Responses received to additional questions

Centre for Translational Data Science
Professor Sally Cripps

Prof Sally Cripps and Dr Roman Marchant
Centre for Translational Data Science

3 September 2019

Response to Joint Selection Committee on Sydney's Night Time Economy question:
Noting the updated BOCSAR report released in August 2019, are these now an agreed set of figures?

The Centre for Translational Data Science (CTDS) does not agree with the results provided by BOCSAR in their latest report [1] (Issue no. 142, August 2019). We present four reference points as evidence for this disagreement. These, broadly speaking, are;

1. Visual examination of the actual data.
2. An analysis of BOCSAR's reported average 4% reduction in NDAs in the CBD. According to CTDS's analysis of [1], this average reduction can range from 12% to -5.8%. The statement of a 'significant' 4% reduction is misleading, as it implies the reduction is significantly different from zero, which it is not.
3. The stated drop in 'level' drop is a statistical artefact of the assumption of linearity. When a decreasing trend slows down, there will always be a negative value (regression coefficient) for the level change as a result of the underlying assumption of the model, even if the level change were positive.
4. BOCSAR's analysis made several assumptions. When these assumptions are relaxed, then we find no evidence of a decrease in NDAs in the CBD following the introduction of the lockout laws, as per our original submission [3].

Our recommendations are:

I. Quantify Uncertainty
Do not rely on point estimates of percentage changes in crime to drive policy decisions, unless they have a confidence interval attached to them to support significance.

II. Improved transparency in generation of crime counts
CTDS requested the raw data of crime locations (individual records including addresses) to verify and help improving geo-referencing of uncertain events, which was not disclosed by BOCSAR. BOCSAR only provided monthly counts of crimes without address information, therefore it is not possible for us to evaluate the quality of data aggregation, which was subject to double counting in previous analysis [2].

III. Improve transparency of methods
It is usual practice to make both the mathematical model and code available to researchers so that results can be reproduced. While we thank BOCSAR for providing a description of the settings used in a software package called SAS, this is too imprecise.

IV. A redefinition of the areas of analysis
The CBD Entertainment Precinct should not include Potts Point, Wollomoollo, Darlinghurst or Elizabeth Bay, as it is shown to do in [2].

Sincerely yours,

Sally Cripps and Roman Marchant

Detail on reasons for disagreement:

1. Visualization of Data

Figures 1(a) and 1(b) are monthly non-domestic violent assaults (NDAs) for the CBD and Kings Cross respectively. CTDS is of the opinion that complicated statistical techniques are not needed to see that NDAs in the CBD did not decrease from 2014 to 2019 while those in Kings Cross did. Crime in the CBD began to decline in 2008 and indeed the rate of decrease in NDAs in the CBD slowed down following the lockout laws. There are several possible explanations for the change in this rate, including a potential migration from visitors in Kings Cross (KC) to the CBD, but more data is needed to establish this. Understanding the causes behind this change could help reduce NDA further. We include the same plot for Kings Cross which shows a marked decrease in NDA's while the rate of decline in NDAs remains constant.

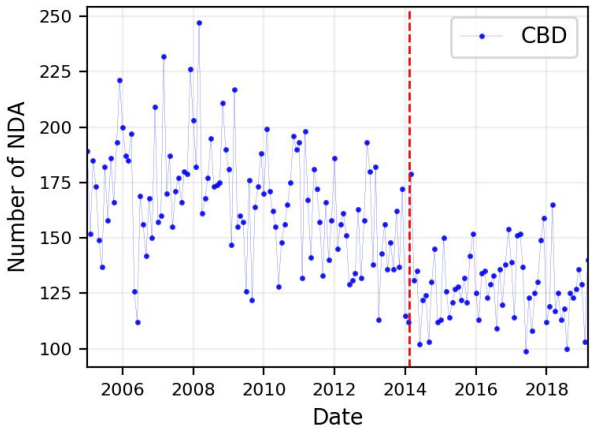

Figure 1(a) Monthly NDAs at CBD

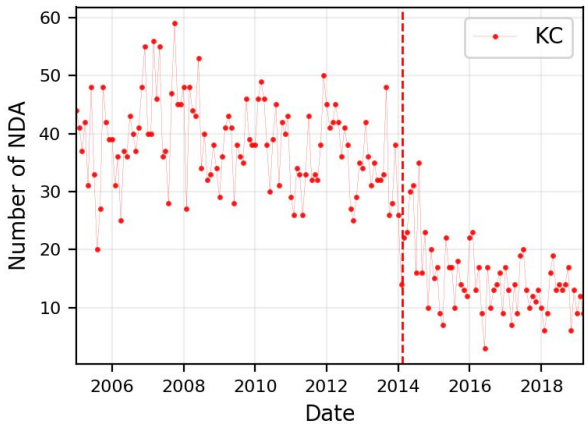

Figure 1(b) Monthly NDAs at KC

2. 4% reduction in NDAs in the CBD

The 4% figure is just one possible outcome, and is not statistically significantly different from zero. The report [1] presents the main result as percentage decrease/increase in NDAs on each area based on a forecast of 62 months into the future since 2014. The ARIMA models used by BOCSAR are not recommended for prediction because, while they fit well in-sample data, they make predictions on predictions and the amount of uncertainty in these predictions increases the further into the future that the forecasts are made.

Reporting a single value of 4% is misleading. The role of statistics is to quantify and report uncertainty. There are several sources of uncertainty in this number, one of which is the standard errors of the regression coefficients. Taking into account only this source of uncertainty (and there are several others) we put this reduction to be in the range (-5.8, 12), that is NDAs could have been reduced by as much as 12% or they could have increased by as much as 5.8%. These intervals will get much wider if the uncertainty in the actual statistical model is taken into account.

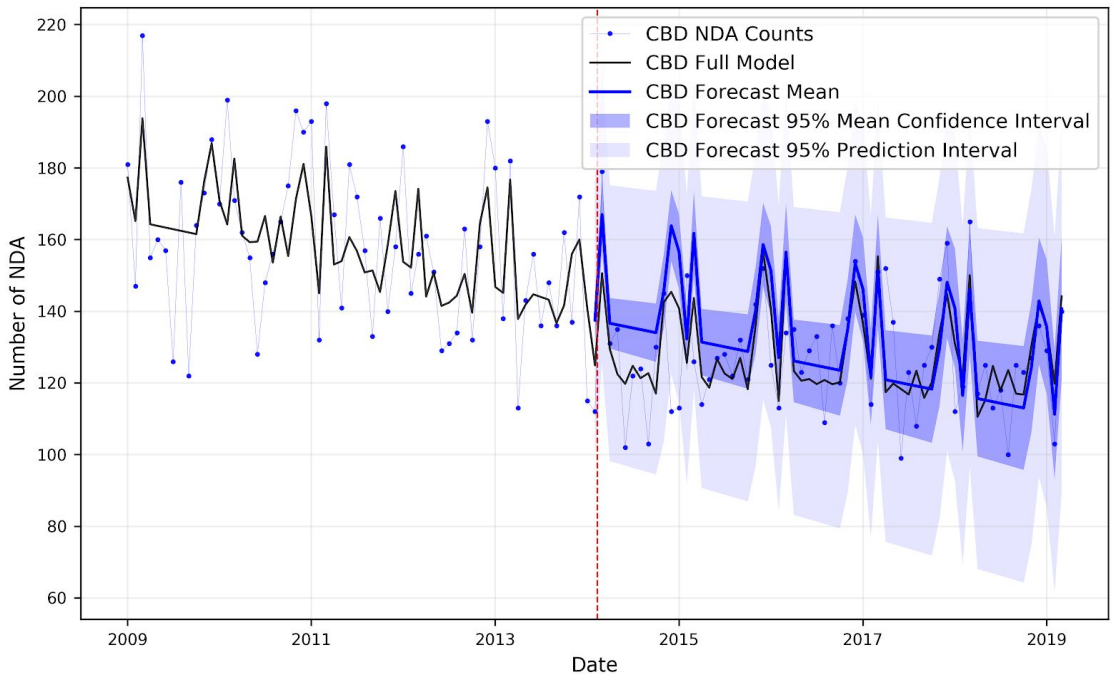

Figure 2:
CBD full model and Forecast assumed as baseline, including the Confidence Intervals (dark blue) and the Prediction Intervals (light blue) calculated by CTDS based on [1].

Figure 2 shows the forecast uncertainty, that we quantified by replicating BOCSAR results in [1]. We show that using the mean of the forecast (blue line) and not taking into account the uncertainty results in misleading results.

We requested the computational source code (in SAS) that BOCSAR used and the specific mathematical model to further quantify uncertainty. BOCSAR only provided a written description of the code, and CTDS inferred the mathematical representation and used that replicate the results in [1].

3. **The stated level drop does not acknowledge the artefact of assuming linearity.**

The model used by BOCSAR assumes that the time trend is piecewise linear, that is if we group the data into before and after the lockout laws, the model fits a straight line to the trend term. As Figure 3 shows the rate of decrease in NDAs changes at the time of the lockout laws, as it also does in 2008 when we see a change from an increasing number of NDAs to a decreasing number, a fact ignored in [1]. The report claims: *"In the CBD ... there was a significant step reduction in the level of non-domestic assaults after January 2014 by around 15 per month (p = .006)".*

This p-value refers to a change in the intercept. That is it refers to the hypothesis; *Is the average number of NDAs in 2014 different from 2009, if the linear trend from 2014-2019 were extrapolated back to 2009*. If the rate of decrease slows down then this number will always be negative. The more the rate of crime slows down the more likely this number is to be significantly different from zero. It is wrong to

infer that the lockout laws had decreased crime in the CBD. It is an artefact of the assumptions in the model.

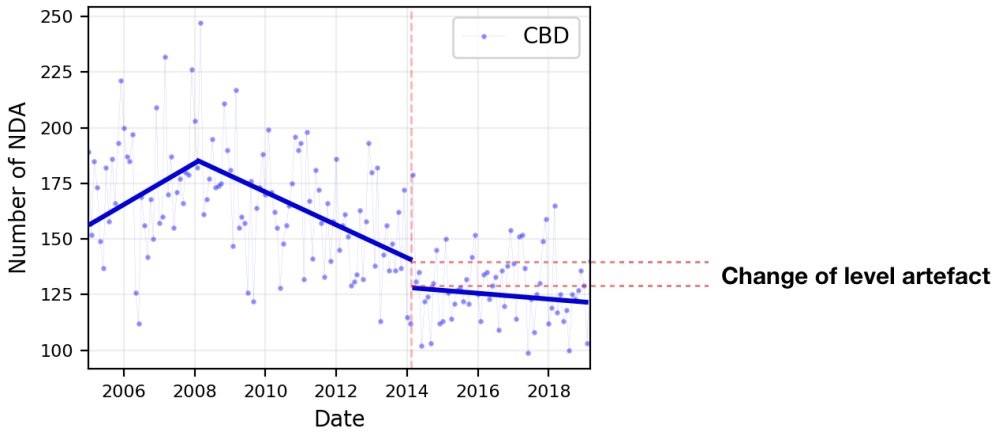

Figure 3: Artifact of level change
(only for illustration purposes for depicting the artifact of level change)

4. Relaxing Assumptions

All statistical models are simplifications of the phenomenon being studied. As George E.P. Box stated "all statistical models are wrong, but some are useful". And some are more useful than others. In this context a useful model is one which makes less assumptions, such as in our report [3]. The model developed by BOCSAR makes a number of assumptions, which are:

(a) There is at most one change point from 2009-2019.
(b) If a change point occurred then it occured in 2014. No other change points were investigated. We are yet to hear from Bocsar why all the data (back to 2005) was not used.
(c) The trend is linear.
(d) The errors follow an ARIMA (Auto-Regressive Moving Average) process.

These assumptions are relaxed in [3] where no change point was found in 2014 for the CBD.

CTDS's overall finding is that NDAs have been decreasing in NSW since 2008, indicating that a reduction in violent assaults are driven by more complex and diverse factors than the 2014 lock out laws alone. This is a good news story, however if we want to understand what is driving this downward trend in violent public behaviour we need to look beyond the lock out laws and analyse the impact of different engagement strategies and policy responses.

Bibliography

[1] N. Donnelly, S. Poynton, "The effect of lockout and last drinks laws on non-domestic assaults in Sydney: An update to March 2019," *Issue paper no 142*, August 2019
[2] N. Donnelly, S. Poynton, D. Weatherburn, "The effect of lockout and last drinks laws on non-domestic assaults in Sydney: An update to September 2016," *Crime and Justice Bulletin*, 201, February 2017
[3] N. James, R. Marchant, S. Cripps, "Technical Report: A Case Study on the Sydney Lockout Laws", Submission to the Joint Select Committee on Sydney's Night Time Economy, July 2019