

研究ノート

相撲八百長のインセンティブ再検証：千秋楽 7-7 対 8-6 の異常勝率は 2011 年以降どう変わったか

渡 辺 雅 仁

東京国際大学論叢 経済学研究 第 10 号 抜刷
2026 年（令和 8 年）3 月 20 日

研究ノート

相撲八百長のインセンティブ再検証：千秋楽 7-7 対 8-6 の異常勝率は 2011 年以降どう変わったか

渡 辺 雅 仁

Revisiting Incentives for Match-Fixing in Sumo: How Did the Anomalous Win Rate for 7-7 vs 8-6 on Day 15 Change After 2011?

WATANABE, Masahito

Abstract

Did the 2011 match-fixing scandal and subsequent reforms in professional sumo change outcomes at the incentive threshold? Using Sumo Reference (SumoDB), I construct a reproducible dataset of Makuuchi Day-15 bouts and focus on the strict 7-7 vs 8-6 matchup, excluding walkovers. Reproducing the pre-2000 window, the 7-7 side wins 72.4% in 1989/01-2000/01, broadly aligning with the high-win pattern reported by Duggan and Levitt (2002). Comparing pre- and post-reform periods, the win rate declines from 58.6% in 2000-2010 to 51.7% in 2011-2025. A two-proportion test yields $z = 1.39$, $p = 0.164$, not significant at 5%, yet the yearly series and pooled period means document a convergence toward 50% after 2011. All acquisition and parsing procedures are documented for replication. Even in this culturally sensitive setting, the classic incentive-discontinuity pattern appears to have weakened in the 2010s.

Keywords: Sumo, Match-fixing, Incentives, Boundary effects, Institutional reform, Reproducibility

要 旨

2011年の大相撲の八百長問題とその後の改革は、実際に取組結果を変えたのか。本稿は、Duggan and Levitt (2002) 以降の知見を踏まえ、Sumo Reference (SumoDB) のデータを用いて検

証する。幕内・千秋楽における7勝7敗（7-7）の力士と8勝6敗（8-6）の力士の取組を抽出し、2011年以前と以降の比較を行う。まず、1989年1月場所～2000年1月場所の再集計では、7-7側の勝率は72.4%（113/156）で、Duggan and Levitt（2002）の「高勝率」パターンを概ね追認できる。さらに、制度改革後の動向に焦点を当て、2000～2010年と2011～2025年を比べると、勝率は58.6%（51/87）から51.7%（61/118）へ低下した。二標本比率検定では5%水準で有意でない（ $z = 1.39$, $p = 0.164$ ）が、年別系列と期間平均は2011年以降に勝率が概ね50%近傍へ収束していることを示す。大相撲における勝ち越し閾値（8勝）周辺のインセンティブ不連続は、2010年代以降弱まっている可能性を示唆する。

キーワード：大相撲、八百長、閾値不連続、インセンティブ、SumoDB、再現可能性

目 次

- I 導入
- II データと分析方法
- III 結果
- IV 結論
- V 限界と今後の課題

I 導 入

大相撲は歴史と伝統をもつ日本の国技だが、その公正性はたびたび議論の対象となってきた。Duggan and Levitt（2002）は、千秋楽に7勝7敗（以下、7-7）の力士が8勝6敗（8-6）の力士と対戦する局面で、7-7力士の勝率が異常に高いことを示し、勝ち越し閾値（8勝）付近にインセンティブの不連続があると数量的に明らかにした。7-7力士にとっては「勝ち越し」をかけた大一番であるのに対し、8-6力士はすでに勝ち越しを確定させているため、両者のインセンティブが大きく非対称であることによると考えられ、八百長の存在を強く示唆するとされた。これは、昇進制度が行動をどう歪めるかをデータで示した先駆的研究として、スポーツ経済学で広く引用されている。

他方、日本国内では、7-7力士の高勝率はファンや関係者の間である程度認識されていたものの、長らく学術的分析の対象にはなりにくかったとされる。相撲が「国技」として強い象徴性を帯びるなか、文化的・社会的な事情が公開データに基づく検証を難しくしてきた側面がある。その後、国内でも数量的分析が現れ始め（例：Hori and Iwamoto, 2012；岩本・堀, 2013）、近年は公正性とは別の応用領域、たとえばOta and Kimura（2023）による傷害予測（スポーツ医学・安全管理）も見られる。しかし、勝ち越し閾値の不連続が時間とともにどう推移したかという長期的な視点は、なお十分に検討されていない。

Duggan and Levitt（2002）の指摘を受け、Dietl, Lang and Werner（2010）はサンプルを拡張して2000年代の推移を再検証し、2000年以降にこうしたインセンティブの歪みの兆候がいったん弱まり、2003～2006年にこの傾向が一時的に再度強まったと報告した。これらは、昇進・番付制度

が不正のインセンティブを生み得るという Duggan and Levitt (2002) の示唆を、より長期のデータにより拡張・追認するものである。もっとも、先行研究には「7-7と相手が既に勝ち越し(≥8勝)」まで含める広い定義を用いる設計もあり、「厳密な7-7対8-6」に限定した定義とは分母・分子が異なる点に注意が必要である(定義の違いによる結果の相違については本稿で後述する)。

2011年には相撲協会が一連の八百長問題に直面し、春場所の中止を含む異例の対応がとられた。これを契機に、コンプライアンス体制の強化や調査委員会の設置、力士教育の徹底など「浄化」を標榜する改革が進められた。

本稿の問いはシンプルで、2011年の不祥事とその後の改革が実際に競技結果に反映されたかを数量的に検証することである。具体的には、2011年を境に、千秋楽の7-7対8-6の局面で観察される勝率に統計的な変化があったかを調べる。もし改革後に7-7側の勝率が通常水準(50%前後)へ収束していれば、不正慣行が行われにくくなった可能性を示す傍証となる。逆に、50%を大きく上回る勝率が維持されていれば、形式的な改革にもかかわらず不正慣行が残存している可能性を示唆する。

相撲に限らず、昇進・降格・賞金配分など制度的閾値を伴う競技では、境界近傍でインセンティブが非線形になることが、スポーツ経済学や行動経済学において広く議論されてきた。¹⁾ こうした制度的閾値に伴うインセンティブの非線形性は、不正や戦略的行動と結びつきやすいとされ、本稿もその系譜に位置づけられるが、焦点を大相撲の2011年の不祥事とその後の改革に置く。

本稿の貢献は二つである。第一に、Duggan and Levitt (2002) および Dietl *et al.* (2010) の知見を、公開データを用いて再現した点にある。第二に、2011年以降の勝率が概ね50%近傍へ収束する事実を示し、勝ち越し閾値におけるインセンティブ不連続の弱まりと整合的であることを明らかにした点にある。

II データと分析方法

1 使用データ

本稿は、公開データベース Sumo Reference (通称 SumoDB) を用いる。²⁾ SumoDBは、日本相撲協会の公式発表(番付・取組結果)や専門誌『相撲』など複数の一次資料を基に編纂された民間のデータベースである。幕内・十両の取組は1909年まで遡って完備されるなど網羅性が高く、国内外の研究で利用されている(例: Hori and Iwamoto (2012); Brunello and Yamamura (2023); Ota and Kimura (2023))。

対象は本場所の幕内(Makuuchi)に限定し、1989年1月場所~2025年(暫定)までを収集した。Duggan and Levitt (2002) との整合性確認のため1989年1月場所~2000年1月場所を再集計した(この期間を以下、D&L期と呼ぶ)。2011年の八百長事件とそれに続く日本相撲協会の改革を分岐点と捉え、主要な前後比較は2000~2010年と2011~2025年で行う。なお、2025年は本稿執筆時点で利用可能な一部場所(7月場所まで)のみの暫定値である。

2 サンプル抽出基準

サンプルの抽出は事前に定めた明示的なルールに基づき実施した。対象は千秋楽(15日目)の「厳密な7-7対8-6」のみとし、不戦(*fusen*)は除外する。SumoDBの星取表は取組後の最終成績が表記されるため、千秋楽で7-7が勝てば場所終了時の最終成績は(8-7, 8-7)となり、7-7が負ければ(7-8, 9-6)となる。この最終成績のペアから、千秋楽直前の「7-7対8-6」が自動化した

一貫した手続きで特定できる。

- ・対象番付：幕内
- ・取組日：千秋楽（15日目）
- ・対戦条件：取組前が7-7対8-6（最終が（8-7,8-7）または（7-8,9-6））
- ・除外：kimariteにfusenを含む取組（不戦勝・不戦敗）

SumoDBの一覧では、同一取組が東西の表示順の違いで二重に取得される場合があるため、本稿では「場所ID」×「対戦組み合わせ（力士ID）」×「日付」で一意化し、1取組=1観測に正規化した。

以上の基準から得られた年別件数と勝率は、年別系列として図1に示す。主要な従属変数は、千秋楽の当該取組で「7-7側が勝ったか」(1/0)である。各取組に年（西暦）を付与し、期間集計はプール平均（件数重み付き）で算出した。年別の勝率と95%信頼区間は、Wilson法による区間を用いる。³⁾

3 統計的手法

千秋楽における7-7側の勝敗を二値変数として扱い、その平均（勝率）を期間ごとに比較する。具体的には、(i) D&L期（1989年1月場所～2000年1月場所）の再集計、(ii) 改革前後（2000～2010年と2011～2025年）の勝率の比較を行い、改革前後の差は二標本比率検定（両側）で評価する。

頑健性の確認として、番付差などの共変量を入れたロジット回帰で勝率の変化を点検する。勝つ確率は

$$\Pr(\text{Win}_i = 1) = \Lambda(\alpha + \beta_1 \text{Post2011}_i + \beta_2 \text{RankDiff}_i) \quad (1)$$

となる。

ここで $\Lambda(z) = \frac{1}{1 + \exp(-z)}$ はロジスティック関数である。Win_iは当該取組で7-7側が勝てば1、負ければ0をとる指標である。Post2011_iは2011年以降の取組であれば1をとるダミー変数、RankDiff_iは番付差であり、

$$\text{RankDiff}_i = \text{Rank}_{\text{self},i} - \text{Rank}_{\text{opp},i}$$

と定義する。Rank_{self,i}は取組iの「7-7側」の番付を、Rank_{opp,i}は同取組の「8-6側」の番付を示す。番付はSumoDBの文字列（例：Y1e, O1w, S1e, K1w, M3e）を以下の規則で序数化する：

$$f(\text{rank}) = \begin{cases} 1 & \text{横綱 (Yokozuna)} \\ 2 & \text{大関 (Ozeki)} \\ 3 & \text{関脇 (Sekiwake)} \\ 4 & \text{小結 (Komusubi)} \\ 4+j & \text{前頭 (Maegashira } j), j=1, 2, \dots \end{cases}$$

ただし、番付の東西の別は区別しない（East = West）。したがって

$$\text{RankDiff}_i = \text{Rank}_{\text{self},i} - \text{Rank}_{\text{opp},i} = f(\text{self}) - f(\text{opp})$$

となる。この定義ではRankDiff_i > 0は「相手が格上（自分より番付が上位）」であることを意味するため、β₂ < 0が自然な符号である。その上で、係数β₁ < 0で統計的に有意であれば、2011年以降に7-7側の勝率が低下したと整合的である。

推定は、同一場所 (basho) 内で生じうる観測間の相関 (同一の行司, 報道環境, 取組編成, 取り締まりの強弱など) を考慮するため, 場所単位でクラスタ化したロバスト標準誤差を用いる。これは, 同一場所内の観測間の相関を許容しつつ, 場所間は独立であると仮定して分散を推定する方法であり, 相関を無視した通常の標準誤差と比べて一般に大きく (保守的に) なる。効果はオッズ比 ($\exp(\beta_1)$) と平均限界効果 (AME) の両方を報告する。

本稿は研究ノートであるため, 主要推定 (厳密7-7対8-6, 不戦除外) の報告に絞る。定義変更 (相手の範囲), 不戦の扱い, 期間の取り方, 小標本年の影響といった頑健性については, 本稿では詳細な検討を行わないが, 分析コードにおいて切り替え可能な形で設計している。

4 再現可能性 (データ取得と処理の再現手順・要約)

前節で述べたとおり, 本稿では分析の拡張は行わないが, 本節では主要結果が再現可能であることを担保するため, データ取得と処理の手順を簡潔に要約する。

データは Sumo Reference (SumoDB) から取得した。公開フィールドである basho, day, kimarite, および各力士の rank 等を用い, 1989年1月場所から2025年 (暫定) までの期間を対象とした。取得後, (i) 幕内, (ii) 千秋楽 (Day 15), (iii) 取組前成績が7-7対8-6, (iv) kimarite が fusen である取組を除外する, という基準に基づいて抽出を行った。このルールの下で, 場所ID × 対戦カードID × 日付を一意化キーとして, 1取組を1観測に正規化した。

番付は Y, O, S, K, Mj をそれぞれ {1, 2, 3, 4, 4+ j} の序数に写像し, Rank_{self} および Rank_{opp} を構成した。これらを用いて, 番付差 RankDiff を自身と相手の番付の差として構成した (具体的な定義は統計的手法の節を参照)。年次勝率は件数重み付きのプール平均として算出し, 信頼区間には Wilson 法を用いた。主要な比較は2000~2010年と2011~2025年の期間で行い, ロジット推定 (式 (1)) では場所クラスターに基づくロバスト標準誤差を用いた。推定結果については, 係数, オッズ比, および平均限界効果 (AME) を併記している。

分析用スクリプトでは, 分析期間や定義 (広義/厳密, 不戦の扱い) をフラグにより切り替え可能な設計としており, 第三者による再実行および頑健性検証が可能である。

III 結 果

1 Duggan and Levitt (2002) の再現

まず, 1989年1月場所~2000年1月場所で, 幕内・千秋楽の取組が厳密に7-7対8-6 (不戦除外) を再集計したところ, 7-7側の勝率は72.4% (113/156) であった (表1)。Duggan and Levitt (2002) の報告値は79.6% (181/227) であり, 本稿のほうが水準は低いが, 同論文が指摘した「勝ち越し閾値局面における7-7側の高い勝率」という基本パターンは再現された。差の要因としては, 定

表1 D&L期 (1989/01~2000/01) における7-7側勝率の比較 (幕内・千秋楽, 不戦除外)

比較対象	件数 (勝ち/総数)	勝率
Duggan & Levitt (2002)	181 / 227	79.6%
本稿 (SumoDB, 厳密7-7対8-6)	113 / 156	72.4%

義(先行研究は「相手が既に勝ち越し(≥8勝)」まで含む)やデータソースの相違などが考えられる。

2 改革前後の比較

2000～2010年の勝率は58.6% (51/87), 2011～2025年は51.7% (61/118)で、差は6.9パーセントポイント低下した(表2)。二標本比率検定では $z = 1.39, p = 0.164$ であり、5%水準では統計的に有意ではない。ただし、図1の年別系列と、三期間(1989～1999年/2000～2010年/2011～2025年)のプール平均を併せて確認すると、2011年以降は勝率が概ね50%近傍へ収束していることが確認できる。

表2 7-7力士の勝率(幕内・千秋楽, 厳密7-7対8-6, 不戦除外)

期間	件数	7-7 勝利数	勝率
1989-1999	155	112	72.3%
2000-2010	87	51	58.6%
2011-2025	118	61	51.7%

3 年別系列の推移

図1は1989年以降の年次における7-7側の勝率(Wilson 95%信頼区間, 不戦除外)である。1990年代から2000年代前半は60.70%台で推移したが、2011年を境に水準が下方にシフトし、その後は概ね50%前後に収束するパターンが視覚的に確認できる。⁴⁾

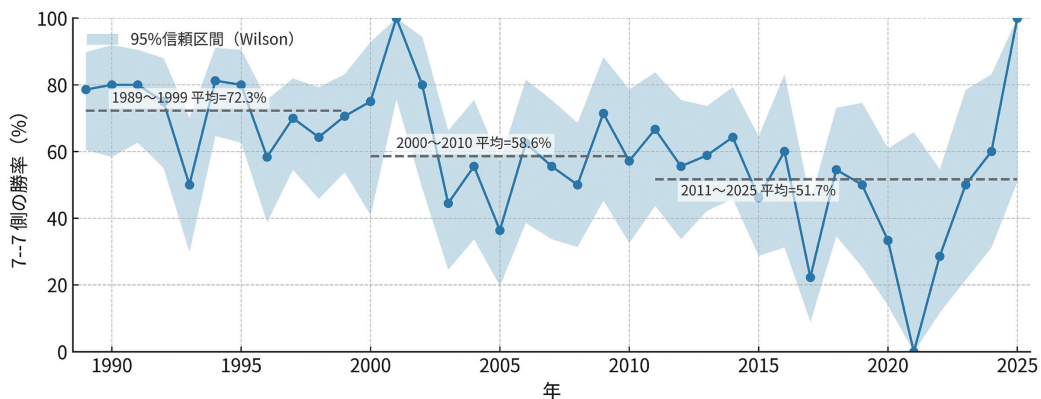


図1 年別の7-7側の勝率(不戦除外, 95%信頼区間, 三期間平均)

注) 幕内・千秋楽, 厳密7-7対8-6(不戦除外)。
95%Wilson 信頼区間を表示し, 破線は各期間のプール平均を示す。
出所) SumoDB。

また、図2は各年の観測数(該当取組件数)を示したものである。年ごとのばらつきはあるものの、2011年以降も毎年一定の観測数が確保されており、主要な結果が観測数の著しい減少に起因

するものではないことが示唆される。

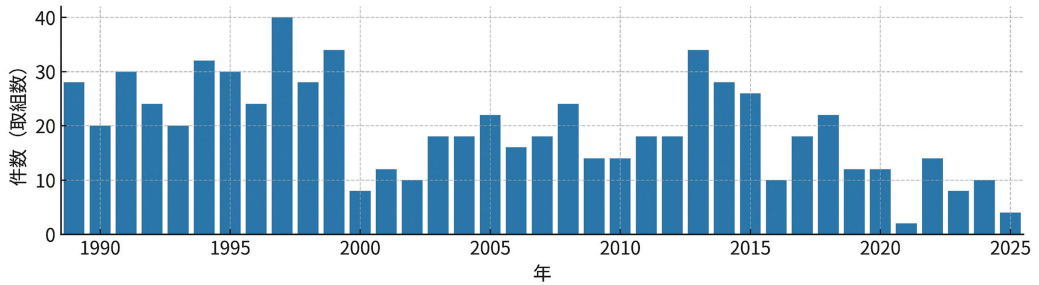


図2 年別の対戦件数 (幕内・千秋楽・7-7対8-6)

注) 幕内・千秋楽, 厳密7-7対8-6 (不戦除外)。
出所) SumoDB。

4 ロジットによる補足検証

最後に, 共変量を考慮したロジット (式 (1)) により, 主要な記述結果の頑健性を確認する。

表3は, 千秋楽・幕内・厳密7-7対8-6 (不戦除外) を対象に, Post2011 ダミーと番付差 (RankDiff) を説明変数としたロジットを推定した結果である。標準誤差はいずれも場所 (basho) 単位のクラスター・ロバスト標準誤差による。

表3 ロジット推定結果：幕内・千秋楽, 厳密7-7対8-6 (不戦除外)

(1) Post2011 + 番付差	
Post2011 (係数)	-0.889
(ロバスト SE)	(0.214)
オッズ比 (Post2011)	0.411
[95% CI]	[0.270, 0.626]
AME (平均限界効果: Post2011)	-0.195
RankDiff (係数)	0.023
(ロバスト SE)	(0.024)
定数項	0.952
(ロバスト SE)	(0.147)
観測数	337
場所クラスタ数	176
擬似 R^2	0.034
対数尤度	-211.977

注: 標準誤差はいずれも場所 (basho) 単位のクラスター・ロバスト標準誤差。オッズ比と 95% 信頼区間は係数の指数変換に基づく。AME は全体平均での平均限界効果。RankDiff は (自分-相手) の番付差 (正 = 相手が格上)。

β_1 (Post2011) は -0.889 (ロバスト SE = 0.214) で有意に負であり、これは2011年以降に7-7側の勝率が低下したと整合的である。オッズ比は 0.411 (95%CI: $0.270, 0.626$)、平均限界効果(AME)は -0.195 と推定された。すなわち、2011年以降に7-7側の勝利確率は平均で約19.5パーセントポイント低下した可能性が示唆される。

一方、RankDiff (自分-相手; 正=相手が格上) の係数は 0.023 (SE = 0.024) となり、統計的に有意ではない。観測数は $N = 337$ 、場所クラスタ数は176である。⁵⁾ なお、年固定効果 (年ダミー) を導入するとPost2011ダミーと完全に多重共線となり識別できないため、本稿の主推定では年固定効果は用いない。

IV 結 論

本稿は、大相撲の幕内・千秋楽における7-7対8-6の対戦に焦点を当て、2011年の八百長問題とその後の改革が、勝ち越し閾値(8勝)周辺のインセンティブ不連続に与えた影響を検証した。得られた主な結果は以下のとおりである。記述統計(二標本比率検定)において5%水準での有意差は確認されない一方、共変量調整と場所クラスター相関を考慮したロジット推定では、2011年以降に7-7側の勝率が低下する方向の効果が一貫して得られた。1989年1月場所~2000年1月場所の再集計では、7-7側の勝率が72.4% (113/156) となり、Duggan and Levitt (2002) の基本的パターン(79.6%)を概ね再現した。改革前後の比較では、2000~2010年の58.6% (51/87) から2011~2025年の51.7% (61/118)へ6.9パーセントポイント低下し、年別系列(図1)でも2011年以降は概ね50%近傍への収束が確認された。二標本比率検定の差は5%水準で非有意($z = 1.39, p = 0.164$)であるが、方向性は一貫しており、少なくとも千秋楽の直接的な閾値局面において、勝ち越し閾値に伴うインセンティブ不連続が弱まった可能性を示す。補完的に、共変量調整と場所クラスター相関を考慮したロジット推定(表3)においても、Post2011の効果は負方向であり、主要な解釈は維持される。本稿の政策的・運営的含意として、明示的な規律・監督の強化と整合的な形で、勝ち越し閾値近傍における不正行動のインセンティブが弱まった可能性が示唆される。

V 限界と今後の課題

本稿には、主として四つの限界がある。第一に、定義および抽出範囲の限定である。先行研究の一部は「相手が既に勝ち越し(≥8勝)」まで含める広い定義を用いる。本稿は厳密な7-7対8-6に限定したため、水準差が生じうる。第二に、小標本年の影響である。年によって該当件数が少なく、Wilson区間を用いても信頼区間が広がる(図1)。第三に、観測できない交絡変数の可能性である。例えば力士世代交代、ケガ、編成方針の変化、対戦経験などは完全には統制できていない。番付差を含むロジット推定では方向性は頑健であるが、共変量を拡充し(東西・身体情報・過去対戦成績等)、再検討する余地がある。第四に、Duggan and Levitt (2002) およびDietl *et al.* (2010) の結果と完全に一致していない理由が明らかにできていない点である。特に2003~2006年の「再燃」は、定義の違い(広い定義では上振れが再現されやすい)、分析期間の取り方、および当時の監督環境の違いが寄与した可能性がある。

今後の課題として、(i) 分析設計の違いの体系的検証(厳密定義/広義定義、fusenの扱い、幕内以外の階層)、(ii) 推定の高度化(年・場所固定効果、力士パネル化、過去対戦・同部屋回避の統制)、(iii) 外的妥当性(他競技の昇降・賞金カットオフ付近)、(iv) 行動メカニズムの解明(番

付利得の推定と行動反応の構造化) を挙げる。これらについては、データ公開後に随時更新していきたい。

注

- 1) 例えば、サッカーの昇降格制度における残留ライン周辺の努力行動や、野球・バスケットボールにおけるプレーオフ進出ライン付近での戦略的行動について、同様の議論が行われている。
- 2) <http://sumodb.sumogames.de/>。取得日：2025年8月26日。
- 3) 小標本年における正規近似の偏りを避けるためである。
- 4) 2011年は八百長問題を受けて春場所（3月）が中止となるなど制度・運営面で例外的な年である。主要結果は2011年を含む定義（2000～2010年対2011～2025年）に基づくが、頑健性の確認として2011年を除外し2012～2025年のみを改革後期間とした場合でも、7-7側の勝率が概ね50%近傍へ収束するという結論は変わらない。
- 5) 参考として、Post2011のみの単回帰ロジットでも効果方向は一致するが、有意水準は満たさなかった。

参考文献

- Brunello, Giorgio and Eiji Yamamura (2023) IZA DP No. 16536, Institute of Labor Economics (IZA).
- Dietl, Helmut M., Markus Lang, and Stephan Werner (2010) Corruption in Professional Sumo: An Update on the Study of Duggan and Levitt, *Journal of Sports Economics* **11**(4): 383-396.
- Duggan, Mark and Steven D. Levitt (2002) Winning Isn't Everything: Corruption in Sumo Wrestling, *American Economic Review* **92**(5): 1594-1605.
- Hori, Masahiro and Koichiro Iwamoto (2012) CIS Discussion Paper Series No. 580, Center for Intergenerational Studies, IER, Hitotsubashi University, <https://cis.ier.hit-u.ac.jp/Common/pdf/dp/2012/dp580.pdf>.
- Ota, Shuhei and Mitsuhiro Kimura (2023) Statistical injury prediction for professional sumo wrestlers: Modeling and perspectives, *PLOS ONE* **18**(3): e0283242.
- 岩本光一郎・堀雅博（2013）『大相撲「八百長」の誘因構造と実証分析』、『経済研究（岩波書店）』、第64巻、第2号：132-146頁。