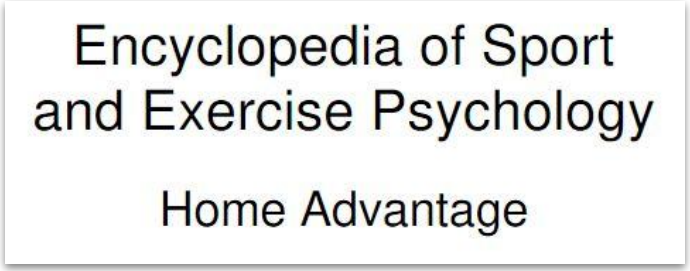
COVID-19 Impact on NHL Home-Ice Advantage

By Ian Keller

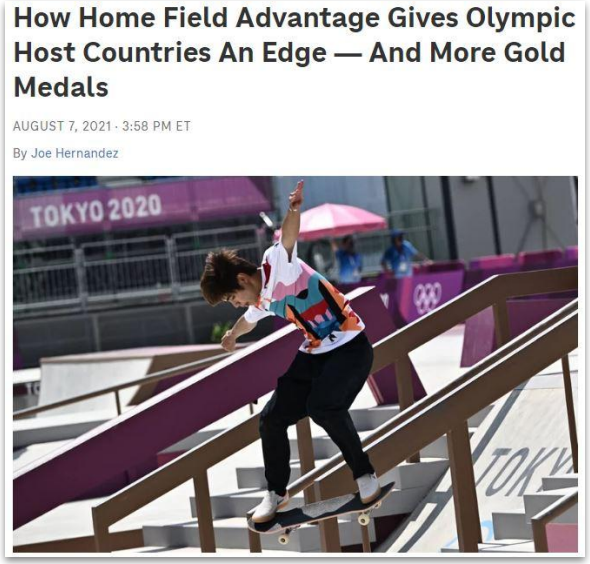
MOTIVATION

- COVID-19 allows for a natural experiment to better understand variation
- Widespread impact of COVID-19 applied to sports
- Historical theme of Home Team Advantage



RELEVANT LITERATURE

- “The Home Advantage” ~ Carron et. Al
  - Findings include significant home advantage in team/individual sports explained by:
    - Crowd Support
    - Referee Bias
    - Comfort / Familiarity
    - Travel
- Guérette et. al research on referee bias being eliminated during COVID-19 in the NHL
- “The Behaviour of Home Advantage during the COVID-19 Pandemic in European Rink Hockey Leagues” ~ Arboix-Alió et. al
- Extends paper “Do Fans Impact Sports Outcomes? A COVID-19 Natural Experiment” ~ Cross et. al
  - Adds to literature by exploring NHL Context

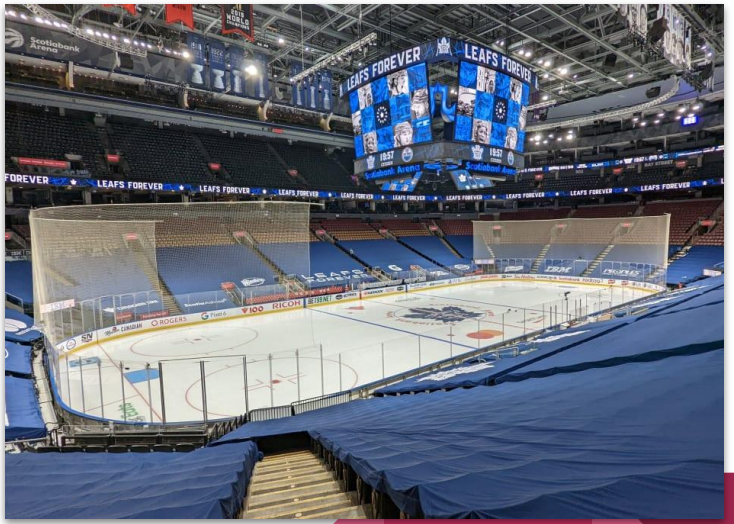


NHL COVID-19 Environment

- March 12, 2020 season shutdown
  - ~70 games played
- August 1, 2020 season resumes with no fans
- 2021 Season largely played without fans
  - 56 regular season games
- Allows for Difference-in-Difference analysis of COVID-19's Impact

NHL to pause season due to coronavirus

'Goal is to resume play as soon as it is appropriate and prudent,' Commissioner Bettman says



COVID-19 Impact on NHL Home Ice Advantage

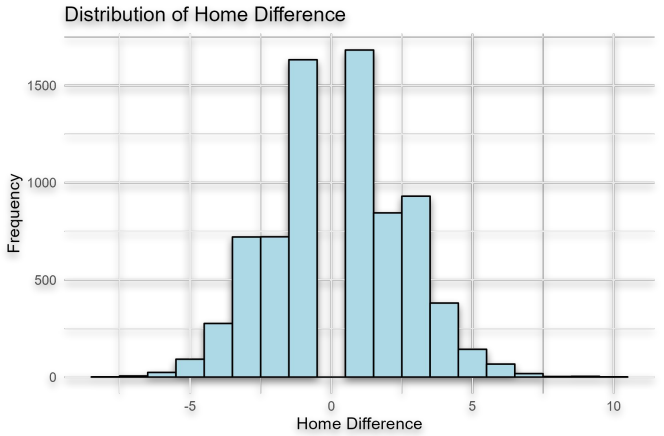
- Home-Ice advantage exists
- COVID-19 allows for causal analysis of home team performance
  - Isolating crowd support variable
- Controls used for extra precision:
  - Close Matchups (within 300 miles of travel for fans)
  - Team vs Team matchup Fixed Effects
  - Season
  - Playoffs

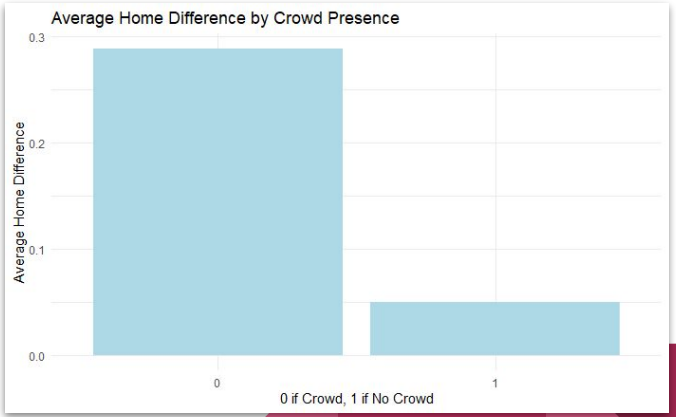
Data Collection

- Scraped individual game data from 2016-2021 from www.hockey-reference.com
  - Includes 7,551 games played (n)
    - Home/Away team and game score
    - Attendance
- Downloaded NHL team distance data from Reddit user u/mbstone
  - Created 32x32 matrix of distance between team's city and closest local airport using gMaps function in Sheets

Summary Statistics

- Feature Engineered Variables:
  - No crowd (n = 707)
  - Playoff game (n = 581)
  - Close (distance) game (n = 445)
  - Home win (n = 4,076)
  - Home Differential
  - Matchup
    - Used for Fixed Effects
      - ex) “Boston Bruins @ Pittsburgh Penguins”
- 7,551 Total Games





Date	Visitor	G	Home	G.1	Att.	playoffs	home_difference	no_crowd
10/7/2015	Vancouver Canucks	5	Calgary Flames	1	19289	0	-4	0
10/7/2015	New York Rangers	3	Chicago Blackhawks	2	22104	0	-1	0
10/7/2015	San Jose Sharks	5	Los Angeles Kings	1	18230	0	-4	0
10/7/2015	Montreal Canadiens	3	Toronto Maple Leafs	1	19241	0	-2	0

Empirical Strategy

- Follows framework laid out in “Do Fans Impact Sports Outcomes? A COVID-19 Natural Experiment”
  - Regressing Y (goal differential) on various factors, primarily no crowds.
  - All variables are binary except goal differential (continuous) and matchup (string)

$$Y = \mathcal{B}_0 + \boldsymbol{\beta}_1(\text{no\_crowd}) + \boldsymbol{\varepsilon}$$

$$Y = \mathcal{B}_0 + \boldsymbol{\beta}_1(\text{no\_crowd}) + \boldsymbol{\beta}_2(\text{close\_team}) + \boldsymbol{\beta}_3(\text{no\_crowd*close\_team}) + \boldsymbol{\varepsilon}$$

$$Y = \mathcal{B}_0 + \boldsymbol{\beta}_1(\text{no\_crowd}) + \boldsymbol{\beta}_2(\text{close\_team}) + \boldsymbol{\beta}_3(\text{playoffs}) + \boldsymbol{\beta}_3(\text{no\_crowd*playoffs}) + \boldsymbol{\beta}_3(\text{no\_crowd*close\_matchup}) + \boldsymbol{\varepsilon}$$

$$Y = \mathcal{B}_0 + \boldsymbol{\beta}_1(\text{no\_crowd}) + \boldsymbol{\varepsilon} \mid \boldsymbol{\lambda}(\text{matchup FE}) + \boldsymbol{\varepsilon}$$

$$Y = \mathcal{B}_0 + \boldsymbol{\beta}_1(\text{no\_crowd}) + \boldsymbol{\beta}_2(\text{playoffs}) + \boldsymbol{\beta}_3(\text{no\_crowd*playoffs}) \mid \boldsymbol{\lambda}(\text{matchup FE}) + \boldsymbol{\varepsilon}$$

$$Y = \mathcal{B}_0 + \boldsymbol{\beta}_1(\text{no\_crowd}) + \boldsymbol{\beta}_2(\text{playoffs}) + \boldsymbol{\beta}_3(\text{no\_crowd*playoffs}) \mid \boldsymbol{\lambda}(\text{matchup FE}) + \boldsymbol{\varepsilon}$$ (2020-2021 Seasons)

RESULTS

$Y = \mathcal{B}_0 + \beta_1(\text{no\_crowd}) + \beta_2(\text{playoffs}) + \beta_3(\text{no\_crowd*playoffs}) \mid \lambda(\text{matchup FE}) + \varepsilon$ (2020-2021 Seasons)

```
Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
no_crowd            0.3476     0.1769   1.964   0.0497 *
playoffs            0.2835     0.3711   0.764   0.4451
no_crowd:playoffs  -1.1466     0.4918  -2.332   0.0199 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Most robust model (1304 df)
- Shows significant effects
  - no_crowd
  - no_crowd*playoffs
- Uses relevant data from seasons affected by COVID-19
- Interprets to a negative scoring differential for the home team when there is no crowd in the playoffs

RESULTS

$Y=\mathcal{B}_0+\beta_1(\text{no\_crowd})+\beta_2(\text{close\_team})+\beta_3(\text{playoffs})+\beta_3(\text{no\_crowd*playoffs}) + \beta_3(\text{no\_crowd*close\_matchup}) + \varepsilon$

```
Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)          0.29113    0.03150   9.243   <2e-16 ***
no_crowd            -0.13234    0.11289  -1.172   0.2411
close_matchup       -0.08206    0.12930  -0.635   0.5257
playoffs             0.02012    0.12151   0.166   0.8685
no_crowd:playoffs   -0.56435    0.25860  -2.182   0.0291 *
no_crowd:close_matchup  0.11629    0.34704   0.335   0.7376
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.452 on 7545 degrees of freedom
Multiple R-squared:  0.00162,   Adjusted R-squared:  0.0009579
F-statistic: 2.448 on 5 and 7545 DF,  p-value: 0.03175
```

- Full Model with no Fixed Effects
- Shows highly significant intercept
  - Represents Home Goal differential with fans and not a close matchup or playoffs
- Significant term of no_crowd*playoffs
  - Shows no crowd effect only important during playoff games
  - Once again creates a negative home goal differential
- Close Matchup not significant
- R-squared shows small percent of explained variation

RESULTS

$Y = \mathcal{B}_0 + \boldsymbol{\beta}_1(\text{no\_crowd}) + \boldsymbol{\varepsilon}$

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.2161     0.1286   1.680  0.09337 .
no_crowd      -0.5694     0.1974  -2.885  0.00405 **
```

$Y = \mathcal{B}_0 + \boldsymbol{\beta}_1(\text{no\_crowd}) + \boldsymbol{\varepsilon} \mid \boldsymbol{\lambda}(\text{matchup FE}) + \boldsymbol{\varepsilon}$

```
Coefficients:
           Estimate Std. Error t value Pr(>|t|)
no_crowd -0.06112    0.11073  -0.552    0.581
```

- Initial model shows strong significance of no crowds on home differential
- Fixed effects model shows this impact is mitigated when controlling for team vs team matchup

LIMITATIONS

- Could control for more factors
  - Team distance travelled before game
  - Days rest between games
  - Player injuries can affect game outcomes
- COVID-19 seasons irregularities
  - Hockey in August
  - Short offseason to 2021 season
  - Health Protocols
- Unlikely to have similar conditions again
- Clustering Standard Errors

CONCLUSIONS

- Having no crowd had a significant negative impact on home performance
- Having no crowd during the playoffs had the strongest effect
    - Flipped the narrative and gave the away team the advantage
- Playing a close team had no impact on home goal differential
- Quantifies importance of a strong home crowd for team performance