

Do batters and pitchers vary in their sensitivities to changes in the quality of the opposition?

Daniel Porter

Columbia University

Objective

To come up with the most accurate possible estimate of a given hitter's expected batting average against a given pitcher, by properly weighting all available data.

1. Overall ability of each player.
2. Head to head stats for each pitcher/hitter combination.

Inspirations

- Stern and Sugano (2006)
This approach uses a beta binomial model to infer the expected batting average of a given batter versus a given pitcher, given their historical track records.
- Bill James (log5)
This formula calculates the expected batting average of a batter versus a pitcher given the batter's batting average and the pitcher's batting average allowed.
I combined elements of each approach.

Approach

- Both Stern and Sugano's and my model starts with a *prior* distribution, which is the batting average we would expect in a given batter/pitcher matchup for two players facing each other for the first time.
- The head to head track record of a batter versus a pitcher receives more weight as the sample size increases.
- The resulting distribution, combining the prior information with the head to head stats, is known as the *posterior* distribution.

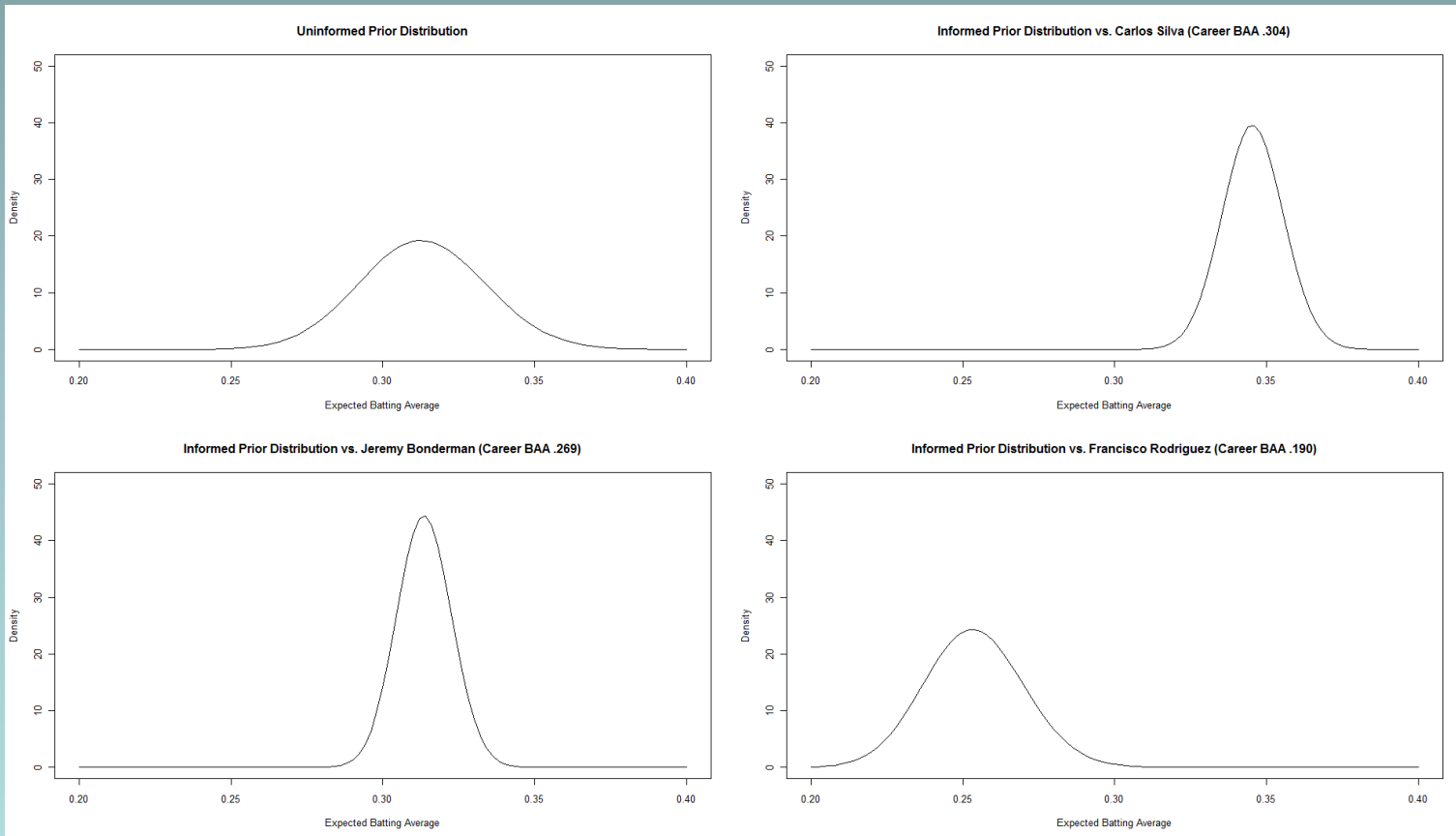
Uninformed versus Informed Prior

The same prior distribution for a given hitter against every possible pitcher.

This particular model that uses an *informed* prior distribution, which accounts for the ability of the opponent.

This informed estimate is made using a process similar to $\log 5$.

Sample Prior Distributions for Derek Jeter



Prior and Posterior Estimates in informed and uninformed models

Pitcher	Rodriguez	Bonderman	Silva
Career BAA	0.190	0.269	0.304
Observed Hits	5	8	7
Observed AB	10	33	15
Uninformed Prior Estimate	0.313	0.313	0.313
Uninformed Posterior Estimate	0.317	0.309	0.318
Informed Prior Estimate	0.254	0.313	0.346
Informed Posterior Estimate	0.257	0.312	0.346

Model Comparisons

- Stern and Sugano emphasized that Batter/Pitcher Specific matchup effects were small.
- In my model, such effects are much smaller!

Comments on head to head

- There is little evidence that performance varies across different batter/pitcher matchups more than would be expected by a simple binomial distribution that accounts for the ability of the hitter and the pitcher.
- Like clutch hitting, it is possible that such effects do exist, but sample sizes are insufficiently large to prove them statistically.

Moving beyond head to head

- So, does this mean that we should just accept log5 and move on?
- Not necessarily. Just because a particular batter isn't expected to perform significantly differently from what log5 expects against a specific pitcher, doesn't mean that there may not be a larger expected difference against a certain *type* of pitcher.

Model to test this phenomenon

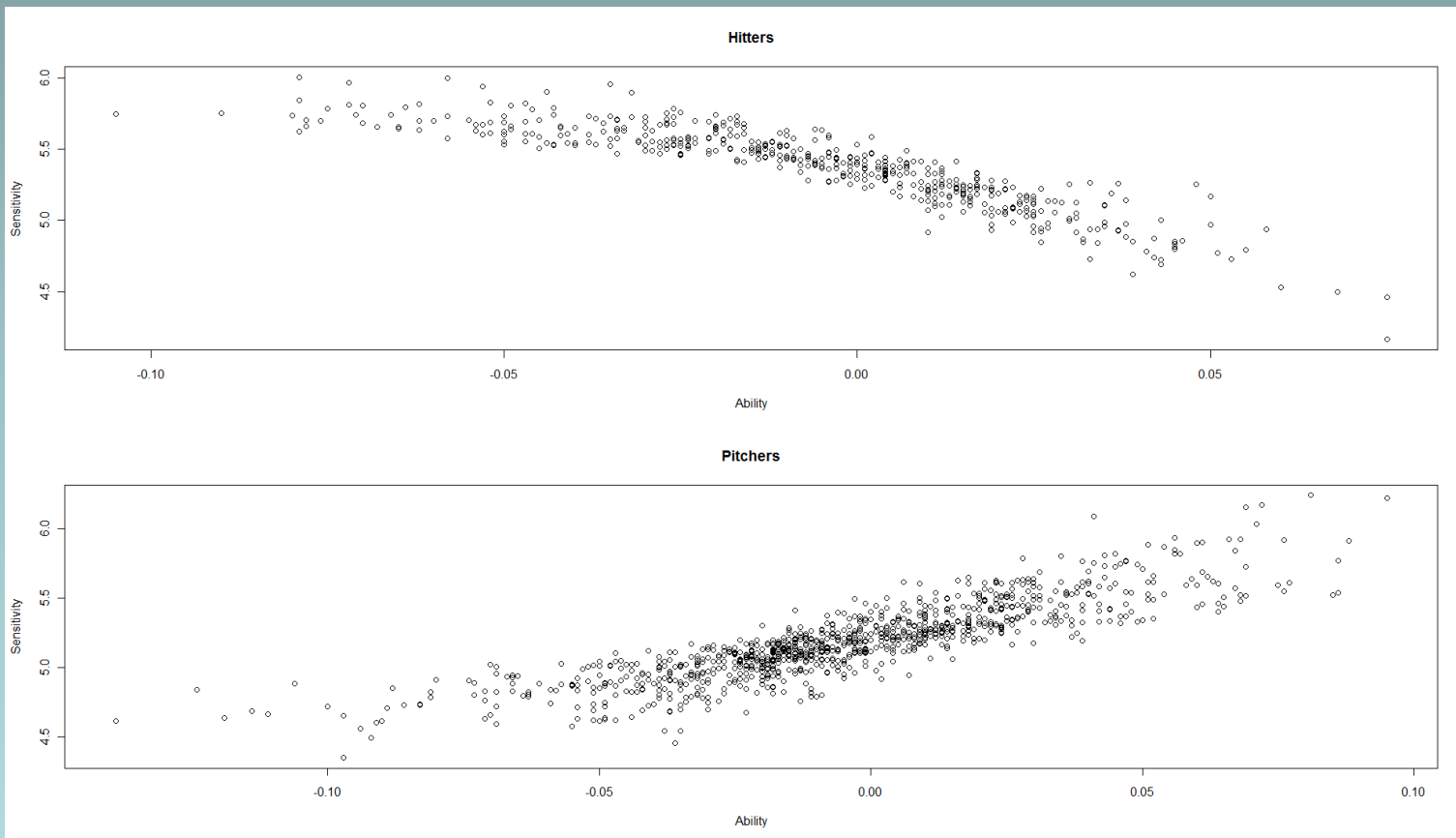
- A varying intercept, varying slope logistic regression can be fit for both pitchers and hitters.
- The slope represents the sensitivity of a player to changes in the quality of the opposition.
- The intercept represents the player's baseline, or his expected performance against an average opponent (the coefficient for opponent's ability is centered around the mean.)

Relationship between intercepts and slopes

We expect the intercepts to vary (if they didn't, we would expect all hitters to have the same batting average against an average hitter/pitcher)!

We are not sure if the slopes will vary

Sensitivity plotted versus ability for hitters/pitchers



Micro Level Answer to Title Question

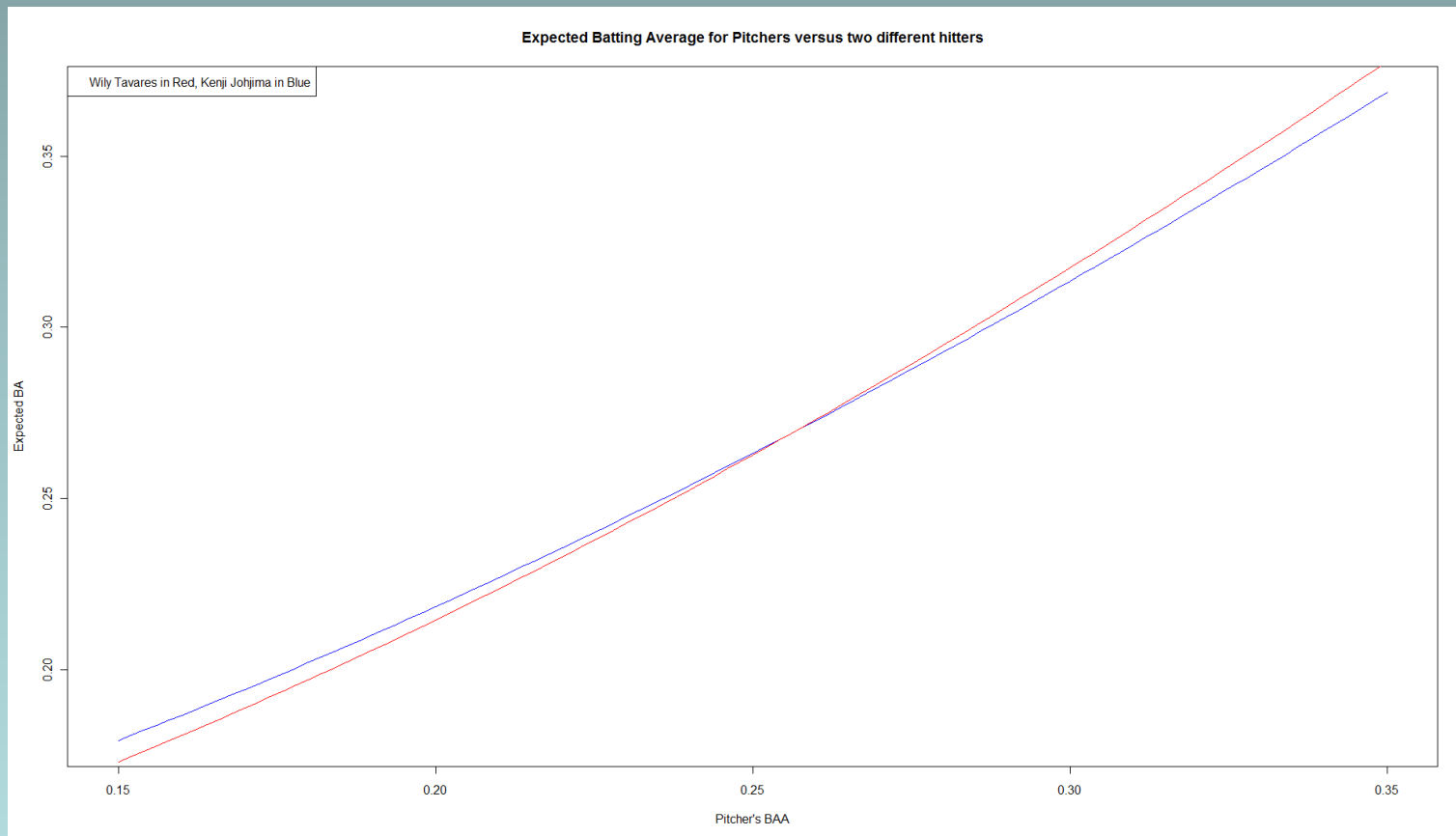
The player who deviates the most from his “Expected Sensitivity,” given his baseline expected BA or BAA.

	Pitcher	Hitter
More Sensitive than expected	Jamie Moyer	Jayson Werth
Less Sensitive than expected	Tim Wakefield	Kenji Johjima

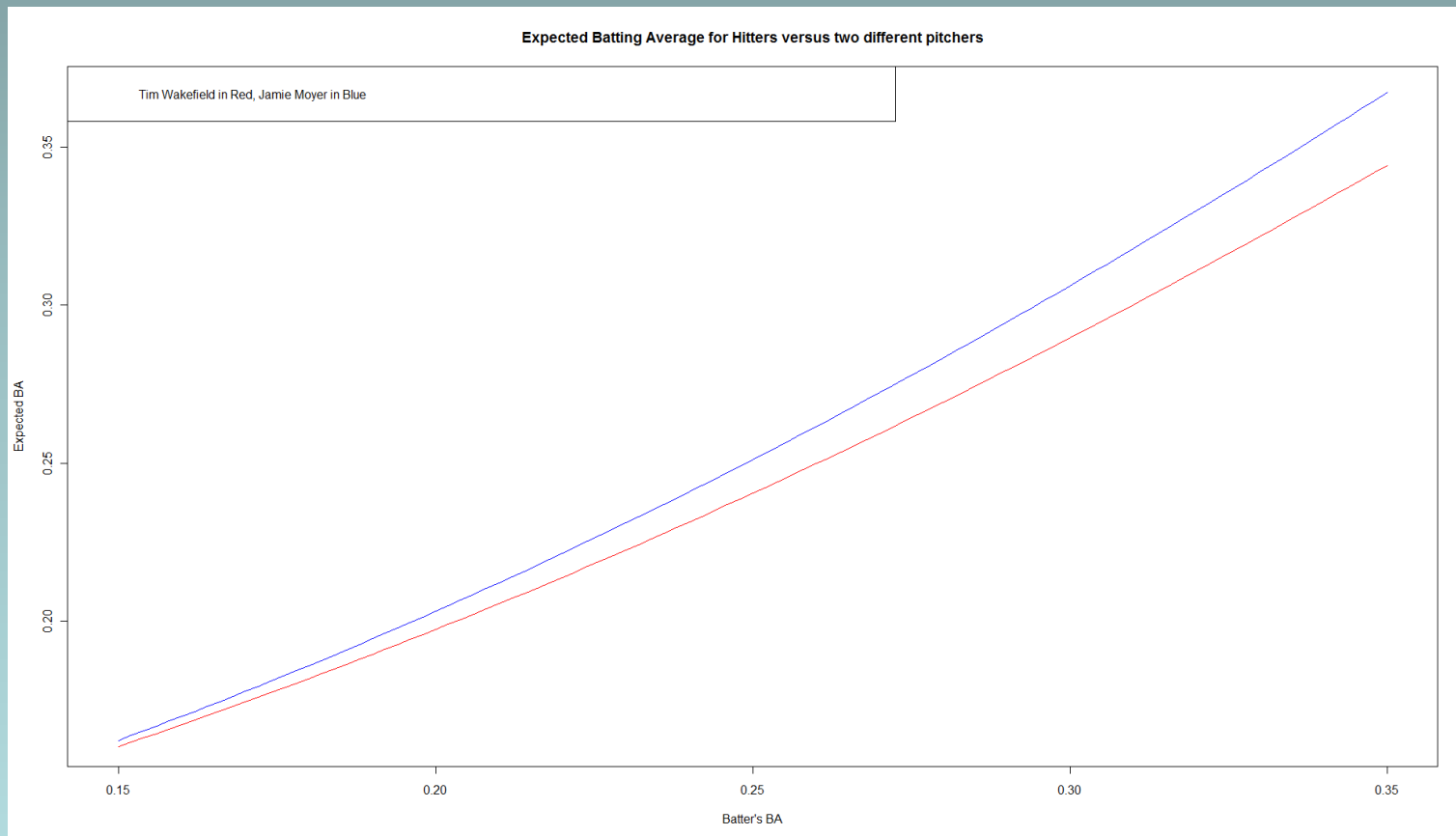
The 5 hitters who are less sensitive than “expected” by the widest margin

1. Kenji Johjima
 2. Victor Martinez
 3. Shea Hillenbrand
 4. Ichiro Suzuki
 5. Hideki Matsui
- All have low strikeout rates
 - Preliminary research indicates that strikeout prone hitters struggle particularly against strikeout pitchers.

Example of expected BA curve for hitters



Example of Expected BA Curve for pitchers.



Further Research Possibilities

- Do historical K rates tell a story for predicting either future K rate or future BA?
- What about moving beyond BA and looking at batter/pitcher specific dynamics of more advanced stats?