The Role of Rankings, Big Shots, and Random Successes

Stefan Legge[†] – University of St.Gallen
Lukas Schmid[‡] – LSE & University of St.Gallen

Preliminary, as of May 24, 2013

Abstract

Does the relative position in a hierarchy or a ranking affect future performance? And to what extent does a one-time great success alter subsequent performance? In this paper, we propose a highly tractable theoretical model and analyze a rich set of data on alpine skiing tournaments for the period of 1992–2013 to evaluate the long-term impact of a higher ranking position, a victory or podium, as well as a top six classification. Using a regression discontinuity approach, the paper provides evidence of substantial performance differences in the long run as a result of the present relative position. Victories, podium finishes, as well as a higher position between rank 6 and 15 are all found to be significantly performance-enhancing.

JEL Classification: D03, D83, D84, J24, L83, M50

Keywords: rankings, success, performance, skiing, risk behavior

[†] University of St.Gallen, Department of Economics, SIAW Institute, Bodanstrasse 8, CH-9000 St.Gallen, Switzerland. stefan.legge@unisg.ch
[‡] Department of Methodology, Houghton Street, London WC2A 2AE. University of St.Gallen, Department of Economics, Bodanstrasse 8, CH-9000 St.Gallen, Switzerland. E-mail: lukas.schmid@unisg.ch

We are grateful to Reto Föllmi and Sam Kortum as well as seminar participants at the universities of St.Gallen and Zurich for helpful comments. Moreover, we thank Daniel Défago, Marcel Looze, and Ulf Seehase (all FIS) who kindly supported the data collection. Dominik Zurbuchen and Rino Heim provided excellent research assistance.

1 Introduction

Is there a long-term benefit to being very successful once? And does a higher relative position in a ranking affect future performance? In this paper, we study these two questions by using data on alpine skiing for the period of 1992–2013. We focus on close races and apply a regression discontinuity design as identification strategy.

One great success (or 'big shot') helps a lot to spur a career in many professions. Previous work has pointed out that first-job experience is important for economists' research productivity (Ginther and Kahn, 2004; Hilmer and Hilmer, 2007), manager's career mobility and salary increases (Cox and Harquail, 1991), and sport athletes' prospects to get a contract as a professional (Rosen and Sanderson, 2001). However, it remains unclear whether it is the very success which leads to a higher performance, or whether the success itself is a consequence of a higher performance level. In particular, being successful is mostly the result of superior skills and effort. As a consequence, the effect of big successes and relative positions on personal performance is difficult to disentangle from the effect of higher skills and effort as they usually come hand in hand.

The paper provides a novel solution to this problem. We examine performance in alpine skiing and use close races as external manipulation of relative position and success. In such races it is often a tiny margin – a few hundredths of a second – that determine whether a racer finishes first or second, third or fourth, or even sixth or tenth. Assuming that such fairly small differences are random and controlling for several measures of past performance, we are able to test whether relative positions and one-time successes have long lasting effects on racers' future performances[1]. The setting is ideal for a regression discontinuity approach because it satisfies the two key identification assumptions. First, the distribution of the assignment variable, the distance to a certain position, is smooth at the threshold. Second, manipulation is unlikely because every racer wants to be as fast as possible.

The implications of our study also relate to organizations in which competition acts as key driver of incentives. As Brown (2011) points out, individual performance can be impaired by the presence of superstars. The introduction of a clear ranking – as in many sports – is likely to bring to mind the differences in performance and ability. Understanding whether

[1]Skiing has a strong advantage as compared to most other sports such as tennis in which winning the decisive point in a tight match represents exactly the superior quality of a particular player. A good example is tennis player Roger Federer who seemed unbeatable in the years 2003–2007.

this has an overall performance-enhancing effect is crucial to optimize organizations with heterogeneous individuals.

In the present study, we use data on 1,141 athletes for the period of 1992-2011 to estimate the effect of relative positions and big shots. Our dependent variables are the number of completed races as well as short and long term absolute performance measured as the average distance to winner.[2] In the first step, we find that a podium finish positively affects the probability of finishing the subsequent race, while we find no such effect for a victory. There is no immediate effect of a podium and victory on short term performance, but a reasonably high effect of the two variables when analyzing long term performance. Podium and victory increase the change in long term performance by 12% and 20%, respectively. However, in contrast to our standard theoretical predictions the effect of a victory (vs. non-victory) is not higher than the effect of a podium (vs. non-podium) or a position #4–6 (vs. non-#4–6). For young racers we find significantly stronger effects. In the second part, we generalize our estimation to relative positions and identify time clusters for which the position of a racer is arguably random. Our results suggest that the relative position only affects future performance for the racers in positions #6–10 and #11–15. We find that these racers need approximately one to two positions to compensate a relatively worse current state of performance.

Our paper provides two novelties. From a theoretical point of view, the question whether relative positions and big shots affect future performance has not been addressed before. Second, to our knowledge, no study has ever used sports data combined with a regression discontinuity approach to estimate the long-run effects of relative positions and one-time great successes. From a policy perspective, our results suggest that one has to bear in mind that rankings and relative positions have consequences on individual performance. The implications extend to a variety of areas because rankings have become widely popular when evaluating students' and workers' performance or when assessing different investment potential of firms.

The question of how different external signals about individual current performance can influence future effort and motivation has been examined in previous research. Most prominently, the literature has stressed the importance of higher self-confidence, setting of more

[2] One argument that could be made is that short-term losers may rather quit in comparison with short-term winners. But in some sense this could also be part of the effect we intend to measure: small and random differences in one race can have substantial effects on performance in the long run of a career.

ambitious targets, revision of the self-image, addiction, and better support by the coaching staff. Camerer et al. (1997) and Farber (2005) analyze target-setting of New York cab drivers and find non-conclusive evidence concerning the question whether car drivers set a daily income target and adjust their labor supply accordingly. Rachlin (2000) describes the importance of keeping track of one's own actions for successful self-regulation. This plays a key role in our paper since skiers and sportsmen in general are usually better informed about their past (relative) performance than ordinary people. This is due to extensive rankings and statistics in this field.[3] With respect to alpine skiing, we expect that a young racer's relative positions contain information about her potential ability which is crucial in order to fully exploit it. Thus, one great success early in their careers can lead to higher targets and motivation. Bénabou and Tirole (2002) as well as Bénabou and Tirole (2004) explore how individuals regard each decision as a possible precedent for future ones. This way of thinking in sequences of choices may explain why winning a ski race once may change future behavior. With each race result the racer will learn more about her own abilities, and consequently makes choices in order to achieve or preserve favorable self-conceptions.[4]

Recent evidence exploring the role of beliefs about individual abilities supports the idea that sport athletes' updated self-image can affect performance (Santos-Pinto, 2008, 2010). Klaassen and Magnus (2001) present evidence for the so-called 'winning mood' in tennis matches: if the previous point was won (lost) the probability of winning a point increases (decreases) substantially. The question about the causes of a positive self-image has been addressed in a seminal paper by Santos-Pinto and Sobel (2005). The results suggest that individuals' positive self-image arises from subjective assessments of their *relative* abilities. In the case of alpine skiing, we argue that race results are perceived as proxies for a ranking of relative abilities. Another mechanism is provided by Kahneman, Wakker and Sarin (1997) who show that positive self-image can also arise from biased memories. Experienced (or remembered) utility often has an effect on current behavior and preferences. In the case of skiers, it is likely that memories of past successes cause a positive self-image that in turn has a positive effect on motivation and thus increases performance.

Bénabou and Tirole (2004) point out that "personal rules" constitute an important driver

[3]This relates to the findings of Caillaud, Cohen and Jullien (2000) who discuss how therapies, such as those offered by Alcoholic Anonymous, work. It is not that participants are provided with new information or commitment devices. On the contrary, the key to success is knowing your own identity.

[4]As Bénabou and Tirole (2004) point out, this idea is prevalent in psychology and also well supported empirically. See for example experiments by Kirby and Guastello (2001) or Quattrone and Tversky (1984).

of human decision-making and can generate very persistent behavior such as various forms of addiction. In this context, early success in skiing may not only increase self-confidence but also cause some kind of addiction. Once having experienced the pleasure of being very successful can cause individuals to become addicted to success. In a similar vein, Becker and Murphy (1988) show that consumption today can increase the likelihood of consumption tomorrow. While there is considerable evidence on the persistence of consumption patterns, less is known about persistence of performance which can be described as addiction to success. Being very successful once can permanently change preferences similar to the one-time consumption of drugs.

An important channel for long-lasting effects of one-time success in skiing is that successful racers receive greater support. Yet there is little evidence on increased support for one-time successful professional sportsmen.[5] The support channel might be especially strong in case of important races such as world championships. This is due to greater public attention and, as shown by Ehrenberg and Bognanno (1990) as well as Sunde (2009), because individual effort is much higher in important races.

Finally, our work also adds to the literature on the role of social status and rankings. Fershtman and Weiss (1998) as well as Rege (2008) show that individuals care about social status because it serves as a signal of unobservable abilities. This is also true for many sports. Neither sponsors nor team managers can observe individual skill levels and take race results and rankings as proxies for the distribution of skill across racers. Tran and Zeckhauser (2012) show that Vietnamese students who receive information about their relative rank in an English course perform better than those who do not receive any information. Azmat and Iriberri (2010) find positive effects for all ability types from providing students information on their relative rank. Kuhnen and Tymula (2012) show that already the announcement of a relative ranking scheme has an effect on workers' performance in solving math tasks.

The paper proceeds as follows. In section 2 we provide the theoretical foundation of our analysis and derive several hypotheses for the estimation. Section 3 describes our identification strategy and discusses some econometric issues. Section 4 presents information about skiing in general as well as on our dataset. In section 5, we provide descriptive statistics as well as the econometric findings. Section 6 concludes.

[5] A notable exception is Anderson (2012) who shows how college athletic success attracts substantial donations.

2 Theory

In this section we first illustrate how random successes can occur in alpine skiing. Then we define our outcome variables and discuss theoretically the impact of one-time successes on subsequent performance. Finally, we set up the propositions which provide the basis for our empirical analysis.

2.1 Random Successes and Subsequent Performance

Assume in a given ski race j there are three variables which determine an individual's race time: θ_i for the time-invariant talent of racer i, $\lambda_{i,j}$ for the training or fitness level, and $n_{i,j}$ as a noise parameter[6]. Note that $n_{i,j}$ captures all kinds of random shocks such as wind and weather conditions that can be heterogeneous or homogeneous across racers.

With these variables we can define $T_{i,j} = f(\theta_i, \tau_{i,j}, n_{ij})$ which states that the time of racer i in race j is a function of a her skill and training as well as some random noise. Moreover, her position, $P_{i,j}$, is a function of her own time as well as her competitors' times:

$$P_{i,j} = g(T_{i,j}, T_{s,j}) = g(\theta_i, \tau_{i,j}, n_{i,j}, \theta_s, \tau_{s,j}, n_{s,j}) \qquad \forall s \neq i \tag{1}$$

By means of this equation, we can illustrate why 'random' successes are possible: Large fluctuations in the noise term $n_{i,j}$ can reduce a racer's race time sufficiently to overcome skill and training deficits. That is, a less skilled racer can be lucky and draw a very low $n_{i,j}$ which enables her to achieve a better race time than a more skilled competitor. As we show in section 5.2, this phenomenon of random successes is indeed found in the data.

In line with Klaassen and Magnus (2001) we argue that any success causes a *short term* winning mood. In addition, previous research suggests that one-time successes can have *long term* effects on individual performance. Racers constantly update their ambition and set new goals based on their experiences (Bénabou and Tirole, 2002, 2004; Camerer et al., 1997; Farber, 2005). Moreover, one great success can enhance a racer's confidence (Santos-Pinto and Sobel, 2005; Santos-Pinto, 2008, 2010). Athletes also receive varying amounts of support depending on their previous performance. However, it takes time for teams to adjust

[6] There is a recent literature on the consequences of noisy performance measurement. Sloof and van Praag (2010) use experimental data to support the hypothesis that noise in performance measurement has a positive impact on productive effort. However, Morgan, Sisak and Vardy (2012) show that noise in performance evaluation has rather ambiguous effects on motivation.

contracts as well as their internal hierarchy. In order to disentangle the long and short term impact we estimate the effect on the next two, five, ten, and twenty races.

2.2 Outcomes

The question we intend to examine is whether random successes as described in section 2.1 have a positive impact on subsequent behavior and performance. In particular, we estimate the effect of racer i's relative position in one race, $P_{i,j}$, on four different types of outcomes: (i) risk behavior, (ii) absolute race times, (iii) relative performance, and (iv) trends in performance. Each outcome is discussed in detail below and will be referred to as dependent variable $Y_{i,j}$ in the empirical section.

2.2.1 Risk Behavior

One problem that may arise when estimating the effect of a ranking position on future outcomes is that racers differ with respect to their probability of survival (i.e. the probability of not crashing). If this probability is related to success our estimates for performance would be biased. It could be, for example, that racers who are very successful once adopt a riskier behavior in subsequent races in order to be successful again. Consider the stratification of the population shown in table 1.

Table 1: Stratification of Population

	Strata		Pot. Outcomes	
	S(1)	S(0)	Y(1)	Y(0)
DD	0	0	$Y_{DD}(0)$	$Y_{DD}(1)$
DL	0	1	$Y_{DL}(0)$	$Y_{DL}(1)$
LD	1	0	$Y_{LD}(0)$	$Y_{LD}(1)$
LL	1	1	$Y_{LL}(0)$	$Y_{LL}(1)$

Following Frangakis and Rubin (2002), we denote racers with a constant low (high) probability of survival by DD (LL). While the survival probability of this set of racers is unaffected by the treatment, other racers adjust their behavior when being treated, in other words after a random success. The racers 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 random success. In the regression of the change in the probability of survival ($\Delta s_{i,j}$) on treatment

$D_{i,j}$ and controls for a racer's own characteristics $\mathbf{X}_{i,j}$ and competitors' characteristics $\mathbf{Z}_{i,j}$

$$\Delta s_{i,j} = \bar{s}_{i,\text{future}} - \bar{s}_{i,\text{past}} = D_{i,j}\tau + \mathbf{X}_{i,j}\,\gamma + \mathbf{Z}_{i,j}\,\delta + \varepsilon_{i,j} \tag{2}$$

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 because we would only observe treated racers in case they are successful in subsequent races. Second, the overall gain from imposing a ranking will be reduced and perhaps negative if one-time successes lead to a strong increase in risky-behavior.[7]

Our data set contains information about whether racers participated in a race and survived. Thus we can estimate the effect of a higher ranking position in race j on the probability of surviving in subsequent races $j + l$:

$$\bar{s}_{i,j,\text{future}} = \frac{1}{N_{i,j,\text{future}}} \sum_{l=1}^{L} s_{i,j+l} \tag{3}$$

where $N_{i,j,\text{future}}$ denotes the number of future races, $L \in \{2, 5, 10, 20\}$, and $s_{i,j+l}$ is a dummy variable which assumes the value one if racer i survived in race $j + l$. We also compute $\bar{s}_{i,j}$ for the past 2, 5, 10, and 20 races in order to estimate the change in the probability of surviving.

$$\Delta\bar{s}_{i,j} = \frac{\bar{s}_{i,j,\text{future}} - \bar{s}_{i,j,\text{past}}}{\bar{s}_{i,j,\text{past}}} \tag{4}$$

This transformation of our dependent variable (from levels to changes) removes a substantial amount of serial correlation in individual performance[8]. Note that $\Delta\bar{s}_{i,j}$ is the percentage change when comparing future and past probabilities of crashing.

---

[7] This finding would be in line with research on addiction to success (Becker and Murphy, 1988).

[8] Note that when using changes instead of levels in our estimation, we still control for performance earlier in the season and in a racer's career.

2.2.2 Absolute Performance

As a first measure of subsequent performance, we use the average absolute time of racer i in races $j+l$:

$$\bar{T}_{i,j,\text{future}} = \frac{1}{N_{i,j,\text{future}}} \sum_{l=1}^{L} T_{i,j+l} \tag{5}$$

where $N_{i,j,\text{future}}$ denotes the number of future races for which we have a race time for both of the racers we compare. For example, if one racer is slightly faster than another one in race j and thus finishes third instead of forth, we compare the two racers' times in the subsequent $L \in \{2, 5, 10, 20\}$ races (given that we observe a time for both racers). As before, we also compute the average race times for the past in order to analyze the change in averages.

2.2.3 Relative Performance

In addition to using the absolute performance in subsequent races, we also use three different measures of relative performance. First, we employ the average position of racer i in races $j+l$ as a measure of future relative performance:

$$\bar{P}_{i,j,\text{future}} = \frac{1}{N_{i,j,\text{future}}} \sum_{l=1}^{L} P_{i,j+l}. \tag{6}$$

Again, we also consider changes in the average position —in the $L \in \{2, 5, 10, 20\}$ races before and after race j— as an outcome variable.

The other two measures of relative performance are based on the relative distance to the winner or median for a given race[9]. In particular, we introduce a new variable $a_{i,j}$ as a standardized measure of racer i's performance in race j:

$$a_{i,j,\text{win}} = \frac{T_{i,j} - T_{win,j}}{T_{win,j}} * \frac{1}{\sigma_{0.9,j}} \tag{7}$$

where $T_{win,j} = \min_s(T_{s,j}),\ \forall s$, denotes the winner's time, and $\sigma_{0.9,j}$ is the standard deviation of the interior 90% of race times.[10]

[9]We discuss the implications with respect to the stable unit treatment value assumption in section 3.2.4 as well as in Appendix B.

[10]We compute the standard deviation based on the interior 90% of race times in order to prevent outliers at the top or bottom from inflating the variance. Otherwise, the presence of a superstar would artificially worsen the performance of other racers. Since some reduction in performance may actually occur (see Brown, 2011), we still capture it by the first fraction. Note that we always cut an equal number of race times at both ends of the distribution. For thirty race times, we trim two at the top and two at the bottom.

One problem with using the winner's time as a reference might be that it is sensitive to outliers caused by superstars. To overcome this issue, we compute a second variable for which we use the median racer as reference:

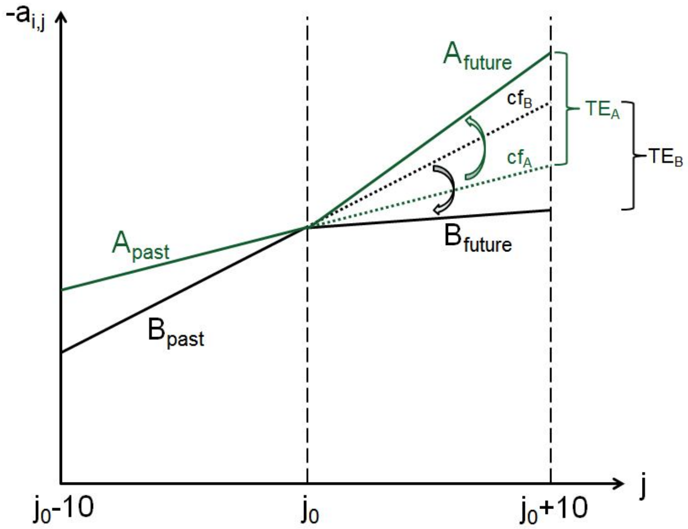
$$a_{i,j\text{med}} = \frac{T_{i,j} - T_{med,j}}{T_{med,j}} * \frac{1}{\sigma_{0.9,j}}. \qquad (8)$$

The two variables, $a_{i,j,\text{win}}$ and $a_{i,j\text{med}}$, provide measures of performance which are comparable across race tracks and over time. They take into account that race tracks vary with respect to length and difficulty. Moreover, we remove the bias that might occur from outliers at the top or bottom. Also, we restrict the sample to races with at least 80% of successful finishers. This is necessary because superior racers are less likely to drop out, thus races with less than 80% of finishers have a very low variance in race times and therefore are less comparable with other races.

2.2.4 Trends

In addition to the outcome variables discussed above, we also intend to estimate the *overall* impact of the ranking. So far, we have mixed together the impact on, for instance, the third and fourth finisher. In order to disentangle the effects on both racers, we suggest an approach that is similar to difference-in-difference (DiD) estimation.

Figure 1: Illustration of DiD-Approach



Note: We use $-a_{i,j} \in \{-a_{i,j,\text{win}}, -a_{i,j\text{med}}\}$ because smaller values represent higher performance with zero being the best possible value. Linearity of trends are only for the purpose of illustration.

In figure 1 we depict two racers which were both improving in the ten races before a particular race j_0. However, racer B was improving more quickly. In race j_0, both racers achieve almost the same time but A was slightly better and thus ranked higher (that is what we call a random success). In the ten races after race j_0, racer A is then improving much faster than racer B.

To estimate the effect of the higher ranking position we follow three steps. First we take the observations (i.e. performances of a racer) before race j_0 and perform a flexible OLS regression of $(-a_{i,j})$ on j to obtain the slope $\beta_{1,i,j}$ for each racer[11]. Then we do the same for the ten observations after race j_0 in order to obtain the subsequent trend in performance, denoted by $\beta_{2,i,j}$.

With respect to figure 1, we can also compare the actual performance trend after race j_0 with the counterfactual one (indicated by the dashed lines). To do so, we estimate the treatment effect for ten races as

$$TE_{i,j,10} = \beta_{2,i,j} - \beta_{1,i,j}. \tag{9}$$

Our hypothesis is that if racer A was lucky to be slightly faster than B, she benefits from the random success and follows a steeper path after race j_0 (compared to the counterfactual path). In contrast, racer B might be disappointed and suffers from the lower ranking position. This disentanglement of the impact on racers A and B allows us to estimate the *overall* impact of the ranking.

2.3 Propositions

The gist of our argument is that all of our outcome variables are affected by a racer's current position, $P_{i,j}$, even if time differences are very small (i.e. random). Thus for our estimation, we hypothesize that the outcomes $(Y_{i,j})$ can be written as functions of three factors: racer i's relative position in race j, her own characteristics summarized as $\mathbf{X}_{i,j}$, as well as her competitors' characteristics $\mathbf{Z}_{i,j}$:

$$Y_{i,j} = h(T_{i,j}, P_{i,j}, \mathbf{X}_{i,j}, \mathbf{Z}_{i,j}) \tag{10}$$

[11]The OLS regression contains polynomials to account for nonlinearities and is only performed if at least four out of ten observations are available.

where the vectors $\mathbf{X}$ and $\mathbf{Z}$ may contain characteristics such as age, gender, nationality, number of podiums in the past, and number of competitors' podiums in the past. For the outcome variable $Y_{i,j}$ we use: $\bar{s}_{i,j,\text{future}}$, $\Delta\bar{s}_{i,j}$, $\bar{T}_{i,j,\text{future}}$, $\Delta\bar{T}_{i,j}$, $\bar{P}_{i,j,\text{future}}$, $\Delta\bar{P}_{i,j}$, $\bar{a}_{i,j,\text{future}}$, $\Delta\bar{a}_{i,j}$, $\bar{\bar{a}}_{i,j,\text{future}}$, $\Delta\bar{\bar{a}}_{i,j}$, or $TE_{i,j}$.

Based on the discussion of the literature as well as our theoretical considerations, we derive several propositions that will be subject to an empirical test in the subsequent sections.

Our first proposition addresses the question whether racers adopt a different risk behavior after a great success. In line with standard economic theory and recent empirical evidence on the risk behavior in professional sports (Paserman, 2010), we expect racers not to adopt a more risky strategy.

Proposition 1. *The probability of survival is not negatively affected by one-time successes.*

The second proposition deals with the effects of relative positions and random successes. Consider two racers, A and B, who finish fifth and sixth in race j. As shown in section 2.1, if the time difference is very small the better ranking can be the result of pure randomness if A is lucky and draws a very low $n_{A,j}$. The same holds true for any other two ranks.

Proposition 2. *Even if differences in ranks result from noise and luck, there is still a positive effect on future performance of achieving a better position in one particular race. Technically, we expect a positive impact on all of our outcomes.*

In addition to the general positive effect of a better rank in one race on subsequent performance, there should be a pronounced effect at important thresholds in the classification. Winning a particular race j, or achieving the podium or top ten should boost a racer's subsequent achievements. Thus when comparing 1st and 2nd as well as 3rd and 4th or 9th and 10th finishers (even if time differences are very small), we expect a considerable impact on future performance.

Proposition 3. *There is a large positive effect of finishing first as compared to finishing second, even if time differences are tiny. Similarly, we expect a pronounced positive impact of finishing third (thus achieving the podium) as compared to finishing forth. To less extent we expect a pronounced positive effect of finishing 9th versus 10th.*

With respect to the impact on subsequent performance, we expect effect heterogeneity. Following Klaassen and Magnus (2001) we assume that one-time great successes have a larger impact if they are achieved by racers with no such successes in the past. For them, a victory or podium finish is indeed 'a big shot'. In a similar vein, we assume that the effects are larger for young racers as they (and their teams) are less aware of their true potential. This is in line with previous research by Oreopoulos, von Wachter and Heisz (2012). In the estimation, we analyze this by reducing the sample to racers with low ages and little experience.

Proposition 4. *The positive effect of a victory or podium finish on subsequent performance is larger if the racer has never achieved a victory or podium finish in previous races and for young racers as compared to old racers.*

Our final proposition concerns the time window we use for the estimation. As outlined above, there are different channels for the effect of a random success on subsequent performance.

Proposition 5. *If the 'winning mood channel' is crucial, we should see a large effect of a better ranking position on the performance in the short term. For the 'goal-setting and support channel', we expect an effect in the long term.*

3 Econometric Framework

3.1 Regression Discontinuity

Regression Discontinuity (RD) designs have the great advantage that they require only mild assumptions for the estimation of causal effects. Lee (2008) argues that randomized variation around some threshold results from agent's inability to precisely control the assignment variable near the known cutoff. In ski races, racers (i) cannot manipulate strategically their race time, and (ii) usually do not know the cutoff in advance. Yet, distinctive cutoff times determine which racer wins the race or who achieves a podium finish. If these successes have a causal impact on subsequent performance, using an RD design to estimate the effect seems appealing: since individuals cannot manipulate their race times around the cutoff, variation in treatment very close to the threshold is arguably randomized.[12]

3.2 Econometric Setup

In any world cup ski race j, racer i achieves a podium finish if her race time $T_{i,j}$ is superior to a cutoff time c_j which is determined by her competitors. This setting generates a *sharp* discontinuity in the treatment (e.g. achieving the podium) as a function of race times. We denote the receipt of treatment by the dummy variable $D_{i,j} \in \{0, 1\}$ such that $D_{i,j} = 1$ if $T_{i,j} \leq c_j$ and $D = 0$ if $T_{i,j} > c_j$.

For the estimation, we assume that except for the treatment there is no reason why subsequent performance $Y_{i,j}$ should be a *discontinuous* function of the race time. Therefore, we regard any discontinuity in $Y_{i,j}$ at the cutoff level c_j as the causal effect of the treatment. Assuming the relationship between $Y_{i,j}$ and $X_{i,j}$ to be linear, we can estimate the treatment effect $\tau_{i,j}$ by fitting the linear regression

$$Y_{i,j} = \alpha_j + T_{i,j}\beta_{i,j} + D_{i,j}\tau_{i,j} + D_{i,j}T_{i,j}\phi_{i,j} + \mathbf{X}_{i,j}\,\gamma_{i,j} + \mathbf{Z}_{i,j}\,\delta_{i,j} + \varepsilon_{i,j} \tag{11}$$

where α_j captures race-fixed effects, $\phi_{i,j}$ allows for different functional forms to the left and right of the cutoff, and $\varepsilon_{i,j}$ is the usual error term[13]. Note that we allow for heterogeneity of

[12]The benefit of a regression discontinuity design over a simple OLS estimation is that the latter would result in biased estimates if some unobservable variable affects both the treatment (higher ranking position) and the outcome (subsequent performance).

[13]It may be problematic to assume the same slope on both sides of the cutoff ($\tau_l = \tau_r$). If this assumption

treatment effects ($\tau_{i,j}$) The inclusion of control variables improves precision of the estimation (Froelich, 2007). Following the work by Lee and Lemieux (2010), we also include the lagged value of the dependent variable in order to lower the variance in the RD estimator.

We expect the outcomes not only to be affected by one-time great successes but also by higher ranking positions in general. Using the previous notation, we then define treatment, $D_{i,j}$, as any difference between two ranking positions.

According to Lee and Lemieux (2010), the use of an RD design allows for the estimation of a 'weighted' average treatment effect where the weights are the relative ex ante probability that the value of an individual's assignment variable will be in the neighborhood of the threshold. However, since we do not observe the density of the assignment variable at the individual level, we cannot estimate the weight for each individual.

3.2.1 Parametric vs. Nonparametric Estimation

Since the seminal work by Hahn, Todd and van der Klaauw (2001), it has been common to rely on nonparametric estimators in regression discontinuity settings. Since kernel estimation is problematic, Hahn et al. suggest to use local linear regression (i.e. linear regressions within the bins on both sides of the cutoff). Lee and Lemieux (2010) also argue that nonparametric estimation (such as local linear regression) should be viewed as a complement to —rather than a substitute for— parametric estimation. A simple way of relaxing the linearity assumption is to include polynomial functions of $T_{i,j}$ in the regression model. One problem with this approach, however, is that data far away from the cutoff might be used to predict the value of Y at the cutoff.

In our estimation, we rely on the nonlinear estimator provided by (Imbens and Kalyanaraman, 2012) which is based on local linear regressions. Following Lee and Lemieux (2010), we test the robustness of our estimates with respect to (i) the inclusion of polynomials and (ii) changes in the bandwidth.

---

was plausible, however, one could get more efficient estimators by imposing this assumption.

3.2.2 Identifying Assumptions

In order to estimate causal effects in an RD setting, four assumption have to be fulfilled[14]. First, the treatment must be determined exclusively by the assignment variable. This is obviously true in our case since a racer's time determines the position (i.e. treatment). Second, there has to be a discontinuity in the level of treatment at the cutoff. Again, this is straightforward in our case (see table 4). Third, we must exclude the possibility of manipulation around the cutoff (McCrary, 2008). Since all racers want to be as fast as possible and do not know about the threshold when racing, manipulation is not an issue. Also the density of observations is smooth around the cutoff. Finally, the distributions of other variables must be smooth around the cutoff (Hahn, Todd and van der Klaauw, 2001; Imbens and Lemieux, 2008). We provide evidence for this in the following section 3.2.3.

3.2.3 Testing the Assumptions

to be added: empirical evidence in favor of the identifying assumptions; as well as a discussion of risky behavior which may affect our estimates

There is a simple two-step procedure provided by (McCrary, 2008). The idea is to check whether the distributions of observed covariates do not change discontinuously at the cutoff. First the frequencies within bins to the left and right of the cutoff are computed. Then the frequencies are used as the dependent variable in a local linear regression.[15]

3.2.4 SUTVA

The stable unit treatment value assumption (SUTVA) is crucial for the estimation of causal effects. In order to justify this assumption, we control for the treatment status and characteristics of the reference racer r (i.e. winner or median racer). With this control, we assume that the outcome for racer i should be independent of the treatment of the reference racer: $(Y_i|X_r, P_{r,j})\perp P_{r,j}$. In the robustness section 5.5, we also reduce the sample to non-treated racers to provide further evidence justifying the SUTVA.

[14]According to Heckman, LaLonde and Smith (1999), RD estimators are a special case of selection on observables. The first crucial assumption for the estimation is that treatment is randomly assigned. Using an RD design, this assumption is trivially satisfied. However, the second key assumption, overlap, is clearly violated. Thus we have to impose the assumption of continuity of all other factors to compensate for the failure of the overlap condition.

[15]This test requires the assumption that the treatment effect is not negative for some individuals.

3.2.5 Bandwidth Choice

We first chose the optimal bandwidth according to Imbens and Kalyanaraman (2012) and then include various lagged performance measures such as a racer's number of victories and podiums, his number of top five classifications, his experience measured by the total of successfully finished races, and his average performance in the last two races[16]. This procedure is conservative because the optimal bandwidth is chosen in order to minimize the mean squared error (MSE), equal to bias squared plus variance. Thereafter, the basic estimator proposed by Imbens and Kalyanaraman (2012) compares the two means on both sides of the threshold, including a linear function of the running variable. In contrast, we are more careful by introducing additional covariates for current and past performance. It is particularly important to control for past performance when using changes in the future performance (as compared to past performance) as outcome. Similar to controlling for level effects in the growth literature, we have to use this control variable since racers who won the last races cannot improve their performance.

[16]Instead of using the data-driven optimal bandwidth method by Imbens and Kalyanaraman (2012) we could also use Cross-Validation (CV) to compute a bandwidth. However, either method delivers smaller or wider bandwidths. Since we vary the applied bandwidth in the estimation anyway, choosing the method by Imbens and Kalyanaraman does not influence our results.

4 Data

4.1 Basics about Alpine Ski Races

The origins of alpine skiing competitions go back to the 1930s when 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 host country, only few racers came from other countries. The *Fédération Internationale de Ski* (FIS) World cup was launched in 1966. In the first years, it included three disciplines: slalom, giant slalom, and downhill races. It was in 1974 when combined races were included, while super G complemented the individual race disciplines in 1982.

Alpine skiing provides a unique real-world setting to examine the effect of a one-time success on subsequent performance. Ski athletes are highly intrinsically motivated, yet the extrinsic motivation, in the form of monetary rewards and international fame, are likely to play a central role for individual performance as well. The prize money of all ten top athletes in the season of 2011/2012 sums up to \$2,740,941 and \$2,695,283 for women, respectively (FIS 2012). However, the distribution of income in prize money is highly skewed as illustrated by figure 2. While the highest income was \$511,875 for men and \$614,437 for women, number ten of the ranking earned only \$182,490 and \$132,409, respectively[17]. A considerable fraction of 76% of male and 80% of female racers earned less than \$50,000. Besides the prize money, success in the races also comes with better sponsorship contracts.[18]

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 racers run the same course. There are five different disciplines, which differ in terms of the vertical and horizontal distance between the gates and the horizontal distance between start and finish. These facts can be best illustrated by looking at the traditional races in Wengen, Switzerland. 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 racers cover a difference in altitude of more than 1,000m from start to finish which is in contrast to the slalom racers' distance that

[17]Note that these prizes are large enough to incentivize racers to exert high effort but not too large to cause one-time winners to reduce their subsequent efforts.

[18]Unfortunately, there is no reliable data on sponsorship incomes. Insiders estimate that, for top athletes, this source of income makes up three to four times the amount of prize money.

is less than 400m. Figure A.4 in the appendix depicts the vertical and horizontal differences between downhill and slalom races in Wengen. Needless to say that the differences in gate distance and altitude 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.

Giant slalom and super G are nested between the two extremes both in terms of gates and speed, whereby giant slalom is closer to slalom and super G is closer to downhill. The results of the traditional combined races are determined by adding up results of separate downhill and slalom competitions. In 2005, the FIS introduced the 'super combined' which consists of individual runs in both downhilll (or super G) and slalom (only one run) to replace former combined races. An exception is the last traditional combined race in Kitzbühl.

Alpine ski competitions enjoy great popularity, mainly in Europe. The downhill race in Wengen (Switzerland) counts on TV quota of over one million in Switzerland for each of the races between 2007 and 2012, which is around 1/8 of the country's population (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. The high importance of alpine skiing in society is illustrated by awards won by ski athletes. Over the last thirty years, ski athletes won twenty-nine awards (out of sixty possible) for the best sports athlete in Austria and sixteen in Switzerland.

4.2 Ski World Cup Results

We use a panel data set of 619 male and 522 female athletes in all World cup ski races for the period of 1992-2013. The data include each racer's exact result (in hundredths of a second), her distance to the winner, as well as gender, age, and the discipline of competition.The panel structure allows us to measure each racer's performance in the races after a certain relative position is reached.

Figure A.5 in the appendix presents a boxplot of the winner's time for each discipline. While combined races tend to be the longest on average, slalom races are considerably shorter. This feature creates problems when comparing performance based on absolute time differences across different disciplines. We overcome this obstacle by creating a measure that corrects for the different course lengths. As was shown in section 2.2.3, we use relative measures of performance which are standardized so that we correct for differences across race types and race tracks.

5 Empirical Results

The empirical results have not yet been updated.

5.1 Descriptive Statistics

5.1.1 Overview

Competition in alpine skiing is fierce, only few junior racers make it to the world cup team and, among them, only a small group is successful. From our total sample of 619 male and 522 female athletes, only 100 men (16.5%) and 203 women (17.9%) have ever won a race during their entire career. And a large fraction of the group of winners, 33.0% for men and 43.0% for women, has only won one race. The fraction of racers with at least one podium finish is 26.5% for the sample of men and 28.4% for women. For a classification in the top five and top ten, the numbers are 33.3% (men) and 33.8% (women) and 45.3% (men) and 46.4% (women), respectively.

Figure 3 presents a histogram of the distribution of different measures of career achievement while discarding the zero entries. The graphs are qualitatively similar for all measures and indicate that career success is positively skewed. This suggests that only few competitors are successful over a lifespan which is in line with empirical evidence concerning the presence of few superstars (Rosen, 1981; Hamlen, 1991). 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. Figure A.6 in the appendix shows different achievement measures as well as the total number of races over time. While the average number of races has fairly increased from 34.8 during the 1990s to 36.3 during the 2000s for men and slightly decreased from 35.3 to 34.1 for women, the average number racers who could win at least one race during an entire season decreased from 18.0 to 16.25 (men) and from 18.1 to 14.2 (women). The average number of races per winning racer has been fairly constant over the course of the 20 years in consideration.[19] The average of the absolute value of the time difference between winner and second place, plotted in panels (a) and (b) of figure A.7, is lower than a second for most races except for combined competitions.

[19]The number of races per winner in the men's competition is 1.95 in the 1990s and 2.0in the 2000s. For women we observe a slight increase from 2.1 to 2.4 races per winner indicating that competition has increased over time. Figure A.7 and A.8 in the appendix depict the average time difference between winner and second place as well as between third and fourth place over our sample period for each discipline.

The relative measure, pictured in panels (c) and (d) shows a slight downward trend from around 0.6% to around 0.5% of the winner's final time–depending on the discipline.

Note that combined races exhibit higher time differences and a higher variance over time for three reasons. First, the races tend to be longer. Second, not the full group of racers participate in this discipline which makes competition less fierce. Third, there are only three to four competitions per year as opposed to the other disciplines. In the analysis part, we will restrict our sample to "close" races which will exclude most combined races.

5.1.2 Performance, Victories, and Podium Finishes

Table 2 reports descriptive statistics for all racers and the subsample of podium and winning racers for the period of 1992–2011.[20] The number of podium racers is less than 10% of the total of 48,607 observations. Not surprisingly, winners account for around a third of podium racers. We have more observations for the assignment variable than for the past and future performance. This comes from the fact that we cannot use observations at the beginning (for the measure of past performance) and at end of the season (for future performance). Together they account for around a fourth of the total observations.[21] Also, there are some racers who did not finish any of the latest two and any of the subsequent two races. Yet this is a only 10% and 13% of the relevant observations, respectively.[22]

Although we use changes of performance in our analysis, we also report the level of past and future performance. The number of successfully finished past races is higher for the sample of winners (difference: 21%) and podium racers (difference: 22%) when comparing it with the full sample of racers. Furthermore, both groups are 12% more likely to finish the subsequent two races. The average of past and future time distances to the winner is around half for podium racers when compared to the full sample. In these measures, winners are around 20% faster than podium racers. Furthermore, they count on 25.6 podiums on average (excluding the current podium) which is more than the average podium racer (20.51) and the average racer (7.42). The same pattern occurs when turning to average victories which are most abundant for winning racers (11.01) as opposed to podium racers (8.33) and all

[20] Note that we shrink our sample to the period 1992-2011 since we have data for the top 30 during this period, while we only have data for the top ten in previous years.

[21] We have missing performance measures for 13,927 observations at the beginning of the season and 11,386 observations at the end of the season.

[22] Note that for this calculation we exclude the above mentioned data at the beginning and at the end of the season.

racers (2.59).

It remains to be shown whether, from a purely descriptive point of view, podium and winning racers have different subsequent changes in performance when comparing them to the sample of all racers. For the full sample, the average change in finished races is almost zero with a relatively high standard deviation of 0.67. This shows that most skiers take high risks. The standard deviation is slightly smaller (0.63) for both podium racers and winners. The average change in short term performance is 0.14 for all racers, 0.08 for podium racers, and 0.03 for winners. The positive mean is probably related to the fact that racers with lower skills are getting better during the season relative to high skill racers who are more constant in their performance.[23]

5.1.3 Performance over Time

Before we present the estimates of the regression discontinuity design, we explore the relationship between past, present, and future performance. Table 3 shows a regression of short and long term future performance on performance in current and past races. All coefficients are very precisely estimated and significantly different from zero. While current performance and short term past performance have similar explanatory power, the introduction of past long term performance leads to an increase of the explanatory power of current performance. This hints at the fact that long term past performance is a better measure for a racer's training and fitness ($\tau_{i,j}$).

5.2 Evidence on Random Successes

The key assumption for our estimation is that random noise can have large effects on individual race times and thus cause random successes. In ski races, the individual noise term, $n_{i,j}$, comprises several components. First, alpine skiing is an outdoor event. As a consequence, wind and weather conditions vary significantly over the course of a race. It is most notably snow, wind, and sight which alter individual prospects of success and can also lead to a cancellation of a race if the event is considered to be a risk for the racers' safety.[24] Yet

[23] A way to see that is the standard deviation in future performance which is almost half for winners compared to all racers.

[24] An illustration of the impact of wind is found in an article about the performance of US racer Bode Miller in 2009. "Miller [...] 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.

unstable external conditions that affect the outcome of the race are not considered to be a reason for cancellation and are broadly accepted as a source of variation among competitors. Second, a racer's performance depends on certain key sections of the course which can lead to considerable time variation. An error not only leads to time lost in a certain section, it also affects speed, and thus time, in the following sections.[25] Our data provides information on both the average time differences between two ranks as well as the individual variation of race times *per racer*. While the average time distance between first and second place is 0.51 seconds, the standard deviation of individual distance to the winner is much larger. If we restrict the sample to all racers with at least one second finish and at least four races at a given place, we get an average distance to the winner (at a given race track) of 0.94 seconds. This is 84.3% larger than the average distance between winner and second. Doing the same calculation for the distance to third place, we find that the individual variation (0.87 seconds) is 262.5% larger than the average time distance between third and forth (0.24 seconds). These numbers show that the assignment to the treatment (victory, podium, or higher relative position) often occurs according to time differences that are much smaller than individual time variation[26]. Thus, for racers with similar skill levels, a worsening of weather conditions or a small error can cost them a good final position. This supports our identifying assumption that the treatment is randomly assigned, especially when focussing on close races.

5.3 The Effect of Victories and Podium Finishes

Our main analysis starts by exploring the effect of victory and podium on future performance. In this specification our treatment in equation (11) is binary, in other words $D_{i,j} \in \{0, 1\}$. Table 4 provides evidence that our discontinuity in the running variable to victory and podium is sharp. There is no racer in the treatment group with a negative distance to the treatment. From a total of 43,523 observations in our sample, 1,355 are winners and 4,095

[25] Didier Cuche's last seconds in the Vancouver 2010 Olympics downhill race exemplify that small errors can result in big consequences. 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.

[26] There are two reasons why we do not use training data as a proxy for individual time variation. First, in training sessions racers usually follow a more risky strategy to find the limit for specific bends. Second, they often do not finish the race track in the fastest possible way since they practice certain parts of the race track.

are podium racers.[27]

Turning to the test of our hypotheses formulated above, we first examine the overall effect of victory and podium on our performance measures. Table 5 reports the results for the full sample of racers. Using the optimal bandwidth, a podium result increases the probability of finishing the next races by 0.04 (t-value: 2.59). This result is robust to extending the window to the double of the optimal bandwidth and adding more polynomial terms of the running variable. In contrast, the estimation examining the effect of victory shows no significant effect of a first place using the optimal bandwidth, and an effect of 0.03 (t-value: 2.06) when using twice the bandwidth. Compared to this, there is no effect of victory and podium when exploring short term performance. However, when analyzing long term performance, we find that a podium result positively changes performance by between 0.12 (optimal bandwidth) and 0.18 (double optimal bandwidth with squared terms). The estimated effect of a victory is even higher, it boosts long term performance by 0.20 (t-value: 2.72) when using the optimal bandwidth. This finding provides empirical evidence for proposition 2 which establishes a positive relationship between victory/podium and subsequent performance.

One immediate question is whether this effect is specific to victory and podium or whether it is just an effect of a higher position in the ranking. We replicate the RDD estimation and code all racers with position four to six as treated and all racers from position seven as control. Note that this procedure excludes all podium racers from the estimation because their inclusion would bias our results downwards.[28] Table 6 presents the results. While the impact of a higher position on the probability of successfully finishing a race is about the same as for the podium, there is considerable evidence that a higher ranking also affects short term performance as opposed to the impact of a podium result which was not statistically different from zero. When we consider long term performance, the effect of a victory and podium result is about the same as the effect of position four to six.

To analyze the question whether a victory or a podium has higher effects for younger racers, we perform the RDD estimation for this group of competitors.[29] Table 7 reports the

[27]Note that the proportion of podium positions to winning positions is slightly above three which would be expected in a race with a winner and three podium racers. This comes from the fact that the third place was given to more than one racer in 78 cases. The same holds for the first place which was given to more than one racer in 36 cases.

[28]Including podium racers would either ignore the fact that they also get the treatment of higher position (if they are included in the control group) or make it impossible to separate the podium effect from the effect of a higher position (if they are included in the treatment group).

[29]We define young racers as those with less than ten race results before the race.

results. While there is no effect of victory and podium on successfully finished races, we find large effects for short term performance after a podium result (0.06) which are significantly higher than the estimates for the sample using all racers. There is no effect of victory on short term performance. Furthermore, the effect of both victory and podium on long term performance is around 0.2 and higher than the effect found for the full sample. All in all, the results suggest that the effect of big shots is considerably higher for younger racers which is consistent with proposition 3.

5.4 The Effect of Higher Ranking Positions

The analysis up to this point has shown fairly robust evidence that a racer's relative position affects his future performance. Yet it is unclear whether this effect is driven by the higher position or by the effect of being top in the ranking which is associated with more media presence and possibly higher support by the coaching staff. In this section we focus on the subset of racers whose position is arguably randomized in a given window of positions.[30] In particular, we estimate equation (11) for racers who could have attained a certain minimal (maximal) position but due to bad (good) luck they did not make it. To illustrate this, consider racer i who is on position seven. In our example, the race time differences between positions five to ten are tiny such that, with a different realization of the random variable $n_{i,j}$, racer i could also have attained position five or position ten. If the margin to both minimal and maximal position is very small, we can assume that the relative position in the ranking is random.

Table 8 presents the results of our estimation for the full sample as well as for the subsample of young racers. While we find no significant effect on the probability of finishing a race in any of our categories, we find that a higher position in the ranking is associated with better short term performance for the position categories #6–10 and #11–15. The effect is around 0.01, meaning that finishing tenth instead of sixth improves short term performance by about 5%. For young racers this effect is even 10%. The effect for the positions #11–15 is also around 0.01 and not different between the full sample and young racers. The estimated effects on long term performance are almost identical to the short term results.

Overall, our results suggest two main conclusions. First, the relative position is particu-

[30] We specify this "closeness" by including only racers whose distance to the two bounds of the category is less than 0.5% of the winner's final time. This is less than any result for our optimal bandwidth estimator in the previous section which are around 0.6% to 0.8%.

larly important for racers in the middle of the performance distribution. It seems that there is no effect of the relative position on racers with very high performance levels. Second, we calculated how many rank gains a racer needs to compensate for lack of current performance. We did this for all categories and found that the effect is largest for the short term performance of young racers in category #6–10 who need less than one position (0.98) to compensate. It is highest for the sample of all racers in category #11–15 who need 2.12 positions to compensate.[31] These findings provide empirical support for propositions 1 and 3 and show that even small changes in the relative performance can considerably alter subsequent performance.

5.5 Robustness Tests

To be added.

[31] Note that for the purpose of this calculation we excluded young racers in category #6–10 because their estimate of current performance is not significant.

6 Conclusion

In this paper, we investigate how current relative positions in rankings affect future performance. While many theoretical contributions emphasize the positive effect of good results on subsequent performance through a better self-image, higher knowledge of one own's potential, and addictive success-seeking behavior, it is difficult to separate the effect of current relative positions from current performance. By applying an RDD approach and regressions for randomized positions in a sample of alpine ski racers, we present novel evidence that even small changes in the relative position can lead to higher performance. The results hold mainly for racers in the middle of the current performance distribution. It is notable that we find a treatment effect for victory and podium which is not significantly different from the effect of positions #4–6.

Understanding the economic consequences of relative positions is useful because it extends beyond alpine skiing. It is likely that relative positions also affect performance in economic fields such as sports, education, or performance evaluations in firms. Imagine, for instance, a pupil who is moved to another class and finds herself in a new relative skill distribution. If relative positions play a role, then the new environment should also have an affect on the absolute performance of the pupil. The same argument is likely to hold for workers who change their employer as well as for sport athletes who compete with different sets of rivals.[32]

Another important policy implication from our paper addresses the implementation of rankings in various occupations. This has been discussed in several papers following the seminal work of Lazear and Rosen (1981). It is also relevant in education as some universities have started to publish not only a student's grade but also her relative position. The intention of this policy is to provide employers with more profound information about applicants' study performances. To our knowledge, however, there has not been any study on the effects of publishing rankings on student's efforts and performance.

[32]This line of reasoning is related to the 'big fish in a small pond'-literature (Marsh and Parker, 1984).

References

Anderson, Michael L. 2012. "The Benefits of College Athletic Success: An Application of the Propensity Score Design with Instrumental Variables." *NBER Working Paper #18196.*

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.

Azmat, Ghazala, and Nagore Iriberri. 2010. "The importance of relative performance feedback information: Evidence from a natural experiment using high school students." *Journal of Public Economics*, 94(7-8): 435–452.

Becker, Gary S., and Kevin M. Murphy. 1988. "A theory of rational addiction." *The Journal of Political Economy*, 96(4): 675–700.

Bénabou, Roland, and Jean Tirole. 2002. "Self-confidence and personal motivation." *The Quarterly Journal of Economics*, 117(3): 871–915.

Bénabou, Roland, and Jean Tirole. 2004. "Willpower and personal rules." *Journal of Political Economy*, 112(4): 848–886.

Brown, Jennifer. 2011. "Quitters never win: The (adverse) incentive effects of competing with superstars." *Journal of Political Economy*, 119(5): 982–1013.

Caillaud, Bernard, Daniel Cohen, and Bruno Jullien. 2000. "Towards a theory of self-restraint." Centre d'Enseignement Recherche Analyse Socio-economique, Paris.

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler. 1997. "Labor supply of New York City cabdrivers: One day at a time." *The Quarterly Journal of Economics*, 112(2): 407–441.

Cox, Taylor H, and Celia V Harquail. 1991. "Career paths and career success in the early career stages of male and female MBAs." *Journal of Vocational Behavior*, 39(1): 54–75.

Ehrenberg, RG, and ML Bognanno. 1990. "Do tournaments have incentive effects?" *The Journal of Political Economy*, 98(6): 1307–1324.

Farber, Henry S. 2005. "Is tomorrow another day? The labor supply of New York City cabdrivers." *Journal of Political Economy*, 113(1): 46–82.

Fershtman, Chaim, and Yoram Weiss. 1998. "Social status and economic performance: A survey." *European Economic Review*, 42: 801–820.

FIS - International Ski Federation. 2012. *Audi FIS Alpine Skiing World Cup prize money ranking Season 2012.* Oberhofen, Switzerland.

Frangakis, Constantine E, and Donald B Rubin. 2002. "Principal stratification in causal inference." *Biometrics*, 58(1): 21–9.

Froelich, Markus. 2007. "Regression Discontinuity Design with Covariates."

Ginther, DK, and Shulamit Kahn. 2004. "Women in Economics: Moving Up or Falling Off the Academic Career Ladder?" *Journal of Economic Perspectives*, 18(3): 193–214.

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." *Review of Economics and Statistics*, 73(4): 729–733.

Heckman, James J., Robert J. LaLonde, and Jeffrey A. Smith. 1999. "The Economics and Econometrics of Active Labor Market Programs." In *Handbook of Labor Economics.* Vol. III. 3A ed., , ed. Orley Ashenfelter and David Card, 1865–2097. Amsterdam, New York, and Oxford:Elsevier Science.

Hilmer, C, and Michael Hilmer. 2007. "Women helping women, men helping women? Same-Gender mentoring, initial job placements, and early career publishing success for economics PhDs." *The American Economic Review (Papers and Proceedings)*, 97(2): 422–426.

Imbens, Guido W., and Karthik Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *The Review of Economic Studies*, 79(3): 933–959.

Imbens, Guido W., and Thomas Lemieux. 2008. "Regression discontinuity designs: A guide to practice." *Journal of Econometrics*, 142(2): 615–635.

Kahneman, Daniel, Peter P. Wakker, and Rakesh Sarin. 1997. "Back to Bentham? Explorations of experienced utility." *The Quarterly Journal of Economics*, 112(2): 375–406.

Kirby, Kris N., and Barbarose Guastello. 2001. "Making choices in anticipation of similar future choices can increase self-control." *Journal of Experimental Psychology: Applied*, 7(2): 154–164.

Klaassen, Franc J. G. M., and Jan R. Magnus. 2001. "Are Points in Tennis Independent and Identically Distributed? Evidence From a Dynamic Binary Panel Data Model." *Journal of the American Statistical Association*, 96(454): 500–509.

Kuhnen, Camelia M., and Agneieszka Tymula. 2012. "Feedback, self-esteem, and performance in organizations." *Management Science*, 58(1): 94–113.

Lazear, Edward P., and Sherwin Rosen. 1981. "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy*, 89(5): 841–864.

Lee, David S. 2008. "Randomized experiments from non-random selection in U.S. House elections." *Journal of Econometrics*, 142(2): 675–697.

Lee, David S., and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*, 48(June): 281–355.

Marsh, Herbert W., and John W. Parker. 1984. "Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well?" *Journal of Personality and Social Psychology*, 47(1): 213–231.

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

Morgan, John, Dana Sisak, and Felix Vardy. 2012. "On the Merits of Meritocracy."

Oreopoulos, Philip, Till von Wachter, and Andrew Heisz. 2012. "The Short- and Long-Term Career Effects of Graduating in a Recession." *American Economic Journal: Applied Economics*, 4(1): 1–29.

Paserman, M Daniele. 2010. "Gender Differences in Performance in Competitive Environments? Evidence from Professional Tennis Players." *Working Paper, Boston University.*

Quattrone, George A., and Amos Tversky. 1984. "Causal versus diagnostic contingencies: On self-deception and on the voter's illusion." *Journal of Personality and Social Psychology*, 46(2): 237–248.

Rachlin, Howard. 2000. *The Science of Self-Control.* Cambridge, Mass.:Harvard University Press.

Rege, Mari. 2008. "Why do people care about social status?" *Journal of Economic Behavior & Organization*, 66(2): 233–242.

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

Rosen, Sherwin, and Allen Sanderson. 2001. "Labour Markets in Professional Sports." *The Economic Journal*, 111(469): F47–F68.

Santos-Pinto, Luis. 2008. "Positive Self-image and Incentives in Organisations." *The Economic Journal*, 118(531): 1315–1332.

Santos-Pinto, Luis. 2010. "Positive Self-Image in Tournaments." *International Economic Review*, 51(2): 475–496.

Santos-Pinto, Luis, and Joel Sobel. 2005. "A model of positive self-image in subjective assessments." *The American Economic Review*, 95(5): 1386–1402.

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

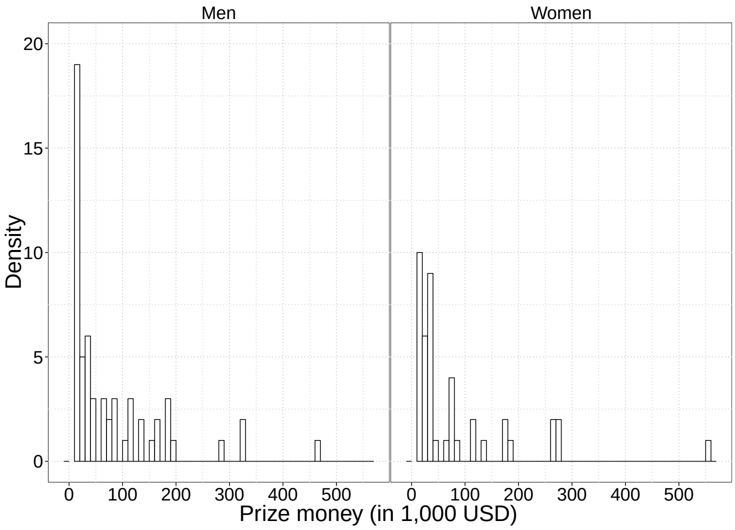
Sloof, Randolph, and C. Mirjam van Praag. 2010. "The effect of noise in a performance measure on work motivation: A real effort laboratory experiment." *Labour Economics*, 17(5): 751–765.

Sunde, Uwe. 2009. "Heterogeneity and performance in tournaments: a test for incentive effects using professional tennis data." *Applied Economics*, 41(25): 3199–3208.

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.

Figures

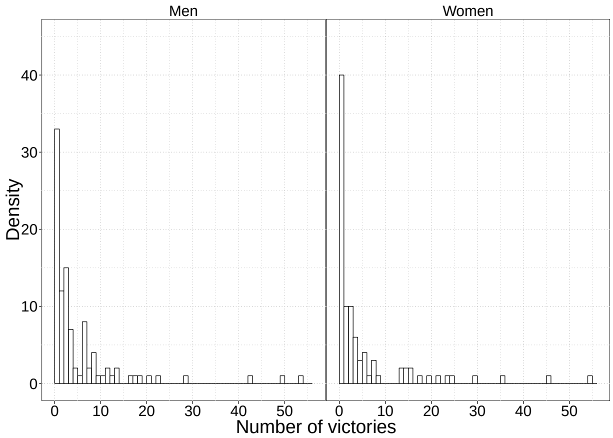
Figure 2: Prize money 2011/2012



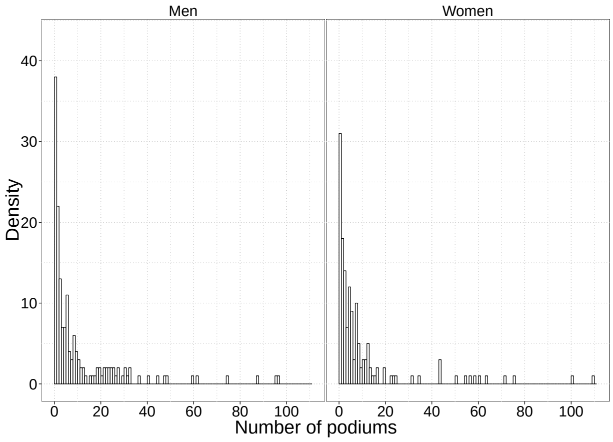
Note: Histogram of prize money in 1,000 USD for the season 2011/2012. Source: FIS 2012.

Figure 3: Distribution of career achievements

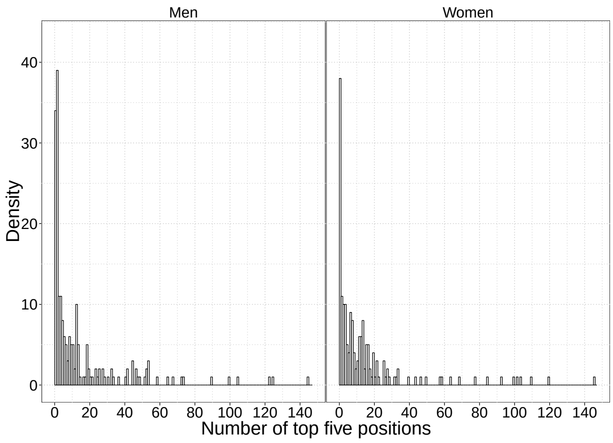
(a) Nr. of victories



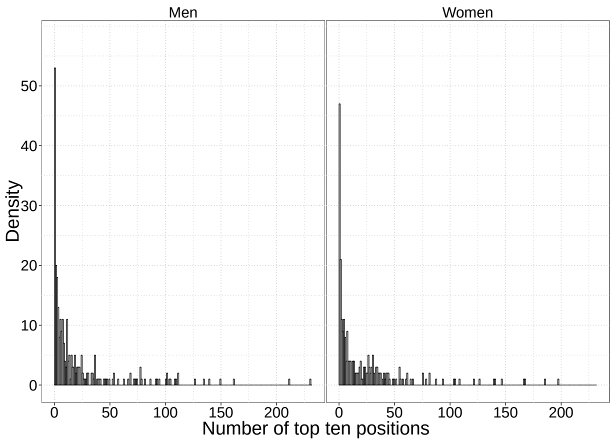
(b) Nr. of podiums



(c) Nr. of top five classifications



(d) Nr. of top ten classifications



Note: Histogram of total number of victories, total number of podiums, total number of top give classifications, and total number of top ten classifications during a racer's active career. Note that we discard zero entries for illustration purposes. The number of racers who had never won a racer during their career is 507 and 428 for men and women, respectively. The number of racers without podium is 446 (men) and 373 (women). For top five and top ten classifications the numbers are 405 (men) and 345 (women) and 332 (men) and 279 (women), respectively.

Table 2: Descriptive statistics

Category	Variable	Mean	Std. Dev.	Min.	Max.	N
(A) All racers						
Outcomes	Change in finished races	-0.01	0.67	-1.1	1.1	48884
	Change in short term in performance	0.14	0.35	-3.26	2.38	26473
	Change in long term performance	0.02	0.35	-3.35	2.35	26481
Treatment	Victory	0.03	0.17	0	1	48884
	Podium	0.09	0.28	0	1	48884
Assignment	Distance to winner	2.42	1.85	0	68.58	48607
	Distance to podium	1.77	1.78	-3.61	67.36	47618
Performance	Future time distance (short term)	2.21	1.55	0	68.58	35814
	Future time distance (long term)	2.2	1.5	0	68.58	35814
	Future time distance (long term)	2.22	1.37	0	33.62	36070
	Future time distance (short term)	2.22	1.43	0	33.62	36070
	No. podiums at time of race	7.42	15.02	0	110	48884
	No. victories at time of race	2.57	6.49	0	54	48884
	Future finished races	1.57	0.71	0	2	34957
	Past finished races	1.66	0.66	0	2	37498
(B) Podium racers						
Outcomes	Change in finished races	-0.02	0.63	-1.1	1.1	4158
	Change in short term in performance	0.08	0.42	-2.24	2.26	2738
	Change in long term performance	0.02	0.4	-2.24	2.06	2740
Treatment	Victory	0.33	0.47	0	1	4158
	Podium	1	0	1	1	4158
Assignment	Distance to winner	0.37	0.46	0	3.75	4119
	Distance to podium	-0.31	0.43	-3.61	0	4095
Performance	Future time distance (short term)	1.04	0.95	0	16.7	3361
	Future time distance (long term)	1.06	0.89	0	16.7	3361
	Future time distance (long term)	1.11	0.82	0	8.49	3432
	Future time distance (short term)	1.08	0.92	0	15.59	3432
	No. podiums at time of race	20.51	22.11	0	109	4158
	No. victories at time of race	8	10.64	0	54	4158
	Past finished races	1.9	0.36	0	2	3030
	Future finished races	1.87	0.42	0	2	2973
(C) Winners						
Outcomes	Change in finished races	-0.03	0.63	-1.1	1.1	1392
	Change in short term in performance	0.03	0.42	-2.24	1.82	926
	Change in long term performance	0	0.4	-2.24	1.82	927
Treatment	Victory	1	0	1	1	1392
	Podium	1	0	1	1	1392
Assignment	Distance to winner	0	0	0	0	1379
	Distance to podium	-0.67	0.48	-3.61	0	1355
Performance	Future time difference (short term)	0.84	0.88	0	13.01	1127
	Future time difference (long term)	0.88	0.84	0	13.01	1127
	Past time difference (long term)	0.88	0.72	0	6.44	1162
	Past time difference (short term)	0.84	0.78	0	6.61	1162
	No. podiums at time of race	24.6	23.32	0	102	1392
	No. victories at time of race	10.01	11.58	0	54	1392
	Past finished races	1.92	0.32	0	2	1017
	Future finished races	1.88	0.4	0	2	995

Note: Table presents descriptive statistics for outcome variables, assignment variables, and performance level measures. Past (future) short term performance is measure in the average distance to the winner in the last (next) two races, present performance is distance to winner in the current race. For the variable past (future) finished races we exclude all observations at the beginning (end) of the season.

Table 3: Regression of performance level over time

	Short term performance				Long term performance			
Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8))
Current performance	0.275	0.169	0.260	0.161	0.267	0.164	0.252	0.155
	(0.014)	(0.012)	(0.016)	(0.013)	(0.013)	(0.011)	(0.014)	(0.011)
Past short term perf.	0.349	0.213			0.351	0.220		
	(0.019)	(0.014)			(0.020)	(0.014)		
Past long term perf.			0.390	0.257			0.392	0.263
			(0.025)	(0.019)			(0.026)	(0.020)
Constant	0.770	1.273	0.713	1.198	0.775	1.264	0.718	1.190
	(0.048)	(0.038)	(0.049)	(0.041)	(0.051)	(0.040)	(0.053)	(0.043)
Racer FE		✓		✓		✓		✓
R-squared	0.264	0.082	0.273	0.090	0.274	0.090	0.285	0.098
Observations	26473	26473	26473	26473	26473	26473	26473	26473

Note: Table shows regression of future performance on present and past performance. Past (future) short term performance is measured by the average distance to the winner in the last (next) two races, present performance is measured by distance to winner in the current race. Standard errors are clustered (at racer level).

Table 4: First stage victory and podium

		Treatment			
		Victory		Podium	
		0	1	0	1
Assignment	0	43,523	0	43,523	0
	1	0	1,355	0	4,095

Note: Table shows assignment to treatment and treatment (victory and podium). Assignment to treatment measures the time difference between a racer's time and victory/podium. It takes the value 1 if this difference is negative or zero, i.e. if a racer is faster (or equally timed) than the first/third. Sample includes seasons 1992-2011.

Table 5: RDD: Full sample

	Podium		Victory	
Effect on finishing race				
Optimal BW (OBW)	0.044***	18'883	0.005	12'103
	(0.017)		(0.024)	
2 times OBW and squared terms	0.04***	30'620	0.033**	27'191
	(0.011)		(0.016)	
2 times OBW and tripled terms	0.043***	30'620	0.032**	27'191
	(0.014)		(0.016)	
Effect on short term performance				
Optimal BW (OBW)	-0.030	5'071	0.110	2'461
	(0.055)		(0.094)	
2 times OBW and squared terms	0.030	9'630	0.063	5'769
	(0.038)		(0.096)	
2 times OBW and tripled terms	-0.149	9'630	0.028	5'769
	(0.096)		(0.484)	
Effect on long term performance				
Optimal BW (OBW)	0.130***	5'515	0.204***	2'728
	(0.047)		(0.075)	
2 times OBW and squared terms	0.175***	10'437	0.181***	6'592
	(0.027)		(0.062)	
2 times OBW and tripled terms	0.119*	10'437	0.131	6'592
	(0.064)		(0.283)	

Note: Table shows regression of short- and long-term performance measures on victory and podium using the optimal bandwidth as suggested by Imbens and Kalyanaraman (2012). All regression include the number of podiums, victories, and top five classifications before the race, the past performance as well as various polynomials of the running variable. Sample includes seasons 1992–2011. Standard errors are clustered (by racer). Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 6: **RDD, Effect of positions #4–6**

	Positions #4–6	
	Estimate	N
Effect on finishing race		
Optimal BW (OBW)	0.038**	17'431
	(0.018)	
2 times OBW and squared terms	0.035*	23'761
	(0.018)	
2 times OBW and tripled terms	0.039**	23'761
	(0.018)	
Effect on short term performance		
Optimal BW (OBW)	0.065***	7'740
	(0.022)	
2 times OBW and squared terms	0.068***	12'276
	(0.012)	
2 times OBW and tripled terms	0.062***	12'276
	(0.013)	
Effect on long term performance		
Optimal BW (OBW)	0.167***	7'623
	(0.021)	
2 times OBW and squared terms	0.147***	12'153
	(0.012)	
2 times OBW and tripled terms	0.144***	12'153
	(0.014)	

Note: Table shows regression of short- and long-term performance measures on a position between four and six using the optimal bandwidth as suggested by Imbens and Kalyanaraman (2012). All regression include the number of podiums, victories, and top five classifications before the race, the past performance as well as various polynomials of the running variable. Sample includes seasons 1992–2011. We discard positions one to three. Standard errors are clustered (by racer). Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 7: RDD: Sample at beginning of career

	Podium		Victory	
Effect on finishing race				
Optimal BW (OBW)	0.047		-0.061	2'660
	(0.046)		(0.092)	
2 times OBW and squared terms	0.052*	13'943	0.108*	9'523
	(0.028)		(0.062)	
	0.033	13'943	0.099	9'523
2 times OBW and tripled terms	(0.044)		(0.167)	
Effect on short term performance				
Optimal BW (OBW)	0.117**	3'046	0.079	1'270
	(0.046)		(0.092)	
2 times OBW and squared terms	0.095***	6'681	0.052	4'346
	(0.027)		(0.059)	
2 times OBW and tripled terms	0.106***	6'681	0.021	4'346
	(0.037)		(0.159)	
Effect on long term performance				
Optimal BW (OBW)	0.218***	5'503	0.205***	3'194
	(0.023)		(0.045)	
2 times OBW and squared terms	0.2***	9'918	0.176***	8'688
	(0.02)		(0.035)	
2 times OBW and tripled terms	0.197***	9'918	0.174***	8'688
	(0.025)		(0.035)	

Note: Table shows regression of short- and long-term performance measures on victory and podium using the optimal bandwidth as suggested by Imbens and Kalyanaraman (2012). All regression include the number of podiums, victories, and top five classifications before the race, the past performance as well as various polynomials of the running variable. Sample includes racers with less than five podiums for the seasons 1992–2011. Standard errors are clustered (by racer). Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 8: Effect of Relative Position

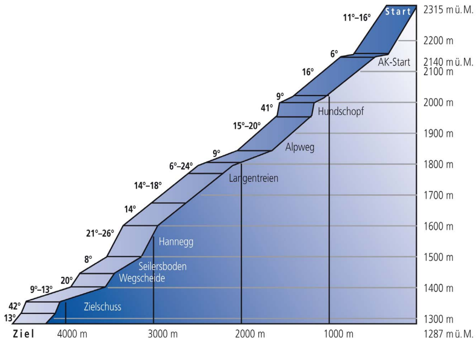
	Full set of racers					Young racers				
	1–5	6–10	11–15	16–20	21–25	1–5	6–10	11–15	16–20	21–25
Effect on finishing race										
Position	0.007	-0.001	-0.001	-0.000	0.003	0.019	-0.001	0.003	-0.004	0.010
	(0.009)	(0.003)	(0.003)	(0.004)	(0.004)	(0.017)	(0.006)	(0.006)	(0.006)	(0.006)
Current performance	-0.044	0.026***	0.019**	0.005	-0.001	-0.116	0.026	0.018	0.011	-0.006
	(0.063)	(0.010)	(0.009)	(0.009)	(0.009)	(0.137)	(0.017)	(0.014)	(0.013)	(0.013)
Effect on short term performance										
Position	-0.006	-0.011**	-0.008**	-0.004	-0.006	0.000	-0.021**	-0.011**	0.004	-0.009
	(0.010)	(0.005)	(0.004)	(0.004)	(0.005)	(0.017)	(0.009)	(0.006)	(0.006)	(0.007)
Current performance	-0.162**	-0.046***	-0.051***	-0.049***	-0.010	-0.165	-0.025	-0.042**	-0.062***	-0.015
	(0.078)	(0.015)	(0.012)	(0.010)	(0.012)	(0.137)	(0.022)	(0.016)	(0.015)	(0.016)
Effect on long term performance										
Position	-0.004	-0.011**	-0.007*	-0.004	-0.005	-0.003	-0.024***	-0.012**	0.001	-0.009
	(0.009)	(0.004)	(0.004)	(0.004)	(0.004)	(0.017)	(0.008)	(0.006)	(0.006)	(0.007)
Current performance	-0.112*	-0.029**	-0.033***	-0.039***	-0.002	-0.173	0.001	-0.038**	-0.056***	-0.006
	(0.070)	(0.014)	(0.012)	(0.010)	(0.011)	(0.134)	(0.021)	(0.016)	(0.015)	(0.016)
Observations	1188	3454	3364	3087	2472	301	1139	1240	1292	1142
Mean of current performance highest rank	0.000	1.114	1.595	2.050	2.397	0.000	1.207	1.686	2.240	2.627
Mean of current performance lowest rank	0.982	1.529	1.927	2.369	2.815	1.042	1.654	2.050	2.552	2.876

Note: Table shows regression of short- and long-term performance measures on position by position category. All regression include a dummy for discipline and gender as well as a number of podiums, victories, and top five classifications before the race, the number of successfully finished races. Also we include the maximum number of podiums of the best racers in the next two (short-term performance) or five (long-term performance) races as a measure for the competitiveness of the next races. Sample includes seasons 1992–2011. Standard errors are clustered (by racer).

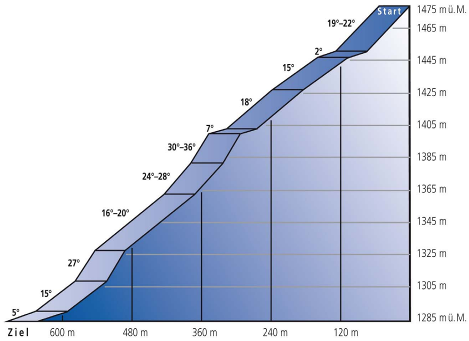
A Additional figures and tables

Figure A.4: Profile of downhill and slalom race

(a) Downhill



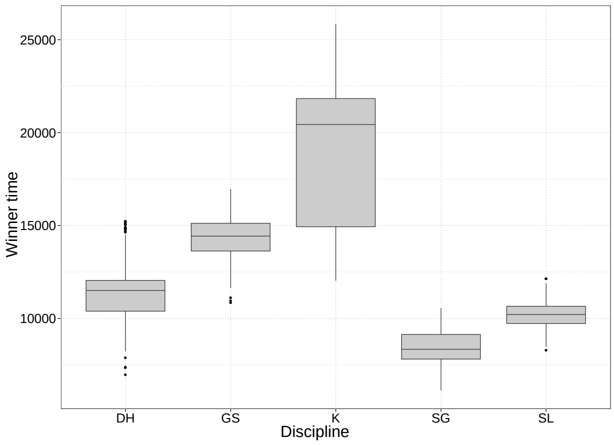
(b) Slalom



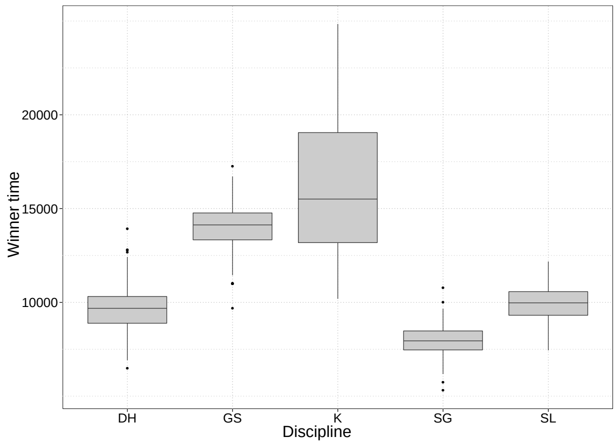
Note: Illustration of the profiles of downhill and slalom race in Wengen. Source: http://www.lauberhorn.ch

Figure A.5: Winner times in different disciplines.

(a) Men



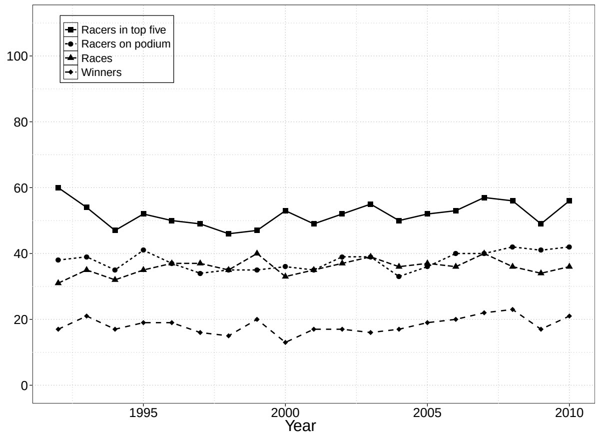
(b) Women



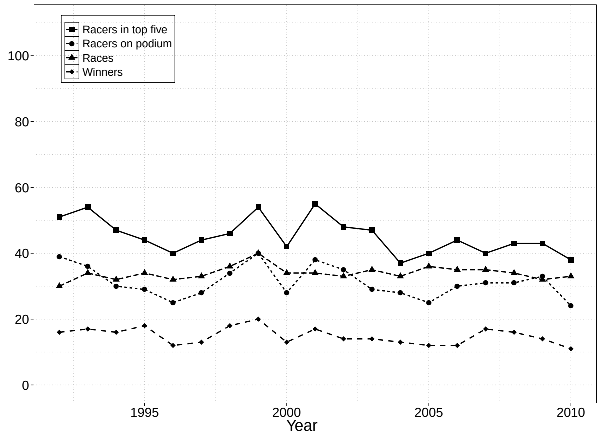
Note: Boxplot of winner times (in hundredths of a second) for each category downhill (DH), giant slalom (GS), comined race (K), super giant slalom (SG), and slalom (SL). Note that 25,000 hundredths of a second (or deciseconds) equal 4.16 minutes, while 15,000 deciseconds equal 2.5 minutes and 10,000 deciseconds equal 1.66 minutes, respectively.

Figure A.6: Racer's classification per season

(a) Men



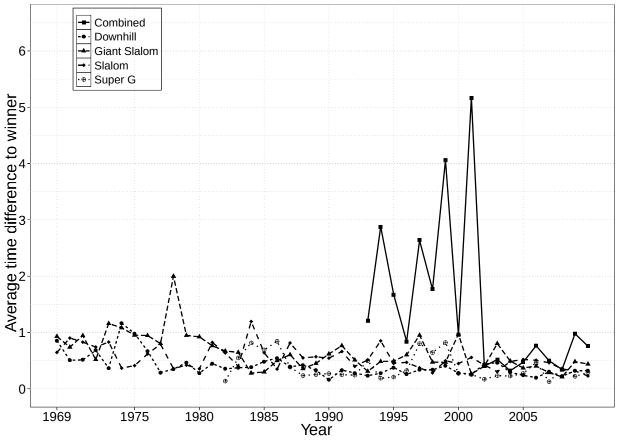
(b) Women



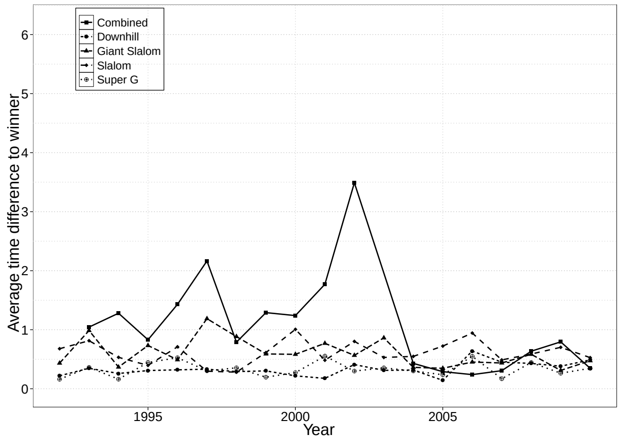
Note: Figure depicts the number of racers with at least one victory in a given year. Furthermore, it also shows the number of races per season as well as the number of racers on the podium, and in the top five. The sample period is 1969–2011.

Figure A.7: Time difference to winner in absolute and relative terms

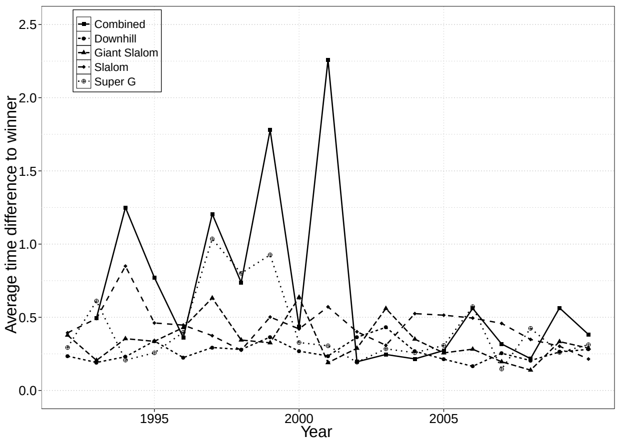
(a) Winner and second (men)



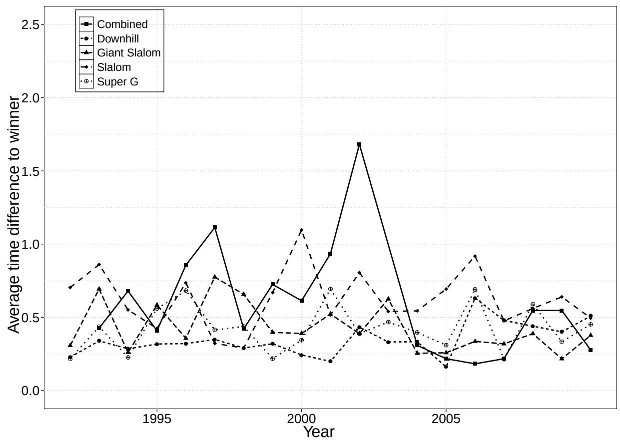
(b) Winner and second (women)



(c) Winner and second (men)



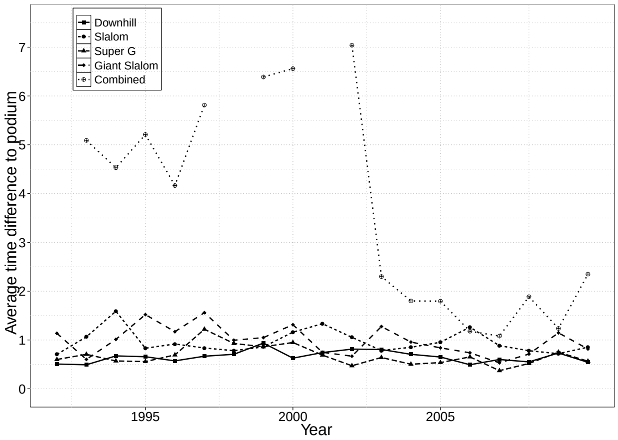
(d) Winner and second (women)



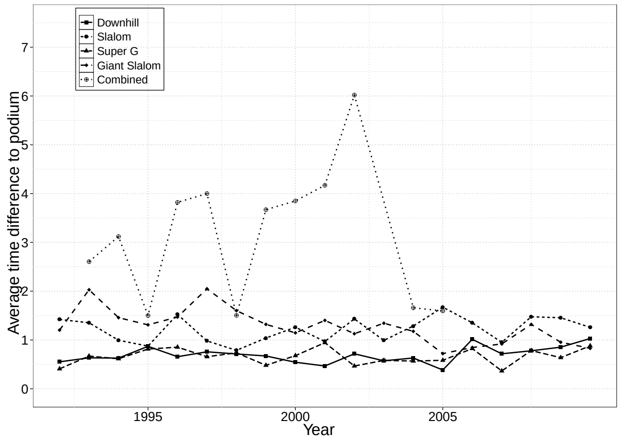
Note: Figure depicts the yearly average time difference between winner and second place racer in both absolute and relative time measures for each discipline. Figures (a) and (b) show the absolute difference in seconds; figures (c) and (d) plot the relative difference in percentages of the winner's final time (as defined in equation (8)). Note that we calculate the difference between first and third racer if two racers take the first place (and so forth).

Figure A.8: Time difference to podium in absolute and relative terms

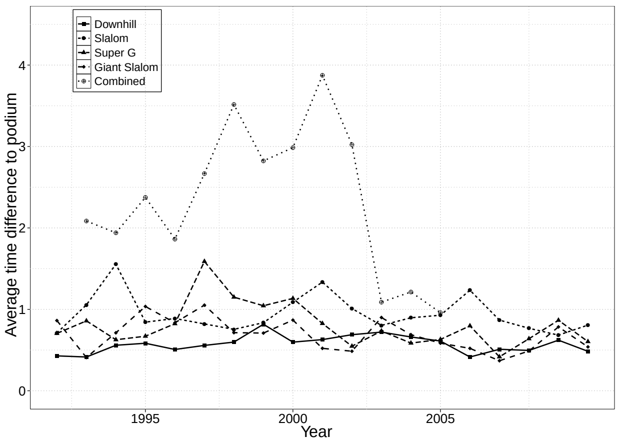
(a) Third and fourth (men)



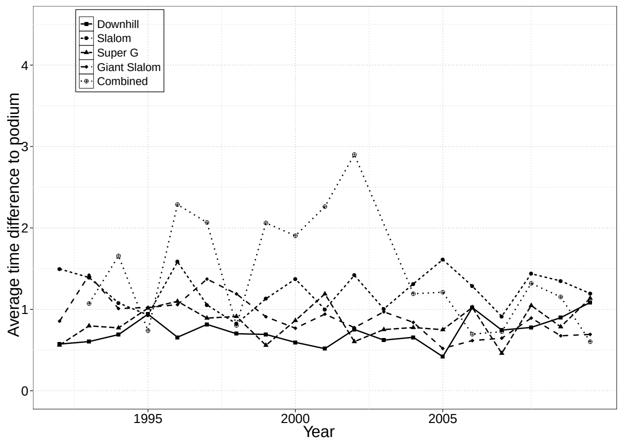
(b) Third and fourth (women)



(c) Third and fourth (men)



(d) Third and fourth (women)



Note: Figure depicts the yearly average time difference between third and fourth place place racers in both absolute and relative time measures for each discipline. Figures (a) and (b) show the absolute difference in seconds; figures (c) and (d) plot the relative difference between third and forth in percentages of the winner's final time (as defined in equation (8)). Note that we calculate the difference between third and fifth racer if two racers take the third place (and so forth).

B SUTVA

B.1 General

The Stable Unit Treatment Value Assumption (SUTVA) is fundamental to most estimators used in the program evaluation literature. It allows (a) to write the treatment status of individual i only dependent on his assignment, and (b) to write the outcome of individual i only dependent on his 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 $Y_i(\mathbf{D}, \mathbf{Z}) = Y_i(D_i, Z_i)$.

B.2 Implications for Our Paper

Let us define the vector of final times in race j looks as follows

$$\mathbf{T} = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ \vdots \\ t_N \end{pmatrix} \tag{12}$$

There is a vector of assignments (to the podium) $\mathbf{Z}$ as well as a vector of treatments $\mathbf{D}$

$$\mathbf{Z} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_N \end{pmatrix} \qquad \mathbf{D} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_N \end{pmatrix} \tag{13}$$

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

$$S_i \begin{cases} 1 \text{ if } S_i^* > 0 \\ 0 \text{ if } S_i^* \leq 0 \end{cases} \tag{14}$$

and

$$S_i^* = \theta_i a_1 + \mu_j + \varepsilon_i \tag{15}$$

where S_i^* is a latent variable, θ_i is a racer's skill, $a_1 > 0$ is a coefficient, μ_j is a race fixed effect, ε_i is an unobserved component. The final time can be written as

$$t_i \begin{cases} \text{NA} & \text{if } S_i = 0 \\ \theta_i a_2 + \delta_j + v_i & \text{if } S_i = 1 \end{cases} \tag{16}$$

$a_2 < 0$ is a coefficient. We consider a specific race with three top racers under two circumstances. First, conditions are equal for all racers. Second, the three top racers ($i \in \{1, 2, 3\}$) suffer bad weather conditions, which makes it impossible for them to attain a place on the podium.

$$\mathbf{Z} = \mathbf{D} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \qquad \mathbf{Z}' = \mathbf{D}' = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{17}$$

In this case, it is obvious that the assignment status of individuals 4–6 depends on the assignment of the three top racers. However, since $D_i = Z_i$, the individual treatment status D_i can 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 $\mathbf{Z} = \mathbf{D}$.

Our time variable is defined as $a_{ij} = \frac{t_{i,j} - t_{\text{win},j}}{t_{\text{win},j}}$ where $t_{\text{win},j}$ is the winner's time. Our outcome for short term performance is the mean of future times

$$\bar{a}_i = \frac{1}{\#k} \sum_{k=j+1}^{j+2} a_{ik} \tag{18}$$

The winner's time can be written as

$$t_{\text{win}j} = \min_{i \in S}(t_{ij}) \tag{19}$$

where S indicates the set of survivors. For the sake of illustration, we specify equation (16) for all survivors as

$$t_{ij} = \theta_i a_2 + \delta_j + \text{Exp}_i a_3 + \text{Exp}_i^2 a_4 + \text{Victory}_i a_5 + u_i \tag{20}$$

where Exp_i is experience and Victory_i is 1 if racer i had won one of the two last races and 0 otherwise. Imagine that the winner in race j is determined by tiny time difference between racer 2 and 4. Both racers have the same skill level θ_i, but while racer 2 is a rookie, racer 4 is an experienced and successful racer. Racer i's relevant outcome $Y_i = \bar{a}_i$ in the next races depends on the performance of the best racer in the next races. So if racer 4 wins today and $a_5 \neq 0$, $t_{\text{win}j+1}$ and $t_{\text{win}j+2}$ are likely to be lower than if racer 2 wins (because racer 4 is more experienced and experience positively affects performance).

Therefore, the outcome of racer i 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)

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

and

$$Y_i(\mathbf{Z}) \neq Y_i(\mathbf{Z}')$$

which violates definition (b) of SUTVA.

We address this potential problem in three different ways. First, we use the absolute times in future races which are not affected by other racers' treatment statuses. Second, when using relative measures of performance ($\bar{P}_{i,j,\text{future}}$, $\Delta\bar{P}_{i,j,\text{future}}$, $\bar{a}_{i,j,\text{future}}$, $\Delta\bar{a}_{i,j}$, $\bar{\tilde{a}}_{i,j,\text{future}}$, or $\Delta\bar{\tilde{a}}_{i,j}$), we control for the treatment status and characteristics of the reference racer (winner or median in subsequent races). Third, in the robustness section we present an estimation for which we restricted the sample to races in which the reference racer is 'new'. That is, we only use data on subsequent races $j + l$ in which the respective reference racer did not participate in race j.