Impact of the revised starting lottery in the alpine speed events on performance of the top athletes

Otto Kolbinger*[1], Elisabeth Kolbinger[1], Jürgen Beckmann[1], Martin Lames[1]

[1] Department of Sport and Health Sciences, Technical University of Munich, Munich, Germany

* otto.kolbinger@tum.de

ORIGINAL ARTICLE

Submitted: 27 May 2022

Accepted: 10 January 2023

Published: 9 May 2023

Editor-in-Chief:

Claudio R. Nigg, University of Bern, Switzerland

Section Editor:

Ansgar Thiel, University of Tuebingen, Germany

ABSTRACT

Before the 2016/17 season, the International Ski Federation conducted major changes to its starting lottery, which defines the starting order in advance of each race of the speed disciplines. As part of this change, the top seeded athletes from then on were able to pick a starting position. Therefore, the aim of this study is two-folded. First, we explored the pick behavior of the top athletes and analyzed if they were able to identify favorable starting positions. Second, we investigated the influence of the rule change on competitive balance. Our sample consisted of 322 races with 17,725 individual performances of the seasons 2011/2012 to 2020/2021 and all our analysis were run separately for Women's Downhill, Women's Super G, Men's Downhill and Men's Super G. We could show that for all disciplines and both genders athletes heavily preferred to start around starting position 7. To start from such positions only led to significantly better rankings in Men's Super G ($z = 2.04$, $p = .041$), but common language effect size showed the same trend for Women's Super G (*CL*: 51.9%) as well as Men's (*CL*: 53.3%) and Women's Downhill (*CL*: 53.6%) competitions. In addition, male starters seeded among the top 7 performed significantly better after the rule change (Downhill: $z = 2.95$, $p = .002$; Super G: $z = 2.89$, $p = .002$), whereas they performed worse in Women's Super G ($z = -2.77$, $p = .005$). Our findings indicate that there is a change in competitive balance for men's competitions but not for women's competitions. In addition, combining these findings with the detected picking preferences of the top athletes, we conclude our work with suggestions for further adaptions of the starting lottery.

Keywords

starting lottery, competitive balance, rule changes, ski alpine

Citation:

Kolbinger, O., Kolbinger, E., Beckmann, J., & Lames, M. (2023). Impact of the revised starting lottery in the alpine speed events on performance of the top athletes. *Current Issues in Sport Science*, *8*(1), Article 005. https://doi.org/10.36950/2023.1ciss005

Introduction

The International Ski Federation (trans: Fédération Internationale de Ski, hereinafter FIS) has to change something, to get better TV ratings (*Änderung in Speed Disziplinen steht bevor [Changes in Speed Disciplines to Debut]*, 2016). With this desperate statement the Chief Race Director World Cup Men, Markus Waldner, justified the changes to the starting lottery for speed disciplines performed by the FIS in advance of the 2016/17 season. To stretch the time window in which the top athletes had to start, the top ten athletes from then on could pick an odd starting position from 1 to 19. Until the 2015/16 season, the top seven athletes were randomly assigned to a starting position between 16 and 22. The goal of this rule intervention, to make a sport more attractive for spectators, is a well-known reason for rule changes (Arias et al., 2011; Mühlberger & Kolbinger, 2021). However, it is also well-established in evaluative research that such interventions can cause unintended side-effects (Kolbinger, 2020). Studies by Lešnik et al. (2013) and Supej et al. (2005) indicate that the starting order might influence the athlete's final ranking. Therefore, it is reasonable to assume that the changes to the starting lottery impact competitive balance, the degree of uncertainty regarding the outcome of a competition, in the speed disciplines in ski alpine. In particular, the change in the starting lottery could lead to an advantage of the very top athletes, as these athletes might win a bigger share of the races.

Whereas competitive balance in team sports is a frequently investigated issue, "[competitive] balance in individual sports receives less attention . . . and is an area clearly in need of more research" (Rodriguez et al., 2020, p. 1). This is also true for research dedicated to shifts in competitive balance affected by changes to the rules or regulations of a competition. Numerous studies have investigated how competitive balance in football alone has been impacted by the transition to the "three points for a win" system (Haugen, 2008, 2016), the introduction of the "back-pass rule" (Kent et al., 2013), adaptations to the numbers of participants in a competition (Geenens, 2014; Groot, 2008) and the Bosman ruling (Flores et al., 2010) among other things. Further studies investigated the influence of rule changes in other team sports like American Football (influence of free agency: Balfour & Porter, 1991; impact of the introduction of a rookie draft: Fort & Quirk, 1995), Basektball (introduction of a salary cap: Fort & Quirk, 1995), Baseball (influence of free agency: Balfour & Porter, 1991; Fort & Quirk, 1995; player allocation: La Croix & Kawaura, 1999), or Handball (tournament designs: Csató, 2021; several rule changes: Haugen & Guvåg, 2018).

Only very few studies looked at the impact of changes of rules and regulations on competitive balance in individual sports so far. Mastromarco & Runkel (2009) evaluated how competitive balance in the Formula One was affected by rule changes from 1950 to 2015 and stated to have detected significant positive impacts, even after interventions that were intended to improve driver's safety solely. Those findings were supported by Judde et al. (2013) for the uncertainty of championship outcome, but not for the uncertainty of race outcome. However, for the latter analysis the authors included several rule changes as separated dummy variables, like the introduction of aerodynamic wings or the permission of refueling, which then again showed significant impacts on the competitive balance. In a study about outcome uncertainty in NASCAR, Berkowitz et al. (2011) also touched on the impact of a rule change, the introduction of the so-called Car of Tomorrow to improve driver's safety, which decreased the uncertainty of outcome.

All these studies did not look at rule changes that affected the order how athletes start in a competition. A study looking on a similar phenomenon was performed by Del Corral (2009), who investigated potential changes of the revised seeding system for Grand Slam tournaments in tennis. In more detail, the International Tennis Federation increased the number of seeded players from 16 to 32, which was expected to lead to a decrease in competitive balance. Del Corral (2009) could verify this for men's tournaments, but not for women's tournaments, which also showed a lower competitive balance overall. In another study, Goller (2022) identified an advantage for the player to throw first in a darts match. However, in darts this first-mover is determined by a preliminary skill-contest (throwing as close as possible to the middle of the board) and not randomly as in alpine skiing. Swartz (2007) suggested a revision for a sport where the starting order gets randomly assigned, Highland Dance, but did so only based on the assumption that participants starting late in the contest have an advantage, which was not empirically verified.

For alpine skiing, there are neither studies that examine the influence of the starting position in the speed disciplines on performance and finishing position in a race nor studies that investigate the influence of rule changes on competitive balance. However, two studies investigated the influence of the starting order in the technical discipline slalom. Interestingly, it was found that performance parameters were affected by the starting slot. Yet, each of these studies relied on small samples. Focusing on 15 athletes participating at one slalom competition at the European Cup, Lešnik et al. (2013) found that for the overall course there was a medium to strong, but not a significant, correlation between the start number and the elapsed time. However, investigating the steep sections of the course separately, they could demonstrate that athletes with a lower starting number achieved significantly better performances. In another study, Supej et al. (2005) set a slalom course for five world cup drivers that was performed three times by each athlete. The authors found significant negative shifts for the trajectory length around gates and the ski's velocity, to name but two examples. As these were considered to be disadvantageously, they concluded that the conditions got worse with each further athlete performing the course. The findings of both studies suggest that it is reasonable to assume that changes to the start lottery impact competitive balance in speed disciplines.

Compared to slalom competitions, where the starting order is just mirroring the world cup standings in this discipline, the new starting lottery protocols in the speed disciplines allow the top athletes to pick a starting position. This type of draft creates at least two further intriguing questions: First, if athletes show similar preferences for starting positions and, second, whether top athletes are in fact able to pick favorable starting positions. Based on the expertise about race conditions at different points of time that world class racers have accumulated, we would expect such racers to be able to recognize the most preferable starting position for them for a certain race. However, the tendency to show collective behavior or even conformity might interfere with this ability. Choosing the best starting position on a given day can be considered as a quite difficult task, which is well-known as a circumstance to increase the tendency to show conformity (e.g. Janis, 1972; Nordholm, 1975), especially when the task is considered as important (Baron et al., 1996). There are no studies about how athlete's pick starting positions in any sport and how this selection might be affected by peers yet. However, studies at least investigated something similar, by investigating pacing strategies in endurance competitions.

Following Renfree et al.'s (2015) conclusions from those studies, we would expect that athletes tend to show conformity when picking positions in a race. In particular, Renfree et al. (2015) pointed at studies about cross country running as well as about marathon competitions (Esteve-Lanao et al., 2014; Hanley, 2014; Renfree & Gibson, 2013). All these found that athletes adopted similar running speeds in the early stages of races, leading to a so-called "positive pacing strategy" for most athletes. This means that these runners performed the first half of the race in a higher speed compared to the second part, which is usually considered as a bad strategy in endurance competitions (Abbiss & Laursen, 2008). Renfree et al. (2015) concluded that this disadvantageous strategy of many athletes was based on the emergence of the collective behavior patterns in the beginning of the race described above.

Whether or not athletes are able to pick the most favorable starting position for them should have an influence on their finishing position if the starting slot, at least partly, determines performance. If this applies, the ability to pick favorable starting positions should influence performance after the introduction of the revised starting lottery. This is why the research questions for this study are explorative. Thus, the first aim of this study is to investigate if athletes show indeed similar patterns when picking their starting positions (pick behavior). Second, we will explore whether the selected starting positions are actually preferable over other starting positions (pick optimality), before we finally measure the effect of the intervention on competitive balance in the speed disciplines in alpine skiing.

Methods

Data collection and sample

We retrieved data for all speed events, Downhill and Super G, and genders from ten world cup seasons (2011/12-2020/21), including world championships and Olympic games, from the official FIS homepage (www.fis-ski.com). This means that the same number of seasons before (period *Before*: 2011/12-2015/16) and after (*After*: 2016/17-2020/21) the rule intervention was included in the data set. Six races had to be dropped as no official result list was available or the data was incomplete. Overall, data of 322 races with 17,725 individual performances were included in this study, consisting of 82 Women's Downhill, 73 Women's Super G, 96 Men's Downhill and 71 Men's Super G competitions. For each participant in a race, we retrieved name, final rank, starting position and draw position. Athletes who withdrew before the start of a competition were excluded. Athletes who have been disqualified or did not finish the race were included and the final rank was set as the final rank of the worst performer who finished the race plus 1 (e.g., the rank was set as 52 for such athletes, if 51 participants did finish the race).

In addition, we retrieved the final standings for each season's World Cup in the speed disciplines for both genders. For each of the 2,186 athletes who appeared in the final standing of one of these 40 individual World Cups, we collected name, final rank and the number of achieved points in the respective discipline.

Statistical analysis

Pick behavior

As described above, since the introduction of the new starting lottery starting positions of the top 10 athletes were no longer drawn. Instead, these athletes could pick an odd starting number between 1 and 19, with the athlete listed at draw position one picking first, the one on draw position two picking second and so on. This draw position is determined by previous World Cup performances. This procedure is also displayed in Table 1, which shows the starting order (BIB) and draw positions (pos) of the first two starting groups for the 8th Women's Downhill competition in 2020 in Crans Montana, Switzerland. Thus, the athlete to pick first in this race was Corinne Suter (Pos = 1) and she selected starting number 3 (BIB = 3).

To check if there was a general preference for certain starting positions, we determined how often each starting position was chosen by the first athlete (BIB_{Pos1}) and compared the distribution of picks to a uniform distribution with Kolmogorov-Smirnov tests. Further, we investigated whether the top starters agreed upon the decision, which starting time would be the best for each race. Therefore, for all eligible athletes, we checked if they took the next closest starting position in relation to the previous athletes (*next*) or not (*skip*). It has to be noted here that athletes know the starting positions selected by the previous athletes. Eligible for this procedure were all athletes between draw position 2 (as the first athlete has no previous athlete) and 8, if the last two remaining spots were the starting positions 1 and 19. Or between draw position 2 and 9, if otherwise, for example if the last two remaining spots were 17 and 19 and the 9th athlete still could choose to not pick the next closest starting position (as it is the case in Table 1). We then calculated a chi-square goodness of fit test for a uniform distribution for each gender and discipline. Cramer's V was calculated as effect size measure.

Pick optimality

For checking the optimality of the pick decisions, FIS (probably unintentionally) provided a quasi-experimental design: After the first ten drivers picked an odd starting number between 1 and 19, the ten drivers from the second drawing group are randomly assigned to the remaining even numbers until 20. This allows to analyze if starters from this second group that were assigned to a spot close to starting positions picked by the higher ranked athletes perform better compared to those starters of group two that were assigned to other starting positions. As displayed in Table 1, the first three athletes to pick all selected early starting spots (3, 1 and 5), whereas the last three remaining athletes of group one had to select the remaining odd spots, which were the three highest starting numbers available for group 1 (15, 17 and 19). Thus, if there is something like a sweet spot for starting positions, those drivers of the second group assigned to position 2, 4 and 6 could have an advantage over those drivers assigned to the starting positions 16, 18 and 20. Or generalized, on average those athletes of starting group 2 with starting positions in or close to the sweet spot should perform better than other athletes of starting group 2.

For our analysis, we determined the attractiveness A_{BIB} for each even starting position BIB between 2 and 18 for each race, by simply taking the average drawing position of the two neighboring starting positions (pos_{BIB-1} and pos_{BIB+1}):

$$A_{BIB} = \frac{pos_{BIB-1} + pos_{BIB+1}}{2}$$

E.g.: in Table 1 the attractiveness A_2 of starting position 2 is 1.5 as the starting position 1 was selected by the driver with drawing position 2 (pos_{BIB-1} = 2) and starting position 3 was selected by the driver with drawing position 1 (pos_{BIB+1} = 1). For A_{20} we did not use this average, as only one of the neighboring positions (19) can be picked by group 1, but the draw position of the athlete picking at number 19 (pos_{19}). Consequently, the sample description space for A_{BIB} is Ω =. {1, 1.5, 2, 2.5, ..., 10}.

As athletes of starting group 2 were not uniformly distributed over the respective starting positions between 2 and 20, we decided to control for the drawing position of each athlete by using the difference between the ranking and the drawing position (see $Rank_{Rel}$ Table 1) for the analysis of the pick optimality.

The actual investigation of the potential sweet spot was performed by dividing athletes into two groups per gender and discipline. Group "*SweetSpot*" included all athletes with an attractiveness value A_{BIB} smaller than 5.5, whereas group "*NoSweetSpot*" included all other athletes. 5.5 was chosen as it is the average value of the sample space of A_{BIB} described above and also led to the most even subsets for all categories: for Women's Downhill, *SweetSpot* consisted of 49.0% of the starters, for Women's Super G of 48.1%, for Men's Downhill of 47.2% and for Men's Super G of 48.5%. The two groups were then compared for gender and discipline using Mann-Whitney-U tests again, as the data was not normally distributed. For each comparison we also report the so-called common language effect size (*CL*), which has originally been introduced by McGraw & Wong (1992) and later adapted for ordinally scaled data by Vargha & Delaney (2000).

Table 1

Starting position BIB, drawing position Pos, final rank and the rank in relation to the drawing position ($Rank_{Rel}$) for athletes of starting group one and two (in italics) for the Women's Downhill competition in 2020 in Crans Montana. For starters of the second group, the attractiveness of the starting position is displayed as defined in formula I.

Name	BIB	Pos	A_{BIB}	Rank	$Rank_{Rel}$
Nicole Schmidhofer	1	2		2	0
Tamara Tippler	*2*	*20*	*1,5*	*7*	*-13*
Corinne Suter	3	1		5	4
Romane Miradoli	*4*	*19*	*2*	*4*	*-15*
Ester Ledecka	5	3		6	3
Nicol Delago	*6*	*13*	*3,5*	*23*	*10*
Federica Brignone	7	4		13	9
Michelle Gisin	*8*	*17*	*4,5*	*11*	*-6*
Stephanie Venier	9	5		1	-4
Breezy Johnson	*10*	*18*	*5,5*	*16*	*-2*
Kira Weidle	11	6		18	12
Joana Haehlen	*12*	*12*	*6,5*	*25*	*13*
Ilka Stuhec	13	7		10	3
Ramona Siebenhofer	*14*	*11*	*7,5*	*3*	*-8*
Elena Curtoni	15	8		8	0
Marta Bassino	*16*	*15*	*8,5*	*14*	*-1*
Lara Gut Behrami	17	9		30	21
Petra Vlhova	*18*	*16*	*9,5*	*DNF (37)*	*21*

Name	BIB	Pos	A_{BIB}	Rank	$Rank_{Rel}$
Francesca Marsaglia	19	10		9	-1
Nina Ortlieb	*20*	*14*	*10*	*12*	*-2*

This measure can be interpreted as the probability that a random athlete sampled from one group (in our case from group *SweetSpot*) has a better ranking than a random athlete sampled from the other group (*NoSweetSpot*).

Competitive balance: Development of the performance by starting group

Traditional measures of competitive balance, like the Gini coefficient or the Herfindahl-Hirschman-Index (*HHI*) are often difficult to interpret in the context of sport and are therefore often supplemented or even substituted by customized measures (see among others Berkowitz et al., 2011; Del Corral, 2009; or Weber et al., 2016). Especially to compare different samples, in our case two different periods, it makes sense to come up with customized measures that allow to apply common statistical inference analysis. This is why we decided to design a customized measure. However, to strengthen our findings we also deploy a traditional measure of competitive balance, the *HHI* for world cup points won (see next sub section).

To investigate the influence of the rule change on different starting groups, we defined five pools of athletes that belonged to the same group before and after the intervention (see Figure 1). Pool *OneOne* includes athletes with a draft position from 1 to 7, pool *TwoOne* athletes from 8 to 10, *TwoTwo* from 11 to 15, *ThreeTwo* from 16 to 20 and *ThreeThree* from 21 to 30. For each pool we calculated the average *rank* of all drivers starting in the respective period:

$$\overline{rank}(e) = \frac{1}{n_P}\sum_{i_P=1}^{n_P} rank_{i_P}(e)$$

Where n_P is the number of starters i_P of the respective pool participating in races in period e. As the distribution of ranks is not normally distributed, we calculated Mann-Whitney-*U* tests to compare the rankings before (group *Before*) and after (*After*) the change of the starting lottery. This was done for both genders separately as well as for each discipline. Again, we stated the common language effect size (*CL*) as the probability that a random athlete selected from period *After* has a better ranking than a random athlete from period *Before*.

Competitive balance: Impact on the distribution of world cup points

To investigate the development of competitive balance among athletes over time we utilized the *HHI*, with *n* being the number of athletes *i* appearing in the final standings of season s and p being the share of all world cup points in the respective speed discipline won by an athlete in a season:

$$HHI_{Athletes}(s) = \sum_{i=1}^{n} p_i^2(s)$$

For the calculation of calculation per period we performed two adaptions to the calculation of the respective share of points. First, we controlled for the different number of points available per season for each discipline. Further, it is not uncommon that an athlete is

Pick Position	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Group *Before*	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Procedure *Before*	Randomly assigned to BIBs 16-22							Randomly assigned to BIBs 8-15								Randomly assigned to BIBs 1-7 and 23-30														
Group *After*	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3
Procedure *After*	Draw odd BIBs between 1 and 19										Randomly assigned to even BIBs between 2 and 20										Randomly assigned to BIB 21-30									
Pick pool	*OneOne*							*TwoOne*			*TwoTwo*					*ThreeTwo*					*ThreeThree*									

Figure 1 Starting group for each drawing position before and after the introduction of the new starting lottery as well as the respective pool for further analysis.

not participating in all seasons of a period. Therefore, for each athlete *i* of all athletes *n*, who appeared at least once in period e, we calculated the average share of points pc achieved per season in this period:

$$HHI_{Athletes}(s) = \sum_{i=1}^{n} p_i^2(s)$$

Further, we investigated the skewness for the number of world cup points gained in the two periods two check our sample for polarization (Lee et al., 2015). To do so, we used the same adaptions for the share of points won as for the *HHI*. All calculations were performed separately for each gender and discipline.

Results

Pick behaviour

All Kolmogorov-Smirnov tests for the picks of the first athlete (BIB_{pos1}) revealed highly significant deviations from a uniform distribution. For Women's Super G 63.9% of the top picking athletes selected starting position 7 and for Women's Downhill 45.0% either picked starting position 7 or 9. Those were also the most selected starting positions with 61.8% for Men's Super G, whereas in Men's Downhill 51.2% picked either starting position 5 or 7 (Figure 2). Differences between the disciplines or genders were not significant (Table 2).

Athletes selecting their starting position second (pos_2) could choose to select any other odd starting number between 1 and 19 that was not picked by the first athlete. The second picking athlete selected the next closest possible starting number to the first athlete 65.0% of the time in Women's Downhill, 80.6% of the time in Women's Super G, 69.8% of the time in Men's Downhill and 76.5% of the time in Men's Super G (Figure 3). For the entity of eligible athletes that could select a starting position, Chi-square tests for goodness of fit proved that in general athletes did rather select the next closest starting number in relation to the previous picking athlete than any other starting number. This applies to both genders and disciplines. Overall, 72.3% and 78.5% selected the next closest number in Women's Downhill (χ^2 = 61.4, p < .001, V = 0.445) and Women's Super G (χ^2 = 90.6, p < .001, V = 0.570) competitions respectively. The results for Men's Downhill ($next_{\%}$ = 75.9%, χ^2 = 89.1, p < .001, V = 0.518) and Men's Super G ($next_{\%}$ = 78.7%, χ^2 = 84.9, p < .001, V = 0.574) showed similar numbers with strong effect sizes in both disciplines. Neither chi-square tests between disciplines nor between genders showed significant differences (Table 2).

Pick optimality

Mann-Whitney U tests revealed significantly better relative rankings for starters of group *SweetSpot* for Men's Super G ($mean_{SweetSpot}$ = 2.66, $mean_{NoSweetSpot}$ = 5.07, z = 2.04, p = .041, CL = 56.4%). For Men's Downhill the mean of group *SweetSpot* (3.88)

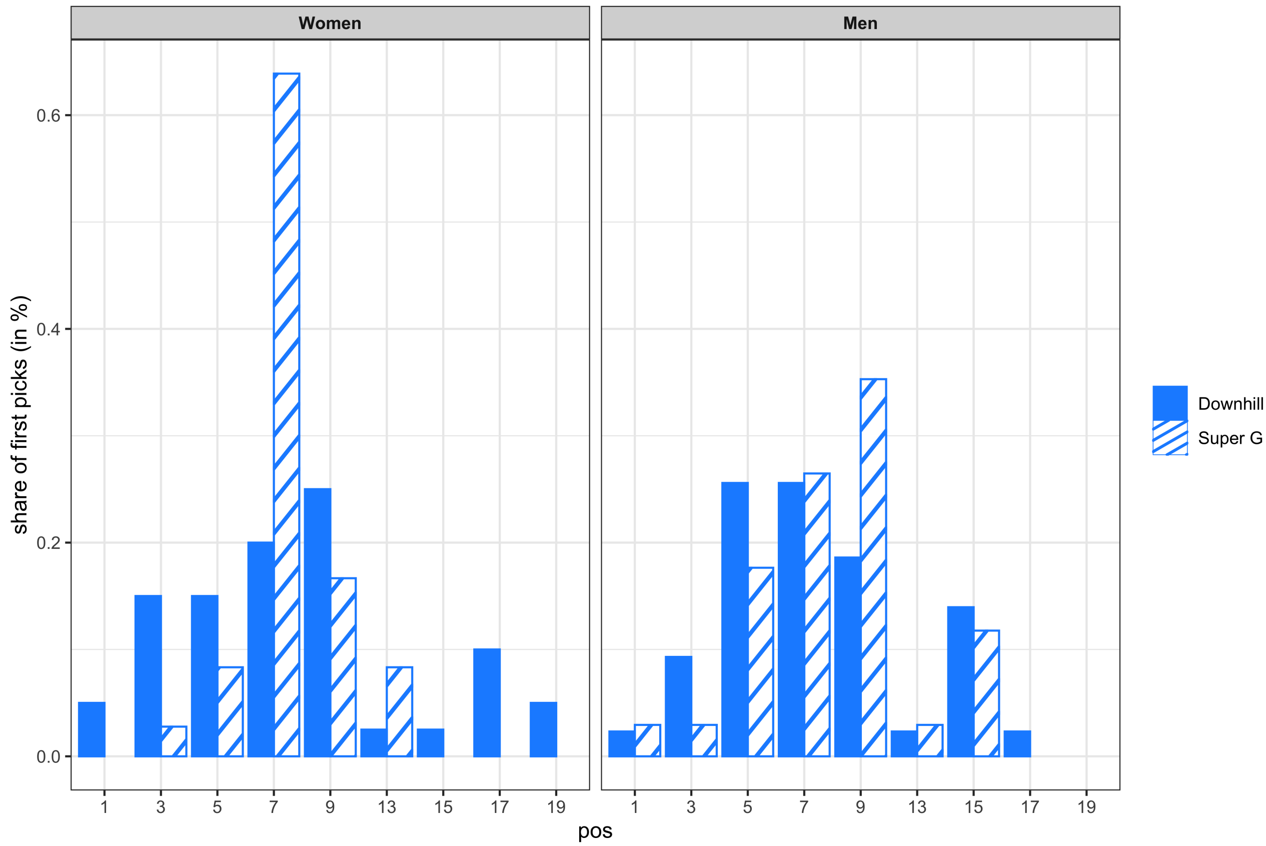

Figure 2 Frequency of starting positions picked by the first athlete by gender and discipline.

was also better compared to group *NoSweetSpot* (5.58) but the difference was not significant ($z = 1.17$, $p = .244$, $CL = 53.3\%$). For Women's Downhill the means of both groups were quite similar ($mean_{SweetSpot} = 1.20$, $mean_{NoSweetSpot} = 1.05$, $z = 0.65$, $p = .518$, $CL = 51.9\%$), whereas for Women's Super G group *SweetSpot* performed better on average, but the difference was not significant as well ($mean_{SweetSpot} = 0.94$, $mean_{NoSweetSpot} = 2.05$, $z = 1.18$, $p = .238$, $CL = 53.6\%$). The differences of the means for three of the four comparisons are relatively large. However, all groups showed a high grade of deviation, which is displayed via the respective boxplots in Figure 4. Still, common language effect size showed that random athletes sampled from group *SweetSpot* had a better ranking than random athletes sampled from the other group *NoSweetSpot* more often than not.

Influence of the starting lottery on the performance by starting groups

No significant differences were revealed by Mann-Whitney U tests for any of the starting pools (Figure 5) for Women's Downhill. For Women's Super G, athletes of pool *OneOne* (Pos_1 – Pos_7) and pool *ThreeThree* (Pos_{21} – Pos_{30}) performed significantly worse under the new drawing procedure (*OneOne*: $mean_{Before} = 10.2$, $mean_{After} = 12.7$, $z = 2.77$, $p = .005$, $CL = 42.9\%$; *ThreeThree*: $mean_{Before} = 22.3$, $mean_{After} = 24.8$, $z = 3.42$, $p <= .001$, $CL = 42.7\%$), whereas those from pool *ThreeTwo* (Pos_{16} – Pos_{20}) performed significantly better ($mean_{Before} = 21.7$, $mean_{After} = 18.1$, $z = 2.92$, $p = .002$, $CL = 58.8\%$).

Table 2

Results of the Kolmogorov-Smirnov tests of the pick of the first athlete (BIB_{pos1}) versus a uniform distribution and of the chi-square test to compare the distribution between different competition types.

	K-S test BIB_{pos1}	χ^2 test relative pick behavior
Women's Downhill	0.356 *(< .001)*	61.4 *(< .001)*
Women's Super G	0.472 *(< .001)*	90.6 *(< .001)*
Men's Downhill	0.370 *(< .001)*	89.1 *(< .001)*
Men's Super G	0.408 *(< .001)*	84.9 *(< .001)*
Women's: Downhill vs. Super G	0.239 *(.230)*	2.74 *(.098)*
Men's: Downhill vs. Super G	0.137 *(.869)*	0.49 *(.485)*
Downhill: Women's vs. Men's	0.127 *(.893)*	0.93 *(.335)*
Super G: Women's vs. Men's	0.250 *(.224)*	0.00 *(.999)*

The results for men were quite different. For Men's Downhill as well as for Men's Super G, starters of pool *OneOne* performed significantly better after the FIS changed the starting lottery (Downhill: $mean_{Before} = 12.9$, $mean_{After} = 10.5$, $z = 2.95$, $p = = .002$, $CL = 56.6\%$; Super G: $mean_{Before} = 15.9$, $mean_{After} = 11.6$, $z = 2.89$, $p = .002$, $CL = 57.5\%$). Similar to Women's Super G and Women's Downhill the ranking for athletes of pool *ThreeThree* became worse, see Figure 6, but the differences were not significant for both speed disciplines for male athletes.

Influence of the rule change on the distribution of world cup points

Figure 7displays the Herfindahl−Hirschman Index for the competitive balance among athletes for each season and the two total periods before and after the introduction of the new starting lottery. For female athletes the difference of the *HHI* of the periods and

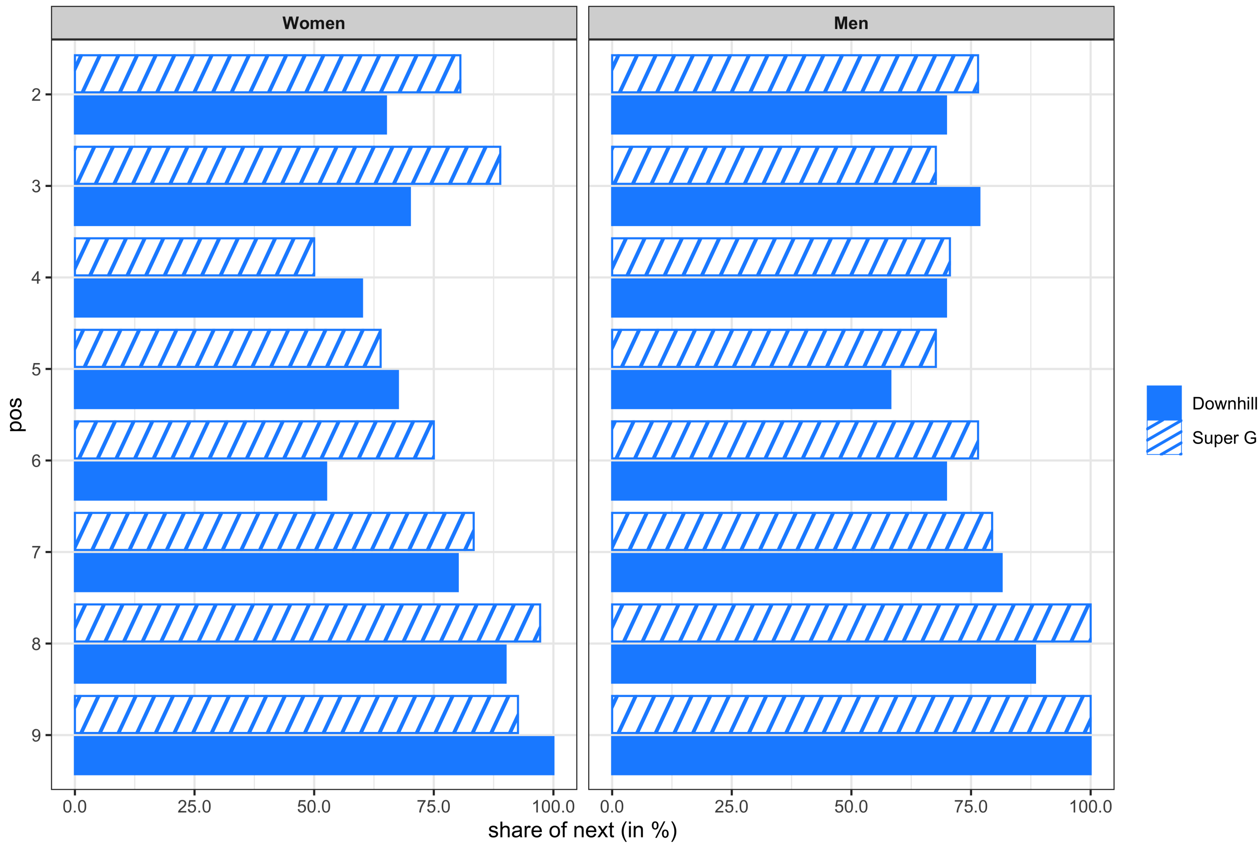

Figure 3 Share of eligible athletes that picked the next possible starting position for each draft position (pos), split by gender and discipline.

the respective seasons show on average lower values for period *After*, whereas for male ones the *HHI* shows a higher concentration for World Cup points for this period compared to *Before*.

We found high positive values for skewness for all groups, which developed similarly over time for all genders and disciplines with one exception. Whereas the skewness decreased for Men's Super G (*Before*: 2.17; *After*: 1.92), Women's Super G (*Before*: 2.00; *After*: 1.56) and Women's Downhill (*Before*: 1.51; *After*: 1.15), it displays a higher polarization for Men's Downhill competitions after the rule change (*Before*: 1.59; *After*: 2.06).

Discussion

This is the first study investigating the influence of the starting lottery on the outcome in ski alpine and the first study exploring how athletes pick a starting position, including the ability to identify favorable points of time to start. As the results of the study show, this is an intriguing area of research. Drivers showed a clear preference for the starting positions 5, 7 and 9, with over 73.9% of the first picking drivers selecting these numbers. In contrast, the first starting position was only selected 2.6% of the time and the last 5 eligible starting positions (11-19) only for 15.7% of the races. The results suggest that on the one hand athletes tend to avoid being among the first starters. This could indicate that athletes hope to gain additional information about the state of the slope at the time of competition, which might have changed since the course has been

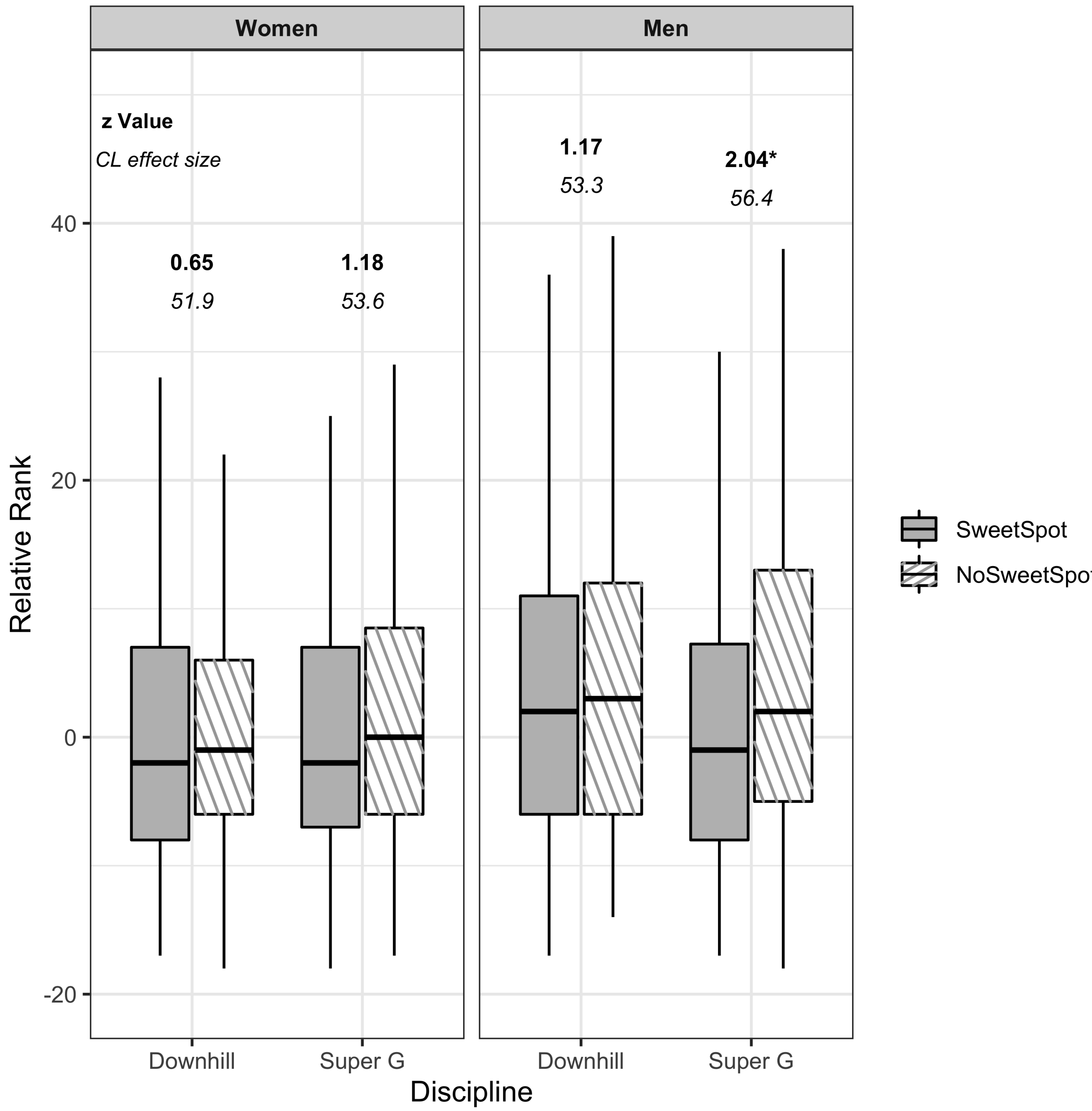

Figure 4 Boxplots of the relative ranks of starting group 2 after the rule change by gender, discipline and attractiveness group. *Z* Value is displayed in bold, *CL* effect size in italics. Significant differences between SweetSpot and NoSweetSpot on a 0.05 level are displayed with an asterisks (*).

inspected. This assumption is further supported with visibility, snow conditions and weather being performance influencing factors in alpine skiing (Gilgien et al., 2018; Supej & Cernigoj, 2006). On the other hand, the very influence of snow conditions could also be a reason why athletes do not want to start too late. It seems to be well-established practical knowledge that the track gets worse with each additional racer passing through, which is reflected in the starting order regulations for technical disciplines in alpine skiing. However, to our best knowledge there is no empirical proof

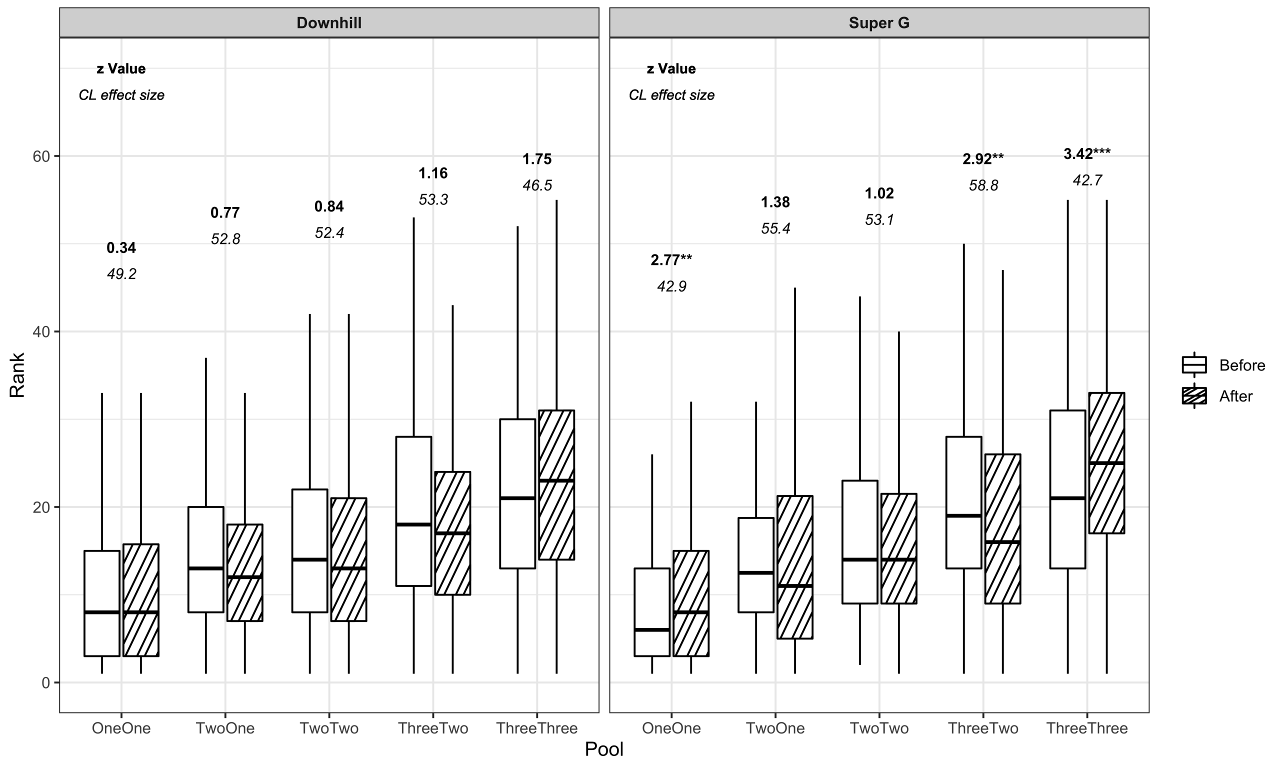

Figure 5 Boxplots of the ranks for each Pickgroup for Women's speed competitions for the periods Before and After. *Z* Value is displayed in bold, *CL* effect size in italics. Significant differences between Before and After on an 0.01 level are displayed with ** and as *** on an 0.001 level.

for this in the alpine speed disciplines. The only two studies addressing this issue were conducted in slalom competitions and suffer from very small sample sizes (Lešnik et al., 2013; Supej et al., 2005).

After the first skier had chosen her or his starting position, 72.5% of the athletes subsequently making their choice picked one of the next closest starting spots. Athletes picking at whatever position choose the next closest starting positions 76.1% of the time. Thus, there seems to be a perceived sweet spot for the best starting position in a race as well as for starting positions in general on which the athletes agree.

To prove if this spot was preferable compared to other starting spots, we investigated the average relative performance of starters seeded at the positions 11 to 20, as those were randomly assigned to the even numbers in-between the odd numbers selected by the 10 top seeded athletes. We found that only for Men's Super G those starters from this group starting close to the perceived sweet spot achieved significantly better results than the other starters of starting group 2. However, for downhill competitions (both genders) and Women's Super G the median relative rank was better for group *SweetSpot* as well. Common language effect size also indicates that it is also more likely that a random athlete sampled from group *SweetSpot* achieved a better ranking in all the four investigated disciplines. Renfree et al. (2015) stated for endurance competitions that athletes agree on strategies that are not beneficial. This seems not to be the case in alpine skiing Downhill and Super G competitions. However, it remains unclear whether athletes, in particular female ones, would be able to benefit even more from the opportunity of picking a starting position.

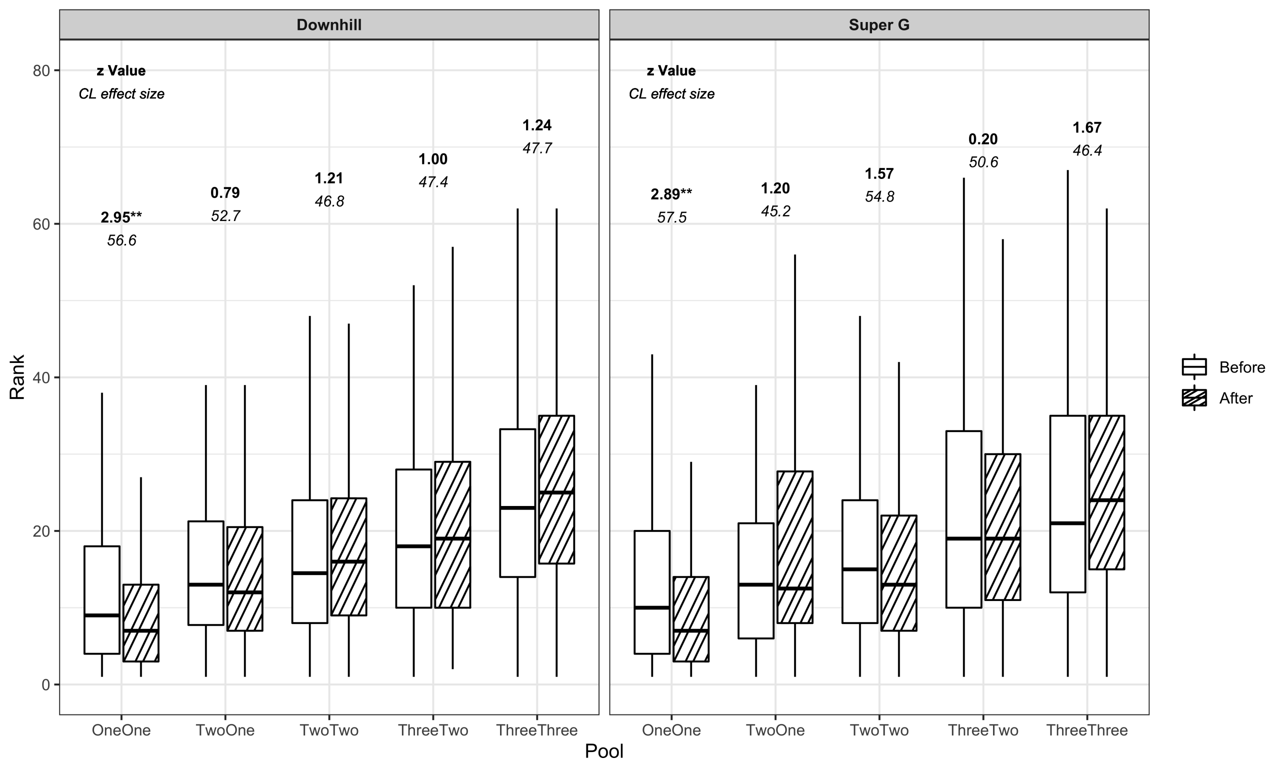

Figure 6 Boxplots of the ranks for each Pickgroup for Men's speed competitions for the periods Before and After. *Z* Value is displayed in bold, *CL* effect size in italics. Significant differences between Before and After on an 0.01 level are displayed with **.

According to these results, the impact of the new starting procedure on the achieved results of top athletes appears to be ambiguous. If the perceived sweet spot could have been actually proven to be a universal and significant sweet spot, one would have expected that the top athletes universally are able to achieve better rankings after the rule change. In particular as this sweet spot very often was different from top athletes' starting slots before the starting lottery was changed. Or to be more precise that in particular the top seven seeded drivers, referred to as group *OneOne* (part of the top group in both periods), perform better starting from the selected starting positions compared to the randomly assigned starting numbers between 16 and 22. This effect could be proven for Men's Downhill, in which the athletes of group *OneOne* improved by 2.4 ranks on average, and Men's Super G, where such athletes even improved by 4.3 ranks on average. Women's Downhill racers from this group did not perform different and for Women's Super G they performed even worse by 2.5 ranks on average.

Those results are further supported by the development of the *HHI*, which showed a higher concentration of world cup points among the top athletes in the speed disciplines after the rule change for men's competitions but not for women's competitions. In contrast, testing for the third statistical moment consistently revealed notable positive skewness values for both genders (and both speed disciplines), meaning that a few top racers tend to have won the major share of world cup points in both periods.

The partly different findings for men and women are in particular interesting as there are no significant differences regarding picking behavior. Neither for the tendency to select similar starting positions than previous

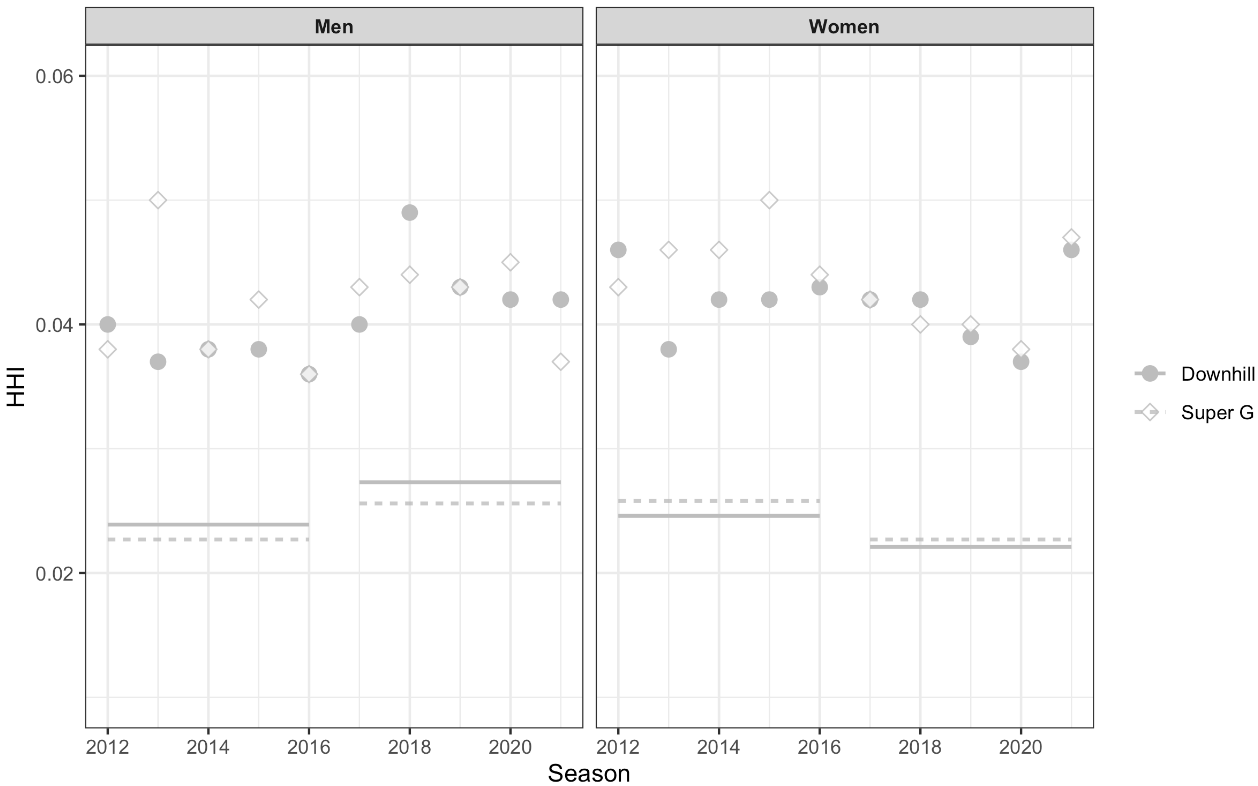

Figure 7 Herfindahl–Hirschman Index for the World Cup points collected by single athletes in one season (points) or one period (lines) by discipline and gender.

picking athletes nor for the preference shown for the starting position chosen with the first pick. For the latter, however, female athletes showed a strong preference for starting position 7 in Super G. They selected this starting number 63.9% of the time, whereas the next frequently picked starting position (9) was only selected by 16.7% of the athletes – a gap of 47.2%. The next biggest gap between the two most preferred starting numbers is 8.8% in Men's Super G (bibs 7 and 9). Conformity is in particular high for difficult and important tasks (Baron et al., 1996). Further, women show a higher tendency for conformity (Eagly & Carli, 1981). This could explain why female athletes more often choose a start position that is universally acknowledged as beneficial, even if another selection could possibly be more beneficial in the certain race. The other way around, research would also support that male athletes would tend to be less cautious in selecting other starting positions (Byrnes et al., 1999). These aspects could be considered as reasons why the top male athletes were able to profit more from the new starting lottery procedure. A caveat to accepting these explanatory aspects is, however, that the results showed high variations and the differences in picking behavior between of female and male racers were not significant.

Findings about the results of other starting groups besides the top seven athletes are also ambiguous. Only for one group of athletes a consistent pattern was found, namely the athletes of pool *ThreeThree*, which are racers with draw positions between 21 and 30. The median rank for this group was at least two ranks worse for each discipline for both sexes after the rule change. In advance of the change of the starting lottery, such athletes could be assigned to a start position between 1 and 7 or 23 and 30. After the change, they were assigned to start numbers between 21 and 30. Thus, instead of having a chance to start from what is perceived to be a beneficial starting position, such racers now only can start from later positions, which is supposed to have negative effects (see Lešnik et al., 2013; and Supej et al., 2005). Even though this explanation may appear plausible, a significant difference was only found for Women's Super G.

Based on the findings of our study we can only very cautiously draw conclusions about the influence of the new starting lottery on competitive balance in general. For male competitions after the rule change, we find significant better rankings for the top seven athletes and worse rankings, which were not significant, for the worst athletes in our sample (draw position 21-30). This could be interpreted as a shift towards less competitive balance, which is also supported by the increased values of the Herfindahl–Hirschman Index after the rule change.

For women, the same trend for athletes with drawing positions between 21 and 30 was found, even with the above-mentioned significant decrease of performance for Super G. However, the top seven female athletes did not perform better on average, actually they even performed worse in Super G, and we thus cannot draw the same conclusion for competitive balance for women's speed competitions as for the male racers. In addition, the *HHI* showed a lower concentration of world cup points among the top female athletes.

Overall, we had notable deviations in each of our subsamples, which is displayed in the boxplots that accompany the different analysis. These huge deviations could also be affected by the many cofounding factors that influence performance in alpine skiing. That we did not control for them is certainly a limitation of this study. Gilgien et al. (2018) as well as Supej & Cernigoj (2006) listed many environmental factors that influence performance in alpine skiing, such as the specific terrain, the course configuration, snow conditions, the weather and visibility. Further, each individual might react differently to the conditions based on their individual capacities as well as equipment. A limitation of the study is that it was not possible to control for these factors. One might imagine a study design for further studies in which for some of them could be controlled, e.g. the type of ski used, the terrain of the course or the weather. But it seems almost impossible to control for visibility conditions as those can change any second, e.g. based on clouds covering the sun or sudden befogging, or a kept secret by the teams like the used wax. We tried to control for these shortcomings by selecting a relatively large number of races and performances for this study.

Conclusion

Even with all these challenges to investigate performance in alpine ski racing in mind, we are confident that our study can expand the relatively little scientific knowledge in this area. The results of the study suggest that the revised starting lottery introduced in 2016/17 brought an unintended side-effect in terms of a shift in competitive balance for male competitions. The top seven racers achieved significantly better rankings on average, whereas the lower seeded drivers performed worse. The latter was also true for women's competitions. However, the top seven female athletes could not profit from picking their starting positions.

Furthermore, this has been the first study to show that top racers prefer early but not too early starting positions. One can assume that athletes want to gain additional information about the current conditions and at the same time want to avoid that the course is already worn off too much. Interestingly, even though this trend was found across speed disciplines for men as well as women, this perceived sweet spot could only significantly been proven for Men's super G.

Considering these different results, it appears reasonable for the FIS to discuss further changes of the starting lottery in the speed disciplines. As racers seem to agree that too early as well as too late starting positions are not beneficial, one could argue that it would be more fair to spread those racers out less and let them for example pick starting positions between 3 and 12. However, the goal of the FIS should not just provide a fair starting lottery to the top ten athletes but to all athletes. A radical approach would be to assign all starting numbers randomly without further restrictions, at least for the top 30. This would go in hand with the intent of the FIS to spread out the time the very top athletes start but could of course lead to very different conditions for athletes of the same performance level. As it cannot be guaranteed that bad luck in terms of disadvantageous starting positions is evening out over the course of a season, the FIS could consider to not assign the starting numbers for each race in a vacuum. The starting lottery can be adapted for each race, e.g. by making athletes that had rather less beneficial starting positions so far only eligible for starting positions that are perceived to be beneficial. Such a procedure could be suitable to fulfil FIS desire for spreading out the starting times for the top athletes but also ensure a certain level of fairness and competitive balance.

References

Abbiss, C. R., & Laursen, P. B. (2008). Describing and understanding pacing strategies during athletic competition. *Sports Medicine*, *38*(3), 239–252. https://doi.org/10.2165/00007256-200838030-00004

Änderung in Speed Disziplinen steht bevor [Changes in Speed Disciplines to Debut]. (2016). https://www.laola1.at/de/red/wintersport/ski-alpin/news/waldner-ueber-startprozedere-speed-bewerbe/

Arias, J. L., Argudo, F. M., & Alonso, J. I. (2011). Review of rule modification in sport. *Journal of Sports Science and Medicine*, *10*(1), 1–8.

Balfour, A., & Porter, P. K. (1991). The reserve clause in professional sports: Legality and effect on competitive balance. *Labor Law Journal*, *42*(1), 8.

Baron, R. S., Vandello, J. A., & Brunsman, B. (1996). The forgotten variable in conformity research: Impact of task importance on social influence. *Journal of Personality and Social Psychology*, *71*(5), 915–927. https://doi.org/10.1037/0022-3514.71.5.915

Berkowitz, J. P., Depken, C. A., & Wilson, D. P. (2011). When going in circles is going backward: Outcome uncertainty in NASCAR. *Journal of Sports Economics*, *12*(3), 253–283. https://doi.org/10.1177/1527002511404778

Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*, *125*(3), 367–383. https://doi.org/10.1037//0033-2909.125.3.367

Csató, L. (2021). A simulation comparison of tournament designs for the world men's handball championships. *International Transactions in Operational Research*, *28*(5), 2377–2401. https://doi.org/10.1111/itor.12691

Del Corral, J. (2009). Competitive balance and match uncertainty in grand-slam tennis: Effects of seeding system, gender, and court surface. *Journal of Sports Economics*, *10*(6), 563–581. https://doi.org/10.1177/2F1527002509334650

Eagly, A. H., & Carli, L. L. (1981). Sex of researchers and sex-typed communications as determinants of sex differences in influenceability: A meta-analysis of social influence studies. *Psychological Bulletin*, *90*(1), 1–20. https://doi.org/10.1037/0033-2909.90.1.1

Esteve-Lanao, J., Larumbe-Zabala, E., Dabab, A., Alcocer-Gamboa, A., & Ahumada, F. (2014). Running world cross-country championships: A unique model for pacing. *International Journal of Sports Physiology and Performance*, *9*(6), 1000–1005. https://doi.org/10.1123/ijspp.2013-0457

Flores, R., Forrest, D., & Tena, J. D. D. (2010). Impact on competitive balance from allowing foreign players in a sports league: Evidence from european soccer. *Kyklos*, *63*(4), 546–557. https://doi.org/10.1111/j.1467-6435.2010.00487.x

Fort, R., & Quirk, J. (1995). Cross-subsidization, incentives, and outcomes in professional team sports leagues. *Journal of Economic Literature*, *33*(3), 1265–1299.

Geenens, G. (2014). On the decisiveness of a game in a tournament. *European Journal of Operational Research*, *232*(1), 156–168. https://doi.org/10.1016/j.ejor.2013.06.025

Gilgien, M., Reid, R., Raschner, C., Supej, M., & Holmberg, H. C. (2018). The training of olympic alpine ski racers. *Frontiers in Physiology*, *9*, Article 1772. https://doi.org/10.3389/fphys.2018.01772

Goller, D. (2022). Analysing a built-in advantage in asymmetric darts contests using causal machine learning. *Annals of Operations Research*. https://doi.org/10.1007/s10479-022-04563-0

Groot, L. (2008). *Economics, uncertainty and european football: Trends in competitive balance*. Edward Elgar Publishing.

Hanley, B. (2014). Senior men's pacing profiles at the IAAF world cross country championships. *Journal of Sports Sciences*, *32*(11), 1060–1065. https://doi.org/10.1080/02640414.2013.878807

Haugen, K. K. (2008). Point score systems and competitive imbalance in professional soccer. *Journal of Sports Economics*, *9*(2), 191–210. https://doi.org/10.1177/1527002507301116

Haugen, K. K. (2016). Point score systems and football coaching secrecy. *Mathematics for Applications*, *5*(1), 11–20. https://doi.org/10.13164/ma.2016.02

Haugen, K. K., & Guvåg, B. (2018). Uncertainty of outcome and rule changes in european handball. *European Journal of Sport Studies*, *6*(1). https://doi.org/10.12863/ejssax6x1-2018x2

Janis, J. L. (1972). *Victims of groupthink*. Houghton-Mifflin.

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. https://doi.org/10.1177/1527002513496009

Kent, R. A., Caudill, S. B., & Mixon Jr., F. G. (2013). Rules changes and competitive balance in european professional soccer: Evidence from an event study approach. *Applied Economics Letters*, *20*(11), 1109–1112. https://doi.org/10.1080/13504851.2013.791010

Kolbinger, O. (2020). VAR experiments in the bundesliga. In M. Armenteros, A. J. Benítez, & M. A. Betancor (Eds.), *The use of video technologies in refereeing football and other sports* (pp. 228–245). Routledge.

La Croix, S. J., & Kawaura, A. (1999). Rule changes and competitive balance in japanese professional baseball. *Economic Inquiry*, *37*(2), 353–358.

Lee, Y. H., Jang, H., & Hwang, S. H. (2015). Market competition and threshold efficiency in the sports industry. *Journal of Sports Economics*, *16*(8), 853–870. https://doi.org/10.1177/1527002514556719

Lešnik, B., Axelsson, E. P., & Supej, M. (2013). Influence of the start number on elite alpine skiing competitors' results. *Kinesiologia Slovenica*, *19*(2), 17–27.

Mastromarco, C., & Runkel, M. (2009). Rule changes and competitive balance in formula one motor racing. *Applied Economics*, *41*(23), 3003–3014. https://doi.org/10.1080/00036840701349182

McGraw, K. O., & Wong, S. P. (1992). A common language effect size statistic. *Psychological Bulletin*, *111*(2), 361. https://doi.org/10.1037/0033-2909.111.2.361

Mühlberger, A., & Kolbinger, O. (2021). The serve clock reduced rule violations, but did not speed up the game: A closer look at the inter-point time at the 2018 US open. *Journal of Human Sport and Exercise*, *16*(3), 528–540. https://doi.org/10.14198/jhse.2021.163.05

Nordholm, L. A. (1975). Effects of group size and stimulus ambiguity on conformity. *Journal of Social Psychology*, *97*(1), 123–130. https://doi.org/10.1080/00224545.1975.9923321

Renfree, A., Crivoi do Carmo, E., Martin, L., & Peters, D. M. (2015). The influence of collective behavior on pacing in endurance competitions. *Frontiers in Physiology*, *6*, Article 373. https://doi.org/10.3389/fphys.2015.00373

Renfree, A., & Gibson, A. S. C. (2013). Influence of different performance levels on pacing strategy during the women's world championship marathon race. *International Journal of Sports Physiology and Performance*, *8*(3), 279–285. https://doi.org/10.1123/ijspp.8.3.279

Rodriguez, P., Kesenne, S., & Humphreys, B. (Eds.). (2020). *Outcome uncertainty in sporting events*. Edward Elgar Publishing. https://doi.org/10.4337/9781839102172

Supej, M., & Cernigoj, M. (2006). Relations between different technical and tactical approaches and overall time at men's world cup giant slalom races. *Kinesiologia Slovenica*, *12*(1), 63–69.

Supej, M., Nemec, B., & Kugovnik, O. (2005). Changing conditions on the slalom course affect competitiors' performances. *Kinesiology*, *37*(2), 151–158.

Swartz, T. (2007). Improved draws for highland dance. *Journal of Quantitative Analysis in Sports*, *3*(1), Article 2. https://doi.org/10.2202/1559-0410.1047

Vargha, A., & Delaney, H. D. (2000). A critique and improvement of the CL common language effect size statistics of McGraw and wong. *Journal of Educational and Behavioral Statistics*, *25*(2), 101–132. https://doi.org/10.2307/1165329

Weber, A. C., Kempf, H., Shibli, S., & Bosscher, V. (2016). Measuring competition in the olympic winter games 1992-2014 using economic indices. *Managing Sport and Leisure*, *21*(6), 399–420. https://doi.org/10.1080/23750472.2017.1304232

Acknowledgements

Funding

The authors have no funding or support to report.

Competing interests

The authors have declared that no competing interests exist.

Data availability statement

All relevant data are within the paper.