Additional Tests of Conformity to Benford's Law By Year and Institution Type

| ID | $\chi^2$ | PV | ULB | Bayes Factor | FDR p |
|---|---|---|---|---|---|
| All Years | 159.47 | 0.00 | 0.00 | 0.00 | 0.00 |
| 2008 | 61.19 | 0.00 | 0.00 | 0.00 | 0.00 |
| 2009 | 54.77 | 0.00 | 0.00 | 0.00 | 0.00 |
| 2010 | 22.96 | 0.00 | 0.05 | 0.05 | 0.00 |
| 2011 | 29.79 | 0.00 | 0.01 | 0.01 | 0.00 |
| 2012 | 27.38 | 0.00 | 0.01 | 0.01 | 0.00 |
| 2013 | 25.38 | 0.00 | 0.02 | 0.02 | 0.00 |
| 2014 | 8.97 | 0.34 | 0.50 | 1.00 | 0.37 |
| **2015** | **19.90** | **0.01** | **0.12** | **0.13** | **0.01** |
| 2016 | 7.41 | 0.49 | 0.50 | 1.00 | 0.49 |
| 2017 | 11.61 | 0.17 | 0.45 | 0.82 | 0.20 |
| Female | 169.47 | 0.00 | 0.00 | 0.00 | 0.00 |
| General Population | 131.44 | 0.00 | 0.00 | 0.00 | 0.00 |
| High Security | 446.36 | 0.00 | 0.00 | 0.00 | 0.00 |
| Reception | 48.24 | 0.00 | 0.00 | 0.00 | 0.00 |

the ULB and Bayes Factors concur, and the conclusions do not change after correcting for false discovery. The new tests do suggest some caution in interpreting the results of 2015 (in bold font in Table D1). The p-value indicates that we can reject the null at the 95% level and that it is just larger than the 99% threshold ($p = 0.011$). However, the ULB indicates that there is at least a 10% chance that the null is true. The Bayes Factor is not as small as the cases that give strong reason to reject the null, nor nearly as big as the instances where we clearly cannot reject it. And the adjusted p-value is 0.013, significant at the 95% level, and just over the 99% level. These results indicate that 2015 is best understood as a marginal case.[4]

[4]We thank an anonymous reviewer for suggesting that we perform these additional tests.