Evaluating Catcher Blocking Models

Jack Rogers and Kai Franke

University of Minnesota
Driven to Discover℠

Introduction

- Catcher defense has been a widely researched topic in the baseball analytics community over the past few years.
- Pitch framing, the most influential portion of catcher defense, is a concept that we have been able to quantify in the statcast era of ball tracking.
- From a public perspective, stolen base prevention is a hard topic to approach with current data limitations, as we don't have data measuring the quality of a throw to second base, or a pitcher's time to the plate.
- Blocking, on the other hand, is a topic that we seem to have all of the data to quantify, and public models aren't quite as good as they should be.

Data Exploration Questions

Why are there more passed balls with runners on third?

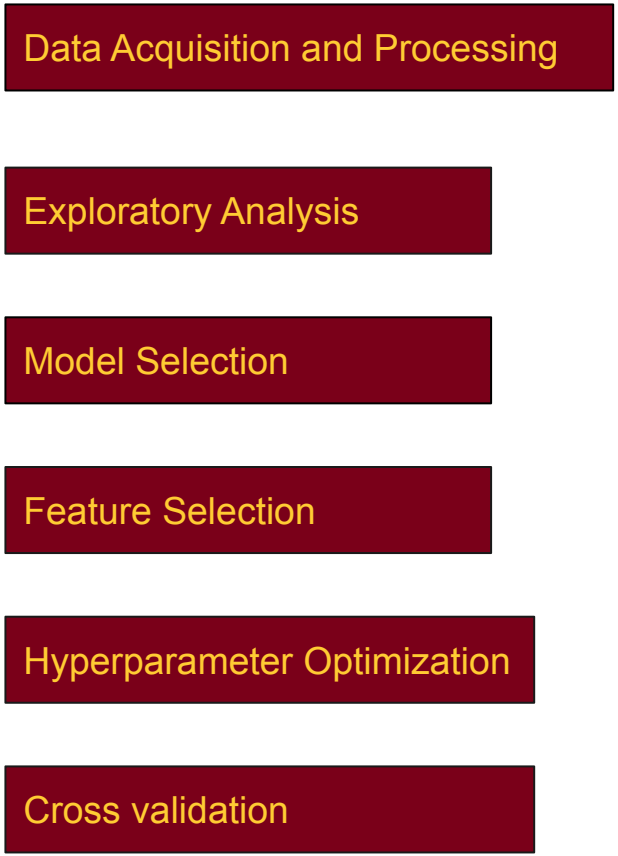
- Passed balls and wild pitches occur disproportionately in 3 scenarios: runner on third, third and second, and runner just on second.
- In game states other than these three only .001 percent of balls get past the catcher, as opposed to .02 in the above game states.
- Shouldn't catchers be worse when there is a stolen base threat?

Potential Answers

- Catchers changing stance
- Runners are being more aggressive
- Pitchers throw less offspeed pitches
- Proportion of pitches more than a foot out of the strike zone has increased by almost two percent.

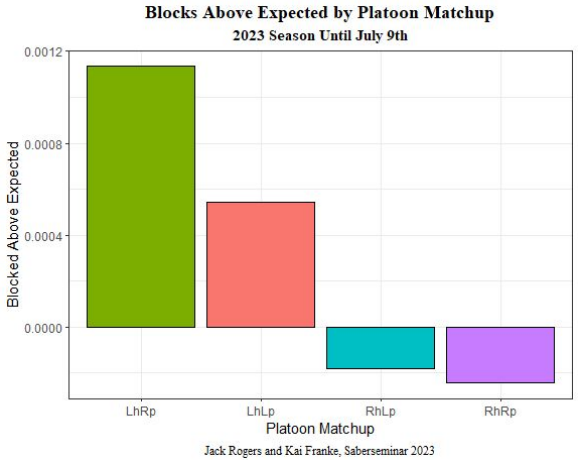
Rules for Effective Modeling

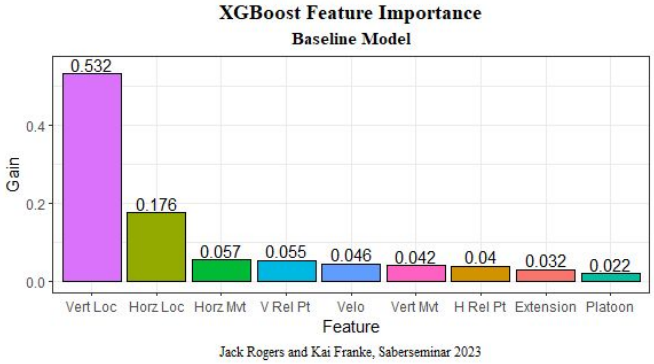
- Interpretable Inputs and Results aimed at coaches and player development
- Emphasis on predictive over descriptive value (ie context neutral run values or skill units)
- An imbalanced dataset requires a variety of evaluation metrics and techniques (F1 score and log-loss
- Hyperparameter optimization on holdout set
- Proper consideration of different model types

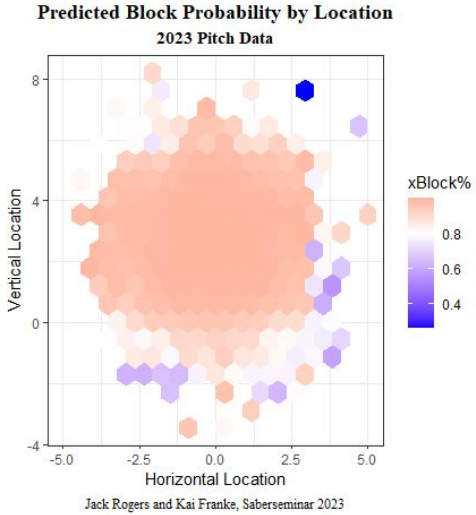


Modeling Results

- No separate feature selection
  - Only 9 total features
  - XGBoost has internal feature selection
- Grid search for optimized hyperparameters
- XGBoost feature importance
  - Pitch location
  - Platoon least important feature
- Precision (cutoff of 0.99 probability for labelling)
  - 0.917
- Recall
  - 0.998
- AUC
  - 0.892
- Log loss
  - 0.028







Leaderboard

- Leaderboard is filled with one knee and “hybrid stances”
- Stragglers tend to be either traditional or hybrid
- Higashioka, Wilson Contreras, and traditional stance catchers are all below average under this metric

BAX Leaders 2023

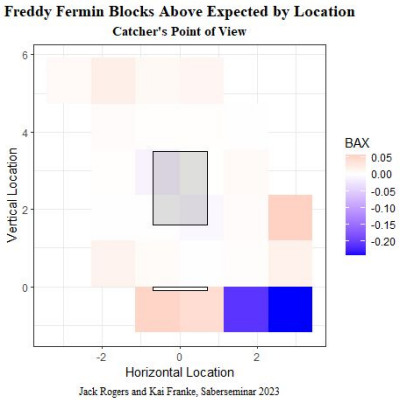
Name	Ops	Blocks	BAX	
Nick Fortes	1838	1833	7.40	
J.T. Realmuto	2634	2625	5.83	
Jose Trevino	1959	1953	5.64	
Austin Hedges	1775	1770	5.08	
Alejandro Kirk	1579	1575	4.19	
Francisco Alvarez	2110	2099	3.82	
Curt Casali	1090	1087	3.52	
Martin Maldonado	1702	1691	2.91	
Yan Gomes	1439	1433	2.82	
Gabriel Moreno	1019	1016	1.99	

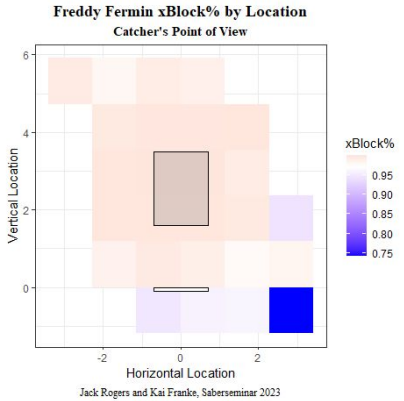
BAX Stragglers 2023

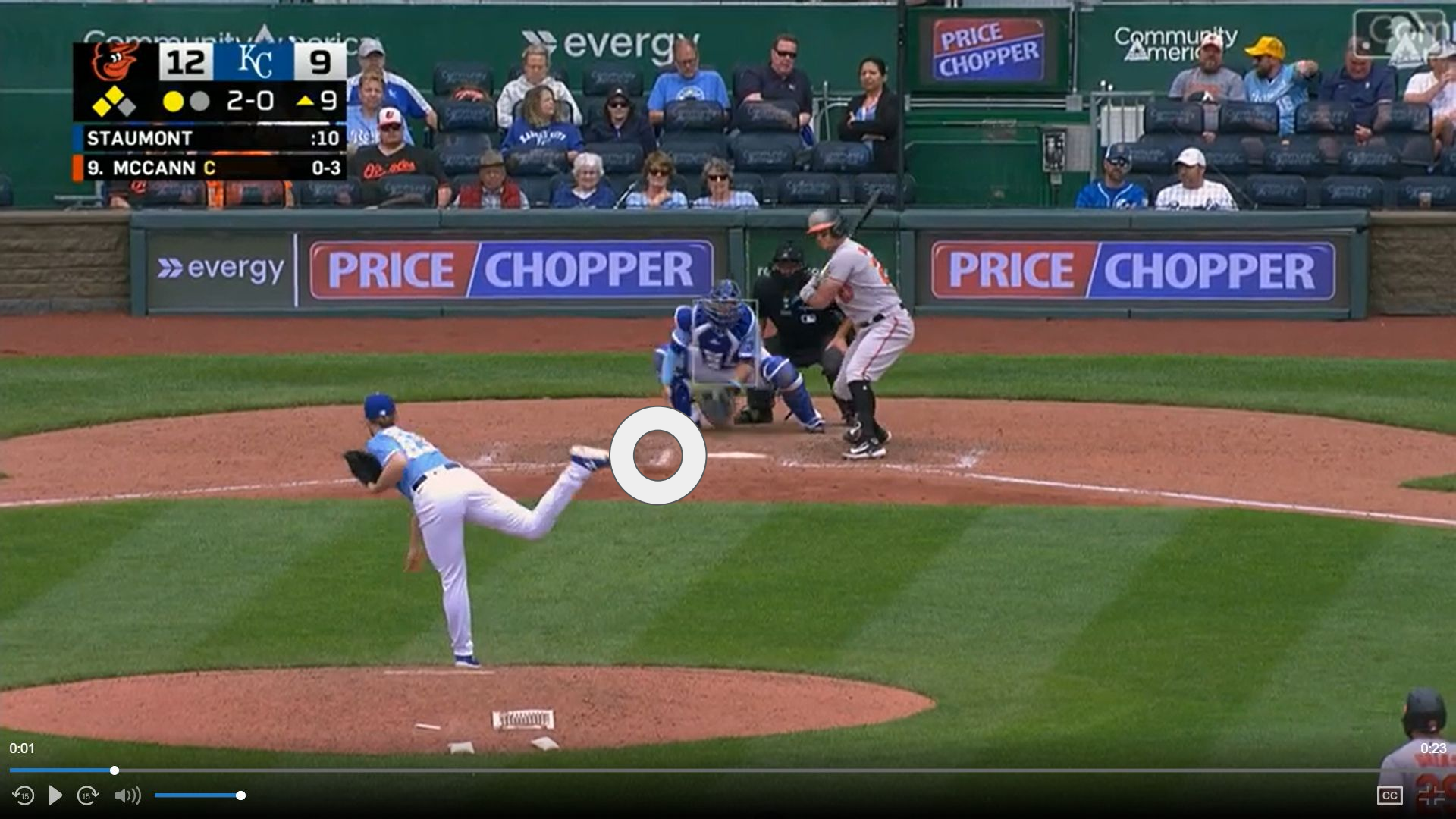
Name	Ops	Blocks	BAX	
Chad Wallach	430	423	-4.23	
Christian Bethancourt	1779	1764	-3.55	
Mike Zunino	1291	1281	-3.37	
Omar Narvaez	614	608	-2.66	
Kyle Higashioka	1521	1510	-2.63	
Travis d'Arnaud	1101	1093	-2.60	
Willson Contreras	2012	1998	-2.06	
Salvador Perez	1868	1853	-1.91	
Bo Naylor	507	503	-1.91	
Freddy Fermin	660	654	-1.90	

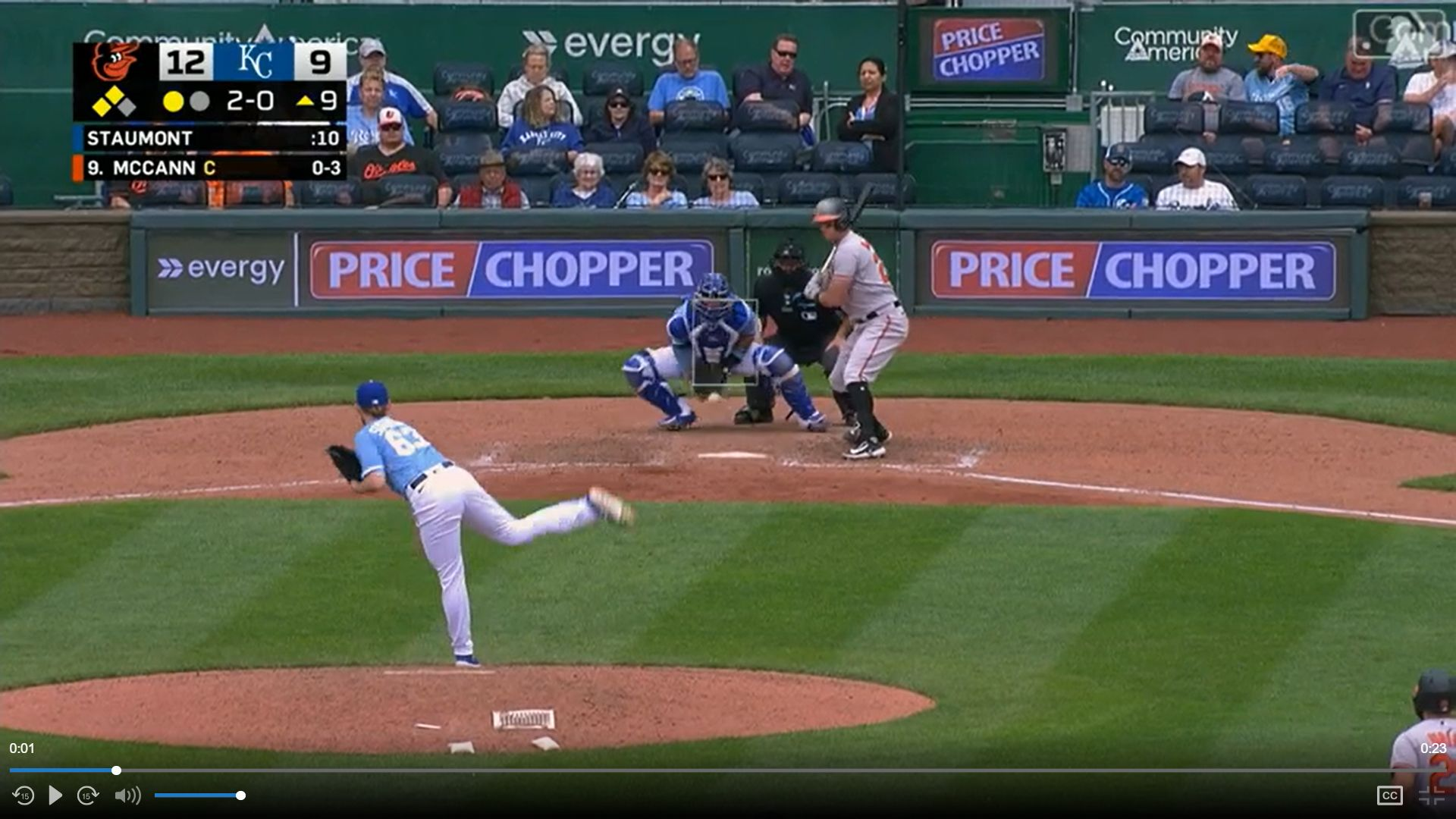
Applications to Player Development

- Freddy Fermin
  - -1.90 BAX
  - Trouble with dirt pitches to throwing side
  - Low and throwing side pitches tough
- 22nd toughest xBlock% out of 80 catchers
- Example Pitch: 92% Block Probability
  - Courtesy of MLB Advanced Media



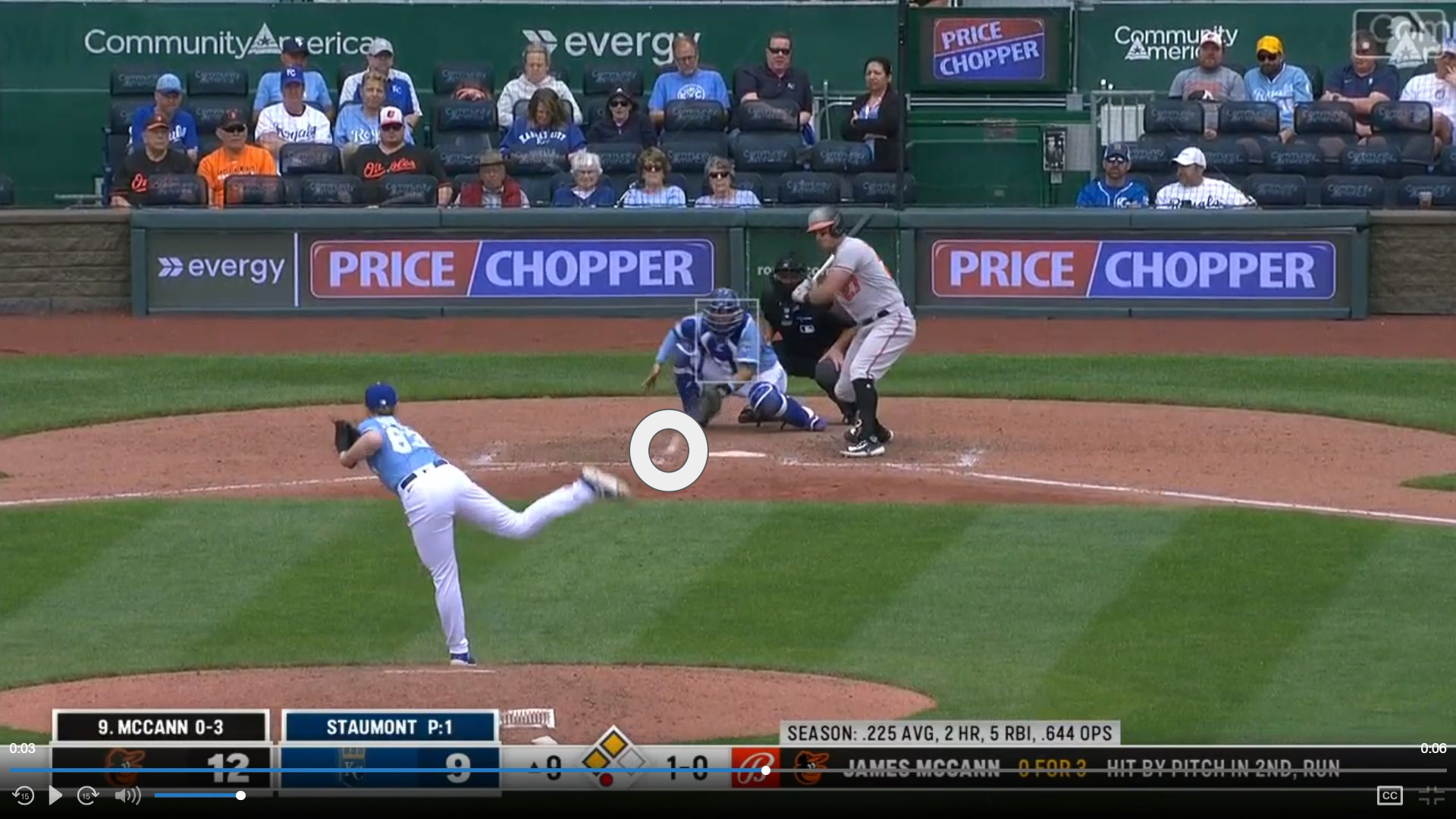


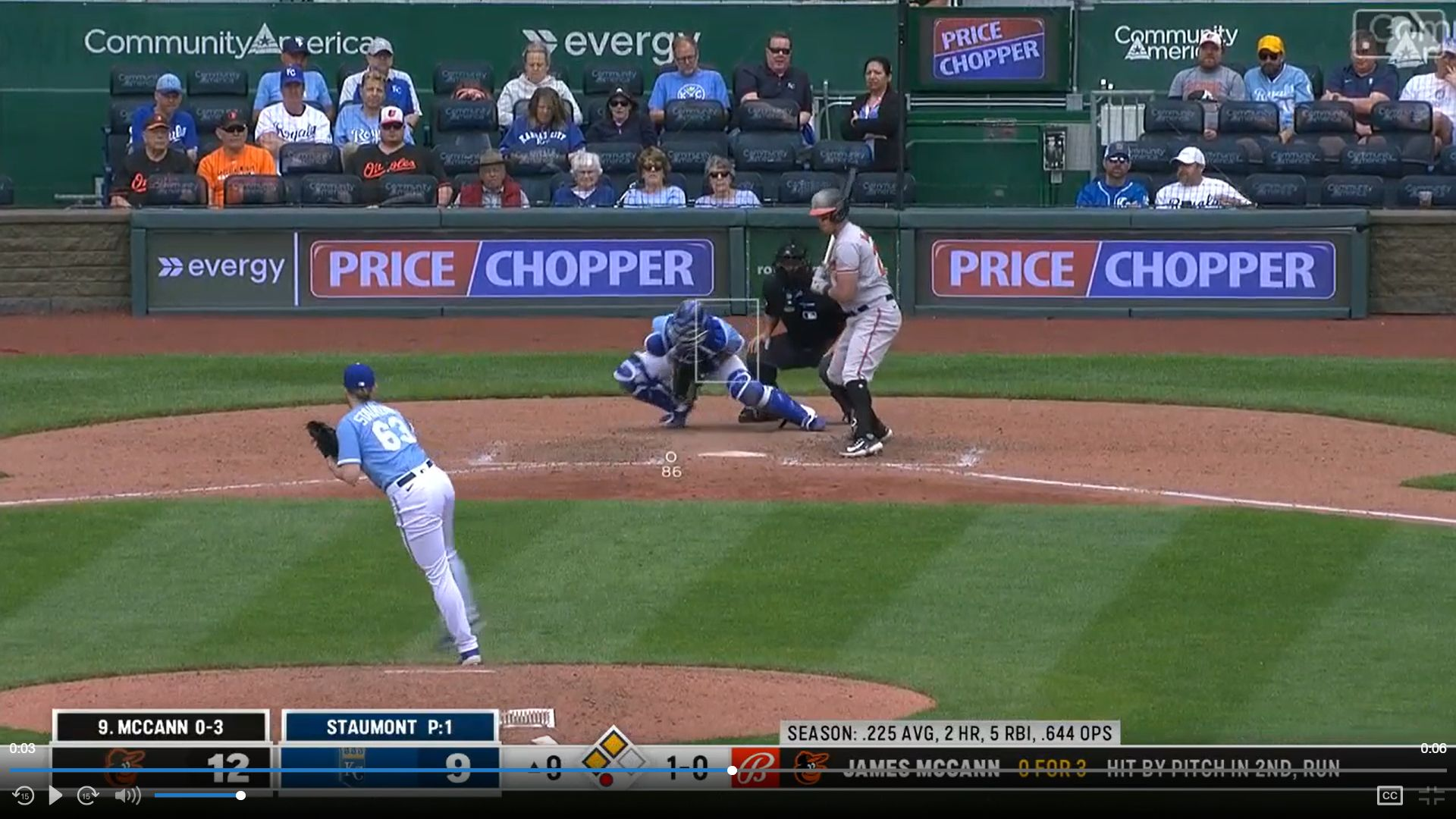

12
KC
9
2-0
9
STAUMONT :10
9. MCCANN C 0-3
evergy
PRICE CHOPPER
PRICE CHOPPER
PRICE CHOPPER
0:01
0:23
CC


12
KC
9
2-0
9
STAUMONT
:10
9. MCCANN C
0-3
evergy
PRICE CHOPPER
PRICE CHOPPER
PRICE CHOPPER
0:01
0:23
CC

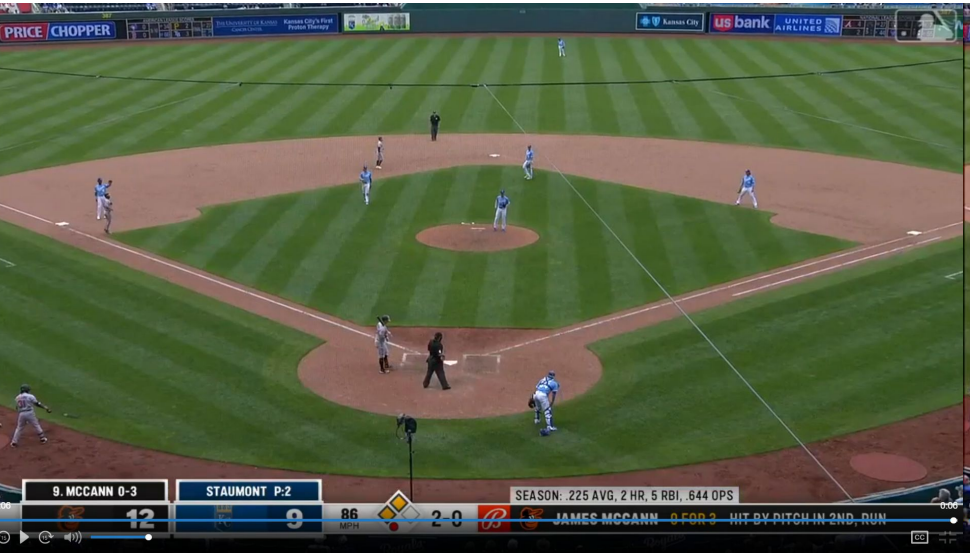
Decent block (the pitch right before!!!)

97% Block Probability


Community America
evergy
PRICE CHOPPER
PRICE CHOPPER
PRICE CHOPPER
9. MCCANN 0-3
STAUMONT P:1
SEASON: .225 AVG, 2 HR, 5 RBI, .644 OPS
0:03
0:06
12
9
0
1-0
JAMES MCCANN 0 FOR 3 HIT BY PITCH IN 2ND, RUN
CC


Community America
evergy
PRICE CHOPPER
PRICE CHOPPER
PRICE CHOPPER
9. MCCANN 0-3
STAUMONT P:1
SEASON: .225 AVG, 2 HR, 5 RBI, .644 OPS
0:03
0:06
12
9
0
1-0
JAMES MCCANN 0 FOR 3 HIT BY PITCH IN 2ND, RUN
CC

“Blocked”



Unblocked

Broader takeaways / outro to questions

- Future
  - Distance away above expected
  - Dashboard
- Not all blocks are created equal
  - Can be kept in front or ball rolls not very far away
- Scorer's discretion and context
  - Runners advancing