

ポスターセッション | ポスターセッション | ポスターセッション (2)

[2PA] ポスターセッション (2)

2024年3月8日(金) 11:00 ~ 12:00 ポスター発表 (デザインコモン2階)

[2PA-13] メディアによるオリンピック成績の予測と注目度

Predictions and Attention to Olympic Results by the Media

*大井 健太郎¹、風戸 真帆¹、福本 わかば¹、星野 雄介¹ (1. 武蔵野大学 経営学部)

*Kentaro Ooi¹, Maho Kazato¹, Wakaba Fukumoto¹, Yusuke Hoshino¹ (1. Musashino University, Faculty of Business Administration)

メディアによるオリンピック成績の予測と注目度 —情報量, 経路依存性, ニュースバリューに注目して—

大井健太郎*, 風戸真帆*, 福本わかば*, 星野雄介*

*武蔵野大学経営学部

Predictions and Attention to Olympic Results by the Media

—Information Volume, Path Dependence, and News Value —

Kentaro OOI*, Maho Kazato*, Wakaba FUKUMOTO* and Yusuke HOSHINO*

*Musashino university, Faculty of Business Administration

yhoshino@musashino-u.ac.jp

Abstract: The purpose of this paper is to conduct an exploratory analysis of (1) whether there is a correlation between media attention and outcomes, and (2) what factors change media attention, using the 2016 Rio Olympics athletes as a case study. We focus on the information volume hypothesis as the cause of the media's constant predictive accuracy, and the path dependence/sunk cost and news value hypotheses as the causes of changing media coverage. As a result of analyzing newspaper articles mentioning the 144 athletes participating in the Rio Olympics, We found that (1) the media's advance publicity has a certain predictive accuracy, (2) the number of articles prior to the event affects the number of articles after the event. (3) The news value hypothesis was partially supported by our sample.

Keywords: News, Rio Olympic, Prediction, Media Coverage

1. イントロダクション

本論文の目的は、①メディアの注目度と結果に相関があるのか、②メディアの注目度がどのような要因で変化するかについて、2016年に開催されたリオ・オリンピック出場選手を事例に探索的に分析していくことである。

ソーシャルネットワークが普及したとしても、メディアは依然として大きな影響力を持っている。メディアによる報道の特徴は早さ・量の多さ・質の高さといえる。それゆえ、新聞の発行部数が減少していたとしても、ニュースサイトなどを通じて、私たちに影響を与え続けている。

それゆえ、メディア報道によって何らかの将来を予測ができるのではないかと、という試みの研究は非常に多い。例えば、新聞記事内容と株価との関係について、数十年にわたり、一定の予測可能性がありうることが指摘されている[1-3]。

他方で、メディアがどのテーマをどれくらい重視して取り上げるかは、テーマによって違いがあると同時に、時間経過に伴って増減する。この変化を説明する理論の1つには、メディアによるフレーミング、アジェンダ・セッティング、プライミングといったメディアのスタンスがあげられる[4]。しかしながら、報道量の増減にはもっとシンプルなニュースバリューや経路依存性[5]といった理由も考えられる。

そこで本論文では2016年に開催されたリオ・オリンピックの出場選手の一部を事例に、メディアが成績を予想できるのか、そして事前の記事数と成績が事後の記事数に影響を与え

るのかを分析していく。

2. 仮説

2.1 情報量仮説

メディアは取材を通して、情報を収集して記事を書くことから、単なる読者よりも判断に関する情報を多く持っている。実際、時折、メディア社員によるインサイダー取引が発覚し、処分されている[6-7]。このように、新聞社は情報を多く持っていることから、新聞社の予測精度は高いかもしれない。実際、新聞記事と株価については一定の予測可能性が指摘されている[1-3]。したがって、多くの情報を持っている場合、オリンピックにおいても予測精度が高まると考えられる。これはメディアの予測可能性に関する情報量仮説といえよう。ただし、メディア報道が選手を勝たせるという因果関係ではなく、メディアの予測可能性を示す相関関係であろう。以上より、仮説1が導出される。

仮説1: 事前の記事数と、オリンピックの成績が正の相関をする

2.2 経路依存性・サンクコスト仮説

社会には非合理的な慣行が残りが残ることがある。その原因を説明する代表的な理論として経路依存性とサンクコストが考えられる。経路依存性とは過去の選択により現在の意思決定が制約される現象を指し、パソコンのキーボード

「QWERTY」配列が有名である[5]. サンクコスト (埋没費用) とは投資, 生産, 消費などの経済行為に投じた固定費のうち, その経済行為を途中で中止にしたとしても, 回収できない費用である. サンクコストであることを理解していても, これまでに投資した分を惜しんで投資を継続してしまう心理的傾向を, サンクコスト効果あるいは, 有名な事例の名を冠してコンコルド効果と呼ばれる[8].

仮にメディアが注目していた選手がオリンピックで活躍できなかったとしても, 事前にメディアがその選手に注目していた事実は変わらない. そのため, 事前の記事数が多かった場合, メディアは経路依存的に, あるいはサンクコスト効果の結果として, 事後の記事数を増加させる可能性がある. 以上より, 仮説2が導出される.

仮説2: 事前の記事数が増加すると, 事後の記事数が増加する

2.3 ニュースバリュー仮説

新聞社も民間企業であるため, 売上や利益が重要である. よって, 新聞社はニュースバリューが高そうなトピックを取り上げる. オリンピックでの好成績は十分にニュースバリューが高いため, オリンピックの好成績は, その選手のその後の記事数を増加させることにつながるだろう. 以上より, 仮説3が導出される.

仮説3: オリンピックの成績が高いと, 事後の記事数が増加する

3. 方法

3.1 サンプル

2016年リオ・オリンピックの内, 10競技 (陸上, 卓球, 競泳, 飛び込み, テニス, 体操, 新体操, ウエイトリフティング, ボート, レスリング) とした. この10競技の参加選手をJOCのHPより取得した[9]. 参加選手は, 合計144名である (表1).

競技	人数	記事数
ウエイトリフティング	7	96
テニス	6	2321
ボート	4	7
レスリング	10	1218
新体操	6	56
水泳・競泳	36	1809
水泳・飛び込み	3	106
体操競技	10	1371
卓球	10	1509
陸上競技	52	1882
総計	144	10375

表1. 競技ごと選手数とデータ件数

注: ただし, 1つの記事で複数の選手が取り上げられていた場合は, その選手の数でカウントしている.

3.2 変数とデータ

仮説1の被説明変数: 最高成績. 最高成績はJOCのHP[9]より選手ごとに取得した. 複数の競技に参加していた場合は, 最も良い成績を用いた. 最高の成績が最も人々の記憶に残ると考えられたためである. 成績は, 金メダルを100, 棄権を0として, 筆者らによってポイント付けされ, そのポイントを変数として用いた (表2).

成績	ポイント
金メダル	100
銀メダル	90
銅メダル	80
入賞	70
決勝進出	60
準決勝進出	50
準々決勝進出	40
予選敗退	30
棄権	0

表2. 最高成績のポイント

仮説1の説明変数: 事前の記事数. 事前の記事数は日本経済新聞の記事数を対象とした. 記事取得期間はリオ・オリンピック開幕3年前の2013年1月1日から, 3年後の2019年12月31日とした. データ期間は, 以下のように分類される.

- ・大会前: 2013年1月1日-2016年8月5日
- ・大会中: 2016年8月6日-2016年8月22日
- ・大会後: 2016年8月23日-2019年12月31日

データを取得するにあたり, 同姓同名の人物の, 企業における人事情報などは目視で削除した. しかし, スポーツに直接関係しない記事は残している.

被説明変数としては, 大会前の3年分の数字について, 2016年を0時点としたうえで, ①0年から開幕まで, 1年前から開幕まで, 2年前から開幕まで, 3年前から開幕までの「累計記事数」を取得した. しかしながら本論文では, 紙面の都合で「1年前から開幕まで」のみを事前の記事数として用いることとする. また, ②累計の影響をコントロールするために, 0年から開幕まで, 2015年 (開幕1年前), 2014年 (開幕2年前), 2013年 (開幕3年前) の「各年記事数」を取得した.

仮説2の被説明変数: 事後の記事数. 上記の大会後の記事数のうち, 2016年を0時点としたうえで, 2016年 (0年後), 2016-17年 (1年後), 2016-18年 (2年後), 2016-19年 (3年後) の4種類の累計記事数を被説明変数とした.

仮説2の説明変数: 事前の記事数. 上記の仮説1の2種類の事前記事数のうち, 2015年から開幕までの1年半の「累計記事数」を採用した. この期間は多くの競技でオリンピック選手の最終選考会が開催されており, オリンピックへの期待が高まる時期だと考えられるためである.

仮説3の被説明変数: 事後の記事数. 上記と同じ.

仮説3の説明変数: 最高成績. 上記と同じ.

コントロール変数として, 生まれ年と競技ダミーを置いた.

生まれ年を置いた理由は、若くしてオリンピックに出場するということが、ニュースバリューを持つと考えられたためである。生まれ年では年をそのまま用いたため、この数字が大きくなるにつれて、選手は若くなる。

3.3 分析

分析は重回帰分析で、statsmodelsを用いた。

4. 結果

4.1 仮説1：事前記事量と最高成績

表3は、仮説1のうち説明変数を「2015年から開幕まで」の累計記事数とした場合の結果である。「2015年から開幕まで」の累計記事数と最高成績との関連は統計的に有意とは言えなかった。表4は、仮説1のうち説明変数を各年記事数とした場合の結果である。モデル3-4ともに大会前0年目の記事数が最高成績と正の相関をしていた ($p<0.01$)。以上から、ロバストとは言えないものの、仮説1については部分的に支持されたといえそうである。

なお、ダミー変数については(モデル4)、生まれ年については負の係数で統計的に有意であった。これはやはりベテランほど最高成績が高くなることを示している。また、新体操ダミーについても正の係数で有意が出ている。これは何らかの理由で、メディアが新体操に注目していたためであろう。

	モデル1	モデル2
被説明変数	最高成績	最高成績
補正R2	-0.006	0.209

説明変数	係数	係数
大会前_1	0.0135	0.0282
生年月日		-0.6472 †
競泳ダミー		20.0315 †
ウエイトリフティングダミー		-13.5207
新体操ダミー		23.4711 †
テニスダミー		-7.6183
ボートダミー		15.6987
陸上ダミー		-0.4739
体操ダミー		11.1236
卓球ダミー		9.6605
レスリングダミー		14.7405

表3. 仮説1 (説明変数：累計記事数) の分析結果

注：† : $p<0.1$, * : $p<0.05$, ** : $p<0.01$. 以下すべてで共通。

	モデル3	モデル4
被説明変数	最高成績	最高成績
補正R2	0.091	0.307

説明変数	係数	係数
大会前_0年目のみ	0.8082 **	0.847 **
大会前_1年目のみ	-0.1994	-0.0967
大会前_2年目のみ	-0.3779	-0.4542
大会前_3年目のみ	0.1929	0.1662
生年月日		-0.675 *
競泳ダミー		18.8624
ウエイトリフティングダミー		-11.7903
新体操ダミー		25.8571 *
テニスダミー		-11.3517
ボートダミー		18.2751
陸上ダミー		-0.8474
体操ダミー		6.0793
卓球ダミー		3.1484
レスリングダミー		8.2081

表4. 仮説1 (説明変数：各年記事数) の分析結果

4.2 仮説2：事前記事量と事後記事数、仮説3：最高成績と事後記事数

表5は、仮説2-3のうち説明変数を「2015年から開幕まで」の累計記事数、被説明変数を大会後0-3年へと変化させた場合(モデル5-8)の回帰分析結果である。被説明変数が大会後0-3年のいずれの場合も、大会前1年間の記事数の記事数は正で統計的に有意($p<0.01$)であった。それゆえ、仮説2は支持されたと考えられる。

	モデル5	モデル6	モデル7	モデル8
被説明変数	大会後_0	大会後_1	大会後_2	大会後_3
補正R2	0.628	0.804	0.823	0.819

説明変数	係数	係数	係数	係数
大会前_1	0.1695 **	0.8763 **	0.8763 **	1.2519 **
最高成績	0.0106	0.2803 †	0.2803 *	0.4494 **
生年月日	0.0474	1.3493 *	1.3493 **	2.2041 **
競泳ダミー	-6.0447	-7.355	-7.355	-6.9777
ウエイトリフティングダミー	-5.0648	-0.7922	-0.7922	2.7884
新体操ダミー	-7.3702	-15.8839	-15.8839	-21.077
テニスダミー	-5.4672	28.462	28.462	31.6955
ボートダミー	-7.4474	-8.5318	-8.5318	-7.5809
陸上ダミー	-6.0901	-0.2793	-0.2793	4.9029
体操ダミー	-4.6527	0.3121	0.3121	-3.5473
卓球ダミー	-5.8965	6.6547	6.6547	20.7488
レスリングダミー	18.6107 **	27.0244 *	27.0244 †	33.2117

表5. 仮説2-3の分析結果

仮説3の大会最高成績は、被説明変数が大会後0-1年の場合(モデル5-6)は統計的に有意ではなかったが、大会後2-3年の場合(モデル7-8)、正の係数で統計的に有意であった(それぞれ $p<0.05$, $p<0.01$)。このことから、最高成績は期間が離れるほど影響力を持つといえ、仮説3は部分的に支持されたと考えられる。

ダミー変数については生まれ年が正の係数で統計的に有意

であった(モデル6: $p < 0.05$, モデル8: $p < 0.01$)。この結果は、若い選手ほど記事量に得供することから、ニュースバリュー仮説が成立すると思われる。

5. まとめ

本論文ではメディアが一定の予測精度を持つ原因として情報量仮説、報道量を変化させる原因として経路依存性・サンクコスト仮説およびニュースバリュー仮説に注目した。

リオ・オリンピック出場選手144名について言及した新聞記事を分析した結果、ある選手の直前の記事数が最高成績と正の相関を示すことが部分的に説明できた。メディアの前評判は一定の予測精度を持つといえる。これは株価予測に関する既存研究の成果とも整合的である[1-3]。

また、事前の記事数は事後の記事数に影響していることが分かった。このことから、経路依存性・サンクコスト仮説は支持されると考える。ただし、経路依存性なのかサンクコストなのかについてはさらなる分析が必要である。

さらに、最高成績と事後の記事数には時点が離れるほど正の関係で統計的に有意が見られた。この点については①なぜ大会後すぐは統計的に有意ではなく、②大会後2-3年で有意になったのか、に分けて検討する。①については、大会後は結果が良くても悪くても、メディアは結果の検証を行っているためかもしれない。大会の成績が良くなかったとしても、「リオ・オリンピックでは予選で敗退したが……」といった表現で選手を評価する可能性がある。つまり、メディアは「答え合わせ」をしているといえる。この解釈は経路依存性・サンクコスト仮説と整合的である。

続いて②について、大会後2-3年ということは、次のオリンピックの1-2年前である。実際には2020年の東京オリンピックは2021年に延期されたが、当時にはそれは予見不能であった。2018-19年当時は2020年オリンピックのための最終選考会が開催されていた。そのため、メディアは過去のオリンピックの成績から注目選手を選んでいく可能性がある。つまり、前大会の最高成績は大会後すぐの記事数に影響与えるのではなく、次大会の前評判に影響を与えるといえる。

最後にダミー変数の中でも生まれ年について検討したい。生まれ年の係数は仮説1については負であったが、仮説2-3では正であった。これは生まれ年が小さい＝ベテランほどオリンピックの最高成績が高くなるが、生まれ年が大きい＝若手ほど、前評判が高くなることを意味している。後者についてダミー変数についてもニュースバリュー仮説が成り立っていたが、最高成績の予測については真逆の傾向であった。これはメディアの偏りであると同時に、ニュースを消費する我々の偏りであるといえる。日本においては、高校野球、高校サッカー、箱根駅伝など若者が活躍するコンテンツの人气が高い。しかし、高校球児のすべてがプロで活躍できるわけではなく、人気とその後の成果にギャップが見られる。つまり、メディアと我々の予測とニュースバリューに関する判断との間に齟齬が発生していることを示唆している。

本論文の回帰モデルは、数多の組み合わせの中のごく一部である。学会当日にはほかのモデルも検討に入れたうえで、結果を議論していきたい。

参考資料

- [1]小川知也, et al. 株価データと新聞記事からのマイニング. *情報処理学会研究報告自然言語処理 (NL)*, 2001, 2001.20 (2000-NL-142): 137-144.
- [2]蔵本貴久, 和泉潔, 吉村忍, 石田智也, 中嶋啓浩, 松井藤五郎, ... & 中川裕志. (2013). 新聞記事のテキストマイニングによる長期市場動向の分析. *人工知能学会論文誌*, 28(3), 291-296.
- [3]細川蓮, 上田健太郎, 諏訪博彦, 梅原英一, 山下達雄, 坪内孝太, & 小川祐樹. (2023). 新聞とインターネット株式掲示板を用いた金融指標の予測. *人工知能学会第二種研究会資料*, 2023 (FIN-030), 94-99.
- [4]Scheufele, D. A., & Tewksbury, D. (2007). Framing, agenda setting, and priming: The evolution of three media effects models. *Journal of communication*, 57(1), 9-20.
- [5]David, P. A. (1985). Clio and the Economics of QWERTY. *The American economic review*, 75(2), 332-337.
- [6]2006/07/26 日本経済新聞
- [7]2008/01/18 日本経済新聞
- [8]Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *science*, 211(4481), 453-458.
- [9]<https://www.joc.or.jp/games/olympic/profile/> (2023年1月9日 閲覧)

【ここに掲載した著作物の利用に関する注意】

本著作物の著作権は日本感性工学会に帰属します。本著作物は著作権者である日本感性工学会の許可のもとに掲載するものです。ご利用に当たっては「著作権法」ならびに関連法規に従うことをお願いいたします。

訂正：

① 表5. 仮説2-3の分析結果は以下に読み替えてください。モデル6にミスがありました。

	モデル5	モデル6	モデル7	モデル8
被説明変数	大会後_0	大会後_1	大会後_2	大会後_3
補正R2	0.628	0.804	0.823	0.819
説明変数	係数	係数	係数	係数
大会前_1	0.1695 **	0.4949 **	0.8763 **	1.2519 **
最高成績	0.0106	0.1321 †	0.2803 *	0.4494 **
生年月日	0.0474	0.6074 *	1.3493 **	2.2041 **
競泳ダミー	-6.0447	-6.8221	-7.355	-6.9777
ウエイトリフティングダミー	-5.0648	-3.0668	-0.7922	2.7884
新体操ダミー	-7.3702	-11.3878	-15.8839	-21.077
テニスダミー	-5.4672	14.7226	28.462	31.6955
ボートダミー	-7.4474	-8.92	-8.5318	-7.5809
陸上ダミー	-6.0901	-3.6481	-0.2793	4.9029
体操ダミー	-4.6527	-5.7561	0.3121	-3.5473
卓球ダミー	-5.8965	2.2853	6.6547	20.7488
レスリングダミー	18.6107 **	22.9937 *	27.0244 †	33.2117

② 3-4 ページの「ダミー変数」はすべて「コントロール変数」に読み替えてください