***Off*season?: The Impact of International Offseason Play on NBA Players' Performance**

By Ishan Sinha '20 (isinha@princeton.edu)

https://github.com/isinha/Offseason

Abstract

Since resting players has gained increasing attention, the impact of offseason play at international tournaments is of great importance. To assess this impact, I analyzed the statistics of the 48 NBA players on the USA National teams for the 2010 FIBA World Championships, 2012 Olympics, 2014 FIBA World Championships, and 2016 Olympics, using t-tests. First, I compared their Advanced Statistics (PER, VORP, OBPM, DBPM, BPM, USG%, TS%) from the seasons immediately before and after summer play. Second, I compared their improvement between the seasons before and after summer play to their rates of improvement earlier in their careers. Finally, I compared the prevalence of rest-related injuries, like soreness, rightness, sprains, and fractures, in the season after summer play to the seasons before summer play.

Ultimately, I found that while there were some changes in Advanced Statistics after summer play, none of these changes were statistically significant. I did, however, find strikingly more injuries in the season after summer play than in the seasons before. This finding should be taken critically and explored further, however, since injury report data is limited online and I only considered what I thought were "rest-related injuries." Thus, while rest is important and the finding about increased injuries after summer play warrants attention, NBA personnel should not be concerned that international play will significantly affect overall player performance.

Data

The Python notebook and key Excel files used in this analysis, as well as a very detailed description of data collection and preparation process, are located at the above Github link. In this paper, I give an overview of the data I am using and of key files involved.

The final dataset I use ("Data_Final.xlsx") includes all 48 players on the 2010 FIBA World Championship, 2012 Olympics, 2014 FIBA World Championship, and 2016 Olympics USA National teams. For demographic statistics, I included player position, height, and weight, during summer play. I then included statistics of their summer play, including the tournament, their age at the time, games played and started, MPG, RPG, APG, and PPG. Then, I included each player's Advanced Statistics - PER, VORP, OBPM, DBPM, BPM, USG%, TS% - from Basketball Reference for the season after summer play, the season before summer play, and two seasons before summer play. Then, I looked at each player's cumulative NBA career before their summer play and took two sets of statistics – the average of their Advanced Statistics, and the average change of their Advanced Statistics from season to season. For each of these different timeframes, I also noted whether or not the player had suffered a rest-related injury, such as soreness, tightness, or sprains. I did, however, include major injuries unrelated to rest that sidelined players for extended periods of time, like fractures. I collected this injury data from Fox Sports and articles online, though I admit that injury data was hard to come by, and my "rest-related injury" criteria could be defined more formally.

I chose to use Advanced Statistics because they give holistic views on player performance, whereas box score stats like RPG/APG/PPG do not attempt to tell the whole story of a player's impact on the floor.

I collected all individual player Advanced Statistics in "Stats.xlsx", all injury data in "Injuries.xlsx", and stored initial summer data in "Data.xlsx". I then wrote a Python script to clean this data and export it to a file called "Fill_in.xlsx", which I copy-pasted into empty parts of "Data.xlsx" to create "Data_Final.xlsx".

Procedure

An effect in player performance could be detected by comparing a player's raw statistics from before and after summer play, as well as comparing rates of improvement (or the opposite) in these advanced statistics from season to season. For example, it would be significant to know that a player's PER dropped from 24 the season before the Olympics to 20 the season afterwards. However, that does not tell the whole story. Let's say a player's PER increased from 20 the season before the Olympics to 20.5 the season afterwards. Though it seems good he improved by 0.5 points, it would be important to know at which rate he was improving before. If his PER was improving by 3 points every season before, his "progress" would actually be a red flag since he was progressing more slowly than usual; if his PER was decreasing every year before this, this progress would seem even more promising.

In total, I had five comparisons. Note that "–" in the following descriptions is an actual minus sign:

1) Raw Statistics: (Season after summer play) vs. (Season before summer play)
    a. This compares the raw statistics from the season before to the season after summer play. It checks if the player performed worse after international summer competition.
2) Rate of Improvement: (Season after summer play – Season before summer play) vs. (Season before summer play – Two seasons before summer play)
    a. This compares the improvement between the seasons before and after summer play to the improvement between the two seasons before summer play. It isolates whether the latest rate of improvement slowed down after summer play.
3) Rate of Improvement: (Season after summer play – Season before summer play) vs. (Average improvement from season-to-season before summer play)
    a. This compares the improvement between the seasons before and after summer play to the average improvement over a player's career. It isolates whether a player's progress was slower than his general career progress rate after playing for a summer.
4) Raw statistics: (Number of injuries the season after summer play) vs. (Number of injuries the season before summer play)
    a. This compares the raw statistics of how many injuries players on the USA National team sustained before and after summer play. It isolates whether the frequency of player injuries spiked or not.
5) Rate of improvement: (Injuries after summer play – Injuries before summer play) vs. (Injuries before summer play – Injuries two seasons before summer play)
    a. This compares the change in amount of injuries between the seasons before and after summer play to the change in amount of injuries between the two seasons before summer play. It isolates whether the rate of player injuries spiked after summer play compared to before summer play.

Using the data I accumulated, I used t-tests to evaluate these comparisons. A t-test shows whether there is a significant difference between two groups of data. The p-value, also called the "significance level," that accompanies a t-test denotes the probability that the difference is randomly due to chance; I use the customary threshold that p-values should be less than or equal 0.05 (there should be a less than 5% chance that results are a fluke). The t-tests, in this case, tell whether there is a significant difference between player performance before and after summer play, which would serve as evidence for a notable correlation between international summer play and worsened performance.

Results

	Comp. #1: Season after vs. Season before	Comp #2: (Season after - Season before) vs (Season before - 2 Seasons before)	Comp #3: (Season after - Season before) vs (Average of Changes of All Seasons Before)
PER	0.220654	0.810376	0.803950
TS%	0.617611	0.849619	0.833774
TOV%	0.308763	0.075446	0.486360
USG%	0.600118	0.694450	0.590157
WS/48	0.282111	0.842358	0.603179
OBPM	0.316614	0.822422	0.476926
DBPM	0.933539	0.545690	0.735585
BPM	0.402894	0.670781	0.444351
VORP	0.283801	0.834848	0.657610

Table 1: p-values of t-tests

Per Table 1, all of the p-values are higher than the standard 0.05 threshold for statistical significance. Therefore, there are no significant differences in any Advanced Statistics after a summer of international play.

However, the p-value for TOV% in the 2nd column, 0.075, is very close to the 0.05 threshold. Upon further analysis, it turns out this is the result of two random outliers – Tyson Chandler around 2010 FIBA and Draymond Green around the 2016 Olympics. After removal of just Tyson Chandler, the p-value spikes to 0.26; also removing Draymond green spikes the p-value to 0.49. Thus, even this one seemingly statistically significant change is actually insignificant. A deep-dive into this anomaly is explained in the Appendix of this paper.

	Total # of Players
Injured in any season before summer play?	35
Injured two seasons before summer play?	21
Injured the season before summer play?	22
Injured during summer play?	1
Injured the season after summer play?	31

Table 2: Number of injuries in different seasons

Looking at injury frequency, there is a striking difference. 22 players had rest-related injuries of some sort during the season before summer play. However, 31 players had rest-related injuries of some sort the seasons after summer play. The p-value of comparing these two distributions of injuries was 0.065, nearing statistical significance.

To see if the rate of injuries increased, I then compared the difference between injuries the seasons after and before summer play ("Injured the season after summer play?" – "Injured the season before summer play?") and the difference between the two seasons before summer play ("Injured the season before summer play?" –

"Injured two seasons before summer play?"). This p-value was 0.192, well above the 0.05 threshold. This higher p-value occurred because different players were injured in these seasons. Nine of the players injured the season after summer play were not injured the season before, while four of the players injured the season before summer play were not injured the season after. Four of the players injured the season before summer play were not injured two seasons before summer play, while four of the players injured two seasons before summer play were not injured the season before summer play.

```
Players injured after but not before summer play: 9
{'Paul George', 'Chauncey Billups', 'Russell Westbrook', 'Draymond Green', 'Stephen Curry', 'Chris Paul', 'LeBron James', 'Kyle
Lowry', 'Andre Iguodala'}

Players injured before but not after summer play: 4
{'Danny Granger', 'Andre Drummond', 'Harrison Barnes', 'James Harden'}

Players injured one season before but not two seasons before summer play: 4
{'Kobe Bryant', 'Andre Drummond', 'Harrison Barnes', 'James Harden'}

Players injured two seasons before but not one season before summer play: 4
{'Deron Williams', 'Paul George', 'Stephen Curry', 'Andre Iguodala'}
```

Figure 1: Difference in sets of players injured

Discussion and Conclusion

The results regarding the impact of offseason international play on Advanced Statistics are decisive. Across the board, there are no large differences in these Advanced Statistics, and whichever differences we see are most likely due to chance. If these comparisons turned out statistically significant, they would have established correlation between summer play and worsened performance. However, the fact that no correlation was established implies that there is no causation. Therefore, NBA personnel can be confident that allowing their players to play at these tournaments will not have significant effects on their performance in future seasons. However, this is not to say that players should not take rest. If players like Stephen Curry and LeBron James had not taken rest after the 2016 season when they felt they needed it, as well as the many other players who have opted out of summer play, these statistics may look very different.

A notable insight warranting further research is the increase in number of injuries suffered after summer play. While it is very difficult to attribute these injuries to summer play directly, the more than 40% spike from 22 injuries before to 31 injuries after summer play is alarming. However, the data I have collected for this part of the analysis need to be verified and judged further. I searched online using Fox Sports and different articles I found to fill in these columns, documenting my reasoning and sources as much as possible in my data files. However, I am not an expert on what truly constitutes a "rest-related injury," and it is possible I missed some injuries due to lack of readily available information online. It would be helpful to have more formal data and standard criteria for categories of injuries to further public research on this topic.

All in all, NBA personnel should look into the apparent spike in injuries the season directly after international play, but can be generally confident that overseas offseason play will not notably harm their players' performance.

Further Research

There are five areas I think warrant further research with this analysis.

First and most importantly, I could use a better framework for collecting and evaluating injury data. It was very difficult finding accurate and holistic injury data for players, so I am not wholly certain my data was completely accurate, though I searched extensively and used my best judgment. Further, this analysis would benefit from professional assessment of which injuries to include. I tried to include only "rest-related injuries," such as soreness, fatigue, tightness, and sprains, but injury reports were not too detailed, and I felt obligated to include

non-rest related injuries like random, severe fractures, which sidelined players for extended periods of time. As I noted in the "Discussions and Conclusions" section, these discrepancies add doubt to the results of comparing number of injuries before and after summer play.

Second, I would like to compare the rates of change in these Advanced Statistics to the rates of change for NBA players who did not play internationally during the offseason for the seasons in question. This would ensure or disprove that apparent statistical trends are unique to players who played internationally over summer and also serve as another benchmark for comparison.

Third, I would like to add more data from FIBA and Olympics tournaments going back farther than 2010. This would add more security to this analysis' findings, and also allow for more granular analyses, such as seeing impacts on different positions.

Fourth, I would like to add NBA players who played for their native national teams, instead of limiting the analysis just to NBA players who played for the USA National Team.

Fifth, I think it might be impactful to normalize statistics for players who played in multiple international tournaments. For example, for a player who participated in both 2012 Olympics and 2010 FIBA, the "two seasons before" statistics for 2012 are equivalent to the "Season after" statistics for 2010. I tried to account for this by comparing the change between the seasons after and before summer play to the average change from season to season earlier in the player's career; if the alternative hypothesis holds and playing in the offseason does harm player performance, then each extra offseason played should lead to a progressively slower improvement rate. However, it is still very possible that directly addressing the cases where a player has played internationally multiple times could make a significant difference. In future versions of this analysis with more data, per points three and four, I would like to take only players' first tournaments.

Appendix: Deep-dive into TOV% p-value

The following graphs show the different values for the change in TOV% for the 48 players on these four rosters. They look at the (Season after summer – Season before) and (Season before – 2 Seasons before) statistics, respectively. They are followed by descriptions of their statistical distributions. Note that these are discrete points, not a continuous spectrum. Also note that the count of players in Tables 3 and 4 is 45 instead of 48 because a few players did not have at least two seasons in the NBA prior to playing in the USA National Team (Stephen Curry in 2010, Anthony Davis in 2012, and Mason Plumlee in 2014).

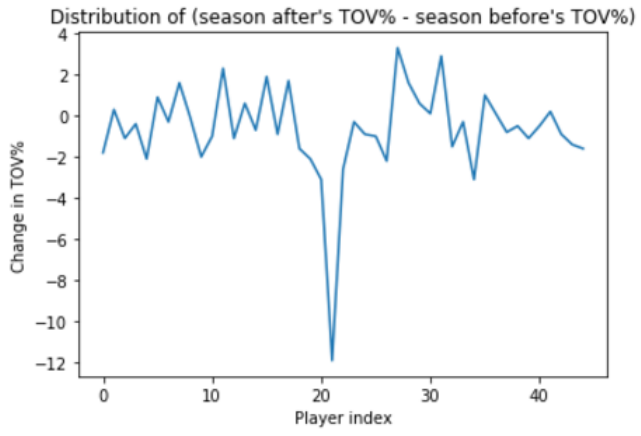

Fig. 2: Change after/before summer

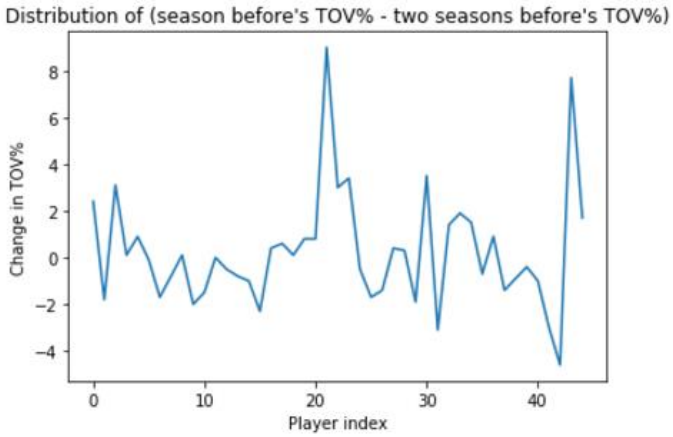

Fig. 3: Change between the two seasons before summer

```
count    45.000000
mean     -0.662222
std       2.271892
min     -11.900000
25%      -1.500000
50%      -0.700000
75%       0.300000
max       3.300000
dtype: float64
```

Table 3: Description of Fig. 2's distribution

```
count    45.000000
mean      0.242222
std       2.492488
min      -4.600000
25%      -1.400000
50%       0.000000
75%       0.900000
max       9.000000
dtype: float64
```

Table 4: Description of Fig. 3's distribution

In Figure 1, there is one notable decrease in TOV%. In other words, someone really cut down on his turnovers! In Figure 2, there are two notable increases in TOV%. Two guys became a lot more turnover-prone quickly!

	Player	Summer	TOV%_SA	TOV%_SB	TOV%_2SB
21	Tyson Chandler	2010 FIBA	14.1	26	17

	Player	Summer	TOV%_SA	TOV%_SB	TOV%_2SB
43	Draymond Green	2016 Olympics	19.8	21.2	13.5

Table 5: Tyson Chandler and Draymond Green

Per Table 2, the culprits are Tyson Chandler and Draymond Green. Tyson Chandler had a massive increase in TOV% from two seasons before to the season before the 2010 FIBA World Championships, which makes sense given his injury-plagued 2009-2010 season. After joining Dallas after FIBA, he had a massive decrease in his TOV%. Draymond Green just had a massive increase in TOV% from 2014-2015 to 2015-2016.

After removal of these outliers, the graphs appear much more stationary, and descriptions of their statistical distributions appear less volatile as well.

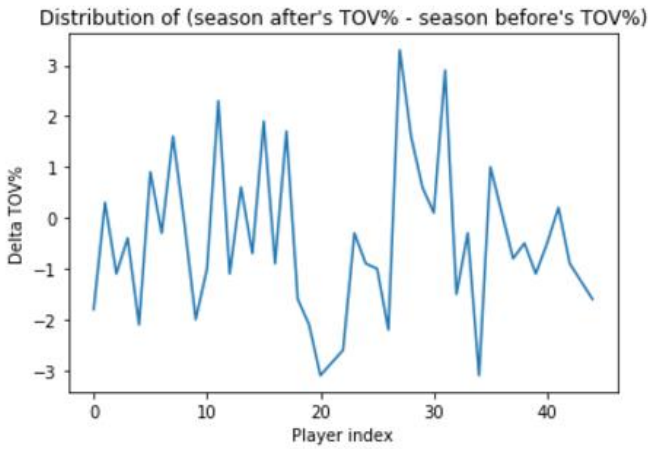

Fig. 4: Figure 2 without the outliers

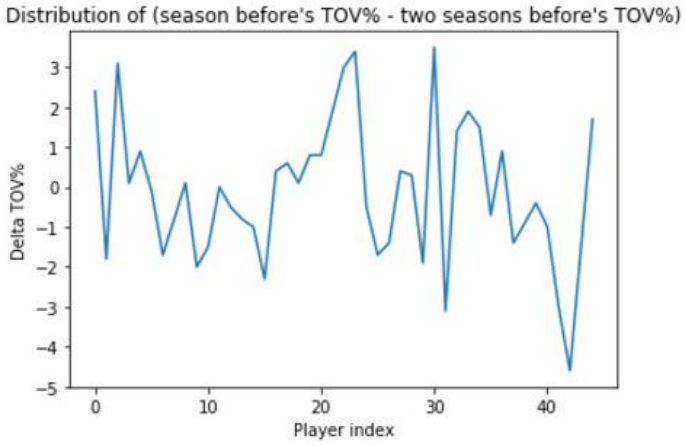

Fig. 5: Figure 3 without the outliers

count	43.000000
mean	-0.383721
std	1.519231
min	-3.100000
25%	-1.300000
50%	-0.500000
75%	0.450000
max	3.300000
dtype:	float64

Table 6: Description of Fig. 4's distribution

count	43.000000
mean	-0.134884
std	1.792296
min	-4.600000
25%	-1.400000
50%	-0.100000
75%	0.850000
max	3.500000
dtype:	float64

Table 7: Description of Fig. 5's distribution

Removing just Tyson Chandler, the p-value spikes to 0.26. Removing both Tyson Chandler and Draymond green spikes the p-value to 0.49. Thus, even this one seemingly statistically significant change actually is not.