Do fans' and media opinions influence NBA player contracts relative to performance?

Teo Pazera `s1159164`

October 2025

Research idea summary

Public opinion may influence NBA executives' contract allocations. As the NBA is primarily an entertainment industry, popularity and narrative may affect a player's market value beyond on-court contributions. This project tests whether pre-free-agency sentiment impacts players' market value.

The research question is: "Do pre-free-agency fan and media sentiments about NBA players influence their contracts relative to statistical performance?" Two subquestions are explored: "Is pre-signing sentiment predictive of being paid above expected value given past performance?" and "Do fans' sentiments change post-signing (e.g., toward perceived 'overpaid' players)?"

Fan opinions will be sourced from Reddit (r/NBA, r/nbadiscussion, r/NBAtalk) and media articles (ESPN, The Athletic, Bleacher Report) in the pre-free-agency window (May–August of the examined offseason) using a PRAW crawler for Reddit and a traditional crawler for media sites. Contract and transaction records, paired with player performance metric win shares (measuring overall contribution), will be gathered from Basketball-Reference.com. The analysis targets players signing in the examined offseason; Named Entity Recognition (NER) will identify relevant player mentions, and statements will be separated to map to just one player to make the analysis sensible .

Sentiment scoring will combine rule-based (VADER) and transformer-based classifiers, validated on a hand-labeled sample. A performance-to-pay ratio (contract value per win share) will be computed, and regression will model the relation between pre-signing sentiment and contract overpayment. Evaluation will include regression effect sizes and statistical significance of fan, media, or combined sentiment.

Top related references

1. Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up? Sentiment Classification using Machine Learning Techniques. In Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP 2002). Association for Computational Linguistics, 79–86. https://doi.org/10.3115/1118693.1118704

2. Guillaume Lample, Miguel Ballesteros, Sandeep Subramanian, Kazuya Kawakami, and Chris Dyer. 2016. Neural Architectures for Named Entity Recognition. In Proceedings of the 2016

Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL '16). Association for Computational Linguistics, San Diego, California, 260–270. https://doi.org/10.18653/v1/N16-1030

3. Hutto, C. and Gilbert, E. 2014. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. Proceedings of the International AAAI Conference on Web and Social Media. 8, 1 (May 2014), 216-225. DOI:https://doi.org/10.1609/icwsm.v8i1.14550

4. Qiwen Li, Jiarui Zhang, Jiayu Guo, Jiaqing Li, and Chenhao Kang. 2022. Evaluating Performance of NBA Players with Sentiment Analysis on Twitter Messages. In Proceedings of the 2021 European Symposium on Software Engineering (ESSE '21). Association for Computing Machinery, New York, NY, USA, 150–155. https://doi.org/10.1145/3501774.3501796

5. Rohan Tummala. 2023. Analyzing NBA Player Sentiment through Reddit. Samford University Sports Analytics. `https://www.samford.edu/sports-analytics/fans/2023/Analyzing-NBA`