On the Causal Effect of Timeouts in NBA Basketball

Alex Mansourati

July 1, 2020

Introduction

June 19, 2019 – The Raptors are 3 minutes away from winning the NBA Championship on their home court. Kawhi Leonard is on a 10 point scoring run to put the Raptors up by 6 points. At this very moment, Raptors fans around the world were shocked when their team's coach, Nick Nurse took a timeout. The fans' claim: Nurse had intervened on the scoring momentum that his team had built. After the timeout, the opponent went on a 9-2 scoring run and the Raptors lost the game. Was the timeout worth taking? With 3 minutes remaining, teams lose all timeouts except two, a recent rule to speed up the pace of play at end of game. Nick Nurse did not want to lose a timeout but would it have in fact been a better decision to lose the timeout and continue playing?

While the Raptors went on to win the 2019 NBA Championships, what might have happened in the counterfactual world where Nick Nurse did not take this timeout? We'll never know, but this raises an interesting question: what is the value of a timeout? Is taking a timeout in anticipation of losing it always worth it? We'll design a dataset to measure the causal effect of a timeout in NBA basketball and compute the avergage treatment effect of the treated.

Design

This analysis is performed over the entire 2018-2019 NBA season. Our objective is to measure the causal effect of a timeout on a team's immediate future performance. Note that we are interested in understanding the effect of a timeout on the team who chose to take the timeout. While one team requests to intervene on the game with a timeout, both teams take the break in game to gather by their team's bench.

Our measure of immediate future performance after treatment is the scoring difference in the 3 minute period immediately after timeout, for team that elected to take timeout. We are interested in the causal effect these treatments have had on immediate future performance, the average treatment effect of the treated (ATT).

We have a clearly defined set of treated units, all timeouts. The play-by-play data contains thousands of events each game, we can consider all non-timeout events as the full set of control units. This is a relatively large set and since we are calculating ATT, we can restrict control units to those that resemble the treated units. Specifically, we can reduce the full set of control units down to a sample of size n_c equal to the sample size of our treated units, n_t, such that the margnial distribution of our covariates resemble those of treated units.

In designing our control units, we also want to ensure our control units are not being sampled too close to the treated unit, introducing bias. To minimize this leak between treatment and control units, we create a buffer around the treated units of 1 minute from which a control unit cannot be sampled. See Figure 1 for a timeline of timeouts taken in five randomly selected games. It is apparent that sampling control units with this limitation will leave a plenty of room for control events to be selected.
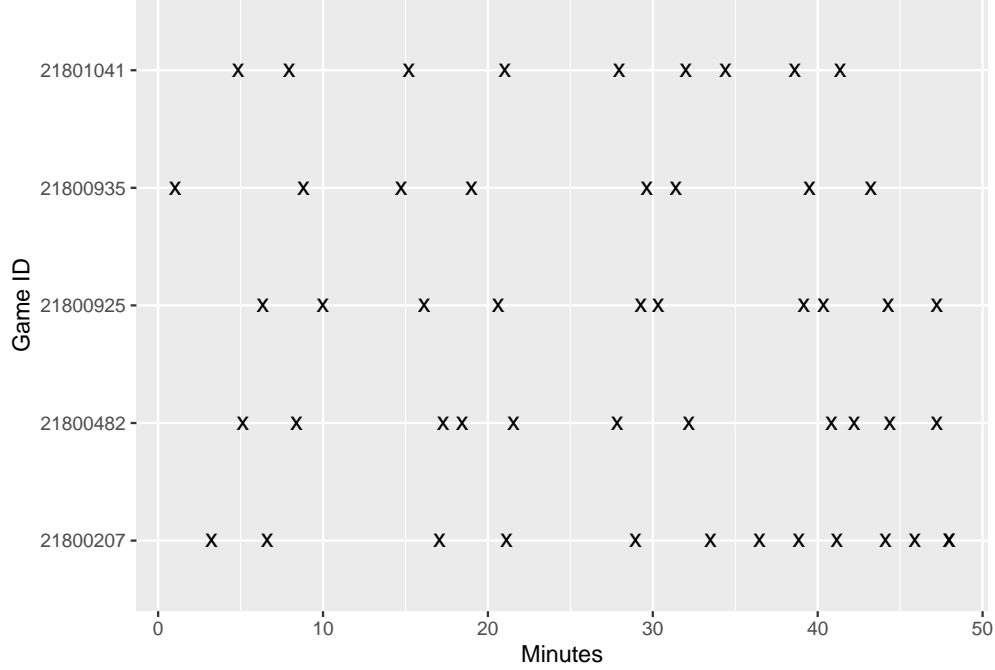
Figure 1: Gametime of timeouts in five randomly selected games.

A further limitation we introduce to both the treated and control units is to remove the first and final 3 minutes of play from our timeout analysis. The first 3 minutes is removed because one of the covariates in satisfying the backdoor criterion is the scoring run in the 3 minutes leading up to treatment. The final 3 minutes is removed because, first, our outcome, Y, is the 3 minute scoring run post-treatment and second because the timeouts in the final three minutes tend to be taken for strategically different reasons.

Applying the buffer around timeouts and at edges of game, we have $n_t = 14343$ and $n_c = 569904$ across the full season. Before we continue to generating a matched control set, let's establish our set of covariates, denoted S, fulfilling the backdoor criterion:

1. Scoring Run - the three minute scoring run leading up to the timeout from frame of reference of the team who elected to take timeout.
2. Coach experience - the number of seasons that the head coach of the team who took the timeout has been a head coach in the NBA.
3. Time since last break - the time elapsed since the last break in play (the minimum of time since the start of the last period and the time since last timeout).
4. Overall point differential - the difference in score at the moment the timeout is taken from the frame of reference of the team who elected to take the timeout.

Our belief is that these covariates achieves conditional exchangeability,

$$X \perp Y | S$$

where X is our treatment, Y is our outcome and S is the set of covariates listed above.

Figure 2 illustrates how are our covariates are balanced between treated and control with sampling buffer. It is clear that our covariates will benefit from balancing to improve ability of achieving a correctly specified outcome model. Our objective is to achieve a balanced, matched, dataset such that $n_t = n_c$.
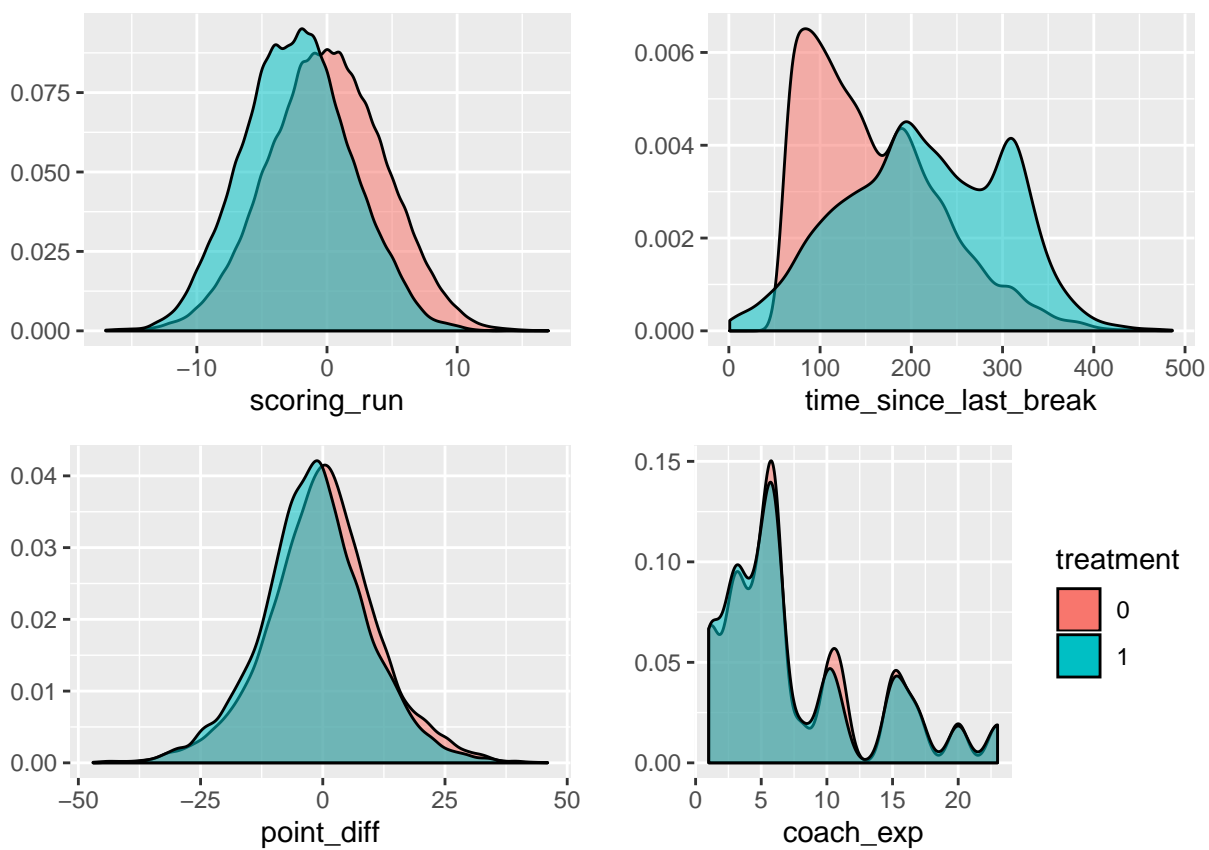
Figure 2: The distribution of covariates in the raw samples.

In sampling down our set of control units to size n_c, we will select control units that match pairwise to treated units, minimizing a scalar distance between covariate values. We use the propensity to be treated, a function of the covariates as our scalar distance and the nearest neighbours algorithm is used via MatchIt (R library) to identify pairs. Note that this algorithm will start with the treated units with highest propensity scores since they will be hardest to match. Our propensity score is modelled as a GAM.

$$g(E(Y)) = log\frac{P(Y = 1)}{P(Y = 0)} = \alpha + s_1(x_1) + ... + s_4(x_4)$$

Here, $x_1, ..., x_4$ denote our four covariates and $s(x_1), ..., s(x_4)$ denote smooth, nonparametric functions.
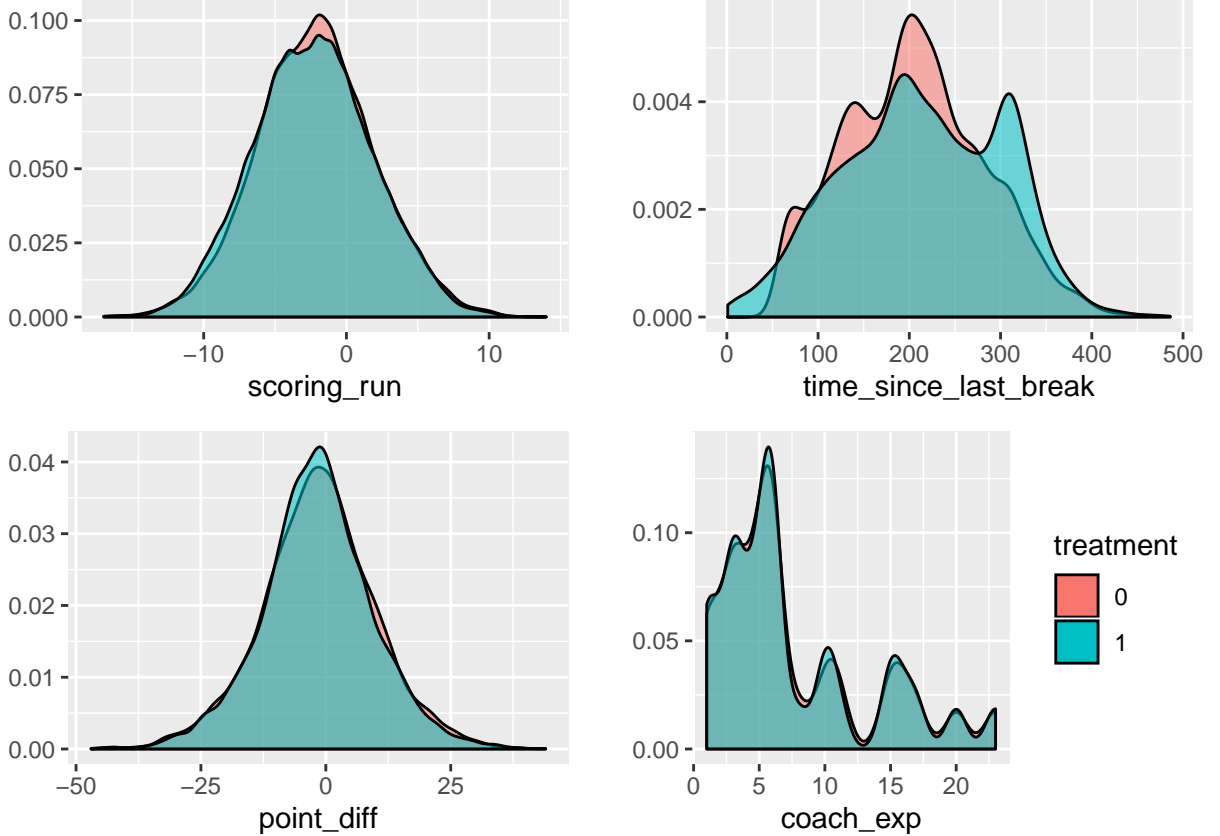

Figure 3: The distribution of covariates in the matched samples.

After the matching procedure, $N_t = N_c = 8690$, we achieve new marginal covariate distributions for the control units. See Figure 3 for matched covariate comparison. This dataset design makes our outcome modelling more robust to model misspecification.

Outcome Modelling

With $E[Y|X = x, S = s]$ abbreviated as $\mu(x_i, s_i)$,

$$ATE = \frac{1}{n}\sum_{i=1}^{n} \hat{\mu}(1, s_i) - \hat{\mu}(0, s_i)$$

With out treatement-matched dataset, this is in fact the ATT. We model μ linearly with a gaussian link function. See Figure 4 for fitted residuals of $\hat{\mu}$. There is no clear systematic errors, evidence against misspecification.

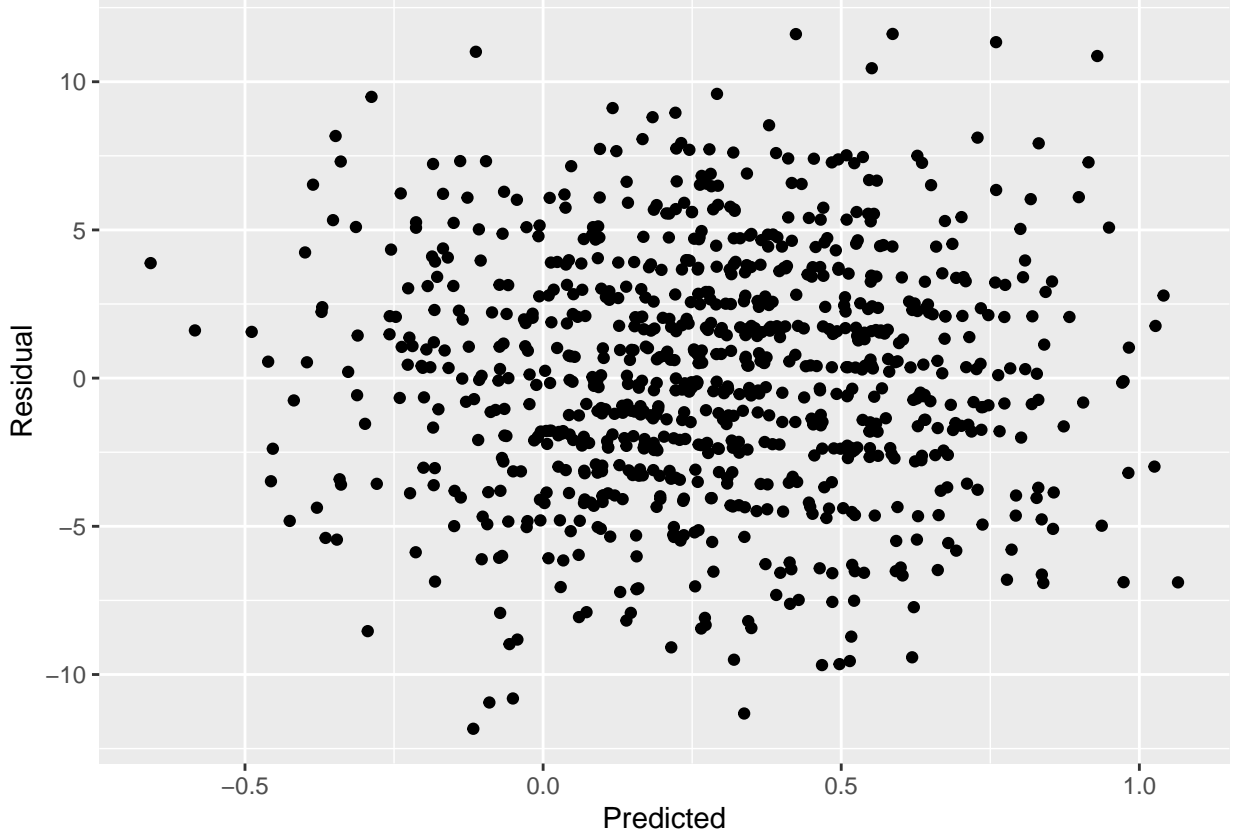$$Y_i = \mu(x_i, s_i) = \alpha + \beta_1 x_1 + ... + \beta_4 x_4$$

Figure 4: Residual plot of unconditional ATT outcome model.

It is quite common to see a coach take a timeout when the opponent is on a significant scoring run with the intent of stopping the opponents momentum and giving his team a chance to mentally reset. For this reason, we believe that it is worth evaluating the average treatment effect of timeouts conditionally, on negative vs. positive scoring runs leading up to the timeout.

Let $E[Y, X = x, S = s, R > 0]$, where R is the scoring run in 3 minutes leading up to timeout, be abbreviated as $\mu_{pos}(x, s)$ and $E[Y, X = x, S = s, R < 0]$ be abbreviated as $\mu_{neg}(x, s)$. Then,

$$ATT_{pos} = \frac{1}{n} \sum_{i=1}^{n} \hat{\mu}_{pos}(1, s_i) - \hat{\mu}_{pos}(0, s_i)$$

and

$$ATT_{neg} = \frac{1}{n} \sum_{i=1}^{n} \hat{\mu}_{neg}(1, s_i) - \hat{\mu}_{neg}(0, s_i)$$

are the conditionally average treatment effects of the treated. To calculate the variance of these three estimates, we use bootstrap estimates.

Method	ATT	CATT (+ve Scoring Run)	CATT (-ve Scoring Run)
Naive	0.46 (0.4, 0.52)	0.32 (0.22, 0.42)	0.4 (0.31, 0.49)
Matching + Standardization	0.34 (0.22, 0.46)	0.25 (0, 0.5)	0.36 (0.21, 0.51)

We find evidence that, on average, timeouts are effective at improving a team's performance relative to their opponent, regardless of the scoring run they are on leading up to the timeout. See Figure 4. A naive method of evaluating ATT, taking the difference between means conditional on treatment, on the unmatched dataset is included for reference.

Conclusion

After game 5 of the 2019 NBA Finals, Nick Nurse was asked about his thought process taking the timeout in the midst of his team's scoring run. He replies,

"...well we had two free ones that you lose at the 3 minute mark...and decided to give the guys a rest...we thought we could use the extra energy push." [2]

Nurse begins with stating that he did not want to lose free timeouts. With the information proposed in this analysis, we have no reason to refute Nick Nurse's experienced read of the game. The Raptors went on to win game 6 at the Warriors' arena to clinch the championship.

Future work might look specifically at timeouts near end of game, leading up to the 3 minute mark when coaches lose all timeouts except reamining two. The theory is that coaches are overusing timeouts to the detriment of their team.

#References

[1] Hernan MA, Robins JM. Causal Inference: What If. Boca Raton: Chapman and Hall/CRC, 2019.

[2] House of Highlights. Nick Nurse Postgame Interview - Game 5 | Warriors vs Raptors | 2019 NBA Finals. https://youtu.be/HwfA9IMqzK0?t=303, 2019.

[3] Imbens GW, Rubin DB. Causal Inference for Statistics, Social and Biomedical Sciences. Cambridge University Press, 2015.