Using Big Data Methods to Analyze the Matching Law Relative to Performance for all Members of a Population

David J. Cox1,2, Bryan Klapes3, and John Michael Falligant4

Affiliations:

1Behavioral Health Center of Excellence

2Endicott College

3?????

4Johns Hopkins University School of Medicine

Conflicts of Interest:

None

Funding:

??????

Abstract:

The generalized matching equation (GME) has been used to describe the behavior of individual organisms in operant chambers, artificial environments, and nonlaboratory human settings. Most of these analyses have used a handful of participants with the aim of determining how well the GME describes choice in the experimental arrangement or how some experimental manipulation influences estimated parameters. No known studies have fit the GME to the behavior of all individuals in a population and how those parameters align with performance. This is likely because the population-level data was not available or because time and computational constraints made population-level analyses prohibitive. In this study, we demonstrate how big data methods can be combined with the GME to: (1) identify the events likely to serve as reinforcement; (2) estimate GME parameters; and (3) associate the estimated parameters with often-used metrics of performance. Importantly, 1-3 were accomplished for every one of *NNNN* individuals within the targeted population. The results suggest different individuals were more sensitive to different degrees of environmental change following behavior. [*Something about pitching performance. Something about the main takeaway.*]

Keywords: matching law; big data analytics; baseball;

*< FROM THE PBS WEBSITE >*

Call for Papers: Applied Quantitative Analysis of Behavior

Foundational work in the Quantitative Analysis of Behavior (QAB) has shaped our conceptualization of behavioral processes and provided insightful ways to describe order in our data. More recently, researchers have begun applying these insights and quantitative approaches to issues of societal importance. Examples of this Applied Quantitative Analysis of Behavior (A-QAB) include the use of behavioral economics to understand health behavior, the use of quantitative models to highlight mechanisms of drug action, leveraging the insights of QAB approaches to describe the resurgence of problem behavior in clinical populations, and more.

We invite the submission of scholarly work providing discussion, examples, technical papers, or summaries of these efforts for a special section of Perspectives on Behavior Science on A-QAB.

Submission Deadline: January 1, 2021

**Using Big Data to Analyze the Matching Law for All Members in a Population**

Introduce choice and matching.

Matching in laboratory settings.

Matching in nonlaboratory settings.

General characteristics of matching studies – small *n*, yes GME works, yes in some contexts. Limitations to these approaches. Absurd to think the reinforcers for one person apply to all.

Above limitations exist because of previous constraints on access to data and ability to efficiently analyze.

Big data methodologies solve those limitations.

Why big data methodologies might tell us something new in matching research. Different reinforcers for different people. Analyze sports performance and how GME parameters associate. Analyze how individual fits in a population of behavior/performance ~ game theory, information theory.

Why big data methodologies may help increase the applied relevance of operant quantitative models.

The purpose of this study was to demonstrate how combining big data methodologies with one operant quantitative model may help make the model more practically applicable. Specifically, we sought to demonstrate how big data methodologies allow researchers and practitioners to: (1) identify the likely environmental changes that are controlling individual responding; (2) associate each individual’s GME parameters with overall performance; and (3) leverage information about one individual’s behavior relative to the population to strategize behavior change.

**Methods**

**Data.** The data were obtained from [*website*] by using the available dropdown menus to include all information about the game context, pitch type, pitch characteristics, and pitch outcome for every pitcher in MLB during the 2016-2019 Major League Baseball (MLB) seasons. These seasons were chosen because [*reason*]. All data can be found in the “./MLB Data/Team Data/” folder at Cox (2020a).

**Quantitative Model.** We used the generalized matching equation (GME; Baum, 1974) to describe each pitcher’s allocation of pitches to the different pitch types. The GME can be written as:

. (Equation 1)

In this equation, *Bi* refers to the behavior of interest (i.e., the specific pitch type we might focus on), *Bo* refers to all other behaviors (i.e., all other pitch types thrown by that pitcher), *Ri*and *Ro* refer to the reinforcement contacted by *Bi* and *Bo*, respectively. And, *a* and *b* are estimated free parameters where *a* refers to the organism’s sensitivity to changing reinforcement schedules and *b* refers to the organism’s bias toward *Bi* or toward *Bo* that is not captured by the measured reinforcement schedules. Based on past research by Cox et al. (2017) and Falligant et al. (2020), *Bi* was always the count of pitches per game that were of the fastball variety (e.g., four-seam fastball, two-seam fastball) and *Bo* was always the count of pitches per game that were of any other pitch type (e.g., curveball, slider, eephus).

Past research that has examined matching of pitch type in baseball contexts has used strikes/outs as the reinforcers in Equation 1. However, it is possible that different pitchers are sensitive to more subtle changes in game states than simple strikes and outs. For example, throwing a strike on the first pitch to the first batter in an inning is unlikely to have a large impact on the outcome of the game when it is the first inning and the score is 0-to-0. However, throwing a strike on the first pitch to the first batter in an inning is more likely to impact the outcome of the game when it is the last inning and the pitcher’s team is winning by one run. Stated differently, contacting strikes or outs for throwing a pitch might be more or less meaningful depending on the context of the game in a manner akin to changing the amount of reinforcement delivered following a level press by rats in an operant chamber (e.g., *citations*).

To quantify the importance of different game states, baseball analysts have developed a metric termed run expectancy 24 (RE24). [*Description of* *RE24*]. RE24 only accounts for the position of runners on the bases and the number of outs recorded to that point in the inning. It is possible that pitchers are sensitive to changing counts on the batter (number of balls and strikes). Combining the logic of RE24 with the addition of count, we can create RE288 to refer to the unique 288 possible combinations of position of runners on the bases, number of outs recorded to that point in the inning, and the current count on the hitter.

For the present analysis, we wanted to determine whether pitchers independently differed in their overall sensitivity of different outcomes following their pitches. Thus, for every pitcher, we fit the GME once using strikes/outs as the assumed reinforcer to replicate past research, once using change in RE24 as the assumed reinforcer, and once using change in RE288 as the assumed reinforcer. Here, the assumption is that the assumed reinforcer that us likely to be maintaining a pitcher’s individual responding is identified by a better fit of the GME to the pitcher’s data (i.e., *R*2 closer to 1.0).

**Calculating Different Reinforcers**. To use strikes/outs as the assumed reinforcers, we counted the number of times per game that each pitch type resulted in a called or swinging strike, or the number of times that each pitch type was put into play and resulted in an out. The number of strikes/outs contacted in a game following a fastball were summed to create *Ri*, and the number of strikes/outs contacted in a game following all other pitches were summed to create *Ro*.

To use RE24 as the assumed reinforcers, we [*where did we get this again?*]. To quantify the “amount” of assumed reinforcer following a pitch, we subtracted the RE24 value after a pitch was thrown from the RE24 value before the pitch was thrown. Thus, a positive RE24 change score reflected an improvement in the game context (potential reinforcer), a negative RE24 change score reflected a worsening of the game context (potential punisher), and a change score of zero reflected no change in the game context. Because Equation 1 only includes reinforcement, only positive RE24 change scores were summed to quantify *Ri* and *Ro* using RE24 as the assumed reinforcer. *Ri* was the summed positive RE24 change scores per game following all fastballs thrown. And, *Ro* was the summed RE24 change scores per game following all other pitches.

No known RE288 tables were available for this analysis. As a result, we combined all the data from the 2016-2019 seasons into a single dataframe consisting of 2,891,247observations. We then labeled every pitch with a corresponding game state number between 1 and 288 (see *Appendix A* for game state labels). Once labeled, we calculated the number of runs that were scored after each pitch was thrown up until the end of that inning. Finally, we calculated the average number of runs scored until the end of an inning for each of the 288 different game states which became the RE288 value used for the analyses in this study. To quantify the “amount” of assumed reinforcer following a pitch, we subtracted the RE288 value after a pitch was thrown from the RE24 value before the pitch was thrown. Similar to RE24, a positive RE288 change score reflected an improvement in the game context (potential reinforcer), a negative RE288 change score reflected a worsening in the game context (potential punisher), and a change score of zero reflected no change in game context. As with RE24, only positive RE288 change scores per game were summed to quantify *Ri* and *Ro* for fastballs and all other pitches, respectively.

Python notebooks containing the code used to create RE288 are located in the “Python\_Scripts\_Notebooks” folder at Cox (2020). The final dataframe that contains all data related to the game context, pitch type and characteristics, pitch outcomes, RE24 change score, RE288, and RE288 change score can be found at Cox (2020b). Readers interested in obtaining these data can contact the corresponding author.

**GME Scripts and Computational Environment.** All programming was completed using the Python (version 3.7) programming language and run in Google Colaboratory (Colab). Colab is a free product from Google Research that allows anyone to write and execute Python code through any Internet browser. Colab is a hosted Jupyter notebook service that provides the programmer with access to GPU or TPU computing resources which allows him or her to conduct computational analyses beyond what is possible on most personal computers due time available hardware or the time it would take to run on a personal computer. The notebooks for this project were downloaded and are available at Cox (2020a).

For each fit of the GME, sensitivity (*a,* Eq. 1) and bias (*b*, Eq. 1) were fit using the *curve\_fit* function from the SciPy Python package (Millman & Aivazis, 2011). *R*2 values were calculated using the *r2\_score* function from the Sklearn Python package (Pedregosa et al., 2011). All statistical analyses were conducted using the corresponding statistical function from the Statsmodels Python package (Seabold & Perktold, 2010).

**Results**

**Overall GME Fits.** Highlight the plots we’ve seen where, overall, strikes and outs seem to be doing better than RE24 or RE288.

**Correspondence of GME Parameters with Performance.** Words.

**GME Parameters by Starter vs. Reliever**. Words.

**Identifying Likely Reinforcer.** Words.

**Population Parameters and Pitching Performance.** Words.

**Leveraging Reinforcers and Location of Individual within Population Distributions.** Words.

**Discussion**

Reminder of what we did and main findings.

**Applicability of Using Big Data Methods to Identify Population Trends**. Limited resources, practitioners might be unable to tailor matching analyses to every individual for whom their analyzing data. Big data methodologies allow us to identify what is going on at the population level so – if we have to choose one approach because of constraints – at least we’ll know what will work for the most people.

Correspondence of each approach with performance data. Possible that what is a reinforcer for one person deviates from the reinforcers for most people. For sporting performance, is this good? Do we need to work on change what is reinforcing for someone’s behavior so that they perform better? Or, is it beneficial to be variable?

**Applicability of Using Computational Techniques to Identify Individual Trends**. Significant time and effort to analyze these data for any single player and for any single reinforcer. Computational techniques allow us to iterate rapidly over many hypothesized behavior-environment relationships to identify which combination most closely matches what we know from the basic operant literature (i.e., that the GME should fit).

Where in each distribution does this individual fit in terms of bias, sensitivity, likely reinforcer, and performance metrics? How can we leverage that information to inform feedback and game strategy?

**Conclusion**

Words.

References

Alferink, L. A., Critchfield, T. S., Hitt, J. L., & Higgins, W. J. (2009). Generality of the matching law as a descriptor of shot selection in basketball. *Journal of Applied Behavior Analysis, 42*, 595–608. https://doi.org/10.1901/jaba.2009.42-595

Baum, W. M. (1974). On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior, 22*, 231–242. https://doi.org/10.1901/jeab.1974.22-231

Cox, D. J. (2020a). MLB Data. Man-vs-Machine Github repository [data sets]. https://doi.org/10.5281/zenodo.4088070

Cox, D. J. (2020b). MLB Data for GME. Open source public dataset. https://doi.org/

Cox, D. J., Sosine, J., & Dallery, J. (2017). Application of the matching law to pitch selection in professional baseball. *Journal of Applied Behavior Analysis, 50*(2), 393-406. http://doi.org/10.1002/jaba.381

Critchfield, T. S., Meeks, E., & Stilling, S. T. (2014). Explanatory flexibility of the matching law: Situational bias interactions in football play selection. *The Psychological Record, 64*(3), 371-380. https://doi.org/10.1007/s40732-014-0064-5

Falligant, J. M., Boomhower, S. R., & Pence, S. T. (2016). Application of the generalized matching law to point-after-touchdown conversions and kicker selection in college football. *Psychology of Sport and Exercise, 26*, 149-153. https://doi.org/10.1016/j.psychsport.2016.07.006

Falligant, J. M., Cero, I., Kranak, M. P., & Kurtz, P. F. (2020). Further application of the generalized matching law to multialternative sports contexts. *Journal of Applied Behavior Analysis.* https://doi.org/10.1002/jaba.757

Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior, 13*, 243–266. doi:10.1901/jeab.1970.13-243

Millman, K. J., & Aivazis, M. (2011). Python for scientists and engineers. *Computing in Science & Engineering, 13*, 9-12.

Pedregosa, F., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning, 12*, 2825-2830.

Poling, A., Weeden, M. A., Redner, R., & Foster, T. M. (2011). Switch Hitting in Baseball: Apparent Rule‐Following, Not Matching. *Journal of the Experimental Analysis of Behavior, 96*(2), 283-289. http://doi.org/10.1901/jeab.2011.96-283

Reed, D. D., Critchfield, T. S., & Martens, B. K. (2006). The generalized matching law in elite sport competition: Football play calling as operant choice. *Journal of Applied Behavior Analysis, 39*, 281–297. https://doi.org/10.1901/jaba.2006.146-05

Romanowich, P., Bourret, J., & Vollmer, T. R. (2007). Further analysis of the matching law to describe two and three-point shot allocation by professional basketball players. *Journal of Applied Behavior Analysis, 40*, 311–315. https://doi.org/10.1901/jaba.2007.119-05

Schenk, M. J., & Reed, D. D. (2020). Experimental evaluation of matching via a commercially available basketball video game. *Journal of Applied Behavior Analysis, 53*(1), 209-221. https://doi.org/10.1002/jaba.551

Seabold, Sk., & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. *Proceedings of the 9th Python in Science Conference.* https://conference.scipy.org/proceedings/scipy2010/pdfs/seabold.pdf

Seniuk, H. A., Williams, L. W., Reed, D. D., & Wright, J. W. (2015). An examination of matching with multiple response alternatives in professional hockey. *Behavior Analysis: Research and Practice, 15*(3-4), 152–160. http://dx.doi.org/10.1037/bar0000019

Seniuk, H.A., Vu, J.P. and Nosik, M.R. (2020), Application of the matching law to Mixed Martial Arts. *Journal of Applied Behavior Analysis, 53*. 846-856. https://doi.org/10.1002/jaba.653

Stilling, S. T., & Critchfield, T. S. (2010). The matching relation and situation-specific bias modulation in professional football play selection. *Journal of the Experimental Analysis of Behavior, 93*(3), 435–454. https://doi.org/10.1901/jeab.2010.93-435

Vollmer, T. R., & Bourret, J. (2000). An application of the matching law to evaluate the allocation of two and three-point shots by college basketball players. *Journal of Applied Behavior Analysis, 33*, 137–150. https://doi.org/10.1901/jaba.2000.33-137