Sports Economic Review 2 (2023) 100007

Contents lists available at ScienceDirect

Sports Economic Review

journal homepage: www.journals.elsevier.com/sports-economic-review

Competitiveness in Formula One

Ronald Peeters *, Dennis Wesselbaum

Department of Economics, University of Otago, PO Box 56, Dunedin, 9054, New Zealand

ARTICLE INFO

JEL classification:
D63
L83

Keywords:
Competitiveness
Measurement
Formula One (F1)

ABSTRACT

We define a measure of competitiveness that is based on full rankings and corrects for partitioning of the parties in exogenously defined clusters. We use the measure to study the change in competitiveness of Formula One racing during the period 1993–2019.

1. Introduction

When outcomes of sports lack parity and hence are predictable, sports fans lose interest and depress their attendance (El-Hodiri & Quirk, 1971; Quirk & Fort, 1992; Schmidt & Berri, 2001). In addition to the purpose of guaranteeing safety and fairness to the athletes, professional sports federations also amend rules to make sports more competitive, thereby increasing excitement and drama (Szymanski, 2003). This effect is not only found in league sports, but also within the context of racing sports. Berkowitz et al. (2011) provide evidence suggesting that a decline in competitive balance may be a causal factor for reduced television audiences and race attendance in NASCAR racing during the 2007–2009 seasons. Race attendance is affected by season-level uncertainty of outcome, while television attendance is affected by both race-level and season-level uncertainty of outcome. In the context of Formula One racing, using data on 400 Grand Prixs between 1993 and 2014, Schreyer and Torgler (2018) find that race outcome uncertainty has a positive effect on television viewership in Germany.

In the literature related to competitiveness in racing sports several measures have been used to assess competitiveness. Most measures used in the literature are based on the distribution of the number of points scored (or number of wins) by drivers or teams, and evaluate these using the standard deviation (Mastromarco & Runkel, 2009), the Gini index (Budzinski & Feddersen, 2019; Krauskopf et al., 2010) or a normalised Herfindahl-Hirschman Index (Judde et al., 2013). As argued in Berkowitz et al. (2011), these measures seem not adequate to assess competitiveness, or uncertainty of outcome, in racing sports. In the context of Formula One racing, Schreyer and Torgler (2018) used a measure based on differences in the qualification times in the top 3 qualifiers, which is an alternative to the margin of victory, and Judde et al. (2013) considered the average number of lead changes. Berkowitz et al. (2011) also criticize these measures as they only focus on the race leaders and ignore the rest of the field. Further, the identity of the drivers are ignored, such that lead changes across races are not picked up. Instead, in the context of NASCAR racing, Berkowitz et al. (2011) uses the adjusted churn measure developed by Mizak et al. (2007). The churn method considers changes in rankings, which can be adopted on race level (comparing start versus finish) and across races (finish one race versus finish next race). Instead of the churn method, we suggest the use of the Kemeny distance. The Kemeny distance counts how many switches (or, in the context of racing sports, overtakes) are needed to transform one ranking into the other.

We illustrate the advantage of the Kemeny distance relative to the churn method via an example. Let there be four drivers (a, b, c, d) and let their start positions be (1, 2, 3, 4) while their finish positions are (1, 4, 2, 3). To achieve this outcome, driver c and d both must have overtaken driver b. These are 2 overtakes, which is the number that is focal in the Kemeny distance (normalised on 0–1 scale: 1/3). The churn method, instead, counts the positions that have changed, which is 2 for driver b and 1 for drivers c and d, and uses 4 as focal number (normalised on 0–1 scale: 1/2). If the finish position would instead have been (1, 4, 3, 2), the focal number would still be 4 for the churn method (normalised on 0–1 scale: 1/2), while it would be 3 in the Kemeny distance, since now driver d must also have been overtaking driver c (normalised on 0–1 scale: 1/2). Hence, the elegance of the Kemeny distance is that it equals the minimum number of overtakes that must have been taking place during a race in order to get from a start ranking to a finish ranking.

* Corresponding author.
E-mail addresses: ronald.peeters@otago.ac.nz (R. Peeters), dennis.wesselbaum@otago.ac.nz (D. Wesselbaum).

https://doi.org/10.1016/j.serev.2022.100007
Received 29 September 2022; Received in revised form 16 November 2022; Accepted 18 November 2022
Available online 21 November 2022

In order to explore the competitiveness across seasons, we use (1) the average Kemeny distance over all pairs of races within the season – based on both the drivers positions at the start and at the finish of the races –, and (2) the average Kemeny distance between start and finish positions over all races within the season. In all cases we adjust the distance on team-level. In order to control for differences in the number of drivers/teams across races/seasons and the number of races across seasons we benchmark the empirically obtained Kemeny distance with the simulated distribution over Kemeny distances based on fully random race outcomes. Using this measure, we assess the competitiveness of Formula One during the seasons 1993–2019 and relate it to the major rule changes within this period.[1]

The measure of competitiveness that we develop has been tailored to the Formula One situation, but can easily be generalised to evaluate the impact of public policy on the competitive balance across exogenously partitioned clusters within landscapes of public facilities, such as universities across countries, hospitals across states, and schools across different pedagogical paradigms.

2. Methods

2.1. The data

A Formula One World Championship season is a sequence of Grand Prixs within a calendar year. Each Grand Prix consists of three parts: practice sessions, qualifying, and race, with the qualifying pinning down the starting order for the race. Teams can race two cars and a driver can win points for himself and for his team. For most of the history of Formula One, the first six drivers would receive points (10-6-4-3-2-1). From 2003 on, the first eight drivers received points (10-8-6-5-4-3-2-1) and since 2010, the first ten drivers obtain points (25-18-15-12-10-8-6-4-2-1). These points are added up at the end of the season to crown the driver World Champion and the constructor World Champion.

We collected all data on start and finish positions of all 486 races in the seasons 1993–2019. Table 1 provides for each season the number of teams that have been racing and the number of races. While the number of Grand Prixs has been increasing since the early 1990s, the number of teams (and drivers) has remained relatively stable at around 11 (22 drivers). For all these seasons teams consistently competed with two cars throughout the season.[2]

In total, our data set contains 10,734 individual finish positions. Of these, 3062 are retired drivers, 225 drivers did not start the race, and 7447 are classified drivers. Retirements can occur due to, for example, technical problems (most common), accidents, or disqualifications (for example, for breaching technical regulations). For drivers who do not complete the qualifying, start the race, or retire during the race, we have to manipulate their start and finish positions. If a driver does not complete the qualifying, we assign to him the last start position. If a driver does not start the race, we assign finish position 26 to him. Finally, if a driver retires from the race, we assign finish position 25 to him. Further, if a team withdraws from the World Championship (e.g. Super Aguri in 2008), we still assign the drivers a start (last two possible ones) and finish position (26) to keep the sample balanced throughout the season.

2.2. Measure of competitiveness

In season y, m_y teams – each consisting of 2 cars/drivers – have been competing in k_y races. Let R_ℓ be the ranking of the drivers in race $\ell = 1, \ldots, k_y$.

The building block for comparing a pair of rankings is the Kemeny distance, which is defined as the minimum number of interchanges of two adjacent elements required to transform one ranking into another. Since we are interested in the ranks on team-level, we have to make a correction to this distance.

Denote by $R_\ell^{j(i)}$ the rank of the $i^{\text{th}}(i = 1, 2)$ ranked driver of team $j = 1, \ldots, m_y$ in race ℓ. The team-corrected Kemeny distance between the rankings in two races ℓ and ℓ' is given by:

$$kemd(R_\ell, R_{\ell'}) = \frac{1}{2} \cdot \sum_{j=1}^{m_y} \sum_{i=1}^{2} \sum_{j'=1}^{m_y} \sum_{i'=1}^{2} \left| sign\left(R_\ell^{j(i)} - R_\ell^{j'(i')}\right) - \mathrm{sign}\left(R_{\ell'}^{j(i)} - R_{\ell'}^{j'(i')}\right) \right|.$$

The average Kemeny distance over all pairs of races in a season y, $kemd(y)$, measures the level of competitiveness of the season. It takes the least possible value of 0 if all races in the season resulted on team-level in the same outcome. While the maximum Kemeny distance between a pair of race rankings equals $m_y(m_y - 1)$, we are not aware of a general expression for the maximum value of $kemd(y)$ if $k_y \geq 3$.[3] However, the maximum distance between two races already indicates that it is sensitive to the number of teams competing. Since the seasons considered differ in the number of teams and the number of races, for the sake of a fair comparison across seasons, we have to normalize $kemd(y)$. Due to the absence of a general expression of the maximum value of $kemd(y)$, the normalisation is based on numerical simulations.

For given m_y and k_y, we draw a random season of race results of which we compute the associated average Kemeny distance. By repeating this for a large number of random seasons, we numerically approximate the distribution of $kemd(y)$ over all possible seasons, $\widetilde{f}_{(m_y,k_y)}$.[4] Thus, the expected value of $kemd(y)$ according to the distribution $\widetilde{f}_{(m_y,k_y)}$ is the average Kemeny distance that we can expect of a random season.

We define our two measures of competitiveness, by benchmarking the true average Kemeny distance against the numerical simulated distribution:

$$comp(y) = kemd(y) \Big/ \widetilde{\mu}_{(m_y,k_y)} \quad \text{and} \quad comp'(y) = \widetilde{F}_{(m_y,k_y)}(kemd(y)),$$

where $\widetilde{\mu}_{(m_y,k_y)}$ refers to the expected value of $kemd(y)$ given the distribution $\widetilde{f}_{(m_y,k_y)}$, and $\widetilde{F}_{(m_y,k_y)}$ is the cumulative distribution of $\widetilde{f}_{(m_y,k_y)}$.

Now, $comp(y) < 1$ means that the season has been less competitive compared to what can be expected from a fully competitive season where the outcome of each race is random, and the lower $comp(y)$ the less competitive the season. The second measure, $comp'(y)$, indicates the probability that a randomly drawn season outcome would be less competitive than the considered outcome. This measure takes as minimum value 0 and as maximum value 1.

While the two measures are designed to pick up the same effects, there are advantages of using both of them. First, $comp$ may provide a different ranking than $comp'$ when comparing seasons that differ in the number of teams and/or the number of races; though, this is mainly a concern if there is a high level of competitiveness. Second, the alternative measure $comp'$ may not pick up differences when the level of competitiveness is not too high. For the case of Formula One, with substantial

[1] The data set is created from individual race outcomes obtained from the Formula1.com webpage (see Archives 1950–2022). All data and *Matlab*-codes to replicate the analysis in this article are available via https://osf.io/yx65s/.

[2] The only exception is the 1994 Grand Prix of Monaco, where the Simtek and the Williams team only raced one car each after the fatal accidents of Roland Ratzenberger and Ayrton Senna at the previous Grand Prix at Imola.

[3] For a pair of rankings the maximum Kemeny distance is obtained by one ranking being the reverse of the other ranking. For a situation with more than two rankings ($k_y \geq 3$), finding the tuple(s) of rankings with maximum average Kemeny distance is a problem of high combinatorial complexity. To give an impression, ignoring the team structure, if we would have d drivers and k races in a season, there are $\frac{(d!+k-1)!}{k!(d!-1)!}$ unique (that is, ignoring permutation-equivalent season outcomes) season outcomes. This high number also complicates finding the exact maximum numerically.

[4] Although real races produce equally ranked drivers due to cars retiring, we only consider strict rankings for our normalisation.

Table 1
Number of teams and races over all seasons.

season	1993	1994	1995	1996	1997	1998	1999	2000	2001
# Teams	13	14	13	11	11	11	11	11	11
# Races	16	16	17	16	17	16	16	17	17
season	2002	2003	2004	2005	2006	2007	2008	2009	2010
# Teams	11	10	10	10	11	11	11	10	12
# Races	17	16	18	19	18	17	18	17	19
season	2011	2012	2013	2014	2015	2016	2017	2018	2019
# Teams	12	12	11	11	10	11	10	10	10
# Races	19	20	19	19	19	21	20	21	21

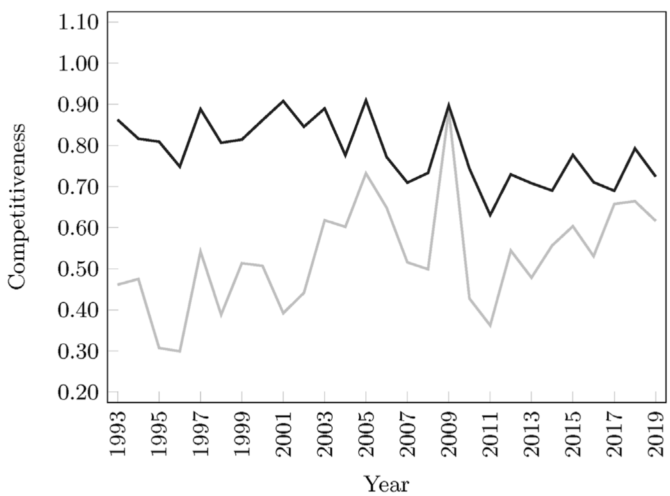

Fig. 1. Mean competitiveness (*comp*) across seasons, based on start position (gray) and based on finish position (black).

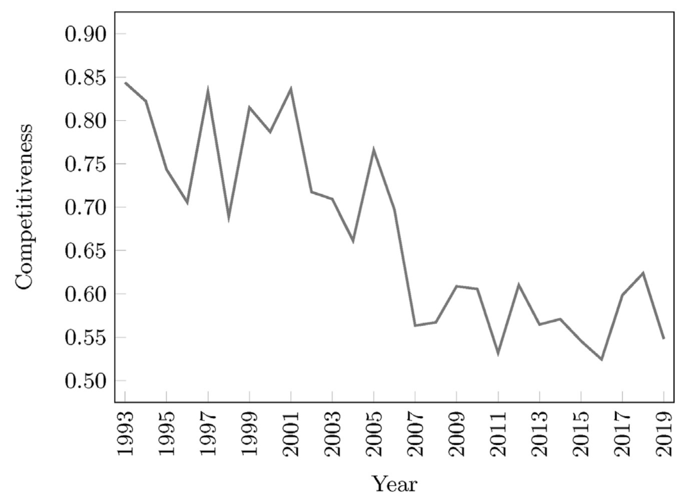

Fig. 2. Competitiveness within races (difference between start and finish positions).

differences in the quality of the cars, *comp*′ is not able to pick up any effect, leaving *comp* the more insightful measure.

3. Results

First, we consider competitiveness across seasons. Fig. 1 shows for the seasons in our sample the levels of competitiveness as measured by *comp*, both based on start (gray) and finish (black) position.

Throughout all 27 seasons, the finish positions imply a higher level of competitiveness than that implied by start positions, which is not too surprising given that the retirement rate of almost 30% will add a lot of randomness (hence, competitiveness) to finish rankings. The time series based on start positions shows increasing competitiveness between 1993 and 2005, and between 2011 and 2019. We speculate that the shock to competitiveness that is visible around the years 2005 and 2011 coincides with major changes to Formula One rules: since 2005 engines have to be more and more reliable and teams are not allowed to frequently change them without a grid penalty; the introduction, abolishment, and reintroduction of the kinetic energy recovery system (KERS) in 2009, 2010, and 2011; the return of the in-race refuelling ban in 2010; and the ban of the so-called double diffuser in 2011. The time series based on finish positions also shows some turbulence during 2005 and 2011, but shows a more constant level of competitiveness outside this period.[5] Overall, throughout the last 27 years competitiveness has been increasing when considering start positions, but Formula One racing has become less competitive based on finish positions.

[5] The spike in the 2009 season in Fig. 1 can be explained by the innovation of the double diffuser by Brawn GP, the simultaneous ban of other aerodynamic parts, the introduction of KERS, and the return of slick tyres.

Fig. 2 shows for each season the average (normalised) Kemeny distance between start and finish positions. This distance shows the predictability of the finish position giving the start position; hence, captures the competitiveness within the races: the lower the value, the smaller the distance between start and finish position and the more predictable the race.

Over the years, races became more predictable. The majority of the increase in predictability occurred between 2004 and 2007. A prominent rule changes that may have caused this is the in 2004 introduced ten place grid penalty for engine replacements during a race weekend. This may have led to teams producing more reliable engines, resulting in less retirements: in seasons 1993–2003 the retirement rate was about 42% while it was with about 18% considerably lower in seasons 2008–2019. Further, since 2007 there is only one tyre supplier rather than two, which, according to our results, coincides with an increase in predictability. With only one tyre supplier all teams will likely face a dominant tyre strategy, reducing the scope for different strategies.

4. Conclusion

Most measures to evaluate competitiveness in sports are developed for sports in which different parties compete for championship via sequences of matches in which two parties directly interact. Racing sports are different; all parties competing for championship are directly interacting in each match. As a result, measures developed for sport league structures like football are not all (such as measures related to winning probability) found useful for racing sports. In this article we develop a measure suitable for racing sports that makes full use of the property that each match results in a ranking over the competing parties and has strong axiomatic foundation. We normalize the measure numerically in order to allow for comparisons over championship seasons that vary in the number of matches (races) and competing parties.

We study competitiveness of Formula One racing during the period 1993–2019. We find substantial decline in competitiveness in the period 2005–2011 which possibly is in response to rule changes fostering drivers' and mechanics' health and safety. This decline is visible for competitiveness based on both start and finish positions. The various rule changes, including no refuelling, one tyre provider and increased aerodynamics complexity, has caused start positions to be more predictable for finish positions. Striving to maximize the attractiveness of Formula One without impairing health and safety, the FIA may want to reconsider those rules that were implemented in 2005–2011 that were not aimed to increase safety and may possibly have been detrimental for competitiveness; one example may be the move to one tyre supplier in 2007.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Berkowitz, J. P., Depken, C. A., II, & Wilson, D. P. (2011). When going in circles is going backward: Outcome uncertainty in NASCAR. *Journal of Sports Economics, 12*(3), 253–283.

Budzinski, O., & Feddersen, A. (2019). Measuring competitive balance in Formula One racing. In P. Rodriguez, S. Kesenne, & B. R. Humphreys (Eds.), *Outcome uncertainty in sporting events*. Edward Elgar Online.

El-Hodiri, M., & Quirk, J. (1971). The economic theory of a professional sport league. *Journal of Political Economy, 79*(6), 1302–1319.

Judde, C., Booth, R., & Brooks, R. (2013). Second place is first of the losers: An analysis of competitive balance in Formula One. *Journal of Sports Economics, 14*(4), 411–439.

Krauskopf, T., Langen, M., & Buenger, B. (2010). *The search for optimal competitive balance in Formula One*. CAWM Discussion Paper. No. 38.

Mastromarco, C., & Runkel, M. (2009). Rule changes and competitive balance in Formula One motor racing. *Applied Economics, 41*(23), 3003–3014.

Mizak, D., Stair, A., & Neral, J. (2007). The adjusted churn: An index of competitive balance for sports leagues based on changes in team standings over time. *Economics Bulletin, 26*(3), 1–7.

Quirk, J., & Fort, R. (1992). *Pay dirt: The business of professional team sports*. Princeton, NJ: Princeton University Press.

Schmidt, M. B., & Berri, D. J. (2001). Competitive balance and attendance: The case of major league baseball. *Journal of Sports Economics, 2*(2), 145–167.

Schreyer, D., & Torgler, B. (2018). On the role of race outcome uncertainty in the TV demand for Formula 1 Grands Prix. *Journal of Sports Economics, 19*(2), 211–229.

Szymanski, S. (2003). The economic design of sports contests. *Journal of Economic Literature, 41*(4), 1137–1187.