Finding the Open Receiver: A Quantitative Geospatial Analysis of Quarterback Decision-Making

Jeremy Hochstedler[a,b]
[a]Telemetry Sports & [b]Massachusetts Institute of Technology

1. Motivation

Within football and media organizations, significant resources are consumed through the analysis of player performance and decision-making, especially at the quarterback (QB) position. Since 2014, radio-frequency identification (RFID) tracking technology has been used to continuously monitor the on-field locations of NFL players. [1 Zebra Technologies] Using geospatial American football data, this research quantitatively evaluates receiver openness, player elusiveness, and QB decision-making.

In addition to enhancing win probability, using NFL injury data, we have discovered how QB decisions and passing ability impact the likelihood of receiver concussions. Specifically, Fig. 1 demonstrates how better passers reduce team losses and receiver concussions. [2 Mrkvicka and Hochstedler] By making better decisions and finding the open receiver, QB's can put their receivers and their teams in better positions to succeed.

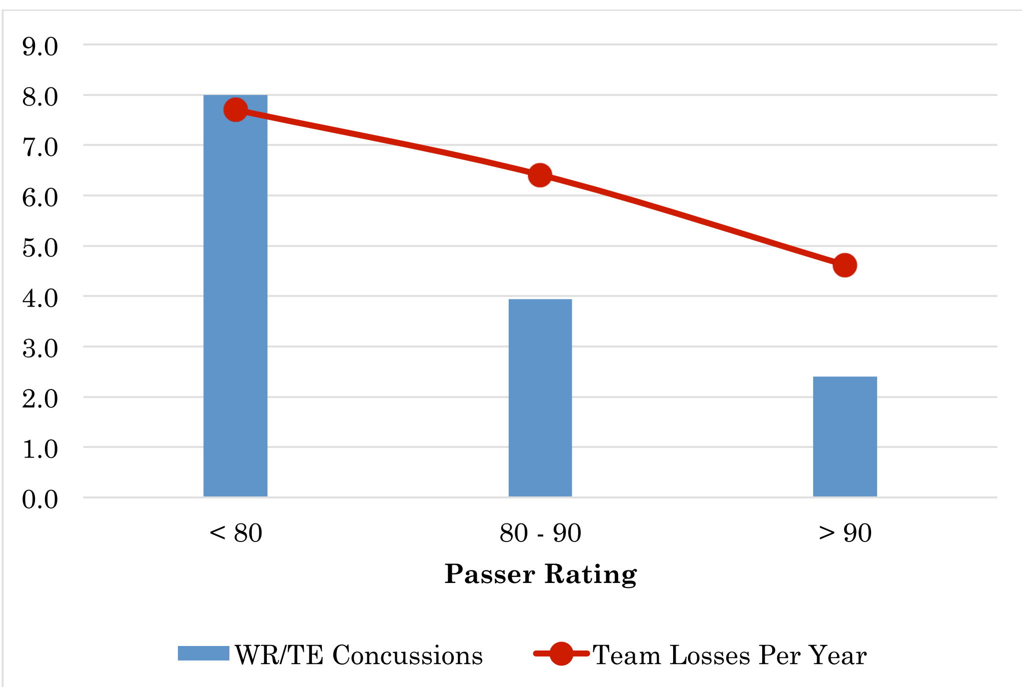


Figure 1. *NFL QB Passer Rating vs. Concussions and Team Losses from 2012 – 2015*

2. Data Collection

NFL RFID data is not publically available for analysis at this time. In order to perform similar analyses, we captured spatial coordinates of all twenty-two on-field players based on game video at three frames per second to the nearest 0.25 yard for every base offensive pass play for the 2014 Indianapolis Colts (231 total plays). Here, the "base offense" is defined as 1st & 2nd down, less than 15-point differential, greater than five minutes remaining in a half, and between the 20 yard lines.

In order to evaluate decision-making, utility must be defined. The success outcomes that this analysis considers are completions and yards gained. The base offensive pass plays for the 2014 Indianapolis Colts provide a significant sample that allows for an analysis of Andrew Luck's decision-making. Because strategy changes towards the end of a half and near the end zones, these play types are removed from consideration. Additionally, this dataset focuses on 1st and 2nd down plays since the marginal team benefit of additional yardage is more direct, which doesn't hold true on 3rd and 4th downs. In summary, a significant benefit is earned when an offensive team achieves more yardage than the yard to gain (i.e. when the offensive team "picks up the first down") on 3rd and 4th downs.[1] Fig. 2 provides a visual representation of team utility.

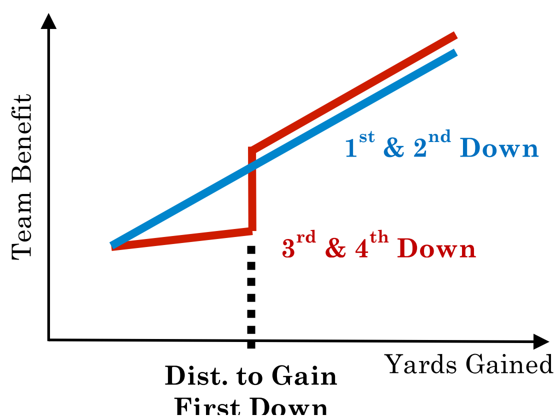


Figure 2. *Down and Distance Utility Function*

3. Receiver Openness

The traditional Voronoi Tessellation has been used to help quantify rebounding skill in the NBA. [3 Maheswaran et al.] This work extends upon the idea to use the tessellation in sports by quantifying receiver "*openness*" in the NFL. However, player velocities significantly impact receiver routes and the defense of those routes. Because the player velocity is important in addition to positional data, a "predictive" tessellation has been developed to quantify receiver *"openness"*. That is, because the game relies on "where a player will be" not necessarily "where he currently stands," predictive methods more accurately reflect player decision-making as it relates to geospatial analysis. *Figures 3A* and *3B* outline the distinctions between a traditional and "predictive" tessellation.

[1] While this utility function is assumed, in truth, there are slight benefits in achieving a first down on 1st and 2nd down plays, however they are minimal compared to picking up a first down on 3rd or 4th down.

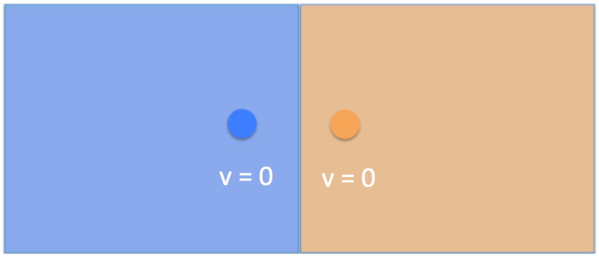


Figure 3A. *A traditional Voronoi Tessellation where both players are immobile and assumed to possess identical acceleration profiles. The blue cell maintains the shortest Euclidean distance to every point shaded blue, whereas the orange player maintains the shortest distance to every point shaded orange.*

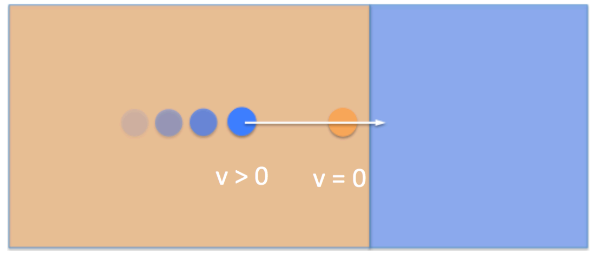


Figure 3B. *A "predictive" Voronoi Tessellation where the blue player possess a non-zero velocity toward an immobile orange player. The blue player is moving fast enough that he now "owns" the ground behind the orange player since he can reach those points more quickly (in time) even though the orange player is currently the closest player (in distance) to the points in that cell.*

3.1. *Zone Size*

By analyzing all twenty-two players' instantaneous positions and velocities, the Voronoi Tessellation is performed after projecting each player's position forward two frames (2/3 of a second) and finding the respective zone area (yards^2) for each player. Zone areas are then established for each eligible receiver for each frame throughout the play. *Figure 4* displays the zone area distribution for every eligible receiver in 2014 Colts' pass plays at the moment of QB release. To further understand the "predictive" tessellation, refer to Appendix I, which displays the evolution of the Indianapolis Colts' first offensive play from scrimmage from the 2014 season. Using the geospatial coordinates captured, this analysis was performed on each base offensive pass play from the 2014 season.

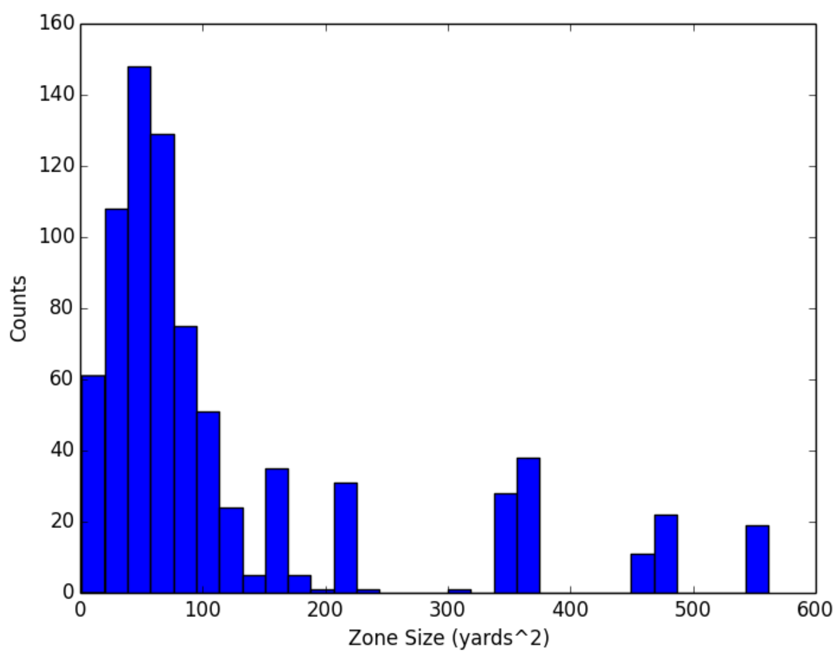


Figure 4. *Distribution of predicted Voronoi zone size for all eligible receivers at QB decision point*

3.2. Zone *Integrity*

While zone size describes how much of the field a receiver "owns," a defender may still be lurking nearby to ultimately break-up the intended pass. Therefore, projected zone *integrities* are calculated for each eligible receiver and play frame. Zone *integrity is measured as the* ***projected*** *distance to nearest defender from an eligible receiver.*

3.3. *Openness* Classification

To enable classification of the *openness* of an eligible receiver, *zone size* and *integrity* are combined to simplify the analysis and explain the spirit of what can be measured using the geospatial data. Receiver *openness* is classified as: *wide-open*, *open*, *defended*, or *well-defended*. Specifically:

- *Wide-open* = zone area > 200 yards^2 or integrity > 8 yards
- *Open* = zone area between 100-200 yards^2 or integrity between 4-8 yards
- *Defended* = zone area between 50-100 yards^2 or between 2-4 yards
- *Well-defended* = zone area < 50 yards^2 or integrity < 2 yards

Table 1 displays Andrew Luck's 2014 completion percentage for each targeted receiver's openness. Note: four plays were not included in the data set, as they did not have a clearly defined targeted receiver.

Openness of Targeted Receiver	Plays	Completions	Completion %
Wide-open	69	49	71%
Open	70	49	70%
Defended	67	39	58%
Well-defended	21	11	52%
N/A	4	0	0%
Total	231	148	64%

Table 1. *Andrew Luck's 2014 completion percentage as a function of targeted receiver openness.*

As observed, Luck completed a higher percentage of passes to *open* and *wide-open* receivers than those who were *defended* or *well-defended*.

4. Expected Yardage and Player Elusiveness

4.1. Expected Gain

For each frame throughout the play, expected yardage is also derived by observing the maximum y-value of that receiver's predicted zone. In a theoretical "pure play" where an ideal throw (on-time and on-target) is matched with an ideal catch and a defender making an ideal tackle (including reaction and pursuit angle), a receiver's maximum gain would occur at the point in his zone which is the furthest point possible down field.

Figure 5 displays Luck's decision point from the Colts first offensive play from scrimmage from the 2014 season. Specifically, Hilton possesses a smaller zone (red) while Wayne maintains a large zone (blue). If Luck were to make an ideal pass at this moment, Wayne should be expected to obtain 26 yards on the play. On this particular play Luck slightly underthrows Wayne, causing Wayne to flatten out his route and carrying him out of bounds with a 21-yard gain. *Appendix I* provides further detail from this play.

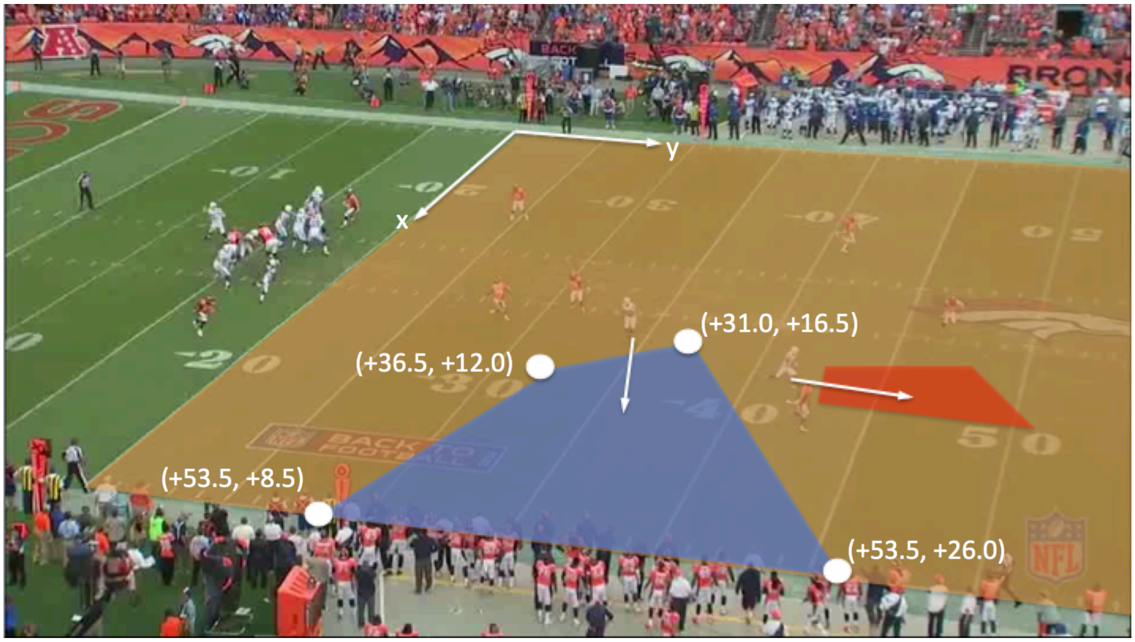


Figure 5. *Colts' 2014 first play from scrimmage. Receiver Wayne is classified "wide open" (blue zone) with predicted gain of +26 yards.*

Figure 6 demonstrates the comparison of the expected gain for every completed pass (148 in total) against the play's actual gain.

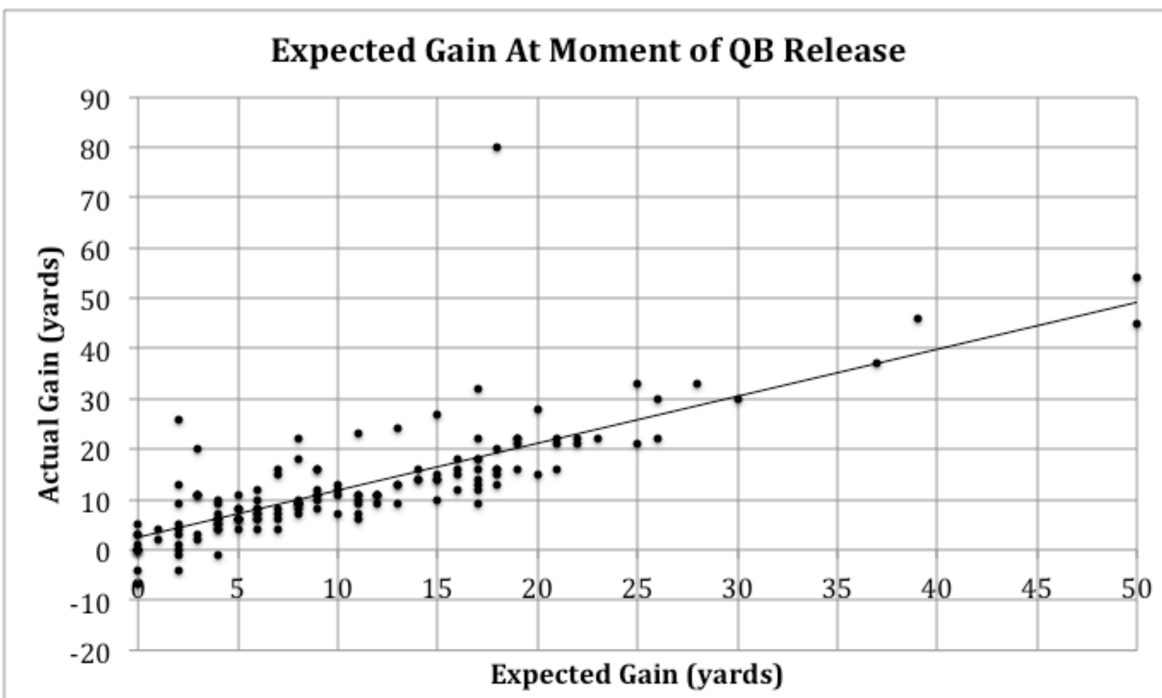


Figure 6. *Expected play gain as measured on 148 completions from the Colts' 2014 base offense.*

4.2. Player Elusiveness

While each completion possesses an expected gain in yardage (whether positive or negative), the actions of the QB, receiver, and defenders will ultimately define the actual yardage achieved. For example, a poorly executed pass that forces a receiver to dive for the catch will likely reduce the actual yardage gained. Alternatively, several broken tackles after the catch will likely lead to a higher than expected actual gain. For each play, the difference in actual yardage gained is measured for each individual receiver. In sum, the additional yardage gained from the expected yardage defines a player's *"elusiveness*." That is, the more yards a player gains than expected, the more elusive he is. *Table 2* displays the player *elusiveness* for each player on the 148 completions. Notice how running backs (RB) are more elusive than tight ends (TE) and wide receivers (WR).

Position	Receiver	Additional Yards Gained From Expectation	Catches	Mean
RB	Herron	34	7	4.86
RB	Richardson	55	12	4.58
RB	Bradshaw	31	10	3.10
TE	Fleener	36	15	2.40
WR	Moncrief	9	4	2.25
WR	Wayne	60	27	2.22
TE	Doyle	8	5	1.60
TE	Allen	8	14	0.57
WR	Hilton	20	44	0.45
WR	Nicks	-1	10	-0.10
All Colts RB's		120	29	4.14
All Colts TE's		52	34	1.53
All Colts WR's		88	85	1.04

Table 2. *Receiver elusiveness for 148 completions from 2014 Colts base offense.*

5. QB Decision Analysis

Although plays typically designed to have primary, secondary, and tertiary targets, the decision of which receiver to target ultimately comes down to the QB. A QB can check down his options and decide for whom to target with his pass. This analysis attempts to model how those decisions are made based on geospatial elements created through *zone size* and *integrity* of eligible receivers. Here, the *zone size* and *integrity* are combined to quantify a receiver's *openness* along with his expected gain.

Using these factors, Andrew Luck's decision-making can be analyzed. For a given play, receiver options are skill players who do not block. Considering every eligible receiver at each frame (taken every 1/3 of a second during play development) prior to the pass release frame (final QB decision point), each receiver frame is assigned an *openness* (*wide-open*, *open*, *defended*, and *well-defended*) and an *expected yardage* of gain. These two factors are then combined to produce an expected utility if that option was chosen at that play frame.

The expected payoff is calculated as:

$$P(\text{CMP\% for openness factor}) * \text{Expected Yards} = \text{Expected Payoff Yds}$$

The expected payoff yardage is then captured for all options prior to the target selection, and the target payoff is compared to the play population options to determine the percentile of the target receiver expected payoff.

Decision types as a percentile of optional targets for given play:

- A percentile of 80 was classified as an *ideal target* decision.
- A percentile of 50 was classified as a *preferred target* decision.
- A percentile of 20 was classified as a *neutral target* decision.
- A percentile of below 20 was classified as an *undesirable target* decision.

In order to isolate QB decision-making, plays in which the intended receiver is unidentifiable (e.g. QB is hit as he throws, four plays in total) are removed from this analysis. Table 3. displays the results of this analysis. Specifically noting how QB Andrew Luck made an *ideal* or *preferred* decision on more than 75% of the pass plays analyzed.

<table>
<tr><th>Decision Type</th><th>Plays</th><th colspan="2">Percentage of Total Plays</th></tr>
<tr><td>Ideal</td><td>92</td><td>40.5%</td><td rowspan="2">75.8%</td></tr>
<tr><td>Preferred</td><td>80</td><td>35.2%</td></tr>
<tr><td>Neutral</td><td>45</td><td>19.8%</td><td rowspan="2">24.2%</td></tr>
<tr><td>Undesirable</td><td>10</td><td>4.4%</td></tr>
<tr><td>N/A</td><td>4</td><td colspan="2" rowspan="2"></td></tr>
<tr><td>Total</td><td>231</td></tr>
</table>

Table 3. *Andrew Luck 2014 Decision Analysis from the "Base Offense"*

6. Summary and Future Work

While this research provides a foundation to base future methods, it does not come without several limitations, including:

- This research projects instantaneous velocity as a constant (no acceleration), and as such, it assumes all players are assumed to perform at their current velocities (no individualized acceleration profiles).
- QB decision-making does not account for receiver distance from line of scrimmage in the completion probability equation [Goldsberry 4].
- While the data collection process was stringent and quality verified, it is only accurate to within 0.25 yards at 3 Hz. It is expected that the existing RFID tracking technology provides more accurate, more frequent data points for analysis.

In summary, this research attempts to show the types of analysis possible with geospatial data available in the NFL. Hopefully the data becomes available for teams, fans, and other media organizations to enable the development of insights into players, strategy, and decision-making.

Additional analyses may include:

- A defense's ability to restrict openness for receivers and how cornerbacks, linebackers, and safeties effectively pass off receivers while in different coverages.
- Analyzing multiple quarterback and team strategies can help identify how QB tendencies affect completion percentage and propensity for target decisions.

The NFL and American football have lagged behind other sports when it comes to analytical player and team analysis, however the geospatial tracking allows for significant opportunities as the data is analyzed both privately and publicly.

7. Acknowledgments

I would like to thank EA Sports for their help in creating and sharing their data, which served as a training set for our methods until the authentic data collection process was complete. I am also indebted to Neil, Troy, Katy, and Craig Mrkvicka for their significant efforts throughout the data collection process.

References

[1] Zebra Technologies. Accessed December 10, 2015. https://www.zebra.com/us/en/nfl.html

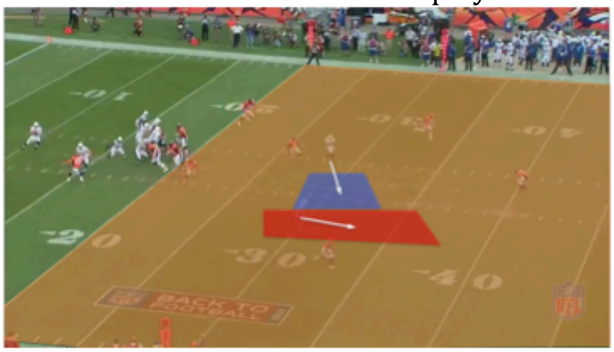
[2] Mrkvicka, Neil and Jeremy Hochstedler. Independent study using NFL injury report data from Data acquired from PBS concussion study. http://www.pbs.org/wgbh/frontline/film/league-of-denial and http://www.nfl.com/injuries

[3] Maheswaran et al., "The Three Dimensions of Rebounding." (paper presented at MIT Sloan Sports Analytics Conference, Boston, Massachusetts, February 28 - March 1, 2014).

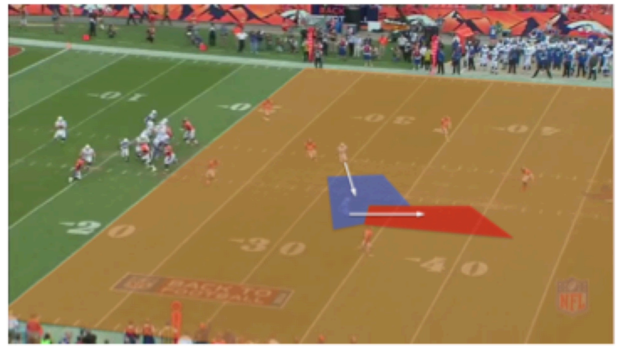
[4] Goldsberry, Kirk. "Pass Atlas: A Map of Where NFL Quarterbacks Throw the Ball." Last modified September 6, 2013. http://grantland.com/the-triangle/pass-atlas-a-map-of-where-nfl-quarterbacks-throw-the-ball/.

Appendix I

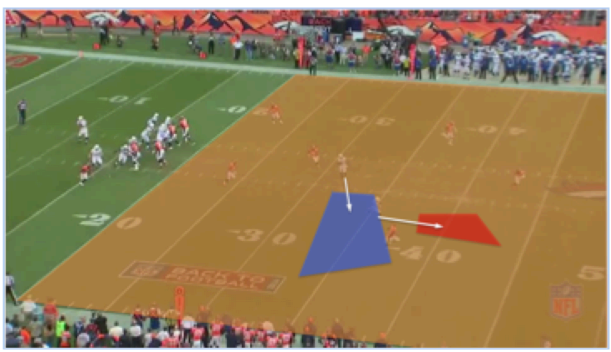
Five frames from the Colts first play from the 2014 season are illustrated below.



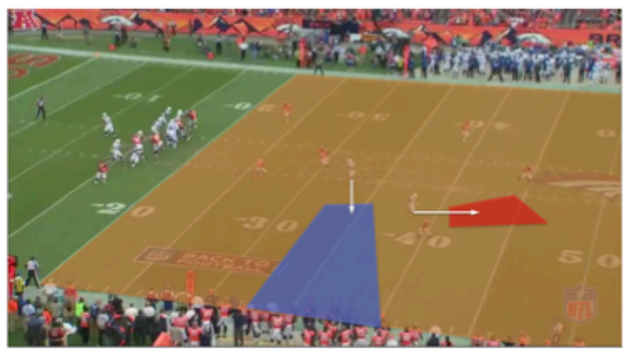
Frame 5. With this play designed to the offenses' right side of the field (Wayne on the "over" or deep crossing route), Luck looks left to hold off the free safety.



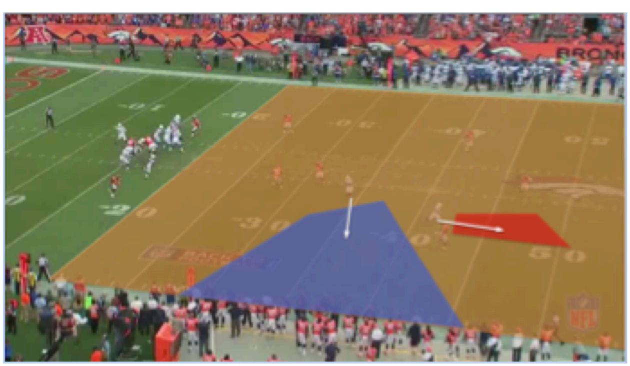
Frame 6. With two receivers maintaining small zones (i.e. "covered"), Luck remains patient, allowing the play to develop.



Frame 7. With Hilton's inside release go-route, he successfully occupies the cornerback responsible for the right one-third of the field.



Frame 8. With the corner staying true to his assignment, Luck observes Wayne's zone starting to open.



Frame 9. Luck makes his decision to throw to Wayne just as his zone reaches its maximum (i.e. when he becomes "wide-open"). Play result: good decision, good outcome. That is, Luck threw to a wide-open Wayne (good decision) for a reception and gain of 21 yards (good outcome).