Universität St.Gallen

Media Attention and Betting Markets

Stefan Legge, Lukas Schmid

October 2015 Discussion Paper no. 2015-21

School of Economics and Political Science,
Department of Economics

University of St. Gallen

Editor:	Martina Flockerzi University of St.Gallen School of Economics and Political Science Department of Economics Bodanstrasse 8 CH-9000 St. Gallen Phone +41 71 224 23 25 Fax +41 71 224 31 35 Email seps@unisg.ch
Publisher:	School of Economics and Political Science Department of Economics University of St.Gallen Bodanstrasse 8 CH-9000 St. Gallen Phone +41 71 224 23 25
Electronic Publication:	Fax +41 71 224 31 35 http://www.seps.unisg.ch

Media Attention and Betting Markets[1]

Stefan Legge, Lukas Schmid

Author's address:

Stefan Legge
Swiss Institute for International Economics and Applied Economic Research (SIAW-HSG)
Bodanstrasse 8
CH-9000 St. Gallen
Phone +41 71 224 23 38
Fax +41 71 224 22 98
Email stefan.legge@unisg.ch

Lukas Schmid
Swiss Institute for International Economics and Applied Economic Research (SIAW-HSG)
Bodanstrasse 8
CH-9000 St. Gallen
Phone +41 71 224 23 38
Fax +41 71 224 22 98
Email lukas.schmid@unisg.ch

[1] For helpful comments, we are grateful to Stefano DellaVigna, Henry Farber, Reto Föllmi, Roland Hodler, Sam Kortum, Michael Lechner, Alexandre Mas, Christopher Neilson, Christian Rusche, Ferenc Szucs as well as seminar participants at the NBER Summer Institute, the Annual Meeting of the European Economic Association, the World Congress of the International Economic Association, the Annual Congress of the Swiss Society of Economics and Statistics, the Spring Meeting of Young Economists, the Warsaw International Economic Meeting, Princeton University, the University of St.Gallen, and the University of Zurich. Moreover, we thank Daniel Défago, Marcel Looze, and Ulf Seehase (all FIS) who kindly supported the collection of skiing data. We are also grateful to Judith Keel (Swissdox AG) and Annkathrin Heidenreich (University of St.Gallen) who supported the collection of data on media attention. Rino Heim, Jonathan Wyss, and Dominik Zurbuchen provided excellent research assistance. The paper was partly written while Legge visited UC Berkeley and Princeton University. He gratefully acknowledges their hospitality.

Abstract

This paper investigates whether biased media attention affects perceptions about future events. We use data on World Cup tournaments in alpine skiing for the period of 1992-2014 and exploit close races as a source of randomness for ranking positions. Our results document that ranking positions generate sharp discontinuities in media attention even in close races. However, both regression discontinuity and instrumental variables estimates reveal that biased media attention neither affects prices nor quantities in the betting market. We conduct a series of robustness tests to explore the sensitivity of our results.

Keywords

Betting, Rankings, Media Attention.

JEL Classification

D03, L83, M50.

1 Introduction

There is a large and growing body of literature documenting the media's impact on economic, political, and social outcomes (Della Vigna and La Ferrara, 2015; Strömberg, 2015). In particular, prior research has pointed out that media outlets play a key role in reducing complex information for consumers with limited time resources (Falkinger, 2007). Hence the media focus on certain events, groups, and individuals which in turn has been shown to affect decisions and behavior in various domains, namely which political candidates are elected (Epstein and Robertson, 2015), how politicians react to disasters (Eisensee and Strömberg, 2007), what family size people prefer (La Ferrara, Chong and Duryea, 2012), or what consumption bundles consumers demand (Bursztyn and Cantoni, 2012).

While this evidence suggests that the media can influence what individuals think of the past and the present, we know surprisingly little about its impact on what people think about the future. In order to study individual perceptions about future events, economists typically use betting markets because they provide an incentive-compatible way for individuals to truthfully reveal their expectations (Wolfers and Zitzewitz, 2006). However, several empirical studies document that behavioral biases or risk-loving preferences might lead to a situation in which prediction markets do not reveal true probabilities (Friedman and Savage, 1948; Kahneman and Tversky, 1979; Snowberg and Wolfers, 2010). Another important and unexplored explanation for why betting odds deviate from true probabilities, however, is that bettors make decisions using information they receive from the media. If media outlets allocate significantly more attention to some individuals than to others, expectations about future events related to these individuals might change. For example, bettors might overestimate future prospects of politicians who receive the lion's share of media coverage after a debate. Similarly, athletes who succeed in one tournament and are covered extensively by media outlets may be perceived more likely to succeed in subsequent tournaments. The methodological challenge researchers face when estimating such an effect of media attention on expectations is that the allocation of attention across individuals is not randomized but correlated with unobserved factors like skill or talent.

This paper offers a novel identification strategy to estimate the causal effects of media attention on betting markets by analyzing close competitions in World Cup alpine skiing from 1992–2014. For several reasons this setting provides a unique real-world natural experiment. First, there are considerable differences in how much media attention athletes receive. Those individuals at the top of the classification typically receive the bulk of attention. Second, unlike in most other settings with ranking schemes, our data set enables us to observe the same individuals numerous times. Hence we can compare the amount of attention they receive when they achieve or miss a top ranking position by including athlete-fixed effects. Third, in alpine skiing small time differences resulting from random shocks, like weather and snow conditions, can manipulate ranking positions. Hence luck is more prevalent in skiing than in other sports such as, for example, golf (Connolly and Rendleman, 2008). In close races it is often a tiny margin —a few hundredths of a second— that determines whether an athlete finishes first or second, third or fourth, or even sixth or tenth.[1] Assuming such small time differences to reflect random noise allows us to argue that those athletes who achieved a higher rank in a close race quasi-randomly received more media attention. Hence we can explore the causal effects of media coverage on betting market outcomes for these athletes.

Our findings suggest that top-ranked individuals receive 31.5% more media attention compared to athletes with an arguably similar performance who barely missed the highest ranking positions. We also document that betting odds *generally* decrease with both ranking positions and media attention. However, when focusing on *close races* for which ranking positions and thus media attention are allocated randomly, we find no discontinuity in betting market outcomes. This suggests that media attention does not cause betting market participants to change their beliefs about the future prospects of athletes. In contrast, we find that betting markets are efficient in the sense that observed odds match estimated true probabilities. Only when extending the sample to include athletes relatively far away from top ranking positions, we find a significant difference in both true probabilities and betting odds between top ranked and not top ranked athletes.

We conduct a series of robustness checks to assess the validity of our findings. First,

[1] Didier Cuche's last seconds in the Vancouver 2010 Olympics downhill race exemplify that small time differences can result in large differences in ranking positions. Lagging only 0.06 seconds behind leader Didier Défago, Cuche was second at the beginning of the last ten seconds. But Cuche did not optimally pass the last gate and finished sixth, trailing winner Défago by 0.36 seconds.

we show that there are no systematic differences in pre-determined covariates between those who achieve, for instance, a podium finish in a close race and those who barely miss it. This supports our empirical strategy that uses random top ranks as exogenous manipulation of media attention. Second, our results are shown to be robust to the choice of bandwidth for the definition of random top ranks. Third, we document that higher ranking positions do not causally change an athlete's risk-taking behavior or performance in subsequent races. Fourth, we point out that our main results from reduced form estimations do not change if we estimate instrumental variable regressions. Finally, we show that our core findings on media attention are not confined to the Swiss newspaper archive 'Swissdox' but generalize to the American-based 'NewsLibrary'. Furthermore, we document that the key results on media attention are similar in other fields of sports by analyzing Formula 1 races in the period of 1992–2014.

Our paper provides several novelties. First, to our knowledge, we are the first to provide causal evidence that rankings generate significant differences in media attention. After a close race, newspapers focus on those athletes who achieved the top positions in the ranking even if performance differences were tiny. A central explanation for why media outlets focus on individuals who achieved a top rank is that they compete for consumers in an information-rich economy (Falkinger, 2007, 2008). Recent advances in information technology have dramatically increased the supply of information. Newspapers and TV programs respond to this by focusing particularly on the most relevant and successful individuals. This form of biased media attention is relevant because previous research documents that it can affect election outcomes (Della Vigna and Kaplan, 2007; Epstein and Robertson, 2015) as well as consumer preferences (Gentzkow and Shapiro, 2010).

Second, we show that media attention has no effect on betting behavior. This finding adds to prior research on how individuals deal with rankings and ratings. Salganik, Dodds and Watts (2006) find that the availability of music ratings increases downloads of already successful songs. Similarly, Feenberg et al. (2015) show that NBER working papers which are listed first in the weekly newsletter receive substantially more views and citations. Our findings, however, suggests that the allocation of attention does not affect betting market outcomes such as odds or the number of bets placed. We argue that the incentive structure and efficiency of the market provide an explanation for this observation.

Third, our results show that professional sport athletes exhibit substantial serial correlation in their performance but do not change their risk behavior or performance following a top ranking position in one tournament. Previous research has pointed out that lagging behind in a ranking increases risk-taking and lowers final performance (Genakos and Pagliero, 2012). Furthermore, our findings contribute to a sizable literature on the so-called hot-hands effect (Gilovich, Vallone and Tversky, 1985; Green and Zwiebel, 2015; Miller and Sanjurjo, 2015). In particular, we find that one-time successes do not have a causal positive effect on subsequent performance. Finally, several studies show that the mere provision of a relative ranking position can affect student performance (Tran and Zeckhauser, 2012; Kuhnen and Tymula, 2012) as well as employee satisfaction (Card et al., 2012). The key difference between our work and these earlier studies is that we analyze the effect of rankings using a setting in which ranking positions are arguably randomized in close competitions. In addition, having more than twenty years of data allows us to observe the same individual multiple times in both the treatment and control group.

The paper proceeds as follows. Section 2 presents general information on World Cup alpine skiing as well as descriptive statistics on our dataset. This comprises a description of the data on media attention as well as the betting market. Section 3 discusses the problem of identifying the causal effect of media attention on betting market outcomes. We explain the concept of the quasi-random allocation of top ranks and how this can be used to overcome the identification problem. Section 4 presents our empirical findings on how media attention affects betting market outcomes. This includes a series of robustness checks and a discussion of the external validity of our findings. Finally, Section 5 concludes.

2 Data

Our data set provides an opportunity to study the behavior of World Cup athletes, media outlets, and bettors in an environment with large stakes and fierce competition.[2] In this section, we first provide background information on World Cup skiing competitions. Moreover, we describe our data set which includes information on race results, media attention and betting market outcomes for all tournaments.[3]

2.1 World Cup Alpine Ski Tournaments

The origins of alpine skiing competitions go back to the 1930s when European ski clubs, most prominently in Switzerland, Austria, and Germany, decided to organize races. However, these first attempts were individual events and most participants used to be from the respective host country, with only few athletes coming from abroad. In 1967, the *Fédération Internationale de Ski* (FIS) decided to bring these separate events together and launched the FIS World Cup. During the first years, it included only three disciplines: slalom, giant slalom, and downhill races. It was in 1974 when combined races were included, while super G was added to the FIS World Cup in 1983.

Today, alpine skiing competitions enjoy great popularity, particularly in Europe. The downhill race in Wengen (Switzerland), for instance, has been followed by a TV audience of over one million viewers in Switzerland (one eighth of the country's population) for each of the races between 2007 and 2012 (Ski World Cup Wengen, 2012). A similar appeal comes from the downhill and slalom races in Kitzbühl, each of which is watched by more than 1.3 million Austrians. In addition to large audiences, sizable prize money is also par for the course. Of all top-ten athletes in the season of 2012/2013, the prize money sums up to $4.4 million for men and $4.2 million for women (FIS 2013). However, the distribution of income in prize money is highly skewed. The highest income among male athletes was

[2] Klaassen and Magnus (2009) discuss the usefulness of sports data to examine behavioral questions. DellaVigna (2009) provides a summary of research documenting that behavioral biases may disappear among experienced individuals.

[3] While our data set on World Cup tournaments and media attention includes all races from 1992–2014, the analysis of betting markets is restricted the period 2006–2014 because we only have the respective data for this period.

$589,009 and among females it was $771,289. Number ten of the prize money ranking earned only $109,010 and $126,858, respectively. A considerable fraction of 76% of male and 80% of female athletes earned less than $50,000.[4]

The goal of alpine skiing is to slide down a race course in the fastest overall time. Each course consists of a series of gates. All of them have to be passed correctly, so that all athletes run the same course. The five disciplines differ in terms of the vertical and horizontal distance between the gates as well as the horizontal distance between start and finish. The differences can be illustrated by the traditional race in Wengen, Switzerland (see Figure A.1 in the Appendix). While the slalom has a total of about 130 gates on a horizontal distance of 1,200 meters (in the total of two runs), the 50 gates on the downhill course are distributed over more than 4,000 meters. Furthermore, the downhill race descends more than 1,000m from start to finish, in contrast about 400m for the slalom race. Differences in gate distances and descending gradients translate into differences in speed. The average speed of a downhill racer is about 100 km/h, while a slalom athlete usually achieves about 40 km/h.

2.2 Data on World Cup Skiing

We use a panel data set on 473 male and 428 female athletes in all 1,587 World Cup ski races for the period of 1992–2014. The data set includes information on whether an athlete finished the race, the exact result in hundredths of a second, as well as gender, age, and the discipline of competition.[5] The panel structure allows us to measure each athlete's performance in subsequent races. In total, our data set contains 23,761 observations when the unit of observation is an athlete in a specific race.

— Table 1 about here —

[4]Note that these prizes are large enough to incentivize athletes to exert high effort but not too large to cause one-time winners to reduce their subsequent efforts. Besides the prize money, top ranks in World Cup races can also lead to better sponsorship contracts. While there is no reliable data on sponsorship incomes, insiders estimate that in the case of top athletes, this source of income makes up three to four times the amount of prize money.

[5]Race times are actually measured more precisely than stated in official reports. For any time in hundredths of a second, the measurement was accurate at the millisecond level.

Table 1 reports descriptive statistics for the data we use in the empirical analysis. In part (I), we show information on all athletes competing in World Cup tournaments for the period of 1992–2014. Since we focus on top ranks, the sample is restricted to all athletes in the top fifteen. The share of athletes on the podium is 20.3% of the total number of observations. Observations for today's race are more numerous than observations for past and future performance because we only use outcomes within season, which results in missing values for the first and last race of the season for each race discipline. Furthermore, we cannot use observations at the end of the season because we would lack future performance and betting odds. Taken together, these account for around a fourth of the total observations. Finally, the number of observations is reduced because only 82% of those athletes who compete in the next race finish the race and get a positive race time. It is important to note that competition in alpine skiing is fierce. Only few junior athletes make it to the World Cup team and, among them, only a small group is successful. From our total sample, only about seven percent of the athletes ever won a race during their entire career. This suggests that only a small set of competitors is very successful over a lifespan which is in line with empirical evidence concerning the presence of superstars in music, entertainment, and academia (Rosen, 1981; Hamlen, 1991).[6]

2.3 Data on Media Attention

We complement our data set of individual World Cup tournament results with information on media attention and betting odds. For the former, we scraped data from the Swiss news database "Swissdox" for various time horizons before and after the race. Overall, the Swissdox database covers more than 200 newspapers, almost all of which for the entire time period of 1992–2014. Our search queries included an athlete's name and the time horizon of the search. In particular, we measure media attention by the number of articles mentioning the athlete's name. Part (II) of Table 1 shows descriptive statistics for the number of articles published at various points in time. Not surprisingly, there are more articles about a particular athlete the more we extend the time window.

[6]The most successful athletes in the history of alpine skiing are Ingemar Stenmark (Sweden, 1973–1989) with 86 victories and Annamarie Moser-Proell (Austria, 1969–1980) with 62 victories.

In Figure 1 we show the distribution of media attention across ranking positions. Panel (a) depicts the average number of newspaper articles that mention an athlete's name on the day after the competition. While winners are mentioned in 17.8 articles on average, athletes on position two and three get an average media presence of 13.0 and 11.2 articles, respectively. The average number of articles is considerably lower for other athletes, namely 5.4 for athletes on positions four to ten and 3.7 for athletes on position eleven to fifteen. The pattern in media attention is notably similar when focusing on media attention during the week or month following the race.

— Figure 1 about here —

Because the source of our media data, Swissdox, may be biased towards Swiss and German-speaking athletes, we repeat the scraping procedure for "Newslibrary", a US-based online news database. It includes more than 4,000 outlets. In the robustness section we use this data to show that we obtain similar regression results for the media attention, irrespective of which source of media data we use. In addition, Figure A.2 in the Appendix illustrates that the distribution of media attention across ranking positions is almost identical among newspapers covered by Newslibrary when compared to newspapers covered by Swissdox in Panel (a) of Figure 1. The reason why we use Swissdox data for our main analysis is that the Newslibrary search often finds mainly articles with ranking lists. In contrast, using newspapers in Swissdox, we find mostly specific reports on World Cup skiing. Since mere ranking lists are not the kind of media attention that we are interested in, we employ Swissdox data for our main analysis. Nevertheless, it is important to stress that we obtain similar distributions of media attention from both sources.

2.4 Data on Betting Behavior

In order to obtain information on betting market outcomes, we collected data from the world's largest internet betting exchange "Betfair". We focus exclusively on bets for a specific athlete to win the next race. This corresponds to about two thirds of total bets.[7] This set of data

[7] A fraction of 23.8% are bets that a specific racers achieves a top 3 finish, while the rest is a bet that pays off if athlete A achieves a better position than athlete B.

includes a total of 77,202 individual bet observations and is available for the period 2006-2014. Note that each individual observation corresponds to a bet offered for a specific event (e.g., athlete A to win the next tournament) with a specific odd. For our analysis, we use three different odds for each athlete-event combination.

— Figure 2 about here —

As illustrated in Figure 2, the betting market for the subsequent race opens after a given race t. At the beginning, a bet for each athlete who is likely to compete in race $t+1$ is offered at an *initial odd*. Over time, this odd changes due to new information (e.g., a competitor being injured) or because of altered demand for bets. If, for example, the demand for a bet on athlete A surges, the odd offered by the betting agency for this athlete will go down. At some point prior to the race $t + 1$, the betting market closes and we record the *final odd* for each athlete. Hence, we have an initial and a final odd for each athlete and event. In addition, we aggregate all individual odds for a specific event and athlete to obtain an *average odd*. In the robustness section, we use all three odds to investigate how media attention affects the betting market at different points in time.

The total betting volume in our data set is £1,132,230 (or about $1.7m). Panel (b) in Figure 1 illustrates the distribution of the average betting odds across ranking positions. Betting odds are what a bettor gets paid for a bet of one unit if the specific event of the bet realizes. For example, a betting odd of 1.2 on athlete A means that the bettor collects 1.2 pounds for a one-pound-bet if athlete A actually wins the competition. Similar to the distribution of media attention, we note substantial differences across ranks. While winners of the current race have an average odd of 6.5 in the subsequent race, a bet on one of the other two athletes on the podium pays off 12.1 times your bet. If you put your money on an athlete on position four to fifteen, your average odd is twice as high (23.0). As expected, those finishing on the top ranks in a tournament receive significantly more bets—both in terms of the number of bets and total money volume—in the subsequent tournament.

3 The Identification Problem

3.1 Selection on Observables

The central research question of this paper is whether media attention affects betting market outcomes. A naive way to test this hypothesis is to run an ordinary least squares regression, assuming selection on observables. If we possessed all variables that affect outcomes, we could simply use our data set and fit the empirical model:

$$Y_{i,t+1} = \phi_i + \tau \text{TOP\_RANK}_{i,t} + \mathbf{X}_{i,t}\, \phi + \varepsilon_{i,t} \tag{1}$$

where $Y_{i,t+1}$ denotes the outcome variable which can be media attention, performance, or betting market outcomes of athlete i in race $t+1$. Note that media attention is measured by the number of articles on the day after race t. The coefficient of interest, τ, indicates the impact of a top rank, either a victory or podium finish. Finally, $\mathbf{X}_{i,t}$ denotes a vector of athlete i's observed characteristics, ϕ_i is an athlete-fixed effect, and $\varepsilon_{i,t}$ the standard error clustered at the athlete level.

— Table 2 about here —

The results of estimating this model are shown in Table 2. We observe that both top ranks, podium and victory, are positively correlated with media attention on the day after a competition. Moreover, the estimates show a negative relationship between high ranking positions and average odds. For the risk and performance measures, we find that victory is associated with a higher probability to win the subsequent race. All these correlations are robust to the inclusion of several control variables as well as athlete-fixed effects. However, these results rely on the selection on observables assumption and should thus not be interpreted as causal effects.

3.2 Identification Strategy

The fundamental problem with estimating equation (1) is that ranking positions—including victories and podium finishes—are not randomly assigned across athletes. Although we can use a wealth of information to proxy for unobserved variables, including fixed effects to net out skill differences, we do not have information on, for example, injuries prior to the competition. In this section, we suggest a novel identification strategy to overcome this problem. The key idea of our approach is that in close races it is often a tiny margin that determines athletes' positions. If, for example, the time difference between two athletes ranked third and fourth is only a tenth of a second, we document that this can be attributed to random weather shocks. Hence, we can use those athletes that achieve a top rank (i.e., victory or podium finish) in a close race to estimate the causal effect on various outcomes. Throughout our analysis, we restrict the sample to races within season and discipline. This is necessary because times vary substantially across disciplines and seasons are separated by more than half a year.[8] When limiting our sample to close races, in which top ranks are arguably randomly assigned, we have to exclude most combined competitions and focus on slalom, giant-slalom, super-G, and downhill races.[9]

The focus of our analysis is on top ranking positions, namely victories and podium finishes. We motivate this by the fact that individual tournaments reward mainly athletes on the podium, and more specifically the winner of a race. This is reflected by FIS World Cup points, substantially higher prize money and increased media attention. In comparison with World Cup victories, however, podium finishes have a couple of advantages. On the one hand, we can draw on significantly more observations. Furthermore, when estimating the effect of a quasi-random victory, implicitly all observations in the control group finished on the podium. Hence, those athletes are also treated with a top rank, although the "treatment dose" is arguably lower. Thus, when using a victory as treatment, we only estimate the additional effect compared to a podium finish.

[8]Typically, the last race of a World Cup season is in March, while the first race of the next season takes place in October.

[9]In such races, time differences are larger and there is a higher variance over time. This is mainly because combined races tend to be longer and have a smaller group of starters, which makes competition less fierce. Furthermore, there are only about five combined races per year as opposed to the other disciplines with about eleven.

3.2.1 Random Top Ranks in Alpine Skiing

In contrast to other fields of sport, World Cup alpine skiing offers a unique feature that allows us to determine quasi-random top ranks. We illustrate this by a simple thought experiment. Assume there are only three variables that determine athlete i's final race time, $T_{i,t}$, in a given race t: the time-invariant skill of athlete i, denoted by θ_i, her training or fitness level, denoted by $\lambda_{i,t}$, and a noise parameter $n_{i,t}$ that captures all kinds of random shocks such as weather or snow conditions that can be heterogeneous or homogeneous across athletes.[10] This setting allows us to write the time of athlete i in race j as a function of a her skill and training levels as well as some random noise:

$$T_{i,t} := f(\theta_i, \lambda_{i,t}, n_{i,t}). \tag{2}$$

Moreover, her position in the final ranking, $P_{i,t}$, is a function of her own time as well as her competitors' times:

$$P_{i,t} := g(T_{i,t}, T_{s,t}) = g(\theta_i, \lambda_{i,t}, n_{i,t}, \theta_s, \lambda_{s,t}, n_{s,t}) \qquad \forall s \neq i \tag{3}$$

By means of this equation, we can illustrate why quasi-random top ranks are possible. Usually, skill differences explain most of the variation in ranking positions. This does not, however, imply that ranking positions are entirely driven by skill levels. Figure A.3 in the Appendix depicts a histogram of winners' and third-ranked athletes' ranking positions in the previous race. The fact that 40.3% of current winners and 22.6% of current third-ranked athletes achieved a podium in their past race documents positive serial correlation of our success measures. Yet the spread of the distribution reveals substantial variation in ranking positions. This challenges the idea that skill differences entirely determine ranking positions. In particular, if two athletes have almost identical race times, random fluctuations in the noise term become critical. Variations in $n_{i,t}$ can reduce athlete i's race time sufficiently to overcome skill and training deficits. In this way, a less skilled athlete can be lucky and draw

[10] We assume skill to be persistent. However, we take into account that injuries may affect athletes' ability to exploit their skills by including the time-variant fitness parameter.

a very low $n_{i,t}$ which enables her to achieve a better race time than a more skilled competitor.

For our estimation, the key identifying assumption is that the noise parameter $n_{i,t}$ has sufficiently large effects on individual race times in order to randomly assign relative ranking positions in close races. In skiing, the individual noise term, $n_{i,t}$, is comprised of several components. First, alpine skiing is an outdoor event and thus wind and weather conditions vary significantly over the course of a single race. Most notably changes in snow, wind, and sight alter individual prospects of success and can also lead to cancellation if race conditions are considered to be a serious risk for the athletes.[11] Yet, the mere presence of unstable external conditions does not lead to cancellation and is broadly accepted as a natural source of variation among competitors. The impact of random wind, weather, and snow conditions is amplified by the fact that individual race times critically depend on the performance in key sections of the course. An error in these sections not only leads to an immediate time loss but also affects speed, and thus time, in the following sections.

It is crucial for our analysis is whether there is any bunching of data around the thresholds which determine who wins a race or finishes on the podium. Following McCrary (2008) there should be a smooth distribution of observations around the cutoff. Otherwise there might be a distorting factor we need to address. Figure A.4 in the Appendix illustrates that the number of observations is in fact smooth at the cutoff for both treatments, victory and podium. This supports the assumption that athletes are not systematically located around the threshold.

Bandwidth Choice — What is a close race and what time difference can be considered random? An important identifying assumption of our research design is that treated and non-treated athletes are not systematically different with respect to pre-determined covariates. If finishing on the podium in a close race is driven by skills instead of luck, our approach would not allow us to assess the causal effects of quasi-random top ranks. To address this

[11] The following excerpt about the performance of U.S. competitor Bode Miller in 2009 illustrates the impact of wind. "Miller, a two-time overall World Cup winner, finished ninth Saturday as the Saslong downhill in Val Gardena, Italy, marked its 40th year. His performance was affected by a strong headwind that whipped up just as he and the other top contenders took the course." New York Times, 20. December 2009.

concern, we compare the characteristics of treated and non-treated athletes. We do this in two steps. First, by means of Figure A.5 in the Appendix, we show that athletes who win or finish on the podium are *in general* different from less successful athletes when comparing prior success. However, we find that there are no significant differences with respect to prior success when considering *close races*. This indicates that those who win or make it to the podium in a close race are not systematically more skilled. Once we restrict the sample to tournaments in which the time difference between successful and non-successful athletes is less than 0.15 seconds, the differences in pre-determined characteristics are insignificant.

— Table 3 about here —

In Table 3, we compare athletes on the podium with those who missed it by up to fifteen hundredths of a second. There is no significant differences in any of the observable athlete characteristics. The top-ranked athletes in close races are not more experienced, successful, or risk-loving than their contestants in the control group. Moreover, we find no difference in their competition, media attention prior to the race, or betting market outcomes for the race which determines who is in the treatment and control group. Importantly, treated and non-treated athletes are also not different in terms of the probability of competing in the following races, which rules out the possibility that lower ranked athletes are discouraged from running in subsequent races. However, it is important to note that treated and non-treated athletes obviously become systematically different if we extend the bandwidth. If we include athletes trailing the podium by a large time difference it is no longer plausible to consider the podium finish a result of 'luck'. Hence, we have to restrict the sample to observations with sufficiently small time differences in order to exploit 'random top ranks'. The decision to choose 0.15 seconds as the bandwidth for our estimations is the result of a trade-off: We can use more observations with a larger bandwidth but the allocation of ranking positions (and thus media attention) is only plausibly random for small bandwidths. As we show in Section 4.3, our specific choice of bandwidth does not affect the empirical findings.

3.2.2 Econometric Specification

For our estimation we assume that, except for the treatment, there is no reason why subsequent outcomes ($Y_{i,t+1}$) like media attention, performance, or betting odds should be a discontinuous function of the race time. We support this assumption using a large set of balance tests (cf. Table 3 and Figure A.5). Hence, any discontinuity in the outcome variable at the cutoff level c_t is identified as the causal effect of the treatment. We estimate the treatment effect τ by fitting the linear regression

$$Y_{i,t+1} = \phi_i + \tau D_{i,t} + \beta(T_{i,t} - c_t) + \gamma[D \times (T_{i,t} - c_t)] + \mathbf{X}_{i,t}\,\delta + \varepsilon_{i,t} \tag{4}$$

where ϕ_i is an athlete-fixed effect, $D_{i,t}$ indicates treatment (victory or podium), $[D \times (T_{i,t} - c_t)]$ allows for different slopes on each side of the cutoff, $\mathbf{X}_{i,t}$ is a vector of control variables, and $\varepsilon_{i,j}$ is the standard error term which we cluster at the athlete level. For the outcome variable, $Y_{i,t+1}$, we use media attention after race t, performance and risk-taking behavior in race $t+1$, and both betting odds and the number of bets in race $t+1$. Note that the inclusion of covariates could in principle improve the precision of the estimation (Frölich, 2007). We explored this possibility but did not find notable differences in the estimates. Moreover, squared and cubic terms of $(T_{i,j} - c_j)$ can be included to allow for a nonlinear relationship.[12]

4 Results

4.1 Effect on Media Attention

We first test whether media attention is affected by a top rank in a close race. Fitting the model specified in equation (4), we use the number of newspaper articles mentioning an

[12]The work by Hahn, Todd and van der Klaauw (2001) as well as Gelman and Imbens (2014) suggests to use local linear regression in an RD setting. In particular, high order polynomials of the forcing variable should not be used. Thus, we omit squared and cubic terms of $(T_{i,j} - c_j)$ in our baseline regressions. However, the results are not sensitive to the specification.

athlete's name on the day and during the week after the race. All our regressions include athlete-fixed effects and thus use within-athlete variation. Hence, in the estimation we compare the same athlete who achieves a high ranking position (i.e., victory or podium) in one race but not in the other. The great advantage of our data set is that we have enough observations to test whether the same person receives different levels of attention by the media if she performs only marginally better than her competitors.

— Table 4 about here —

Table 4 shows that those athletes who finish on the podium in a close race are mentioned about 31.5% more often than those who barely miss the top three ranks. For a close victory we find an additional 15.0% increase. Note that the point estimate for victory is considerably smaller than the estimate for a podium finish because most athletes in the control group also finished on the podium and thus benefited from increased media attention for top-ranked athletes. The positive impact of a high ranking position in close races is still present when counting all articles published during the seven days following the race as indicated in the second column in Table 4. By means of Figure 3, we can visualize the discontinuity in media attention around the cutoff time by plotting the average absolute media attention on the day after the competition.

— Figure 3 about here —

The difference in media attention between top-ranked and not top-ranked athletes is present even when considering only close races in which the time difference was tiny. These significant differences indicate that top ranks introduce a sharp discontinuity in *absolute* media attention. However, we also have to examine whether there are differences in *relative* media attention. Not only do athletes' sponsorship contracts depend on how many times they are mentioned compared to their competitors, it is very likely that relative media attention affects bettors expectation about who is going to win the next race. To test the effect on relative media attention, we define $m_{i,t}$ as athlete i's share of total media attention to all athletes within our preferred bandwidth of 0.15 seconds. For example, suppose racer A wins race t, racer B trails her by less than 0.15 seconds and all other athletes have a larger distance to the winner. Then we count all newspaper articles that mention racer A and we do the same for racer B. If the winner is mentioned in 60 articles and the second in 40 articles, we have $m_{A,t} = 0.6$ and $m_{B,t} = 0.4$, respectively. Formally, $m_{i,t} = A_{i,t}/\sum_{s} A_{s,t}$ with s containing all athletes who won race t (or achieved a podium finish) as well as all those missing the victory (or podium) by 15 hundredths of a second or less. Using this measure of relative media attention, we find a discontinuity of 11 percentage points at the threshold to the podium which corresponds to a 32% increase in media attention compared to the control group (see Table A.1 in the Appendix). The equivalent relative increase in media coverage for achieving a victory is 19%. Overall, these estimates strengthen the claim that top ranks causally affect media attention.

Media Attention Before and After a Race — Despite the fact that allocation to treatment in our case is arguably random for close victories and podium finishes, the question remains whether it is the very success that affects media attention. In order to investigate this, we compare the media attention of top-ranked and lower ranked athletes at various points in time before and after the race.

In the balance tests shown in Table 3, we observe that *before* the race, successful and non-successful athletes receive very similar levels of attention by the media. This is not surprising, since treatment and control group seem to be fairly balanced in terms of experience and prior success. However, *after* the race the control group—those who did not win the race or did not finish on the podium—receive significantly less media attention. Although all athletes in the treatment and control group show a very similar performance, media attention is tilted heavily in favor of those finishing on the top three positions of the ranking. A second observation is that, as expected, the difference in media attention is less pronounced when considering longer time periods. However, even when we examine all articles published in the week after a competition, we observe a significant gap between successful and non-successful athletes. This gap remains significant if we subtract articles published on the day after a race. In the long run, the gap in media attention subsides. This is shown in Table A.1 in the Appendix and driven by the fact that within a month the next tournaments took place.

4.2 Effect on the Betting Market

When estimating the effect of media attention on betting markets it is important to first test whether the true probability of athletes to succeed is altered by achieving a top rank. If one-time great successes have an effect on performance, we also expect a difference in betting odds because bettors update their beliefs about the future performance of athletes based on the ranking positions. In this case, it would not be possible to disentangle the effect of higher media attention from the increased performance effect. However, if rankings have no effect on athletes' performance, the effect on betting odds is likely to depend on the discontinuity in media attention introduced by the ranking scheme.

Effect on the True Probability — To test whether athletes respond to what ranking position they achieve in one tournament, we investigate their risk-taking behavior and performance in the subsequent race. First of all, it is not surprising to find a substantial amount of serial correlation in our data set. The correlation between today's distance to the podium (victory) and the average position in the next race is 0.23 (0.14).[13] The key question, however, is whether there is a discontinuity around the cutoff. We test this by plotting the average probability of achieving a victory in race $t + 1$ around the time cutoff for a victory and podium finish in race t. The variable capturing whether an athletes wins the next race is our preferred measure of performance because it corresponds to the very event for which bettors can place bets.

— Figure 4 about here —

The result indicate that there is no discontinuity at the threshold. Neither a victory nor a podium finish in a close race has a causal effect on the probability of winning the next race. One may argue that this finding is driven by the choice of the dependent variable. Very few athletes win the next race and thus winning the next race might be an imprecise measure of performance. To test whether the results are sensitive to the choice of the dependent variable, we use several alternative measures of performance including the probability of a

[13]The Appendix Figure A.6 provides a graphical illustration of this positive serial correlation of performance.

podium finish and the average position in the next race. The results confirm that professional ski athletes do not change their performance following a top position in the ranking.

We also investigate whether achieving a high ranking position in one tournament changes the risk-taking behavior of athletes in subsequent races. If rankings affect risk-taking, the interpretation of the performance result above would be more difficult because rankings would alter the composition of athletes who obtain a final race time and thus a performance measure. In the context of our study, it may be expected that athletes change their behavior and perception of risk after a quasi-random top rank. Athletes may misinterpret one-time victories as signals of high ability. As a consequence, they might act too ambitiously in subsequent races, leading to an increase in the probability of crashes. Hoelzl and Rustichini (2005), for example, find in an experiment that overconfidence becomes important when monetary payments are at stake. Yet our results indicate that athletes on a top ranking position are not more likely to finish the subsequent race. In addition, their position and the probability of finishing on the podium in the next race do not differ from those athletes who missed the top ranks (i.e., victory or podium) by a small margin. From an econometric point of view, this finding supports our estimation of the effect on performance without taking into account an attrition bias by using a principal stratification framework (Frangakis and Rubin, 2002).[14] Finally, we consider the overall time in race $t+1$ as a plausible outcome variable.[15] Again we find no difference between successful and non-successful athletes. The regression results reported in Columns 3 and 4 of Table 4 confirm that the point estimates of top ranks on risk-taking and performance are very close to zero and far from being statistically significant.

Before turning to the impact of top ranking positions on the betting market we briefly summarize our key findings. Whether an athlete wins or finishes on the podium in a close race has a substantial effect on the amount of media attention she receives. However, it does

[14]Consider, for example, the case in which a one-time success had a negative effect on the survival probability. This would indicate increasing risk-taking among successful (i.e., treated) athletes. We discuss this concern in detail in the Appendix 5.

[15]In contrast to using ranking positions in race $t+1$, the advantage of the race time is that we do not have to address problem of potentially violating the stable unit treatment value assumption (SUTVA) which is necessary for the estimation of causal effects. In Appendix C we provide a detailed discussion.

not affect her risk strategy and performance in the next race. Hence, if betting odds reflect true probabilities there should not be any significant difference between top ranked and lower ranked athletes.

Effect on the Betting Market — Given the large positive effect of top ranking positions on media attention, it appears likely that public expectations about the performance of top-ranked athletes in the next race increase. A natural way to test whether public expectations discontinuously rise as a consequence of an exogenous shift in media attention is the analysis of betting data. Information on betting behavior should reflect prior expectations of bettors in an incentive-compatible way because a betting agency that deviates from bettors' expectations would either incur losses (if betting odds are too high) or attract no bets (if odds are too low). We use betting odds from Betfair, the world's largest Internet betting exchange, for the period of 2006–2014 to explore the impact of top ranks in close races on betting market outcomes.

The distribution of betting odds across ranking positions is shown in Panel (b) of Figure 1. We observe a positive gradient with better ranked athletes facing lower odds in the next race. There is also a pronounced difference for close winners and non-winners as well as between podium and non-podium finishers. However, we want to examine whether an athlete who randomly achieved a top rank faces different odds in the subsequent race.

Fitting the empirical model of equation (4) with betting odds and number of bets as dependent variables, we obtain the results reported in columns 5 and 6 in Table 4. For both outcomes, the estimates are very close to zero and fall short of conventional significance levels. These results are consistent with a graphical inspection provided by Figure 5.

— Figure 5 about here —

Based on these findings there are two conclusions. First, athletes who achieve a top ranking position in a close race receive higher subsequent media attention but not a significantly higher number of bets. The results remain unchanged if we use the volume of bets instead of using the total number of bets. The second conclusion from our estimation is that the absence of any discontinuity in betting odds matches the fact that randomly assigned higher ranking positions do not increase the true probability of winning the next race. In this sense, columns 5 and 6 of Table 4 support the idea of using markets for predictions (Wolfers and Zitzewitz, 2006).

Betting Market Efficiency — We use our data set and investigate in more detail to what extent betting odds reflect true probabilities. In panel (a) of Figure 6, we show the estimated effect of finishing on the podium on the true probability of achieving a victory in the next race. The estimation is conducted for all bandwidths between 0.05 and 4.00 seconds.

— Figure 6 about here —

We observe that once we include athletes trailing the podium by more than 1.5 seconds, those who finished on the podium are significantly more likely to win the subsequent race. Interestingly, panel (b) shows that there is a significant difference in (inverse) betting odds between athletes on the podium and other athletes once the bandwidth is larger than 1.5 seconds. These results suggest that the betting market closely mimics true probabilities of future events. Figure 6 also reveals a relatively constant gap between true probabilities and betting probabilities. This difference includes the betting agency's profit margin as well as the empirical regularity that favorites are underbet in betting markets. This phenomenon is also known as long-shot bias (Snowberg and Wolfers, 2010).

4.3 Robustness Checks and External Validity

In this subsection, we discuss the robustness of our main results. In a first step, we investigate whether top ranks have an effect on betting odds at specific points in time of the betting process. Second, we discuss the bandwidth choice in our regression discontinuity design. Third, we explore whether we obtain the same results in an instrumental variable framework using betting market behavior as outcomes, media attention as treatment, and top ranking positions as instrument. Finally, we address concerns about our source of media data.

Initial, Average, Final Odds — To shed light on the betting agency's behavior and the bettors' corresponding response, we analyze initial and final betting odds. First, it might be that the estimated effect on the average odd is driven by differences in the initial odds which are exclusively determined by the betting agency. These odds change over time until the next race starts. The changes in odds are driven by new information (e.g., news about an athlete being handicapped) as well as the number of bets placed. If more and more bettors want to buy bets that athlete i will win the next race, her odds are likely to decrease. We have data not only on the initial odd but on all odds that were offered for a given individual and race. Hence we can separate the initial odd, the average odd and the final odd. This wealth of information helps us understand whether the betting agency attempts to increase its profit by offering lower odds as a result of higher media attention. We can also investigate whether such a strategy is successful or whether the market adjusts the price. Table A.1 in the Appendix shows that the betting agency offers 'fair' odds in the first place. The point estimates for the average and final odds are very similar to those we obtain using the initial odds. This finding supports the idea that bettors' expectations are not biased as a result of selective media attention.

Bandwidth Choice — Throughout our empirical analysis we use the concept of random top ranks to identify the causal effects of media attention. As we explained in Section 3, it is crucial to focus on close races to overcome the identification problem. A central question in this regard is what time difference between two athletes can be attributed to random shocks. This means that it remains a priori unclear what bandwidth we should use in our estimations. In all regressions so far we have used a bandwidth of 15 hundredths of a second. The bandwidth choice was primarily based on the results of the balance tests in Table 3 as well as on the comparison of prior success for top-ranked and other athletes depicted in Figure A.5 in the Appendix. We can illustrate the magnitude of 0.15 seconds by plotting the distribution of time differences to the podium. Figure A.7 in the Appendix shows the distribution as well as a vertical line for the bandwidth we use in our estimation. When restricting the sample to those athletes trailing the podium by 0.15 seconds or less, only 10%

of the sample are included.

In order to investigate the robustness of our empirical results, we re-run the RDD estimation (equation 4) using different bandwidths ranging from 0.10 to 0.50 seconds. In our preferred specification we include athlete-fixed effects to hold constant all individual-specific covariates. Figure A.8 in the Appendix depicts the effect of a podium finish on media attention, performance, and betting odds using different bandwidths. There are two notable observations. First, the point estimate is very stable irrespective of the bandwidth choice. Second, when decreasing the sample size the confidence intervals become very large. Using a bandwidth of 0.10 seconds, for example, leaves us with only 304 observations in the betting odds estimation that includes athlete-fixed effects. Our preferred bandwidth of 0.15 seconds is the result of the trade-off between bias and precision: On the one hand we can use more observations with a larger the bandwidth. On the other hand the allocation of ranking positions (and thus media attention) is only plausibly randomized for small bandwidths.

Instrumental Variable Estimation — Thus far we have analyzed the effects of media attention on the betting market using separate regressions that relate randomized top ranks to media and betting outcomes. While top ranks create sizable discontinuities in media attention, the results of the intention-to-treat regression (or reduced form) suggest that top ranks do not create a difference in betting odds. This indicates that the effect of interest, the parameter on media attention in the second-stage regression with betting market outcomes as dependent variable, is absent (Angrist and Krueger, 2001). To explore this (absent) effect in more detail, we estimate a two-stage least-squares regression using our betting market variables as outcomes, media attention as treatment, and top ranks as instrument. We focus on close podium finishes and extend the bandwidth to half a second to avoid suffering from weak instrument problems.

— Table 5 about here —

Table 5 reports the results of these regressions using the average odds as well as the total number of bets in the subsequent race as dependent variables. All regressions include the distance to the podium as well as athlete-fixed effects. The estimates highlight that the effect of the media on betting market outcomes is very close to zero and not significant. These findings add to the previous results using reduced form regressions suggesting that there is no effect of media attention on betting behavior.

Media Data from NewsLibrary — In order to examine the robustness of our empirical findings about the effect of rankings on media attention, we address the fact that our source of media data (SwissDox) is a Swiss-based, largely unknown source. An alternative source is the American newspaper archive NewsLibrary. For all top-15 athletes in all races between 1992 and 2014 we use NewsLibrary and count the number of articles mentioning the athletes names. As before, we count the articles before and after the race over different time periods. We obtain a distribution of media attention across ranks that is very similar to the one that we observed for the SwissDox data. There is a notable gap between the winner and the runner-up as shown in Figure A.2 in the Appendix. Even more noticeable is the difference between the amount of media attention received by the third-ranked athlete compared to those athletes who missed the podium. We can use the data from NewsLibrary to repeat the OLS and RDD estimation. The estimates are very similar to the ones we obtained when comparing athletes around the podium cutoff as we document in the Appendix Figure A.9.

External Validity — In order to explore whether our findings in World Cup alpine skiing are specific to skiing, we collected data on all Formula 1 races for the period 1992–2014. For each race, we used Swissdox to obtain the media attention for each racer in the top eleven.[16] Results shown in Figure A.10 in the Appendix show that the distribution of media attention across the ranking positions is very similar in Formula 1 races as in World Cup ski races. When counting the number of articles published on the day after the race, there is a significant gap between the winner and the runner-up as well as between the third- and the fourth-ranked athlete. These gaps persist for at least a week after a tournament. The obvious problem with Formula 1 races, however, is that differences in final race times between ranking positions are usually too large for estimating causal effects of top ranks. In addition,

[16]Note that in Formula 1, usually only about ten contestants actually finish a race. All the other racers do not finish due to crashes or technical defects.

athletes are well aware of time differences towards the end of a Formula 1 race and respond to them. If, for example, the leader is twenty seconds ahead of the runner-up he will slow down in the remaining laps to avoid a crash or technical problems while still finishing first. The distorted time differences in Formula 1 complicate the identification of the causal effect of media attention on betting odds using apparently 'close' races.

5 Conclusion

This paper investigates how media attention affects the betting market. In a first step, we use a novel data on all World Cup tournaments in alpine skiing between 1992 and 2014 to show that media attention is highly skewed in favor of successful athletes. Even if performance differences are tiny, there is a significant gap in media attention between athletes on the podium and those athletes who miss it. We exploit this discontinuity in media attention to estimate the causal effect on betting market outcomes.

Our results reveal that ranking positions significantly affect the amount of media attention individuals receive after a tournament. Although prior theoretical and empirical work suggests that top ranks increase athletes' self-confidence and goal-setting behavior, we find no effect of high ranking positions on subsequent performance or risk-taking behavior. Since the true probability of winning the next tournament is not affected by a top rank in a close race, we expect to find no difference in betting odds for athletes with different amount of media attention if the betting market is efficient. Using data from Betfair, we find that increased media attention has neither an effect on average odds nor on the number of bets. The betting agency offers initial odds that reflect the unchanged true probability of athletes succeeding in the subsequent tournament. Moreover, we document that betting odds do not adjust over time.[17] While Thaler and Ziemba (1988) suggest that bettors can achieve a positive rate of return by placing bets on extreme favorites, we find no such opportunity. Our results suggest, in contrast to Levitt (2004), that bookmakers do not gain from being

[17]The empirical result that betting odds are announced and subsequent adjustments are relatively small is in line with previous findings by Levitt (2004).

more proficient at predicting future events than bettors.

Our findings with respect to the distribution of media attention are consistent with theoretical studies on markets for attention (Falkinger, 2007, 2008). In an information-rich world, media have to concentrate information and sport athletes compete for the scarce resource of media attention. Since athletes in World Cup tournaments draw a large share of their earnings from sponsorship contracts, being among the top-3 is of particular importance to get higher media attention which translates into better sponsorship deals.

Overall, our findings add to the growing literature on the economic impact of the media (Della Vigna and La Ferrara, 2015). We show that media attention is highly biased in favor of successful individuals but this bias neither affects prices nor quantities in the betting market. While our empirical approach adds credible causal evidence on the effects of the media using field data, exploring the relationship between attention and expectations about future events in different settings or in a lab environment appears to be a fruitful area for future research.

References

Angrist, Joshua D., and Alan B. Krueger. 2001. "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments." *Journal of Economic Perspectives*, 15(4): 69–85.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association*, 91(434): 444–455.

Bursztyn, Leonardo, and Davide Cantoni. 2012. "A Tear in the Iron Curtain The Impact of Western Television on Consumption Behavior." *Review of Economics and Statistics*, forthcoming.

Card, David, Alexandre Mas, Enrico Moretti, and Emmanuel Saez. 2012. "Inequality at Work: The Effect of Peer Salaries on Job Satisfaction." *American Economic Review*, 102(6): 2981–3003.

Connolly, Robert A., and Richard J. Rendleman. 2008. "Skill, Luck, and Streaky Play on the PGA Tour." *Journal of the American Statistical Association*, 103(481): 74–88.

DellaVigna, Stefano. 2009. "Psychology and Economics: Evidence from the Field." *Journal of Economic Literature*, 47(2): 315–372.

Della Vigna, Stefano, and Eliana La Ferrara. 2015. "Economic and Social Impacts of the Media." *CEPR Discussion Paper No. 10667.*

Della Vigna, Stefano, and Ethan Kaplan. 2007. "The Fox News Effect: Media Bias and Voting." *The Quarterly Journal of Economics*, 3(122): 1187–1234.

Eisensee, Thomas, and David Strömberg. 2007. "News Droughts, News Floods, and U.S. Disaster Relief." *Quarterly Journal of Economics*, 122(2): 693–728.

Epstein, Robert, and Ronald E. Robertson. 2015. "The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections." *Proceedings of the National Academy of Sciences*, 112(33): E4512–E4521.

Falkinger, Josef. 2007. "Attention economies." *Journal of Economic Theory*, 133(1): 266–294.

Falkinger, Josef. 2008. "Limited Attention as a Scarce Resource in Information-Rich Economies." *The Economic Journal*, 118(532): 1596–1620.

Feenberg, Daniel R, Ina Ganguli, Patrick Gaule, and Jonathan Gruber. 2015. "It's Good to be First: Order Bias in Reading and Citing NBER Working Papers." *NBER Working Paper #21141.*

FIS - International Ski Federation. 2012. *Audi FIS Alpine Skiing World Cup Prize Money Ranking Season 2012/2013.* Oberhofen, Switzerland.

Frangakis, Constantine E., and Donald B. Rubin. 2002. "Principal Stratification in Causal Inference." *Biometrics*, 58(1): 21–29.

Friedman, Milton, and L. J. Savage. 1948. "The Utility Analysis of Choices Involving Risk." *Journal of Political Economy*, 56(4): 279–304.

Frölich, Markus. 2007. "Regression Discontinuity Design with Covariates." *IZA Discussion Paper*, No.3024.

Gelman, Andrew, and Guido Imbens. 2014. "Why High-order Polynomials Should not be Used in Regression Discontinuity Designs." *NBER Working Paper #20405.*

Genakos, Christos, and Mario Pagliero. 2012. "Interim Rank, Risk Taking, and Performance in Dynamic Tournaments." *Journal of Political Economy*, 120(4): 782–813.

Gentzkow, Matthew, and Jesse M. Shapiro. 2010. "What Drives Media Slant? Evidence From U.S. Daily Newspapers." *Econometrica*, 78(1): 35–71.

Gilovich, Thomas, Robert Vallone, and Amos Tversky. 1985. "The Hot Hand in Basketball: On the Misperception of Random Sequences." *Cognitive Psychology*, 17(3): 295–314.

Green, Brett S, and Jeffrey Zwiebel. 2015. "The Hot-Hand Fallacy: Cognitive Mistakes or Equilibrium Adjustments? Evidence from Major League Baseball."

Hahn, Jinyong, Petra Todd, and Wilbert van der Klaauw. 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica*, 69(1): 201–209.

Hamlen, William A. 1991. "Superstardom in Popular Music: Empirical Evidence." *The Review of Economics and Statistics*, 73(4): 729–733.

Hoelzl, Erik, and Aldo Rustichini. 2005. "Overconfident: Do You Put your Money on it?" *The Economic Journal*, 115(503): 305–318.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect theory: An analysis of decision under risk." *Econometrica*, 47(2): 263–292.

Klaassen, Franc J.G.M., and Jan R. Magnus. 2009. "The efficiency of top agents: An analysis through service strategy in tennis." *Journal of Econometrics*, 148(1): 72–85.

Kuhnen, Camelia M., and Agneieszka Tymula. 2012. "Feedback, Self-Esteem, and Performance in Organizations." *Management Science*, 58(1): 94–113.

La Ferrara, Eliana, Alberto Chong, and Suzanne Duryea. 2012. "Soap operas and fertility: Evidence from Brazil." *American Economic Journal: Applied Economics*, 4(4): 1–31.

Levitt, Steven D. 2004. "Why are gambling markets organized so differently from financial markets?" *Economic Journal*, 114(3): 223–246.

McCrary, Justin. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of Econometrics*, 142(2): 698–714.

Miller, Joshua B, and Adam Sanjurjo. 2015. "Surprised by the Gamblers and Hot Hand Fallacies? A Truth in the Law of Small Numbers." *IGIER Working Paper #552.*

Rosen, Sherwin. 1981. "The Economics of Superstars." *The American Economic Review*, 71(5): 845–858.

Salganik, Matthew J., Peter Sheridan Dodds, and Duncan J. Watts. 2006. "Experimental Study of Inequality and Cultural Market." *Science*, 311(5762): 854–856.

Ski World Cup Wengen. 2012. *Fakten und Zahlen.* Wengen.

Snowberg, Erik, and Justin Wolfers. 2010. "Explaining the Favorite - Long Shot Bias: Is It Risk-Love or Misperceptions?" *The Journal of Political Economy*, 118(4): 723–746.

Strömberg, David. 2015. "Media Coverage and Political Accountability: Theory and Evidence." In *Handbook of Media Economics.* , ed. Simon Anderson, David Strömberg and Joel Waldfogel, 1–26. Amsterdam:North-Holland.

Thaler, Richard H., and William T. Ziemba. 1988. "Anomalies: Parimutuel Betting Markets: Racetracks and Lotteries." *The Journal of Economic Perspectives*, 2(2): 161–174.

Tran, Anh, and Richard Zeckhauser. 2012. "Rank as an inherent incentive: Evidence from a field experiment." *Journal of Public Economics*, 96(9-10): 645–650.

Wolfers, Justin, and Eric Zitzewitz. 2006. "Prediction Markets in Theory and Practice." *NBER Working Paper #12083.*

Tables and Figures

Table 1: Descriptive Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
I. World Cup Alpine Skiing:					
Position	7.96	4.32	1	15	23,761
Victory	0.07	0.25	0	1	23,761
Podium	0.2	0.4	0	1	23,761
Time Distance to Victory	148.73	140.34	0	3705	23,761
Time Distance to Podium	71.40	107.45	-968	3220	23,761
Male	0.51	0.5	0	1	23,761
Age	26.42	3.67	16.38	41.68	23,761
Experience (# races)	101.29	80.02	1	443	23,761
# Victories at time of race	3.99	7.62	0	59	23,761
# Podiums at time of race	11.18	17.4	0	109	23,761
Finished next race	0.82	0.39	0	1	19,430
Position in next race	11.66	9.25	1	65	15,845
Time in next race	11199.1	2675.59	5307	25,529	15,845
II. Media Attention:					
Articles the day after a race	6.62	10.38	0	125	23,761
Articles the week after a race	15.07	25.61	0	381	23,761
Articles the month after a race	45.65	82.17	0	974	23,761
III. Betting Market:					
Initial odds in next race	17.15	17.52	1.07	200	2,840
Final odds in next race	17.72	25.42	1.3	1000	2,840
Average odds in next race	17.33	17.45	1.31	200	2,840
Volume of a bet in next race (in £)	252.16	574.35	0.04	6860.12	2,840
Number of bets in next race	12.18	15.18	2	213	2,840

Note: The table presents descriptive statistics for all variables used in the empirical analysis. Panel (I) presents information on the Alpine skiing data set. Time is always measured in hundredths of a second. Data on media attention in Panel (II) is drawn from our Swissdox database covering all athletes with a final rank between 1 and 15. Panel (III) is based on Betfair. Note that one bet observation is defined as a bet on a specific event ("athlete A wins race X") with a specific odd. Hence the number of bets is the total of all individual bets for a specific event. The volume of bets is also defined at the bet observation level. The time period is 1992–2014 for the World Cup and media data and 2006–2014 for the betting data.

Table 2: Ordinary Least Squares Regression

	Media Attention		Performance		Betting Market	
	Day After	Week After	Survival	Victory	Avg. Odds	# Bets
Mean of dep. var.:	11.02	23.56	0.84	0.12	2.12	2.22
I. Podium:						
Podium	5.302***	7.765***	0.015	0.004	-0.087**	0.092
	(0.377)	(0.701)	(0.011)	(0.013)	(0.038)	(0.057)
Distance to Podium	-0.015***	-0.027***	-0.000	-0.001***	0.003***	-0.003***
	(0.004)	(0.008)	(0.000)	(0.000)	(0.000)	(0.000)
Log Experience	0.033***	0.056***	-0.003	0.006	-0.008	0.239*
	(0.008)	(0.021)	(0.012)	(0.010)	(0.099)	(0.134)
Log Prior Podiums	0.100***	0.297***	-0.001	-0.000	-0.008***	0.006*
	(0.038)	(0.100)	(0.001)	(0.001)	(0.002)	(0.003)
Log Competition	-0.000	0.013	0.000***	0.000	0.000	0.000
	(0.002)	(0.008)	(0.000)	(0.000)	(0.000)	(0.000)
Fixed Effects	Athlete	Athlete	Athlete	Athlete	Athlete	Athlete
Observations	7,408	7,408	6,478	6,478	1,625	1,625
R-squared	0.257	0.188	0.003	0.022	0.165	0.067

	Media Attention		Performance		Betting Market	
	Day After	Week After	Survival	Victory	Avg. Odds	# Bets
Mean of dep. var.:	11.02	23.56	0.84	0.12	2.12	2.22
II. Victory:						
Victory	5.151***	10.186***	-0.006	0.026*	-0.255***	0.141*
	(0.387)	(0.903)	(0.014)	(0.015)	(0.049)	(0.077)
Distance to Victory	-0.013***	-0.020***	-0.000***	-0.000***	0.002***	-0.002***
	(0.002)	(0.003)	(0.000)	(0.000)	(0.000)	(0.000)
Log Experience	0.035***	0.059***	-0.004	0.007	-0.012	0.242*
	(0.008)	(0.021)	(0.012)	(0.009)	(0.104)	(0.134)
Log Prior Podiums	0.086**	0.273***	-0.001	-0.000	-0.008***	0.006*
	(0.037)	(0.098)	(0.001)	(0.001)	(0.002)	(0.003)
Log Competition	-0.002	0.011	0.000***	0.000	0.000	0.000
	(0.003)	(0.008)	(0.000)	(0.000)	(0.000)	(0.000)
Fixed Effects	Athlete	Athlete	Athlete	Athlete	Athlete	Athlete
Observations	7,408	7,408	6,478	6,478	1,625	1,625
R-squared	0.227	0.192	0.005	0.020	0.182	0.074

Note: The table shows the results of twelve separate linear regressions using six different dependent variables as indicated in the top rows. The data on media attention is taken from SwissDox. Media attention is measured by the number of articles published on the day and during the week after the race t. In columns 3 and 4, the dependent variable is the probability of finishing (col 3) achieving a victory (col 4) in race $t+1$. In the last two columns, log average odds as well as the log of the total number of bets for the race $t+1$ are used as dependent variable. The sample includes all athletes who finished in the top-5 in race t. Columns (1) to (4) include tournaments from 1992-2014 while columns (5) and (6) are based on 2006-2014. Numbers in brackets indicate standard errors clustered at the athlete level. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 3: Balance Tests for Close Podium Finish

	Mean Value			
	Treatment	Control	Difference	p-value
A: Athlete Characteristics				
Male	0.55	0.56	-0.00	0.98
Experience	114.80	117.50	-2.70	0.48
Average Survival	0.93	0.93	-0.00	0.89
Survival in Last Race	0.86	0.87	-0.01	0.64
Number of Victories	5.63	5.53	0.10	0.76
Number of Podiums	15.58	15.33	0.25	0.76
First Prize Possible	0.60	0.59	0.01	0.50
B: Competition				
Total Podiums among Top 5	76.49	76.72	-0.22	0.92
Total Victories among Top 5	31.33	31.40	-0.07	0.93
Total Podiums among Top 10	135.18	135.38	-0.19	0.95
C: Media Attention				
Day Before Race t	5.79	5.34	0.45	0.23
Week Before Race t	15.26	14.74	0.51	0.60
Month Before Race t	49.69	48.76	0.93	0.78
D: Betting Market				
Average Odd Race t	13.25	14.03	-0.78	0.50
Volume in Race t	252.23	259.03	-6.80	0.90

Note: The table shows mean comparisons (t-tests) for all relevant pre-treatment variables. The sample includes all athletes within a bandwidth of 0.15 seconds around the podium. Age is the exact age in years at the time of the race, experience is measured by the total number of races prior to the race, survival in last race is the indicator for successfully finishing in the preceding race, victory and podium measure the total number of an athlete's victories and podiums prior to the race. Media attention is measured by the total number of articles mentioning an athlete's name in the Swissdox archive. The sample in A–C includes tournaments from 1992-2014 while part D is based on 2006-2014.

Table 4: Regression Discontinuity for Random Top Ranking Positions

	Media Attention		Performance		Betting Market	
	Day After	Week After	Survival	Victory	Avg. Odds	# Bets
Mean of dep. var.:	9.36	20.83	0.85	0.09	2.33	1.96
I. Close Podium:						
Podium	2.951***	4.735***	0.030	0.023	0.115	-0.290
	(0.753)	(1.815)	(0.037)	(0.034)	(0.116)	(0.202)
Distance to Podium	-0.146***	-0.226*	-0.002	-0.001	0.011	-0.018
	(0.045)	(0.115)	(0.003)	(0.002)	(0.010)	(0.019)
Log Experience	0.035***	0.069**	0.000	0.000	0.002**	0.001
	(0.010)	(0.030)	(0.000)	(0.000)	(0.001)	(0.002)
Log Prior Success	0.063	0.198	-0.000	0.000	-0.013***	0.005
	(0.041)	(0.131)	(0.001)	(0.001)	(0.003)	(0.005)
Log Competition	0.002	0.019*	0.000	-0.000	-0.000	0.001*
	(0.004)	(0.010)	(0.000)	(0.000)	(0.000)	(0.001)
Fixed Effects	Athlete	Athlete	Athlete	Athlete	Athlete	Athlete
Observations	2,227	2,227	1,966	1,966	455	455
R-squared	0.185	0.143	0.004	0.002	0.095	0.029

	Media Attention		Performance		Betting Market	
	Day After	Week After	Survival	Victory	Avg. Odds	# Bets
Mean of dep. var.:	15.33	32.33	0.88	0.15	1.90	2.45
II. Close Victory:						
Victory	2.294**	5.586*	0.000	-0.025	0.172	-0.174
	(1.149)	(3.034)	(0.045)	(0.064)	(0.124)	(0.192)
Distance to Victory	-0.172*	-0.357	-0.005	0.001	0.013	-0.019
	(0.091)	(0.253)	(0.004)	(0.004)	(0.012)	(0.025)
Log Experience	0.059***	0.110***	0.001**	-0.000	0.002	0.003
	(0.015)	(0.033)	(0.000)	(0.000)	(0.001)	(0.002)
Log Prior Success	0.021	0.098	-0.004**	0.000	-0.018***	0.004
	(0.066)	(0.117)	(0.002)	(0.002)	(0.004)	(0.006)
Log Competition	0.015***	0.054***	0.001**	0.001**	-0.002**	0.003***
	(0.005)	(0.016)	(0.000)	(0.000)	(0.001)	(0.001)
Fixed Effects	Athlete	Athlete	Athlete	Athlete	Athlete	Athlete
Observations	916	916	811	811	221	221
R-squared	0.255	0.197	0.029	0.005	0.212	0.074

Note: The table shows the results of twelve separate linear regressions using six different dependent variables as indicated in the top row. In Part (I) the treatment variable is a podium finish while in part (II) treatment is defined by victory. The sample includes all athletes within a bandwidth of 15 hundredths of a second and to athletes finishing first or second in Part II. The data on media attention is taken from SwissDox. In the last two columns, log average odds as well as the log of the total number of bets for the race $t + 1$ are used as dependent variable. Columns (1) to (4) include tournaments from 1992-2014 while columns (5) and (6) are based on 2006-2014. Numbers in brackets indicate standard errors clustered at the athlete level. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

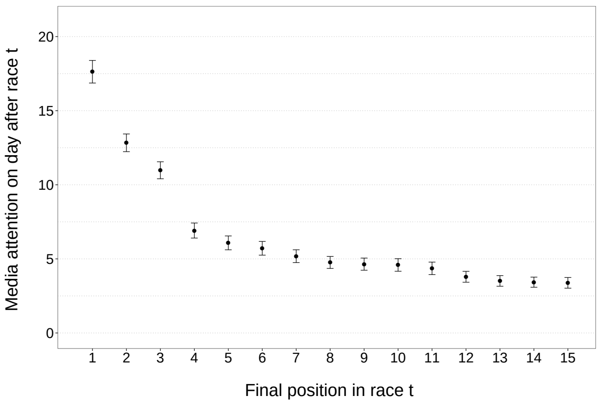
Table 5: Instrumental Variable Estimation

	Avg. Odds		# Bets	
Mean of dep. var.:	2.40	2.40	1.95	1.95
Log Media Attention	-0.012	-0.011	-0.020	-0.018
	(0.010)	(0.009)	(0.016)	(0.015)
Fixed Effects	Athlete	Athlete	Athlete	Athlete
Covariates	No	Yes	No	Yes
Instrument	Podium	Podium	Podium	Podium
Observations	1,356	1,356	1,356	1,356
R-squared	0.130	0.157	-0.032	-0.032
F-value	8.64	10.56	8.64	10.56

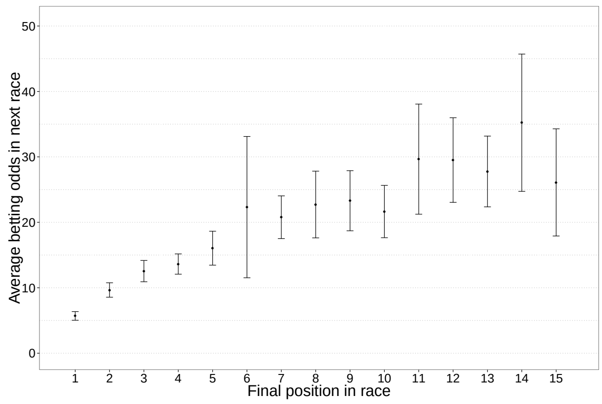
Note: The table shows the results of four separate linear regressions using two different dependent variables as indicated in the top row. The sample includes all athletes within a bandwidth of 50 hundredths of a second. The data on media attention is taken from SwissDox. The dependent variables are log average odds as well as the log of the total number of bets for the race $t + 1$. All columns include tournaments from 2006-2014. Numbers in brackets indicate standard errors clustered at the athlete level. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Figure 1: Media Attention and Betting Odds for each Ranking Position

(a) Articles on the Day after Race t

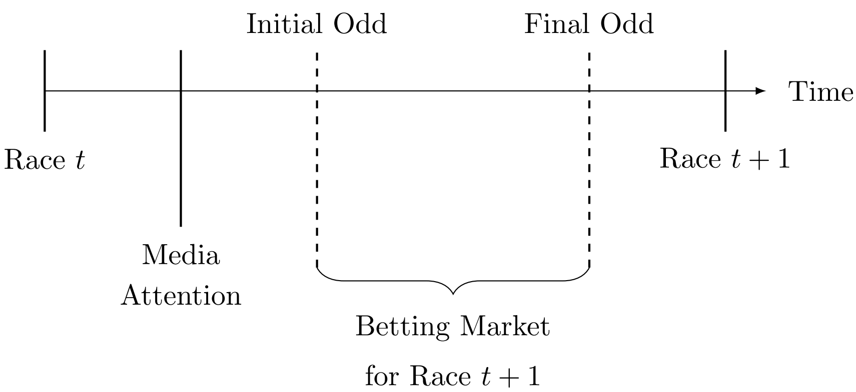


(b) Average Betting Odds in Race $t+1$



Note: The figure in Panel (a) shows the average number of articles that mention an athlete who finished on a specific position in race t. Media attention is measured on the day after the race took place. In Panel (b), we show the average of all betting odds for bets offered before the race $t+1$ took place (before 9.30am) for each ranking position in race t. The bars indicate 95% confidence intervals.

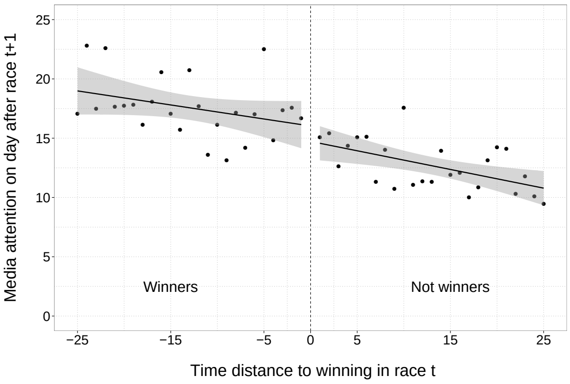
Figure 2: Illustration of the Timing of Events



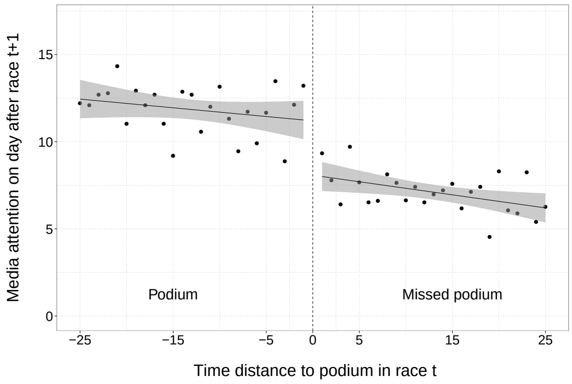
Note: The figure illustrates the timing of events. After the rate t, the betting market for the next race opens with an initial odd for each athlete. We measure media attention on the day after race t took place. The betting market closes with a final odd before the race $t+1$ takes place.

Figure 3: Effect of Top Rank on Media Attention

(a) Close **Victory**



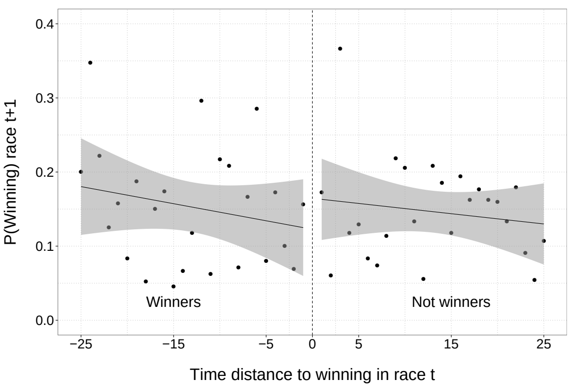
(b) Close **Podium Finish**



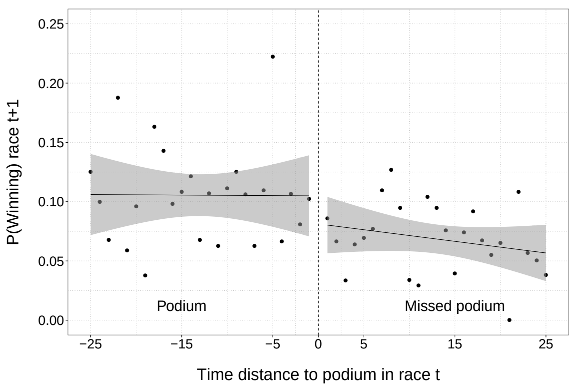
Note: The figures show a linear fit for the number of media articles that mention an athlete (on the day after a race took place) for both treatment (left) and control group (right). In panel (a) we compare victory and non-victory while in panel (b) athletes on the podium are compared with those missing it by 0.25 seconds or less. The grey region indicates the 95% confidence interval.

Figure 4: Effect of Top Rank on Performance

(a) Close **Victory**



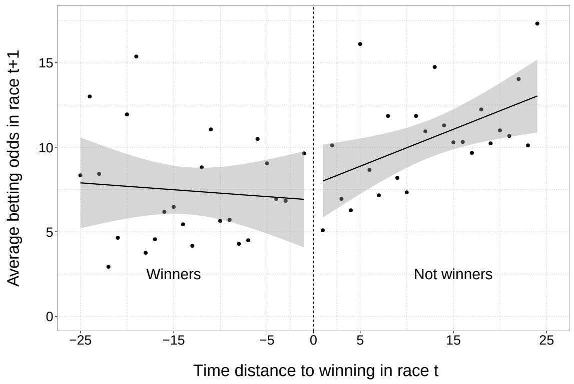
(b) Close **Podium Finish**



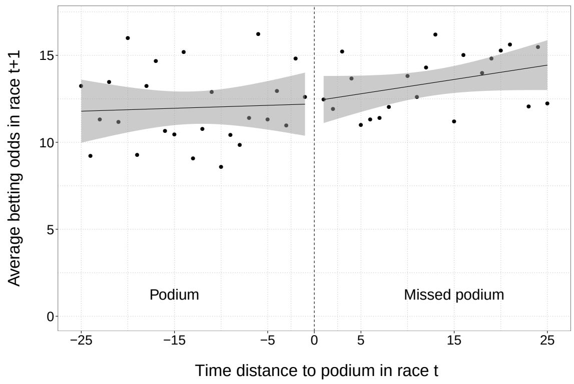
Note: The figures show a linear fit for the probability of achieving a victory in the subsequent race for both treatment (left) and control group (right). In panel (a) we compare victory and non-victory while in panel (b) athletes on the podium are compared with those missing it by 0.25 seconds or less. The grey region indicates the 95% confidence interval.

Figure 5: Effect of Top Rank on the Betting Market

(a) Close Victory



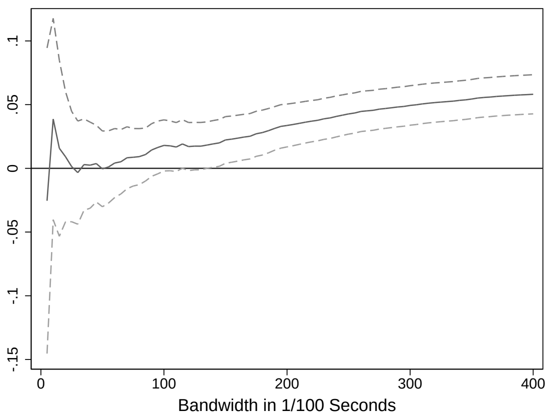
(b) Close Podium Finish



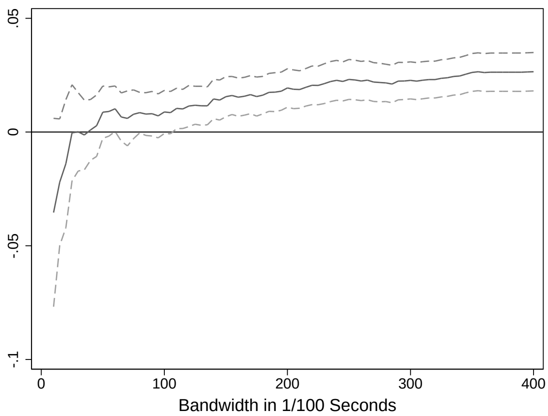
Note: The figures show a linear fit for the average betting odds in the next race for both treatment (left) and control group (right). In panel (a) we compare victory and non-victory while in panel (b) athletes on the podium are compared with those missing it by 0.25 seconds or less. The grey region indicates the 95% confidence interval.

Figure 6: Betting Market Efficiency

(a) Victory in Race $t + 1$



(b) Inverse Odds in Race $t + 1$



Note: The figures show the estimated effect of a close podium finish on the probability of winning the subsequent race (Panel a) as well the inverse average odds in the subsequent race (Panel b). The dashed lines indicate 95% confidence intervals.

Appendix (For Online Publication)

Appendix A: Additional Tables and Figures

Table A.1: Random Top Ranks and Additional Outcomes

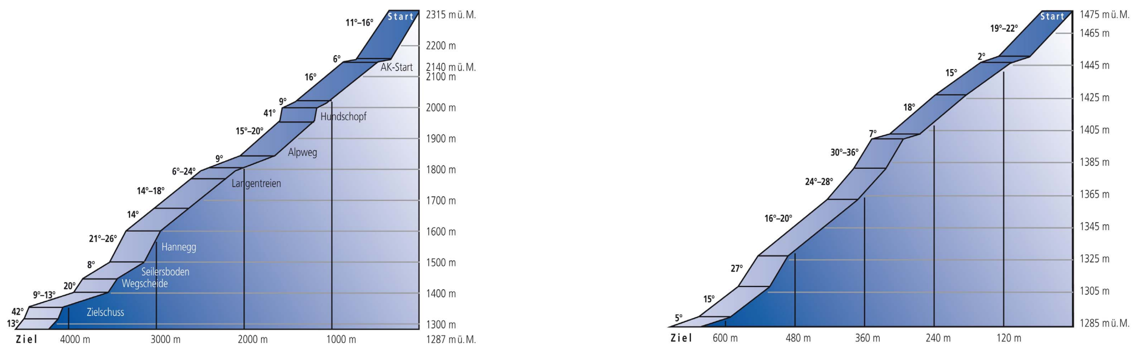
	Media Attention		Performance		Betting Odds	
	Relative	Month	Position	Time	Initial	Final
Mean of dep. var.:	0.34	55.88	9.58	10,790	2.32	2.33
I. Close Podium:						
Podium	0.108***	1.628	-0.775	-233.664	0.116	0.073
	(0.022)	(5.285)	(0.773)	(266.732)	(0.119)	(0.120)
Distance to Podium	-0.001	-0.367	0.029	-29.082	0.009	0.010
	(0.001)	(0.472)	(0.066)	(19.345)	(0.011)	(0.011)
Log Experience	-0.000	0.241**	-0.007	-6.028***	0.002**	0.002**
	(0.000)	(0.102)	(0.012)	(1.921)	(0.001)	(0.001)
Log Prior Success	0.002	0.532	0.006	21.128**	-0.014***	-0.013***
	(0.001)	(0.522)	(0.047)	(9.648)	(0.003)	(0.003)
Log Competition	-0.000***	0.004	-0.002	-3.428*	0.000	-0.000
	(0.000)	(0.031)	(0.006)	(1.859)	(0.000)	(0.000)
Fixed Effects	Athlete	Athlete	Athlete	Athlete	Athlete	Athlete
Observations	2,054	2,227	1,698	1,698	455	455
R-squared	0.171	0.120	0.008	0.016	0.088	0.085

	Media Attention		Performance		Betting Odds	
	Relative	Month	Position	Time	Initial	Final
Mean of dep. var.:	0.26	82.35	7.87	10,764	1.91	1.95
II. Close Victory:						
Victory	0.049*	-9.790	0.528	-9.247	0.240*	0.164
	(0.027)	(8.506)	(1.186)	(429.756)	(0.126)	(0.133)
Distance to Victory	-0.003	-1.469**	0.110	-3.533	0.010	0.017
	(0.002)	(0.568)	(0.084)	(28.459)	(0.010)	(0.013)
Log Experience	-0.000**	0.374***	-0.003	-0.480	0.003*	0.002*
	(0.000)	(0.127)	(0.010)	(3.758)	(0.001)	(0.001)
Log Prior Success	0.002*	0.075	0.017	-3.260	-0.019***	-0.018***
	(0.001)	(0.502)	(0.034)	(14.905)	(0.004)	(0.004)
Log Competition	-0.000	0.173***	-0.014*	-3.888	-0.001	-0.002**
	(0.000)	(0.062)	(0.008)	(2.836)	(0.001)	(0.001)
Fixed Effects	Athlete	Athlete	Athlete	Athlete	Athlete	Athlete
Observations	848	916	724	724	221	221
R-squared	0.196	0.161	0.009	0.009	0.213	0.198

Note: The table shows the results of twelve separate estimations using three different dependent variables as indicated in the top row. In Part (I) the treatment variable is a podium finish while in part (II) treatment is defined by victory. The sample includes all athletes within a bandwidth of 15 hundredths of a second. Numbers in brackets indicate standard errors. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

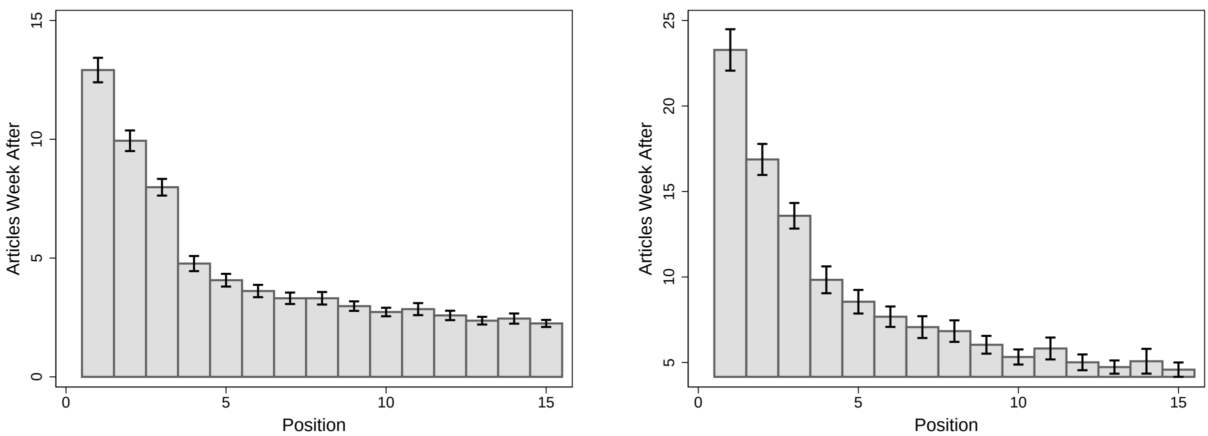
Figure A.1: Profile of Downhill and Slalom Race

(a) Downhill (b) Slalom



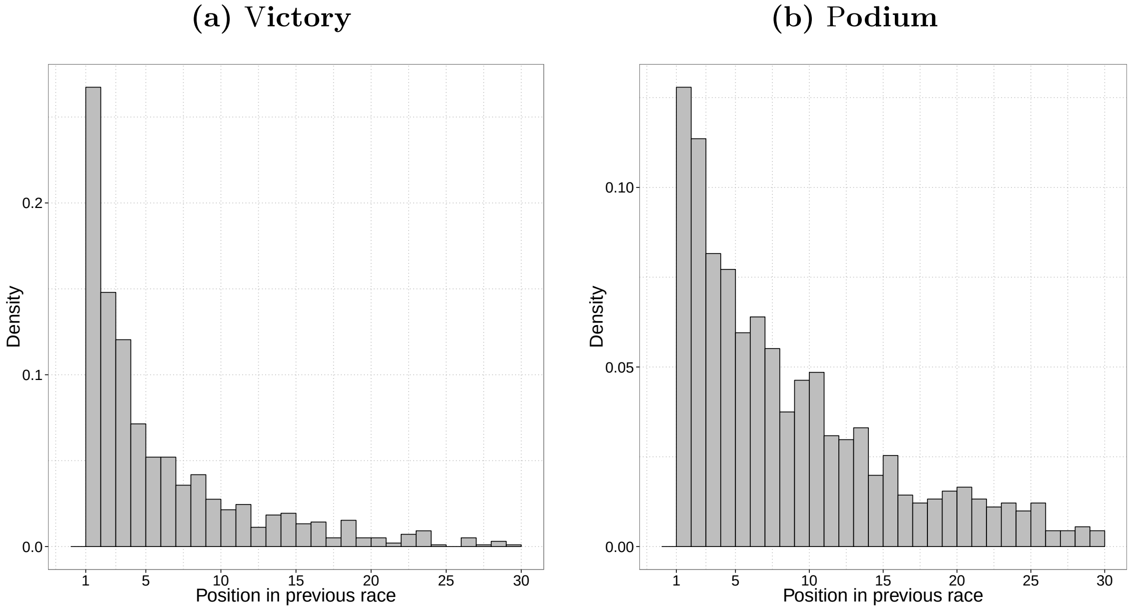
Note: The figure illustrates the profiles of downhill and slalom race in Wengen. Source: www.lauberhorn.ch

Figure A.2: Distribution of NewsLibray Media Attention



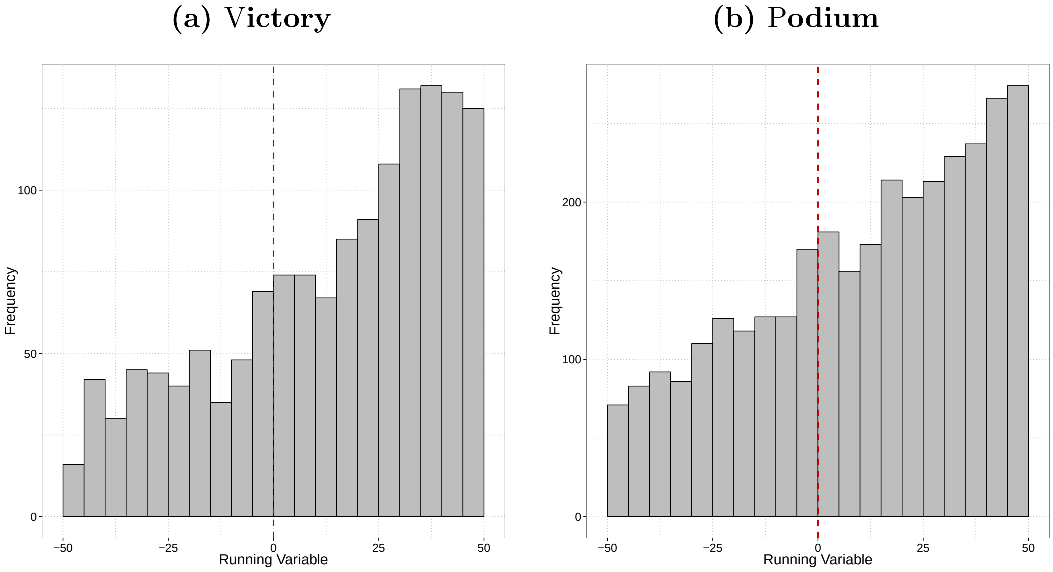
Note: The figure shows the distribution of media attention across the ranking positions in World Cup Alpine Skiing between 1992–2014. The data is based on NewsLibrary. Confidence intervals at 95% are shown.

Figure A.3: Previous Positions of Winners and Third-Ranked Athletes



Note: The figures show histograms of the ranking positions in the previous race of winners and third-ranked athletes in the current race.

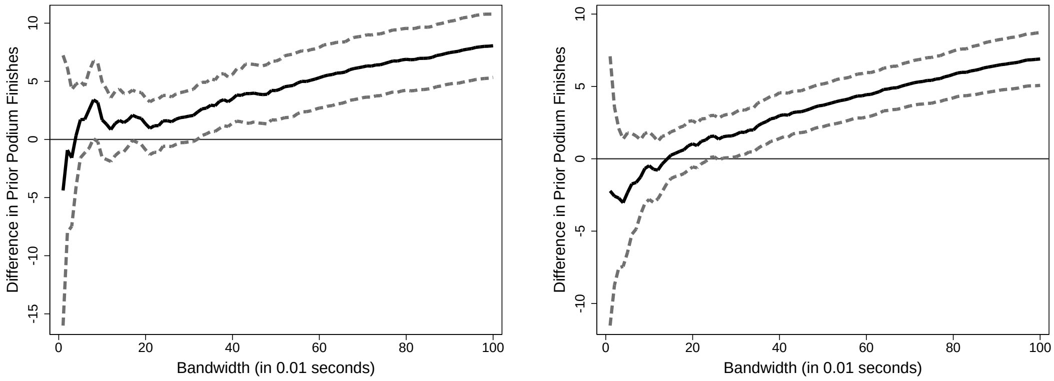
Figure A.4: Observations around the Cutoff



Note: The histograms show the densities of race times around the cutoff for a victory and podium finish. In line with McCrary (2008), this indicates that there is no manipulation around the respective cutoffs. In panel (a) we plot the distance to rank 2 (from left-hand side) and 1 (from right-hand side), respectively. For the distance to the podium in panel (b), we plot the distance to rank 4 and 3, respectively.

Figure A.5: Balance Tests by Bandwidth

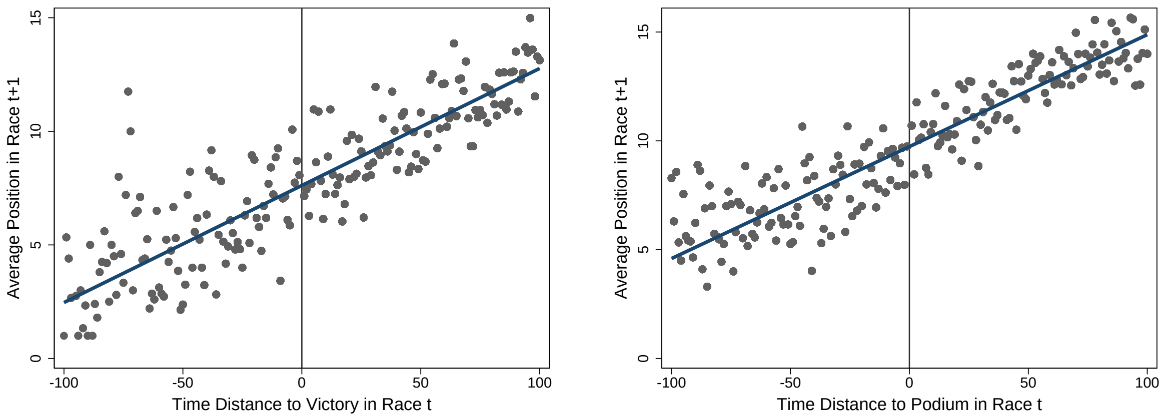
(a) Prior Success by Victory (b) Prior Success by Podium



Note: The figures show the results of t-test on two pre-determined athlete characteristics. We compare athletes' number of podium finishes prior to the race determining treatment. In plot (a), treatment is defined by victory while in plot (b) it is based on whether an athlete finished on the podium. The sample includes all tournaments between 1992–2014. Confidence intervals at 95% are shown.

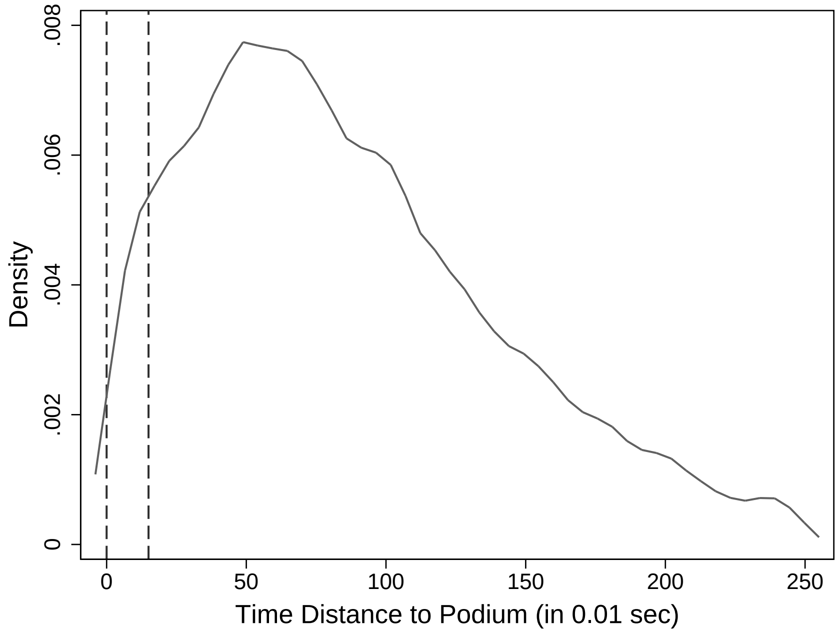
Figure A.6: Serial Correlation in Individual Performance

(a) Distance to Victory (b) Distance to Podium



Note: The figure shows the average position in race $t + 1$ for each distance to the winner (left-hand side) or podium (right-hand side) in race t In addition, we add a linear regression line. The sample includes all athletes between 1992–2014 who finished within a bandwidth of one second to the victory (a) or podium (b).

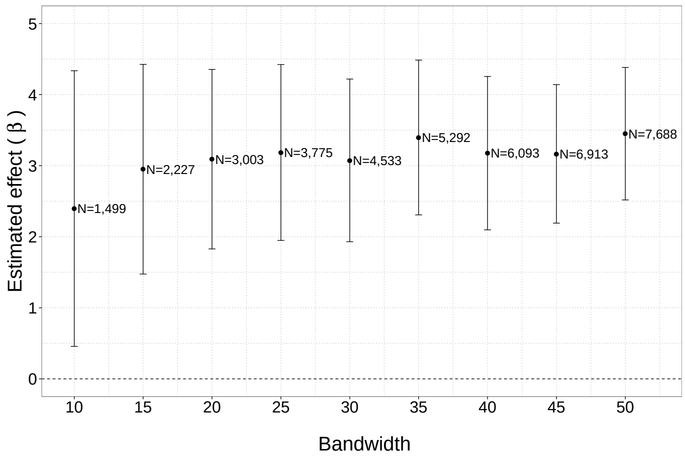
Figure A.7: Distribution of Running Variable and Bandwidth Choice



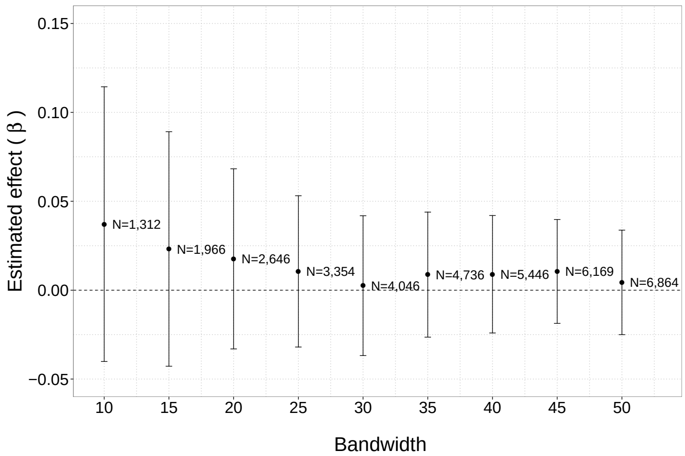
Note: The figure shows the distribution of the time distance to the podium (i.e., running variable). We plot an Epanechnikov kernel function with a bandwidth of 0.05 seconds. The dashed vertical lines illustrate our preferred bandwidth choice of 0.15 seconds.

Figure A.8: Regression Results using Different Bandwidths

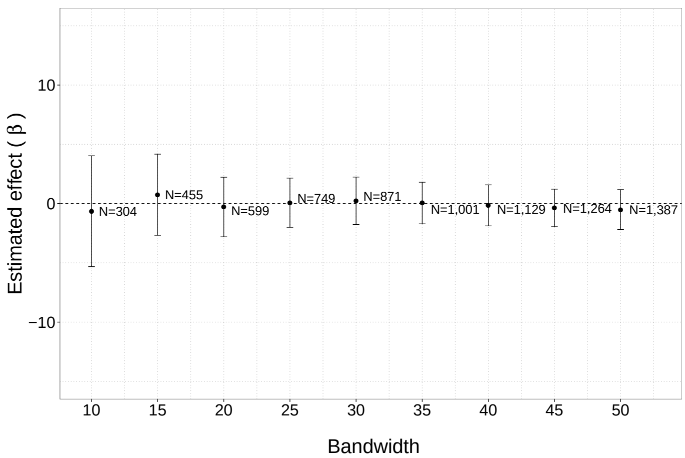
(a) Effect on Media Attention



(b) Effect on Performance

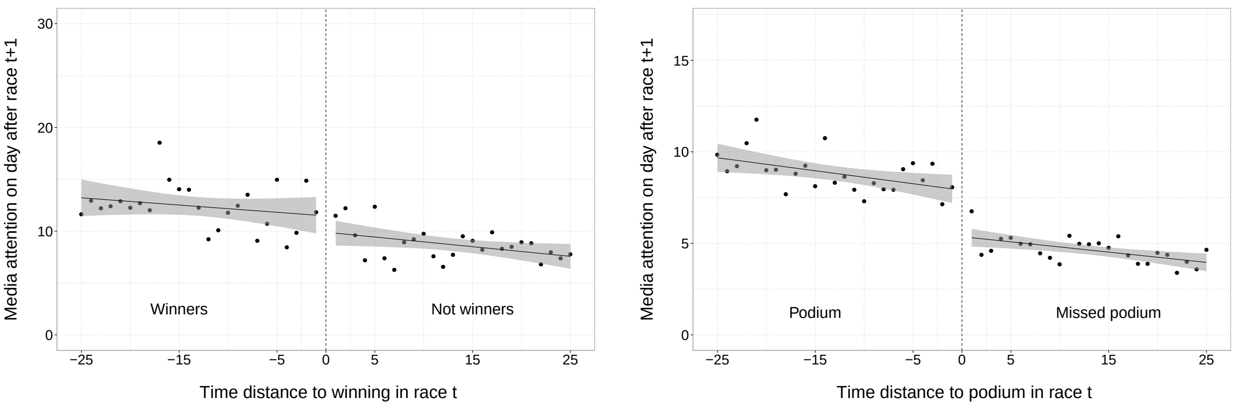


(c) Effect on Betting Odds



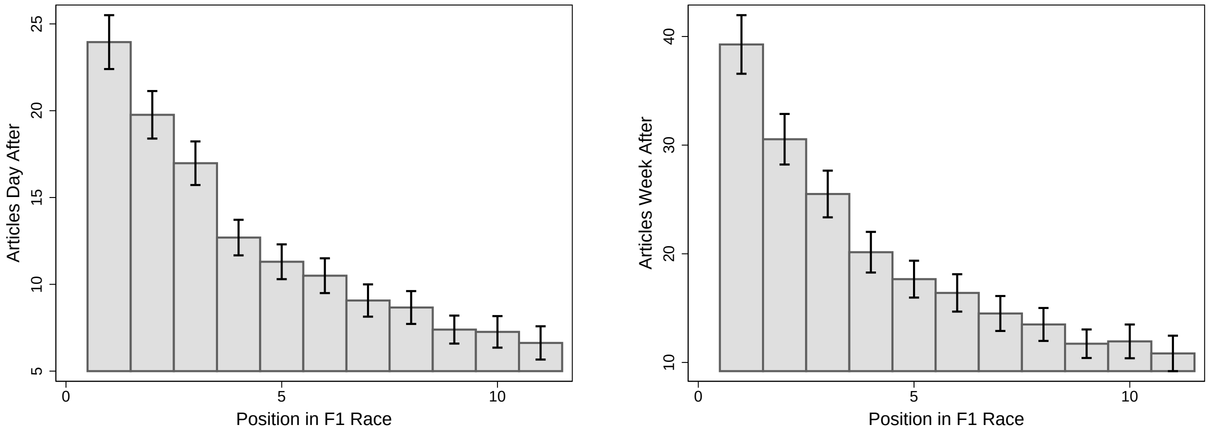
Note: The figure shows the results from estimating the effect of podium on media attention on the day after the competition (Panel (a)), the probability to win the next race (Panel (b)), and average betting odds in the next race (Panel (b)) using bandwidths from 10 to 50 hundredths of a second. The dots indicate the point estimate, while the bars depict the 95% confidence interval.

Figure A.9: Effect of Top Rank on NewsLibrary Media Attention



Note: The figures show a linear fit for the number of media articles (using Newslibrary data) that mention an athlete (on the day after a race took place) for both treatment (left) and control group (right). In panel (a) we compare victory and non-victory while in panel (b) athletes on the podium are compared with those missing it by 0.25 seconds or less. The grey region indicates the 95% confidence interval.

Figure A.10: Distribution of Media Attention in Formula One



Note: The figure shows the distribution of media attention across the ranking positions in Formula 1 races between 1992–2014. The data on media attention is based on SwissDox. 95% confidence intervals are shown.

Appendix B: Attrition Bias

One problem that may arise when estimating the effect of a ranking position on future performance is that athletes differ with respect to their probability of survival, i.e. not crashing. If this probability is related to success our estimates for performance would be biased. It could be, for example, that athletes who are very successful once adopt a riskier behavior in subsequent races in order to be successful again.

Following Frangakis and Rubin (2002), we denote athletes with a constant low (high) probability of survival by DD (LL). While the survival probability of this set of athletes is unaffected by the treatment, other athletes adjust their behavior when being treated, in other words after a quasi-random top rank. The athletes in subset LD adopt a more risky strategy after treatment while those in subset DL follow a low-risk strategy in case they achieve a quasi-random top rank. In the regression of the probability of survival ($s_{i,j+1}$) on treatment $D_{i,j}$ and controls for a athlete's own characteristics $\mathbf{X}_{i,j}$ and competitors' characteristics $\mathbf{Z}_{i,j}$

$$s_{i,j+1} = D_{i,j}\tau + \mathbf{X}_{i,j}\,\gamma + \mathbf{Z}_{i,j}\,\delta + \varepsilon_{i,j} \tag{5}$$

we should expect $\tau = 0$ for the two types with constant behavior (DD and LL). For types DL we expect $\tau < 0$ and for types LD we should see an increase in the probability of survival. Thus we have two problems if the coefficient τ is significantly negative: First, our estimates with respect to subsequent performance would be biased upwards because we would only observe treated athletes in case they are successful in subsequent races. Second, the overall gain from imposing a ranking will be reduced and perhaps negative if top ranks lead to a strong increase in risky-behavior.[18]

[18]This finding would be in line with the above-mentioned research on addiction to success.

Appendix C: Relative Performance Measures and SUTVA

The Stable Unit Treatment Value Assumption is fundamental to most estimators used in the program evaluation literature. It allows to write the treatment status of individual i only dependent on her assignment, and the outcome of individual i only dependent on her assignment and treatment status. More formally, SUTVA is defined as follows (according to Angrist, Imbens and Rubin (1996, 446))

(a) If $Z_i = Z'_i$, then $D(\mathbf{Z}) = D(\mathbf{Z}')$

(b) If $Z_i = Z'_i$, then $Y_i(\mathbf{D}, \mathbf{Z}) = Y_i(\mathbf{D}', \mathbf{Z}')$

This allows us to write $D_i(\mathbf{Z}) = D_i(Z_i)$ and $Y_i(\mathbf{D}, \mathbf{Z}) = Y_i(D_i, Z_i)$.

Applying this assumption to our paper, let us define the vector of final times $\mathbf{T}$ in race j as well as vectors of assignments (to the podium) ($\mathbf{Z}$) and treatments ($\mathbf{D}$):

$$\mathbf{T} = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ \vdots \\ t_N \end{pmatrix} \qquad \mathbf{Z} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_N \end{pmatrix} \qquad \mathbf{D} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_N \end{pmatrix} \tag{6}$$

We assume that there is a positive probability of a crash. We model survival as

$$S_i = \begin{cases} 1 \text{ if } S_i^* > 0 \\ 0 \text{ if } S_i^* \leq 0 \end{cases} \tag{7}$$

with

$$S_i^* = \theta_i a_1 + \mu_j + \varepsilon_i \tag{8}$$

where S_i^* is a latent variable, θ_i is an athlete's skill, $a_1 > 0$ is a coefficient, μ_j is a race fixed effect, ε_i is an unobserved component. With $a_2 < 0$ being some coefficient, the final time can be written as

$$T_i = \begin{cases} \text{NA} & \text{if } S_i = 0 \\ \theta_i a_2 + \delta_j + v_i & \text{if } S_i = 1 \end{cases} \tag{9}$$

We consider a specific race with three top athletes under two circumstances. First, conditions are equal for all athletes. Second, the three top athletes ($i \in \{1, 2, 3\}$) suffer from bad weather conditions, which makes it impossible for them to attain a place on the podium.

$$\mathbf{Z} = \mathbf{D} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \qquad \mathbf{Z}' = \mathbf{D}' = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{10}$$

In this case, it is obvious that the assignment status of individuals 4–6 depends upon the assignment of the three top athletes. However, since $D_i = Z_i$, the individual treatment status D_i can still be written as a function of the assignment Z_i.

Turning to the implication (b) of SUTVA, we first note that our outcome can be written purely as a function of the assignment, i.e. $Y_i(\mathbf{Z}, \mathbf{D}) = Y_i(\mathbf{Z})$. This comes from the fuzzy design where $\mathbf{Z} = \mathbf{D}$.

Any measure of relative performance —such as the position— depends on a athlete's own time as well as the competitors' times: $P_{i,j} = g(T_{i,j}, T_{s,j}) = g(\theta_i, \lambda_{i,j}, n_{i,j}, \theta_s, \lambda_{s,j}, n_{s,j}) \quad \forall s \neq i$. The winner's time is often the benchmark and can be written as

$$T_{\text{win},j} = \min_{i \in S}(T_{i,j}) \tag{11}$$

where S indicates the set of survivors. For the sake of illustration, we specify equation (9) for survivors as

$$T_{i,j} = \theta_i a_2 + \delta_j + \text{Exp}_i a_3 + \text{Exp}_i^2 a_4 + \tau_i a_5 + u_i \tag{12}$$

where Exp_i is experience and treatment τ_i equals one if athlete i won the last race and zero otherwise. Imagine that the winner in a given race j is determined by a tiny time difference between athlete 1 and 4. Assume both athletes to have the same skill level θ_i, but while athlete 1 is a rookie, athlete 4 is an experienced and successful athlete. Athlete i's relevant outcome $Y_i = P_{i,j+1}$ in the next race depends on the performance of the best athlete in that race. So if athlete 4 wins today and $a_5 \neq 0$, $T_{\text{win},j+1}$ is likely to be lower than if athlete 2 wins (because athlete 4 is more experienced and experience positively affects performance).

Therefore, the outcome of athlete i in race $j+1$ is likely to depend on the assignment of the winner (note that treatment and assignment are henceforth defined for the victory treatment and not the podium treatment as above)

$$\mathbf{Z} = \mathbf{D} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad \mathbf{Z}' = \mathbf{D}' = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \qquad \text{and} \qquad Y_i(\mathbf{Z}) \neq Y_i(\mathbf{Z}')$$

which violates definition (b) of SUTVA.

This problem arises to different extents with all kinds of relative performance measures. Thus we limit our analysis to using absolute race times as outcome variable for performance.

Appendix D: Sample Race

Table A.2: Sample Race for Illustrative Purpose

Rank	Name	Nationality	Time	Difference	Articles
1	Benjamin Raich	AUT	1:36.66		13
2	Akira Sasaki	JAP	1:36.83	0.17	9
3	Thomas Grandi	FRA	1:37.17	0.51	15
4	Michael Janyk	USA	1:37.19	0.53	4
5	Ted Ligety	USA	1:37.54	0.88	5

The table above shows the result of the 2006/07 Men World Cup Slalom race in Shigakogen, Japan. We observe that Thomas Grandi achieved a podium finish because he was 0.02 seconds ahead of Michael Janyk. This race result can be used to illustrate how we define treatment and control group in our estimation. First, we take the time of the third-ranked athlete (1:37.17 here). Then we compute a 0.25 seconds window around this race time. All athletes within this time window (1:12.17 to 2:02.17) are part of our estimation sample. Every athlete in this group who finished on the podium is in the treatment group. Every other athlete in the estimation sample serves as part of the control group.