Away-Team Advantage:

Evaluating the Impact of Crowd Absence on NHL Home Ice Advantage during the Covid-19 Pandemic

Abstract:

This study quantitatively examines the impact of crowd absence on home team performance in the National Hockey League (NHL), focusing on the COVID-19-affected 2020-2021 seasons. Using the natural experiment of Covid-19, we ran numerous linear regression models to assess the impact of having no fans in attendance, while also controlling for playoff games, close matchups, and using fixed effects for team matchups. Our analysis reveals a significant negative effect on home team performance during playoff games played without a crowd, indicating a 0.5155 goal advantage for the away team. The impact of regular season game results without fans is less conclusive but saw some negative relationships to the home team's performance. These effects were mitigated when controlling for matchups suggesting historical team advantages, independent of crowd presence. Overall, our findings underscore the practical importance of fan attendance in bolstering home team performance, particularly during playoff environments, contributing valuable insights into the Covid-19 impacted NHL seasons.

Ian Keller

University of Pittsburgh

Economics Capstone (ECON 1700)

Dr. Claire Duquennois

April 26, 2024

Introduction and Literature Review

The impact of Covid-19 has been profound and reverberated widely, leaving a lasting impact that is still being analyzed today. Spanning from healthcare systems to global economies, its repercussions have affected every facet of society, from supply chains to mental health, leaving no aspect untouched. This includes the world of sports and sports economics. My paper answers the question of did Covid-19's fan-less had an impact on home-ice advantage in the National Hockey League (NHL). "Home Ice Advantage" refers to the phenomenon where the home team is typically favored to win a game. My paper quantifies the influence of crowd support by analyzing the 2020-2021 NHL seasons, during which arenas were devoid of fans. This research provides insights into the importance of having a strong, home-biased crowd to enhance home team performance in not just hockey but all professional sports.

To answer this question, we used a difference-in-difference analysis in which we used the Covid-19 pandemic as a reference point for this type of "before-and-after" study. Before Covid-19, NHL arenas typically held 15,000-20,000 fans each game, with a large majority coming to cheer on the home team. Covid-19 led to government policies throughout the United States and Canada that banned or limited large gatherings, directly affecting sporting events as fans could no longer attend games, despite these games still being played. This natural experiment allows us to easily analyze the home team's performance with fans versus without fans. Our research finds a statistically significant negative impact of having no crowd on team performance, with an especially strong negative impact when having no crowd during a playoff game. Our findings show that having no fans in attendance can even lead to an "Away-Team Advantage".

Our paper builds off of previous research done by Cross et al. in their paper "Do Fans Impact Sports Outcomes? A COVID-19 Natural Experiment", which found that the no-fans

policy had an impact on top European soccer leagues home field advantage, reducing the home goal scoring differential by more than 50%. We plan on further contributing to the literature by analyzing the impact of no crowds within the context of the National Hockey League.

To conduct our research, we draw upon prior studies such as "The Home Advantage" by Carron et al. “The Home Advantage” conducted a literature review analyzing the impact of home advantage over the past 5 years in professional Baseball, Soccer, Ice Hockey, American Football, and Basketball. These sports resulted in significant performance improvements when at home, resulting in anywhere from a 61% to 54% home-team advantage. Additionally, they found a strong increase in individual sports by assessing World Cup Alpine Skiing in which athletes competing in their home country performed 16% better from their initial seeding going into the race. Carron explains this advantage in a few ways such as crowd support, officiating bias, travel, and comfort/familiarity. This is an important framework on which we will be basing our analysis and will be discussed in detail shortly.

Another relevant piece of literature is called "Eliminating Home Advantage: The Impact of Fan Absence on Penalties Called by National Hockey League Referees" by Guérette et al. which shows precisely what the title says in that the removal of fans during Covid-19 removes the home penalty call bias that Carron and others (Gong, Boyko) have found in professional sports. This research showed that significantly more penalties were given to the away team compared to the home team when there was a crowd present. This difference was eliminated during the Covid-19 seasons as there were no fans in the crowd, showing that fans can influence a referee’s decision-making by cheering for a penalty call to be made in their favor.

By tying these two papers together we have set a clear structure for causal analysis of the impact of no fans on teams' performance. Starting with Carron et al.'s home advantage

framework being explained by crowd support, referee bias, comfort/familiarity, and travel. We see that in the National Hockey League when no fans were in attendance officiating bias is eliminated but also travel and comfort/familiarity stay consistent allowing us to isolate the crowd support variable to analyze its impact on home team performance.

Similar research on hockey has been conducted before by Arboix-Alió et. al in the paper "The Behaviour of Home Advantage during the COVID-19 Pandemic in European Rink Hockey Leagues". Their results showed that home ice advantage significantly decreased during Covid-19 from 63.99% to 57.41% but, as we can see, there was still an advantage for the home team. This research has also been applied to other sports such as the National Basketball Association (NBA). Price et al.'s paper "The Effects of the NBA COVID Bubble on the NBA Playoffs: A Case Study for Home-Court Advantage" explored this same problem and found that away teams faired much better during the 2020 season playoffs without crowds compared to the 2017-2019 seasons. I plan on adding to this literature by analyzing home team performance during Covid-19 in the context of the National Hockey League (NHL).

Additionally, we plan on adding to the current research by exploring if playing a close team (defined as a team within 300 miles), mitigates the home-ice advantage due to traveling fans, along with exploring the impact during the regular season vs more important playoff games. These results can help quantify the importance of having a strong, home-biased crowd in attendance every night to help the home team win and further explain the causal effects COVID-19 had on sports.

Looking at the Covid-19 NHL environment we see that it differs greatly from any other prior season in numerous ways. The 2019-2020 NHL season was shut down on March 12, 2020, due to Covid-19 concerns in the NHL and around the world. At this point in the year, roughly 70

games were played for each team out of a typical 82-game NHL regular season and on average the attendance in these games was 17,500 fans. Nearly 5 months later, on August 1, 2020, the NHL season resumed play with a modified playoff format in two distinct NHL "bubble" locations Toronto and Edmonton in which no fans were in attendance. Given the unusual circumstances surrounding the regular season's premature end, the playoffs commenced with a unique "play-in" round, featuring best-of-five game series, a format unprecedented in NHL history. Following this round, the traditional NHL playoff format ensued, characterized by a best-of-seven series, wherein teams faced each other repeatedly until one team secured victory in four games.

The following 2020-2021 season also had a different environment. Teams were finally allowed to return to their home stadiums but for much of the shortened 56-game regular season many teams were not allowed any fans in attendance. Additionally, these regular season games were only played versus in-conference opponents which means that 1) they were against rival teams, and 2) this limited travel time for all teams, as conference games tend to be much closer. Here, 558 out of 871 games were played without crowds during the regular season. This changed during the playoffs where only 18 out of 87 games were played without crowds, with restrictions mainly being implemented on the Canadian teams. Understanding this background is imperative to our research design and guides many of our explanations within our results section.

Our framework for the remainder of the paper is as follows: Empirical Approach, Data and Descriptive Statistics, Results and Discussion, Summary and Conclusion, followed by our sources.

Data Collection and Descriptive Statistics

To collect the data necessary for our research we scraped individual game data from 2016-2021 from www.hockey-reference.com which included 7,551 total regular season and playoff games. Each game included valuable information such as home and away team names, along with their corresponding scores, and finally the attendance of each game. Figure 1 shows the distribution of attendance based on the NHL season. Here we see the impact of Covid-19 on

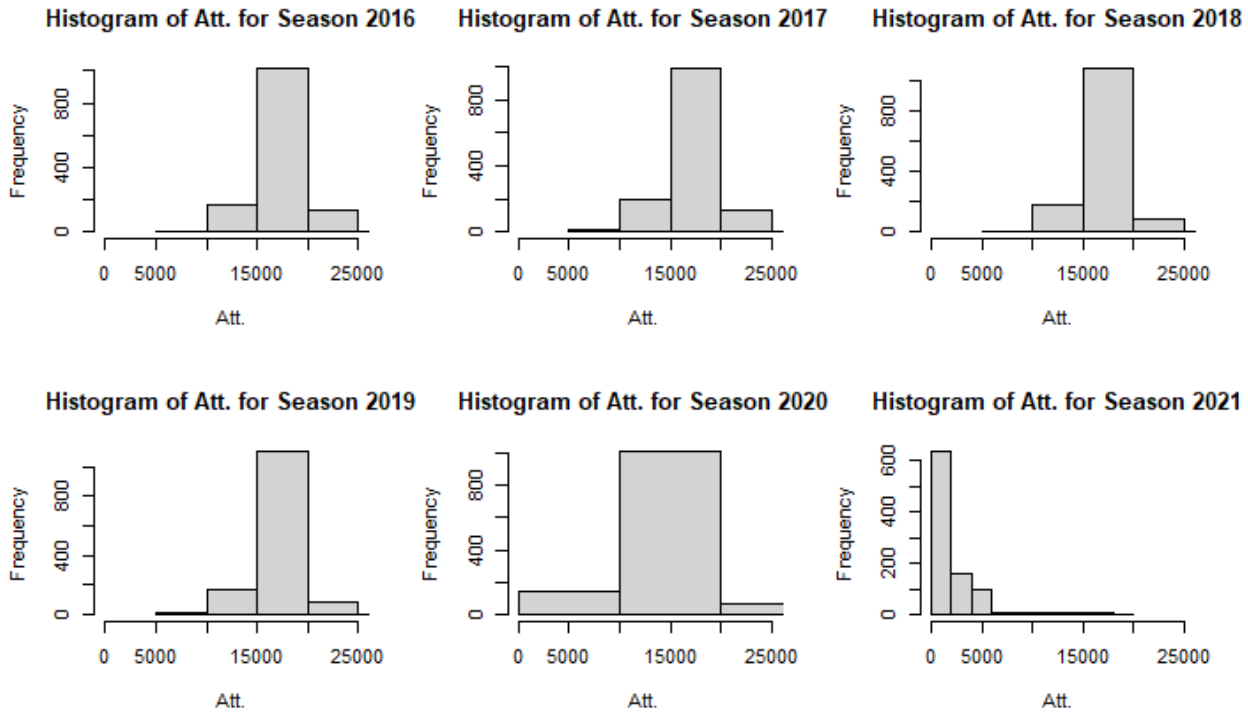

Figure 1: NHL Attendance from 2016-2021

attendance during the 2020 and 2021 seasons as the data begins to skew towards 0. From this data, we collected 707 games (9.3%) that had an attendance of 0. Additionally, binary variables were created within this dataset to establish whether a game occurred during the playoffs, whether a crowd was present, and if the game occurred against a team of close distance, and a continuous variable of home goal differential, which was our dependent variable.

Additionally, to find whether it was a close matchup or not we used data collected from Reddit user u/mbstone who had (conveniently) created a 32x32 matrix in Google Sheets which mapped the distance from each team's closest airport to every team in the league's airport using the gMaps function. We then took this data and filtered it for teams that were within 300 miles of each other to create a list of close matchups which was used to create a binary variable where 1 meant the two teams were within 300 miles of each other and 0 meant otherwise.

During the exploratory data analysis, it became clear that there was a trend of home team advantage as you can see in Figure 2 with the data being slightly left-skewed representing an advantage to the home team as a majority of our values our positive (recall: home difference = home score – away score). Furthermore, we can begin to visualize the difference of home

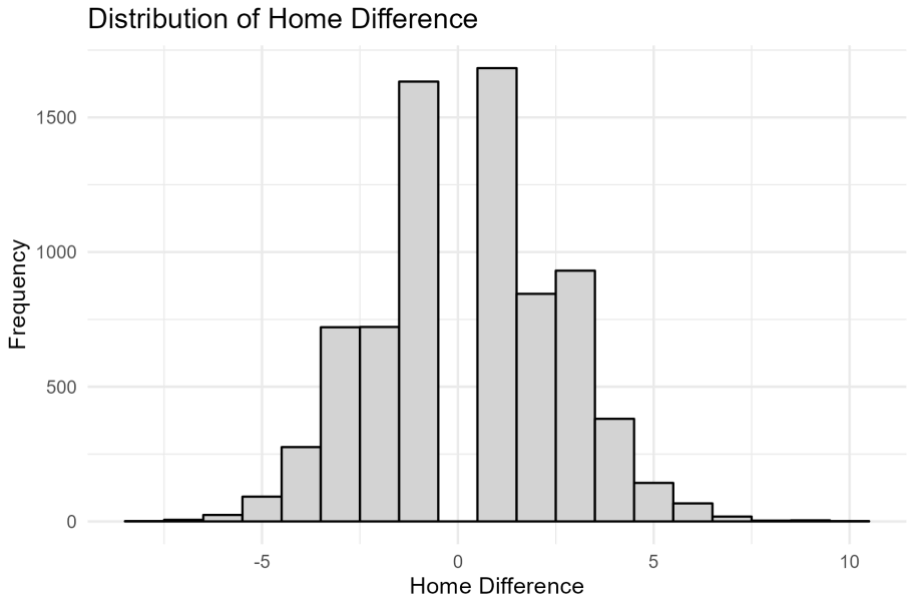

Figure 2: Histogram of Home Goal Scoring Differential

performance when there is a crowd versus when there is not a crowd by looking at a simple bar chart in Figure 3 comparing the mean home scoring differences.

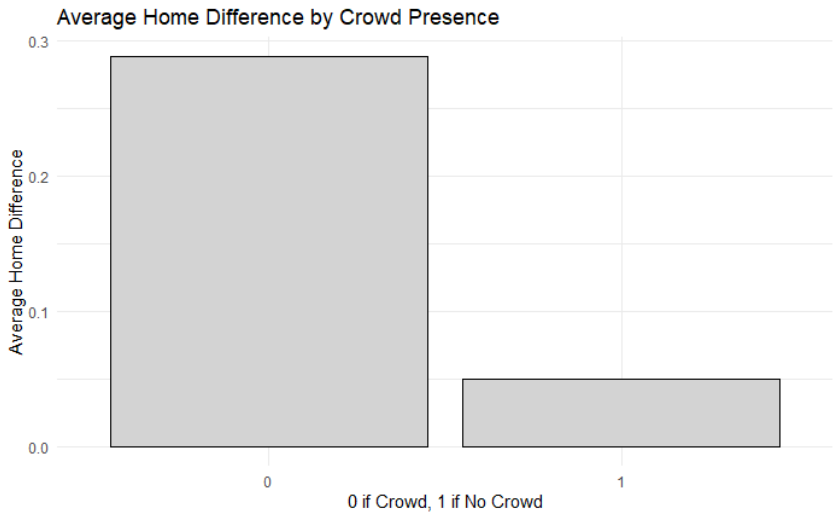

Figure 3: Home Goal Differential with Crowd (0) versus without Crowd (1)

Here we see a clear difference in the means of home difference dependent on whether there is a crowd or not. This is something that we will further explore with our regression models.

Overall, Table 1 summarizes our data structure used for analysis.

Table 1: Variable Summaries

	Min	Mean	Max	Unique Obervations
Numeric Variables				
Home Goals	0	**3.019**	10	7551
Away Goals	0	**2.754**	9	7551
Playoffs (Binary)	0	0.07694	1	581
Season	2016	2018	2021	6
Close Matchup (Binary)	0	0.05893	1	445
Home Difference	-8	**0.2655**	10	7551
No Crowd (Binary)	0	0.0936	1	707
Character Variables				
	Format Example			
Date	10/7/2016			
Visitor	"Boston Bruins"			
Home	"Pittsburgh Penguins"			
Matchup	Boston Bruins @ Pittsburgh Penguins			

Table 1: Summary Statistics of Variables

The important differences in mean home and away goals are highlighted here and we begin to have a better idea of how many data points we are working with for each of our variables.

Empirical Approach

Our empirical approach follows the research design laid out in "Do Fans Impact Sports Outcomes? A COVID-19 Natural Experiment" by Cross et al. In this study, we conducted a multiple linear regression analysis to examine the influence of several independent variables on the dependent variable of home goal differential. The independent variables considered were no_crowd, close_team, and playoffs, all of which were binary variables in which 1 represented no fans in attendance, a team that was within 300 miles, and if it were a playoff game, respectively. The dependent variable, coded as home_difference, was calculated as the difference between the total number of goals scored by the home team and the total number of goals scored by the away team. Additionally, on some of our models, we implemented matchup fixed effects that controlled for team vs team matchups. For example, "Boston Bruins @ Pittsburgh Penguins" was a control group that looked at games when Boston played Pittsburgh in Pittsburgh, but not when Pittsburgh played Boston in Boston. This allowed for extra precision in our analysis by controlling for the impact that differing home crowds may have. Additionally, we looked at isolating the 2020-2021 seasons in our regressions to explore the effects directly on these seasons as they were much different relative to prior seasons. By exploring all these variables, we aimed to assess their impact on the outcome of NHL hockey games.

The following regression models were run to assess variable importance:

1) Home_difference = $\mathcal{B}_0 + \beta_1$ (no_crowd) + ε
2) Home_difference = $\mathcal{B}_0 + \beta_1$ (no_crowd) + β_2 (close_team) + β_3 (no_crowd*close_team) + ε
3) Home_difference = $\mathcal{B}_0 + \beta_1$ (no_crowd) + β_2 (close_team) + β_3 (playoffs) + β_3 (no_crowd*playoffs) + β_3 (no_crowd*close_matchup) + ε
4) Home_difference = $\mathcal{B}_0 + \beta_1$ (no_crowd) + ε | λ (matchup FE) + ε
5) Home_difference = $\mathcal{B}_0 + \beta_1$ (no_crowd) + β_2 (playoffs) + β_3 (no_crowd*playoffs) | λ (matchup FE) + ε
6) Home_difference = $\mathcal{B}_0 + \beta_1$ (no_crowd) + β_2 (playoffs) + β_3 (no_crowd*playoffs) | λ (matchup FE) + ε (2020-2021 Seasons)

Table 2: Regression Models used for analysis

Our empirical approach utilizing numerous regressions across different contexts provides us with a comprehensive understanding, enhancing the robustness of our approach. This holistic view ensures we consider all facets of the issue, leading to more informed decisions and effective solutions. In the results section below we will discuss the results of these various regression models in order of importance.

Results

The results from our most robust model are shown below in Table 3. This model was run using regression 6, the full model with matchup fixed effects, as discussed in the 'Empirical Approach' section.

Table 3: Full Linear Regression Model with Matchup Fixed Effects (2020 -2021)

	Estimate	Std. Error	t-value	p-value	Observations
no_crowd	0.3476	0.1769	1.964	**0.0487 ***	707
playoffs	0.2835	0.3711	0.764	0.4451	581
no_crowd:playoffs	-1.1466	0.4918	-2.332	**0.0199 ***	146

Table 3: Full Regression Model with Fixed Effects on 2020-2021 data (6)

This model shows two significant variables: the first is no crowd and the second, and of higher significance, is no crowd with the interaction term of the game being a playoff game. Interpreting these results, we can see that on average we can expect the home team to outperform the away team by 0.3476 goals on average which is significant and counters what we may expect given prior research shown in our literature review. Interestingly, we find that when there is no crowd during a playoff game home-ice advantage disappears and there becomes an away-team advantage. This is shown by our most significant variable which is the interaction between no crowd and playoffs. Under this scenario, if there is no crowd during a playoff game, we can expect the ***away*** team to outperform the home team by 0.5155 goals on average. This model's use

of matchup fixed effects allows us to better isolate the true effect of our variables of interest. This is essential as it helps us address concerns of omitted variable bias and endogeneity.

The next model considered was a full linear regression model (regression 3) using all of our features without the use of fixed effects. This allows us to test the robustness of our results while exploring other variables such as close matchup individually, as it is controlled for when using fixed effects. The results of this regression are shown below in Table 4.

Table 4: Full Linear Regression Model Without Fixed Effects

	Estimate	Std. Error	t-value	p-value	Observations
(Intercept)	0.2911	0.0315	9.243	**<2e-16*** **	7551
no_crowd	-0.1323	0.1129	-1.172	0.2411	707
close_matchup	-0.0821	0.1293	-0.635	0.5257	445
playoffs	0.0201	0.1215	0.166	0.8685	581
no_crowd:playoffs	-0.5644	0.2586	-2.182	**0.0291***	146
no_crowd:close_matchup	0.1163	0.347	0.335	0.7376	64

Table 4: Full Linear Regression Model on 2016-2021 data (3)

Here we see similar results to those shown in Table 3. Additionally, while not using fixed effects, we gain an intercept term which can help us interpret our analysis in greater depth. The intercept coefficient represents a regular season game with fans in attendance with teams who are greater than 300 miles from one another. The highly significant coefficient estimate represents that under this scenario we can expect the home team to outperform the away team by on average 0.2911 goals. In all of our regressions, this is the most significant variable, which can be expected considering the vast research on home team advantage in sports. Furthermore, although we see that the coefficient for the no crowd variable in this model is not statistically significant on its own, its negative value still offers valuable insight. This shows us that without a crowd the home team still has a 0.1588 goal advantage, holding all other variables constant. This is important considering how no crowd interacts with the playoff variable.

Our only other significant variable in this model is the interaction term between having no crowd during a playoff game. Once again, we have a fascinating interpretation from this variable in which during a playoff game with no crowd, while also implying it is not a close matchup, we can expect the *away* team to win by 0.3855 goals on average. In other words, the home team has a -0.3855 goal (dis)advantage.

Other variables of interest in this model such as close matchup was proved not significant and had little impact on our model due to its small coefficient. Overall, this model performed very similarly to our first model which was a good sign. Despite the no crowd variable being not significant at the 10% level, its negative coefficient provides us with some information regarding its impact and adds to the strength of its interaction with the playoffs.

The last two models we ran were the simplest and most interpretable models in our research. These regressions are shown in Table 2 as regressions 1 and 4, respectively. Here we regressed the no crowd variable against our dependent variable of home difference with and without fixed effects. The outputs from these regressions are shown below in Table 5.

Table 5: No Crowd Effect With and Without Fixed Effects

No Fixed Effects					
	Estimate	**Std. Error**	**t-value**	**p-value**	**Observations**
(Intercept)	0.2161	0.1286	1.68	**0.0934 .**	7551
no_crowd	-0.5694	0.1974	-2.885	**0.0041 ****	707
Fixed Effects					
no_crowd	-0.0611	0.1107	-0.552	0.581	707

Table 5: No Fixed Effects vs Fixed Effects on No Crowd Variable 2016-2021 (1,4)

The first output shows a very straightforward impact of having no crowd home team performance. Our significant intercept variable shows us that on average we can expect the home team to win by 0.2161 goals when there is a crowd in attendance. On the other hand, when there

is no crowd this home team differential drops by 0.5694 to an effective -0.3533, once again giving the away team an advantage.

With that being said, the second model tells us that there is some more explaining to do. Here we see that the no crowd variable is no longer significant when using fixed effects. This can possibly be explained by the impact of having no crowd varies greatly depending on the matchup. To add to the robustness of our results we ran the same regressions on the 2020-2021 subset of our data and the results are shown below in Table 6.

Table 6: No Crowd Effect With and Without Fixed Effects (2020-2021)

No Fixed Effects					
	Estimate	**Std. Error**	**t-value**	**p-value**	**Observations**
(Intercept)	0.3159	0.065	4.861	**1.25e-06*** **	7551
no_crowd	-0.2664	0.114	-2.337	**0.0195 ***	707
Fixed Effects					
no_crowd	0.127	0.1586	0.801	0.423	707

Table 6: No Fixed Effects vs Fixed Effects on No Crowd Variable 2020-2021 Subset (1,4)

Here we see much of the same thing confirming the validity of our results. To start, the no fixed effects model on the 2020-2021 seasons has similar results to that seen in Table 5 on the entire sample. Likewise, the no crowd variable in the fixed effects model is no longer significant which is an interesting finding considering the results from Table 3, on the same subset of data with matchup fixed effects, which found no crowd to be significant in the full model. These findings strengthen the importance of the playoff interaction term in regression 6 as it helps explain a unique influence on home goal differential that is not captured otherwise in a simpler model. Comparing the results from Table 3 and Table 6, we see that in Table 6 the no crowd variable is not significant suggesting that the fixed effects help us control for omitted variable bias. Moreover, adding the playoff term continues to strengthen our analysis by accounting for more variables giving us more precision in our analysis. The approach of using a full model with

fixed effects enhances the robustness of our analysis by maximizing control over potential confounding factors, thereby giving us a more rigorous causal relationship. This allows for a better interpretation of our coefficients, ensuring each variable contributes optimally to our model.

Summary and Concluding Remarks

Our results provide us with a clear interpretation that having no crowd can negatively impact the performance of the home team in the National Hockey League specifically during the NHL playoffs. The interpretation of our results shown in Table 3 and Table 5 provide statistically significant evidence that during an NHL playoff game without a crowd, the away team has an advantage. This is an important finding that can be applicable today showing the importance of having a strong home-biased crowd in attendance for playoff games.

The results from our research on regular season games are not as straightforward. While we have significant negative coefficients in our no fixed effects models in Table 5 and Table 6, these negative coefficients disappear in significance and value when we control for team matchups. While it is hard to know the exact cause of this, we can imagine a story in which a team historically has an advantage over another team regardless of outside factors such as crowd and whether the game was home or away. The non-significant coefficients in Tables 5 and 6 for our fixed effects models help show the importance of the playoff variable, as including this allows for better control over all the outside factors, allowing our coefficients to properly explain their variation.

Overall, our models provide insight into the nature of home team performance during the Covid-19 affected 2020-2021 NHL seasons. Our results show the practical importance of having fans in NHL arenas to bolster home team performance, particularly during the playoff

environment. While the fixed effects help control for a lot of variability there are a lot more outside factors that go into the score of a hockey game. Our sample size, particularly of playoff games without a crowd is relatively small (146) compared to the entirety of our sample. This along with the Covid-19 season irregularities can impact our results. The 2020-2021 seasons experienced hockey in August along with a shorter offseason, and Covid-19 health and travel protocols that can impact team performance. It is unlikely that we will experience similar circumstances soon within the NHL, but our analysis provides great insight into the seasons directly affected by Covid-19.

These findings can be extremely beneficial to help quantify the importance of a home crowd. Relating our findings back to Carron et al.'s framework in "The Home Advantage" where he discussed home crowd, referee bias, travel, and familiarity as important factors in home advantage. We can confidently say that the home crowd is the biggest factor in what creates the home advantage for these teams within the NHL. Given that framework, we isolate the home crowd variable due to the natural experiment Covid-19 allows for us and our research shows that there is a significant, negative impact on home team advantage when there is no longer a crowd in the National Hockey League during the 2020-2021 seasons.

Sources

Arboix-Alió, Jordi, et al. “The behaviour of home advantage during the COVID-19 pandemic in European Rink Hockey Leagues.” *International Journal of Environmental Research and Public Health*, vol. 19, no. 1, 26 Dec. 2021, p. 228, https://doi.org/10.3390/ijerph19010228.

Boyko, Ryan H. “Referee bias contributes to home advantage in English premiership football.” *Journal of Sports Sciences*, vol. 25, no. 11, Sept. 2007, pp. 1185–1194, https://doi.org/10.1080/02640410601038576.

Carron, Albert & Paradis, Kyle. (2014). The Home Advantage. 10.4135/9781483332222.n138.

Cross, Jeffrey, and Richard Uhrig. “Do fans impact sports outcomes? A covid-19 natural experiment.” *SSRN Electronic Journal*, 2020, https://doi.org/10.2139/ssrn.3705085.

Gong, Hua. “The effect of the crowd on home bias: Evidence from NBA games during the covid-19 pandemic.” *Journal of Sports Economics*, vol. 23, no. 7, 17 Jan. 2022, pp. 950–975, https://doi.org/10.1177/15270025211073337.

Guérette, Joël, et al. “The absence of fans removes the home advantage associated with penalties called by National Hockey League referees.” *PLOS ONE*, vol. 16, no. 8, 20 Aug. 2021, https://doi.org/10.1371/journal.pone.0256568.

“NHL Stats, History, Scores, Standings, Playoffs, Schedule & Records.” *Hockey*, www.hockey-reference.com/. Accessed 10 Apr. 2024

Price, Michael, and Jun Yan. “The effects of the NBA Covid Bubble on the NBA playoffs: A case study for home-court advantage.” American Journal of Undergraduate Research, vol. 18, no. 4, 13 Apr. 2022, pp. 3–14, https://doi.org/10.33697/ajur.2022.051.

u/mbstone. R/NHL on Reddit: NHL Travel Distance, July 2023, www.reddit.com/r/nhl/comments/15cev26/nhl_travel_distance/.

.