IFAU
Institute for Evaluation of Labour Market and Education Policy

Is there a gender difference in the ability to deal with failures? Evidence from professional golf tournaments

Olof Rosenqvist

WORKING PAPER 2016:14

The Institute for Evaluation of Labour Market and Education Policy (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala. IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market and educational policies, studies of the functioning of the labour market and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

IFAU also provides funding for research projects within its areas of interest. The deadline for applications is October 1 each year. Since the researchers at IFAU are mainly economists, researchers from other disciplines are encouraged to apply for funding.

IFAU is run by a Director-General. The institute has a scientific council, consisting of a chairman, the Director-General and five other members. Among other things, the scientific council proposes a decision for the allocation of research grants. A reference group including representatives for employer organizations and trade unions, as well as the ministries and authorities concerned is also connected to the institute.

Postal address: P.O. Box 513, 751 20 Uppsala
Visiting address: Kyrkogårdsgatan 6, Uppsala
Phone: +46 18 471 70 70
Fax: +46 18 471 70 71
ifau@ifau.uu.se
www.ifau.se

Papers published in the Working Paper Series should, according to the IFAU policy, have been discussed at seminars held at IFAU and at least one other academic forum, and have been read by one external and one internal referee. They need not, however, have undergone the standard scrutiny for publication in a scientific journal. The purpose of the Working Paper Series is to provide a factual basis for public policy and the public policy discussion.

ISSN 1651-1166

Is there a gender difference in the ability to deal with failures? Evidence from professional golf tournaments[a]

by

Olof Rosenqvist[b]

September 15, 2016

Abstract

Recent experimental evidence suggests that women in general are more discouraged than men by failures which potentially can explain why women, on average, are less likely than men to reach top positions in firms. This paper provides the first quasi-experimental evidence from the field on this issue using data from all-female and all-male professional golf tournaments to see if this result can be replicated among competitive men and women. These top-performing men and women are active in an environment with multiple rounds of competition and the institutional set-up of the tournaments makes it possible to causally estimate the effect of the result in one tournament on the performance in the next. The results show that both male and female golfers respond negatively to a failure and that their responses are virtually identical. This finding gives support to the hypothesis that women's difficulties in reaching top positions in firms are caused by external rather than internal barriers, but does of course not rule that other internal barriers may be present for women.

Keywords: Glass ceiling, failure, gender, regression discontinuity design, golf
JEL-codes: J16, J24, L83, C93

[a] I want to thank Anna Sandberg, Oskar Nordström Skans, Lena Hensvik, Marcus Eliason and audiences at IZA Summer School 2016 and the Department of Economics in Uppsala for helpful comments and suggestions.
[b] Uppsala University, Department of Economics and Institute for Evaluation of Labour Market and Education Policy (IFAU), *olof.rosenqvist@nek.uu.se.*

Table of contents

1 Introduction

Women are underrepresented in top positions in firms and other organizations (Bertrand and Hallock 2001; Wolfers 2006; Bertrand 2009). One important explanation for this observation is that on average, women choose less competition-intensive careers than men. For example, women are typically underrepresented in the private sector (Lanfranchi and Narcy 2015). Experiments have also shown that to a greater extent than men, women prefer piece rate schemes over winner-takes-it-all schemes (Niederle and Vesterlund 2007; Dohmen and Falk 2011).

More puzzling, however, is the observation that there also seems to be a glass ceiling for the women who actually enter highly competitive work environments, i.e. even competitive women struggle to reach top positions in firms (see, e.g, Albrecht et al. [2003] and Arulampalam et al. [2007] for evidence of the glass ceiling effect). Discrimination against women is probably one explanation for this phenomenon, but it could potentially also be driven by remaining gender differences in competitiveness. Individuals that want to make career progress in competitive environments typically participate in multiple rounds of competitions in which they repeatedly compete for new positions and promotions. Most individuals are bound to experience multiple failures in the initial stages of their careers because they are competing against more experienced competitors and such negative outcomes might be detrimental for their confidence and subsequent performance (see Rosenqvist and Skans [2015] for the importance of previous competitive outcomes on current performance). Having a firm belief in one's ability is arguably important for not becoming too discouraged by failures and since previous evidence indicates that men have higher levels of confidence than women (Lundeberg et al. 1994; Barber and Odean 2001; Niederle and Vesterlund 2011), women are potentially more vulnerable to failures in terms of ability to bounce back. Consistent with this hypothesis, recent experimental evidence from a study on university students in the UK suggests that women, on average, respond more negatively than men to failures with respect to subsequent performance, which might explain why, on average, women are less likely than men to make substantial career advancements (see Gill and Prowse [2014] for the experimental study).

While the finding in Gill and Prowse (2014) is highly interesting because of the understanding it provides of the general behavior of men and women, its relevance for the glass-ceiling phenomenon hinges on whether it can be replicated out of the laboratory, in situations where stakes are high, and in particular among men and women who have chosen to pursue careers in competition-intensive work environments. However, identifying causal effects of successes/failures on subsequent performance for competitive men and women on the regular labor market is difficult due to the general scarcity of relevant data and because of systematic ability differences between individuals that fail and individuals that succeed. The situation is, however, more favorable when turning the focus to the world of sports. While performance in an athletic setting relates to very particular tasks, it is a setting in which highly competitive men and women are active and in which performance data are often readily available. As such, sports competitions constitute a useful testing ground for theories about the behavior of competitive men and women.[1]

Wozniak (2012) and Jetter and Walker (2015) both use data from all-male and all-female professional tennis tournaments to study how the probability of winning the current game is affected by previous results. Using selection-on-observables strategies to identify the causal effect of previous results on current performance, they both find that men and women are more likely to win the current game if they have experienced recent successes and that these effects are very similar in magnitude across the genders. Similarly, Banko et al. (2016) study whether female tennis players are more likely than men to lose in straight sets (the hypothesis being that women find it harder to come back after losing the first set), but do not find any gender differences. Overall, these findings would suggest that the result in Gill and Prowse (2014) about women being particularly sensitive to failures does not hold among competitive men and women, who are instead equally sensitive to previous competitive outcomes with respect to current performance. A fundamental problem with these observational tennis studies is, however, that they cannot control for the counterfactual development, which makes it hard to determine whether persistent successes (and failures) are due to causal success/failure effects (i.e. the first success/failure causing the next one) or just time-varying ability. In addition, the result of a tennis game is affected by the performance of

[1] Data from golf tournaments have, e.g., been used to study predictions of tournament theories (Ehrenberg and Bognanno [1990a and 1990b], Orszag [1994] and Melton and Zorn [2000]), peer effects (Guryan et al. [2009]) and loss aversion (Pope and Schweitzer [2011]).

the opponent, making it even harder to cleanly estimate causal success/failure effects on subsequent performance.

Regarding this issue, using data from professional golf tournaments on the male European Tour, Rosenqvist and Skans (2015) made a key contribution by providing quasi-experimental evidence from professional golf tournaments in which same-ability players randomly end up in success or failure states. In these tournaments, players are separated into success and failure halfway through the tournaments by the so-called cut (a qualification threshold). Players close to the cut have performed almost equally well, but will arguably perceive their performances differently in terms of success or failure. By comparing the performance of marginally successful players and their marginally unsuccessful counterparts in the next tournament, the confounding impact of ability can be purged from the analysis and the causal effect of experiencing a success (relative to a failure) can be identified (i.e. a regression discontinuity [RD] strategy is used for identification). Rosenqvist and Skans (2015) found that male golfers on the European Tour substantially enhance their performance after a success, but they did not analyze the corresponding behavior of female golfers. In this paper, I add data from the PGA Tour (main tour for men in the US) and the LPGA Tour (main tour for women in the US), making it possible to use the same RD strategy to examine potential gender differences in the causal impact of a previous success/failure on current performance.

Thus, in this paper I provide the first quasi-experimental evidence from the field on potential gender differences in the productivity response to previous competitive outcomes among competitive men and women. These top-performing male and female athletes are active in an environment with multiple rounds of competition, which resembles the situation for men and women in the corporate sector trying to make career progress.

The results show that the current performance of both male and female golfers is negatively affected by a previous failure and that the effects are virtually identical in magnitude. The results suggest that the confidence of top-performing competitive men and women is affected by previous competitive experiences and that this effect has a substantial impact on subsequent performance. However, women show no tendencies toward being more sensitive than men to previous outcomes. Thus, if the behavior of these professional male and female athletes is similar to the behavior of competitive men and women in the rest of the society, it seems unlikely that women are unable to reach top positions in firms because they are worse than men at dealing with failures. Instead, it seems likely that women's difficulties reaching top positions in firms and other organizations are caused by external barriers, which calls for more research on the structure of these barriers and how to penetrate them.

The remainder of this paper is structured as follows. Section 2 gives a detailed description of the data and Section 3 explains and tests the validity of the identification strategy. In Section 4, I present the main results. Section 5 concludes.

2 Data

2.1 General description

In this paper, I use data from professional golf tournaments. The data come from the European Tour (men),[2] the PGA Tour (men) and the LPGA Tour (women).[3] The typical tournament in these tours is played over four days (normally Thursday–Sunday) and the players play one round of 18 holes each day. The goal is to use as few strokes as possible to complete the holes. All entrants in a tournament play the first two days and based on the scores (i.e., number of strokes) after two completed days of play, a line is drawn in the leaderboard (i.e., the list of players ordered by scoring) that separates the players with the 70 lowest scores plus ties from the rest of the field. This line is called the *cutline* or the *cut*, since players with a higher score are cut (i.e., eliminated) from the tournament at this stage.[4] The cut thus specifies the maximum number of strokes that a player is allowed to have to be qualified for the rest of the tournament. If a player satisfies that criterion, he or she *makes the cut*. Note that the cut is determined after two days of play which makes it hard to predict for the players while they are playing. It is, however, highly predictable at the late stages of the second round. Players that make the

[2] This data was used in Rosenqvist and Skans (2015). The data cover the years 2000–2012.

[3] The data from the European Tour has been collected manually from the European Tour website (*www.europeantour.com*). The data from the PGA Tour has been collected manually from the PGA Tour website (*www.pgatour.com*) and from *https://sports.yahoo.com/*. The data from the PGA Tour cover the years 1997–2014. The data from the LPGA Tour has been collected manually from the LPGA Tour website (*www.lpga.com*) and from *https://sports.yahoo.com/*. Small parts of the data on the female golfers also come from the following sites: *www.golfdata.se*, *http://www.foxsports.com/* and *golfweek.com*. The data from the LPGA Tour cover the years 1999–2015.

[4] The exact rule varies between the different tours and it has also varied within tours over time. However, the most common use of the cut is that players that are tied for the 70th position or better make the cut. The field size in the tournaments varies between approximately 120 and 150 players meaning that the cut often lies in the middle of the leaderboard.

cut continue the tournament during the weekend and the final result is based on the total score after four rounds (72 holes). All players that make the cut and finish the tournament receive prize money; the exact amount depends on the player’s final position. Players that miss the cut must leave the tournament empty-handed after two days of play. Since making the cut brings money, prestige and ranking points, it seems reasonable to assume that players that fail to make the cut experience a sense of failure relative to the players that make the cut. Importantly, at the cut there is only a one-stroke difference between success and failure, making this setting ideal for identifying potential success/failure effects through an RD strategy.

The data is structured in pairs of tournaments played during two consecutive weeks where the first tournament plays the role of a “treatment” tournament and the second the role of an “outcome” tournament. In the first tournament, I have data on the specific cut and the score of each player after two rounds.[5] It is, therefore, possible to determine whether a player made the cut or not, i.e. if he or she is treated. I have access to the same kind of information for the second tournament, which is used to measure potential performance effects of making the cut in the first tournament (conditional on ability). The scoring after two rounds is used because all players participate up to that point. It should be noted that not all players in the first tournament participate in the second tournament, which means that the outcome is missing for some of the players that participated in the first tournaments. To reduce the potential problem of selective participation in the second tournament I only use results from tournament pairs where the participation rate is at least 60 % in the second tournament.[6]

The dataset also contains information on total prize money in each tournament and individual player characteristics in terms of experience measures and ability measures.

There are some institutional differences between the male and female golf tours. First, while men always play over four days (unless the weather forces the competition to be shortened) some tournaments on the LPGA Tour only have three days of play. The cut is still determined after two rounds, but those who make it only play one, instead of

[5] For a large share of the data, I only have access to results for players within a six-stroke difference from the cut. To create consistency across tournaments, this restriction is used throughout the paper, though some tournaments contain data on more players in the list of results.
[6] Making the cut is generally associated with a higher probability of participating in the outcome tournament. But conditional on the empirical RD model, making the cut has a negative effect on participation (statistically significant for men). This is, however, only a problem if it biases the distribution of the skill of the participating players around the cut. I test this in Section 3.2.

two, additional rounds. Since this institutional difference has nothing to do with the cut rule or the importance of making the cut, it seems unlikely that it should matter for the success/failure effects. Secondly, the average prize money in men's tournaments is much higher than the average prize money in women's tournaments. The average prize money for men in my data is roughly $4,000,000 while the corresponding amount for women is $1,400,000. Essentially, this means that making (or missing) the cut has much larger financial consequences for male golfers than for female golfers. Even though the gender difference in prize money is large, there is still substantial prize money involved in women's tournaments as well, suggesting that perceptions of success or failure following a made or missed cut are likely to emerge both for male and female golfers. To make sure that difference in prize money does not interfere with the analysis, I do robustness checks in Section 4.2 on samples that are comparable in terms of prize money.

2.2 Descriptive statistics

The most widely used sample in the paper contains 189 tournament pairs from the European Tour, 251 from the PGA Tour and 202 from the LPGA Tour. The total number of observations pertaining to the European Tour is 21,912 (16,515 participate in the outcome tournaments). The corresponding numbers for the PGA Tour and the LPGA Tour are 28,988 (19,604 participate in the outcome tournaments) and 21,682 (17,014 participate in the outcome tournaments). The number of unique players in the sample of 16,515 observations with non-missing information on outcomes on the European Tour is 1,020. The corresponding numbers for the PGA Tour and the LPGA Tour are 807 and 673.

Table 1 provides summary statistics for the sample that I use to examine treatment effects, i.e. for observations with non-missing data on the outcome variables. The statistics are presented separately for men and women. Two general facts related to the empirical strategy should be highlighted. First, players that are successful in the treatment tournament (i.e., score≤cut) have better results than the unsuccessful players (i.e. score>cut) in the outcome tournament. The successful players have lower scores after two rounds and are more likely to make the cut. This difference in future performance between successful and unsuccessful players is particularly pronounced for women. Second, while the above finding is consistent with a positive impact of a success on future performance, a complicating factor is that successful players were already better than their unsuccessful counterparts before the treatment tournament (see stroke average in the previous year). Thus, a simple comparison of mean future outcomes between successful and unsuccessful players is biased by ability differences. This highlights the need for an empirical model that allows us to estimate the impact of making the cut, relative to missing it, on future performance conditional on ability. A model that does just that is explained in Section 3.

It should be noted that, on average, the female golfers have roughly three years less experience than the male golfers as measured by time as a professional golfer (being a professional golfer means the golfer can compete for money). This implies that the two samples are not completely comparable. On the other hand, as Hensvik (2014) shows, women in the higher ranks of firms are often less experienced than their male peers, making the data in this paper empirically relevant. Nevertheless, in Section 4.2 I report the results from robustness checks on samples that are comparable in terms of experience.

Table 1. Descriptive statistics for the used sample

	Men			Women		
	All	score≤cut	score>cut	All	score≤cut	score>cut
Treatment tournaments						
Average cut	143.01	143.01	143.00	145.45	145.51	145.38
Normalized strokes	0.19	-2.06	3.01	0.07	-2.18	3.00
…standard deviation	3.00	1.64	1.64	3.05	1.67	1.63
Made the cut	0.56	1.00	0.00	0.56	1.00	0.00
Outcome tournaments						
Average cut	142.75	142.77	142.72	145.61	145.72	145.47
Strokes relative to the cut	0.31	-0.14	0.87	0.05	-0.77	1.12
…standard deviation	4.37	4.31	4.38	4.68	4.49	4.71
Made the cut	0.55	0.59	0.50	0.56	0.63	0.47
Player characteristics						
Years as pro	11.98	12.02	11.93	8.63	8.51	8.80
…standard deviation	6.38	6.29	6.50	5.79	5.59	6.04
…nonmissing	0.98	0.98	0.97	1.00	1.00	1.00
Stroke average in previous year	71.67	71.56	71.81	72.96	72.72	73.28
…standard deviation	1.16	1.09	1.23	1.36	1.29	1.39
…nonmissing	0.88	0.90	0.87	0.93	0.95	0.91
Number of tournaments	440	440	440	202	202	202
Number of clusters (strokes by tournament)	5,229	2,627	2,602	2,404	1,208	1,196
Number of observations	36,119	20,100	16,019	17,014	9,597	7,417

Notes: This table contains statistics on players that participated in outcome tournaments with high participation rates (i.e. where at least 60 % of the players in the treatment tournament also participated in the outcome tournament).

3 Empirical strategy

3.1 Empirical model

The fundamental assumption behind the empirical strategy is that players with scores close to the cut ended up on the right or wrong side of it by chance. If so, the ability of players close to the cut should be virtually identical, which means I can estimate the effect of making the cut on the performance in the next tournament conditional on ability, i.e., the causal effect of experiencing a relative success on future performance. The validity of this assumption is, of course, central for this exercise and it will be assessed in detail in Section 3.2.

In the ideal RD setting, the researcher can compare the mean outcome of the treated and the controls that are infinitely close to the threshold, since the covariates in such a sample are balanced (i.e. their distributions do not vary across treatment levels). In practice, however, the number of observations typically goes to zero as we get closer to the threshold, forcing the researcher to adopt a wider bandwidth. With a wider bandwidth comes the problem of unbalanced covariates, which means that the simple comparison of outcomes must be abandoned in favor of a method that approximates the value precisely at the cutoff for treated and controls respectively. When the variable that determines assignment to treatment (hereafter called running variable) is discrete, as in this paper, then by construction the bandwidth is too wide for a simple comparison of mean outcomes. Instead, I use the relationship between the running variable and the outcome variable to approximate the outcome for hypothetical individuals that are just marginally on the success or failure side of the threshold (see Lee and Lemieux [2010] for a thorough description of this method). This is done using the regression model specified in Eq. (1):

$$Y_{ic} = \beta_0 + \beta_1 I[Z_{ic} \leq 0] + \beta_2 Z_{ic} + \beta_3 I[Z_{ic} \leq 0] Z_{ic} + \delta_c + u_i \qquad (1)$$

The outcome, denoted Y_{ic}, is any of the two measures of performance in the second tournament, where the subscript *c* indicates that it is a competition-specific outcome. Z_{ic} is the number of strokes after 36 holes in the treatment tournament normalized by the subtraction of the cut in the tournament. Thus, as Z_{ic} crosses zero from the positive side, the treatment goes from off to on. Since both Z_{ic} and Y_{ic} constitute functions of ability we expect a positive relationship between the two. The terms $\beta_2 Z_{ic}$ and $\beta_3 I[Z_{ic} \leq 0]Z_{ic}$ allow this relationship to be different on the two sides of the threshold. With the help of the estimated relationship between Z_{ic} and Y_{ic} it is possible to predict the values of Y_{ic} as Z_{ic} approaches zero from below and above, respectively. The difference between these two values measures how the outcome changes as the treatment is "turned on" while the running variable is held constant. Thus, the estimate of β_1 approximates the difference in mean outcome for treated and controls that are infinitely close to the cut (i.e. that have virtually the same ability). δ_c captures competition fixed effects and u_i is an error term.

If players close to the threshold really have the same ability, the estimate of β_1 corresponds to the causal effect of making the cut, relative to missing it, on the performance in the outcome tournament. Importantly, β_1 gives the performance difference between marginal winners and marginal losers, not between marginal winners and completely unaffected players. Thus, if marginal winners outperform marginal losers, this potential difference can be driven both by marginal winners improving their performance (relative to their hypothetical unaffected control state) and by marginal losers decreasing their performance (relative to their hypothetical unaffected control state). The data do not allow me to disentangle these two potential mechanisms.

As in Rosenqvist and Skans (2015), I cluster the standard errors at the strokes by tournament level because of potential joint specification errors for each stroke-group (see Lee and Card [2008] for a discussion of standard errors when performing RD analyses with a discrete running variable).

The main purpose of the paper is to investigate whether the value of β_1 is different for men and women and such tests can be performed by interacting all variables in Eq. (1) with an indicator for being a woman. By doing so, success/failure effects for both men and women can be estimated in a joint regression framework and potential gender differences can be directly examined. Formally, I use the statistical model specified in Eq. (2):

$$\begin{aligned} Y_{ic} = {} & \beta_0 + \beta_1 I[Z_{ic} \leq 0] + \beta_2 Z_{ic} + \beta_3 I[Z_{ic} \leq 0]Z_{ic} + \beta_4 Female_i \\ & + \beta_5 Female_i I[Z_{ic} \leq 0] + \beta_6 Female_i Z_{ic} + \beta_7 Female_i I[Z_{ic} \leq 0]Z_{ic} + \delta_c + u_i \end{aligned} \tag{2}$$

In this model, β_1 corresponds to the causal effect of making the cut, relative to failing to make it, on the performance in the outcome tournament for men, while the sum of β_1 and β_5 gives the corresponding effect for women. The difference between the genders is thus given by β_5 which constitutes the main parameter of interest.

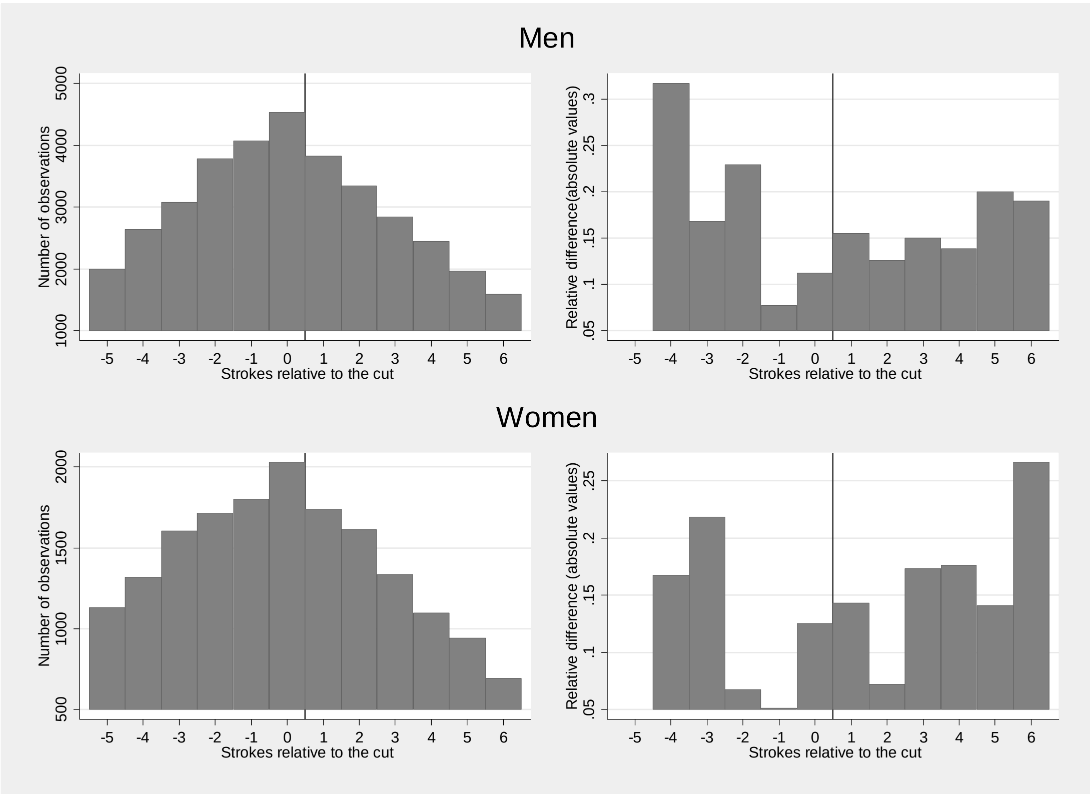
3.2 Validity of the empirical strategy

Two conditions need to be fulfilled in order for the empirical strategy to be valid. First, the ability of players must be continuous in the running variable across the threshold. Second, it is required that the incomplete participation in the outcome tournament does not bias the ability balance at the threshold. Thus, it is not enough that players distribute themselves randomly around the cutoff in the treatment tournament; instead, the test of randomization must be performed on players actually participating in the second tournament.

To test whether players close to the cutline, that actually participated in the subsequent outcome tournament, ended up on the success or failure side of the threshold by random chance rather than due to ability differences, I investigate if the number of observations evolves smoothly over the cutoff and if predetermined measures of experience and ability are continuous around the cutoff, conditional on the empirical model (i.e. Eq. [1]).

Figure 1 describes the distribution of the running variable for men (top panel) and women (bottom panel) respectively. Positive numbers on the running variable indicate that the players failed to make the cut with 1 stroke, 2 strokes and so on. For both men and women the number of observations reaches its maximum at 0, meaning that the most common result after 36 holes is to have the same number of strokes as the cut stipulates. However, this is not an unnatural mass point; rather, it is the distribution one would expect even if there was no cut and all players were allowed to complete the tournament. Since the cut often lies in the middle of the leaderboard and since ability is arguably a normally distributed variable, it is not surprising to see that a score equal to the cut is the most common score. As an investigation of the validity of the empirical strategy it is instead more informative to study the difference in the number of observations between 0 (made the cut) and 1 (did not make the cut). If this difference is extreme relative to differences in the number of observations between other scores it would suggest that we do not have random assignment to success and failure around the cut. I investigate this in the right-hand side of Figure 1, where the distribution of the relative differences in the number of observations between adjacent scores is shown.[7] The value at -4 represents the absolute difference in the number of observations between -5 and -4 divided by the number of observations at -5 and so on. Thus, the relevant bar for our purposes is the one at 1. As can be seen, the relative difference between 0 and 1 does not, in any way, stand out in the distribution of relative differences, which suggests that neither men nor women can "force" themselves into just barely making the cut.

Figure 1. Distribution of running variable for men and women



Notes: The histograms for the men are based on 36,199 observations. The histograms for the women are based on 17,014 observations. On the right hand side the value at -4 represents the absolute difference in the number of observations between -5 and -4 divided by the number of observations at -5 and so on.

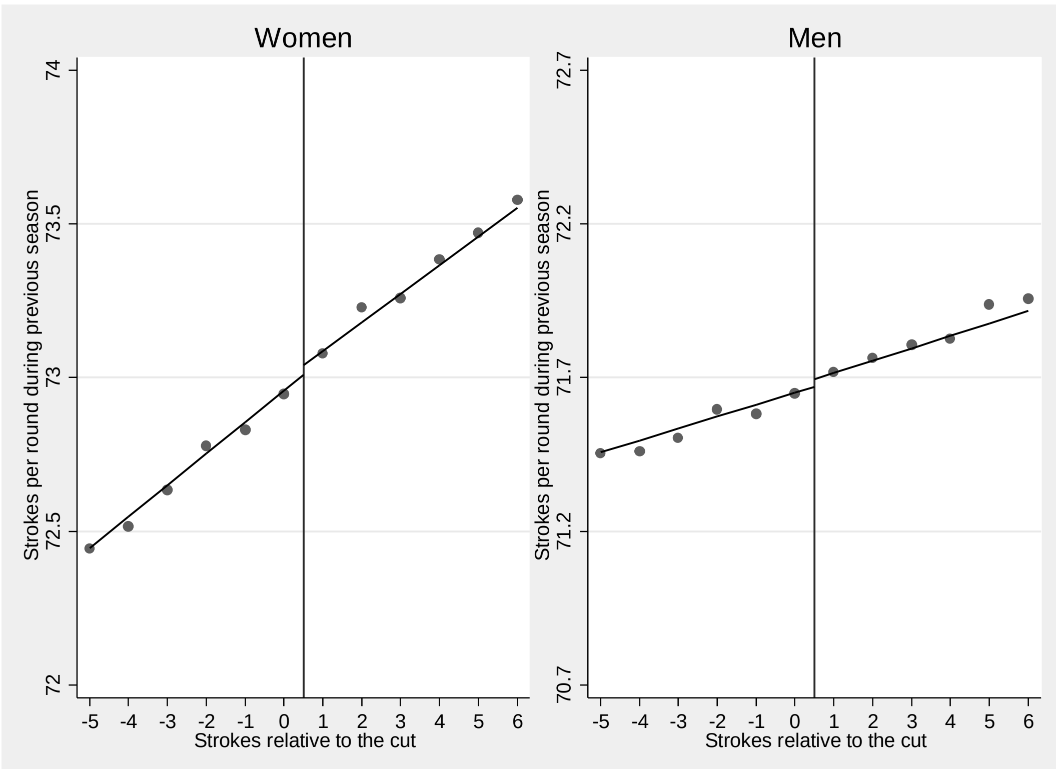
Even if no signs of manipulation at the cutoff can be found by looking at the distribution of the running variable, we still cannot rule out the possibility that marginal winners and marginal losers are different from each other in a systematic way. In Figure 2, I therefore examine how the predetermined ability of the players evolves over the

[7] I graph the relative difference in the number of observations between adjacent scores (i.e. the absolute difference in the number of observations divided by the number of observations at the best score in the pair) to adjust for the level of observations.

threshold for women and men. The predetermined ability is measured by the average scoring during the previous season. This statistic is generally considered a precise measure of a player's underlying ability and it tends to be stable over years. The last point is clear in Figure 2 since we see that female and male golfers that performed poorly in the treatment tournament (e.g. strokes relative to the cut equal to 5 or 6) also displayed a high stroke average in the previous season. The fact that the data show a strong positive association between the running variable and the stroke average in the previous season suggests that the stroke average the previous season is an accurate measure of players' abilities going into the treatment tournament.

Reassuringly for the empirical strategy, there is no jump in this ability measure at the cutoff, meaning that any potential jumps in the outcome variables at the cutoff are not driven by predetermined ability differences.

Figure 2. Strokes per round during the previous season

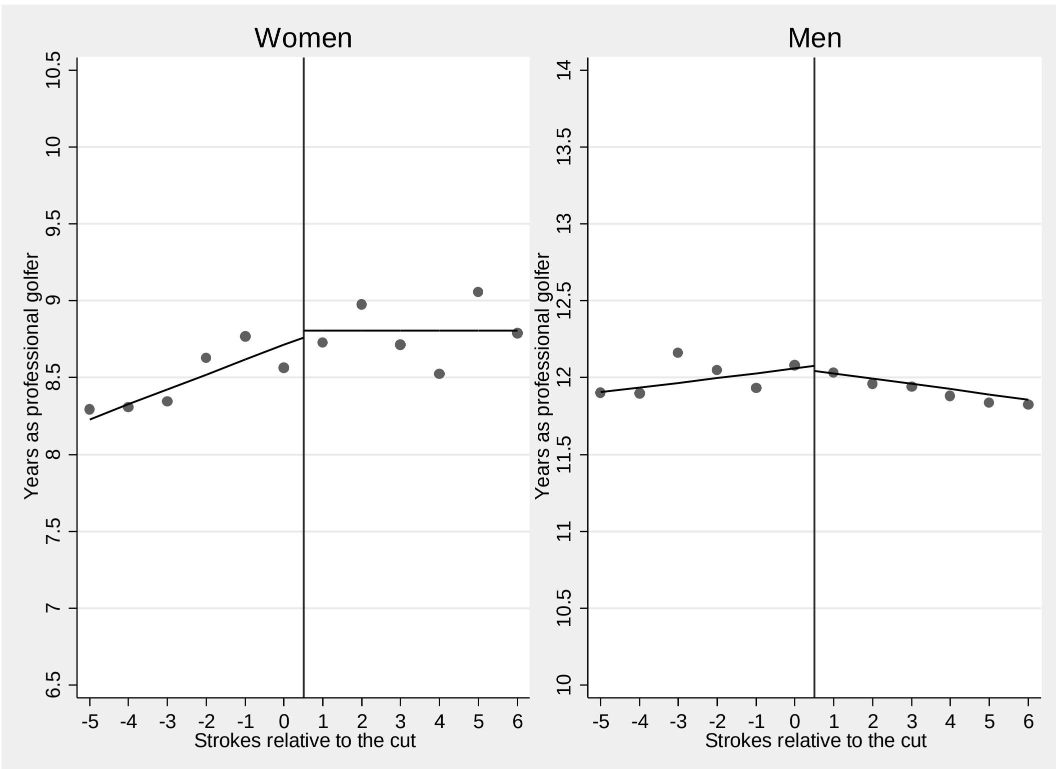


Notes: The figure is based on observations with non-missing data on the relevant characteristic which amounts to 15,878 observations for the women and 31,940 observations for the men.

Another way to test whether marginal winners and marginal losers are comparable is to examine how the experience of the players differs around the threshold. If there is a jump at the cutoff such that more experienced players are on the success side to a greater extent, it would indicate that the cut is predictable and that experienced players can better predict the cut and adjust their play so that they just marginally make it. Figure 3 shows, however, that such worries are misplaced, since the measure of experience displays smoothness at the cutoff. The specific measure of experience is number of years as a professional golfer, which captures how long the player has been competing for money. The relationship between the running variable and experience is unclear. For women, relatively inexperienced players have the best results while inexperienced men display large variation in their results. However, the picture around the threshold is similar for the two genders; there is no discontinuity in experience as the running variable crosses the cut.

Overall, these validity checks confirm that the empirical strategy can give a robust identification of success/failure effects on performance.

Figure 3. Years as professional golfer



Notes: The figure is based on observations with non-missing data on the relevant characteristic which amounts to 16,994 observations for the women and 35,350 observations for the men.

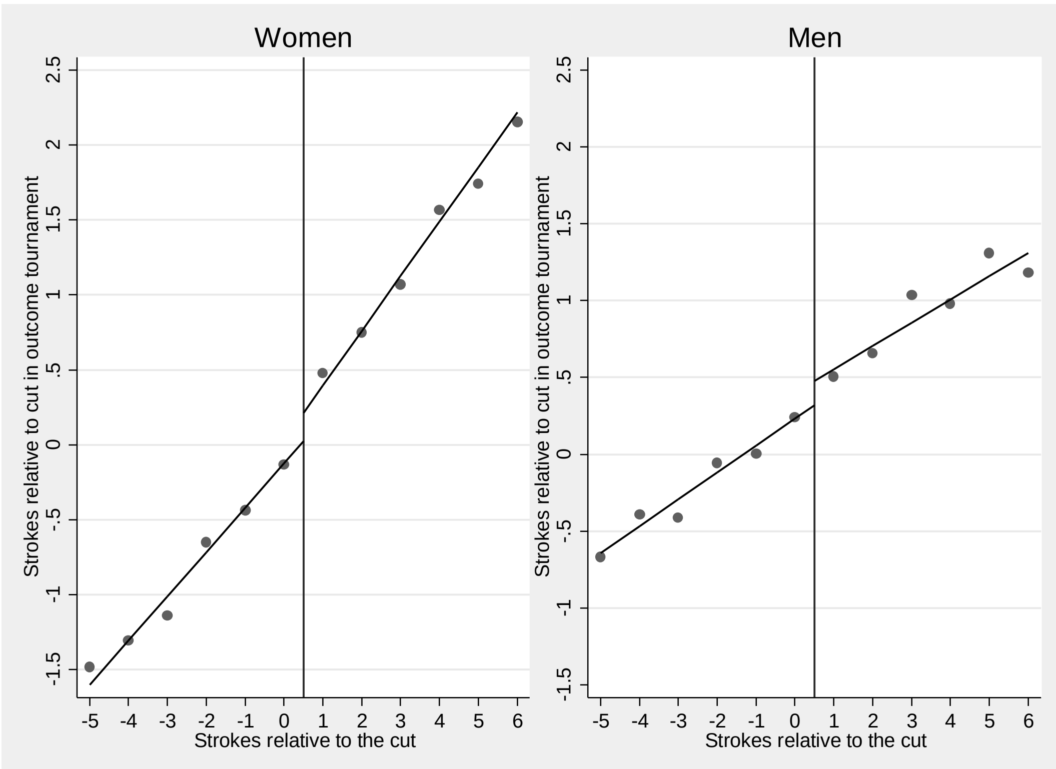
4 Results

4.1 Main results

Figure 4, Figure 5 and Table 2 present the main results, i.e. estimates of the effect of a previous (relative) success on current performance conditional on ability.

Figure 4 shows the number of strokes after two days in the outcome tournament (normalized by the cut in the relevant outcome tournament) on the y-axis and the number of strokes relative to the cut in the treatment tournament on the x-axis. Clearly, a good performance in the treatment tournament (i.e. a low x-value) is associated with a good performance in the outcome tournament as well (i.e. a low y-value). Thus, relatively better players consistently perform well whereas relatively worse players consistently perform poorly. The focus of our attention should, however, be the action at the cutoff where we see that the hypothetical marginal winners (just below 0.5) outperform the hypothetical marginal losers (just above 0.5) among both women and men. The difference at the threshold is virtually identical for the two sexes, amounting to roughly 0.16 strokes.

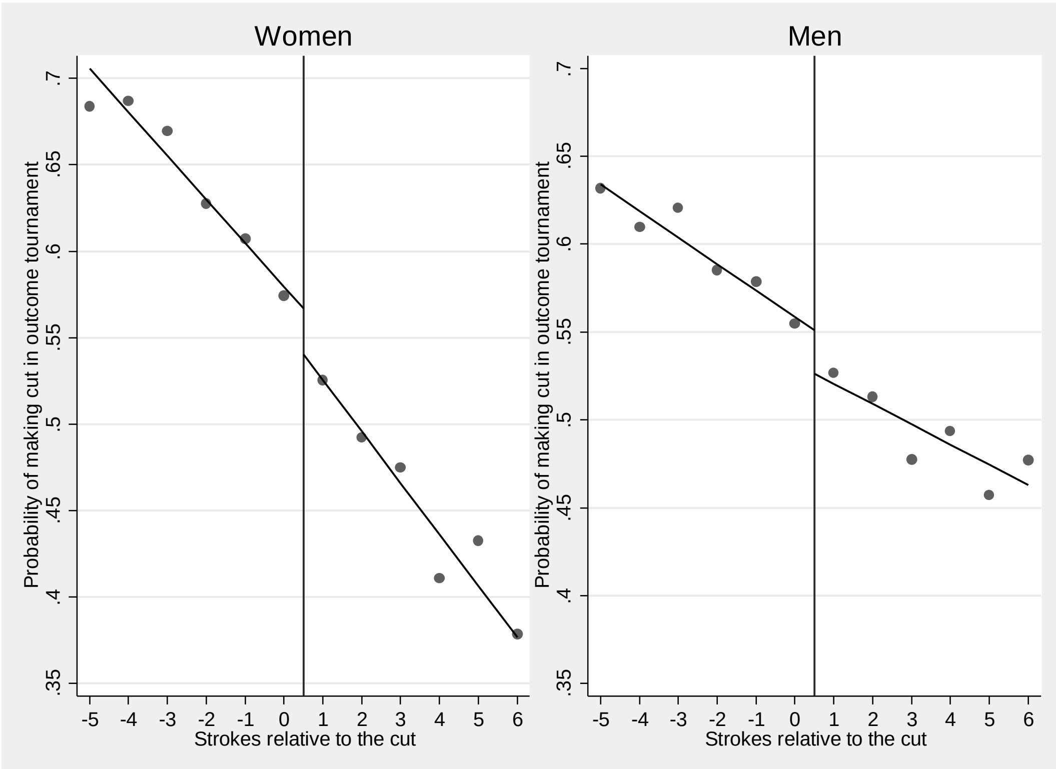
Figure 4. Strokes relative to the cut in the outcome tournament



Notes: The figure is based on observations with non-missing data on the outcome variable which amounts to 17,014 observations for the women and 36,119 observations for the men.

Figure 5 shows the results on making the cut in the following tournament. Again, we see that women and men behave similarly around the cut with marginal winners being about 2.5 percentage points more likely to make the cut in the next tournament than marginal losers. Making the cut is of substantial economic importance for the players, so the performance difference around the cut that we saw in Figure 4 has quite large real effects.

Figure 5. Probability of making the cut in the outcome tournament



Notes: The figure is based on observations with non-missing data on the outcome variable which amounts to 17,014 observations for the women and 36,119 observations for the men.

Table 2 summarizes the results shown so far and reports the tests of the differences between men and women. The estimates in column (1) correspond to estimates of β_1 when I estimate Eq. (1) using the full sample. In columns (2) and (3) I repeat the exercise for men and women respectively. The estimates in column (4) correspond to estimates of β_5 when I estimate Eq. (2) using the full sample. The negative effect on the number of strokes in the outcome tournament is significant on the 5 percent level for the full sample (column [1]) and for men (column [2]). As we saw in Figure 4, women exhibit a similar performance difference at the threshold but due to the smaller sample size, the precision is not enough to establish a statistically significant effect (see column [3]). The difference in effect size between men and women is very small (roughly one tenth of the baseline effect in column [1]) and statistically insignificant (see column [4]). This (small) point estimate also has the opposite sign of what should be expected if women are more sensitive than men to previous competitive outcomes.

With regard to performance differences between marginal winners and marginal losers measured by the propensity to make the cut in the outcome tournament, which are presented in panel B, we also find a statistically significant result for women (column [3]). Again, the difference in effect size between men and women is small and statistically insignificant (see column [4]).

Overall, the estimates suggest that performance of both men and women are causally affected by previous successes. Furthermore, the estimates do not lend support for the notion that female performance should respond more than male performance to previous competitive outcomes. Although an important caveat is that the confidence interval of the estimated interaction term allows for substantial differences between men and women (but in either direction). Hence, the second conclusion primarily rests on the fact that the point estimates remain tiny relative to the main effect.

Table 2. Main results – Marginal winners relative to marginal losers

Column:	(1)	(2)	(3)	(4)
Sample:	All	Men	Women	All
Estimate:	Main	Main	Main	Women-Men
Panel A.				
Outcome:	Strokes after 36 holes in outcome tournament			
Making the cut	-0.1629**	-0.1706**	-0.1558	0.0148
	(0.0673)	(0.0804)	(0.1200)	(0.1444)
Observations	53,133	36,119	17,014	53,133
Mean of dep. var.	143.8936	143.0593	145.6649	143.8936
Panel B.				
Outcome:	Indicator for making the cut in outcome tournament			
Making the cut	0.0255***	0.0266***	0.0240*	-0.0025
	(0.0079)	(0.0097)	(0.0133)	(0.0164)
Observations	53,133	36,119	17,014	53,133
Mean of dep. var.	0.5527	0.5487	0.5612	0.5527
Sample window	[-5,6]	[-5,6]	[-5,6]	[-5,6]
Linear RV	Yes	Yes	Yes	Yes
By treatment RV	Yes	Yes	Yes	Yes
Quadratic RV	No	No	No	No
By tournament RV	No	No	No	No
Covariates	No	No	No	No

Notes: Standard errors are clustered on the strokes * tournament level (in parentheses). */**/*** significant at the 10 /5/1 percent level. RV=running variable.

4.2 Robustness checks: model variations and experience and prize money

In Table 3, the results from Table 2 are subjected to a number of robustness checks in the form of model variations. The estimates in all seven columns are from variations of Eq. (2) and they all correspond to estimates of β_5 in that model (i.e. the interaction between making the cut and being a woman).[8] In columns (1–4) the bandwidth is gradually reduced and in column (5) I introduce a quadratic term for the running variable. In column (6) I add covariates (experience and predetermined ability) to the baseline model and specification (7) allows the linear relationship between the outcome and the running variable to be specific for each tournament. Overall, the point estimates of the interaction effect are fairly robust to these model variations, although they display some sensitivity to reductions of the bandwidth (see especially column [3]). As we can see in Figure 4 and Figure 5, this is mainly driven by the fact that men with a running variable equal to three display quite divergent results relative to the general trend. Given that the other estimates are reasonably similar to the ones presented in Table 2, I interpret the results as suggesting that men and women do indeed respond similarly to

[8] Separate results for men and women are presented in Table A1 in the appendix.

previous failures. As Table A 1 in the appendix shows, making the cut in the treatment tournament (relative to missing it) decreases the number of strokes after 36 holes in the outcome tournament by roughly 0.15 strokes for both men and women. Similarly, the probability of making the cut in the outcome tournament increases by about two percentage points for both men and women.

Table 3. Robustness checks – model variations

Column:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Strokes after 36 holes in the outcome tournament							
Making the cut *	-0.0229	0.0075	-0.2635	-0.1437	-0.0281	0.0382	0.0102
Female	(0.1548)	(0.1705)	(0.1978)	(0.2622)	(0.1402)	(0.1382)	(0.1364)
Observations	47,725	40,870	32,638	22,964	53,133	53,133	53,133
Mean of dep. var.	143.9033	143.8988	143.9144	143.8884	143.8936	143.8936	143.8936
Panel B. Making the cut in the outcome tournament							
Making the cut *	0.0019	-0.0119	0.0216	-0.0001	0.0011	-0.0042	0.0003
Female	(0.0176)	(0.0197)	(0.0229)	(0.0303)	(0.0159)	(0.0161)	(0.0155)
Observations	47,725	40,870	32,638	22,964	53,133	53,133	53,133
Mean of dep. var.	0.5513	0.5505	0.5470	0.5474	0.5527	0.5527	0.5527
Sample window	[-4,5]	[-3,4]	[-2,3]	[-1,2]	[-5,6]	[-5,6]	[-5,6]
Linear RV	Yes	Yes	Yes	Yes	Yes	Yes	Yes
By treatment RV	Yes	Yes	Yes	Yes	No	Yes	Yes
Quadratic RV	No	No	No	No	Yes	No	No
By tournament RV	No	No	No	No	No	No	Yes
Covariates	No	No	No	No	No	Yes	No

Notes: Standard errors are clustered on the strokes * tournament level (in parentheses). */**/*** significant at the 10 /5/1 percent level. RV=running variable.

As pointed out in Section 2, the male and female golfers studied in this paper differ in terms of experience and they are active in environments with different financial conditions. These differences might interfere with the analysis in such a way that I detect institutional differences instead of gender differences (cf. Rosenqvist and Skans [2015] who found that the prize money in the outcome tournament affected the results). To examine this potential problem, I reestimate the model in Table 4 using samples that are fairly comparable in terms of the golfers’ experience and the prize money in the tournaments. Since I have far more observations for male golfers than for female golfers, I make sample restrictions on the male sample to achieve comparability across the sexes. I drop all men that have more than 18 years of experience and that participated in tournaments with total prize money of more than $3,300,000. These restrictions leave me with a sample of male golfers that, on average, have 9.5 years of experience (compared with 8.6 for women) and that participate in tournaments with an average prize sum of $1,800,000 (compared with $1,400,000 for women). Thus, with these restrictions, the samples are substantially more similar than before, while they still allow me to keep roughly 15,000 observations for male golfers. The results from this exercise, which are presented in Table 4, are similar to the corresponding estimates in Table 2. The success/failure effect for men on the number of strokes increases somewhat in absolute terms (see column [2] of panel A) while the effect on making the cut instead decreases (see column [2] of panel B). But overall, the results are strikingly robust to these substantial sample restrictions, suggesting that the gender differences in experience and prize money do not interfere with the main conclusion that competitive men and women respond similarly to previous results with respect to current performance.

Table 4. Robustness checks – similar experience and prize money across genders

Column:	(1)	(2)	(3)	(4)
Sample:	All	Men	Women	All
Estimate:	Main	Main	Main	Women-Men
Panel A.				
Outcome:	Strokes after 36 holes in outcome tournament			
Making the cut	-0.1837**	-0.2151*	-0.1558	0.0593
	(0.0868)	(0.1241)	(0.1200)	(0.1726)
Observations	32,164	15,150	17,014	32,164
Mean of dep. var.	144.6088	143.4227	145.6649	144.6088
Panel B.				
Outcome:	Indicator for making the cut in outcome tournament			
Making the cut	0.0230**	0.0220	0.0240*	0.0021
	(0.0099)	(0.0146)	(0.0133)	(0.0197)
Observations	32,164	15,150	17,014	32,164
Mean of dep. var.	0.5608	0.5605	0.5612	0.5608
Sample window	[-5,6]	[-5,6]	[-5,6]	[-5,6]
Linear RV	Yes	Yes	Yes	Yes
By treatment RV	Yes	Yes	Yes	Yes
Quadratic RV	No	No	No	No
By tournament RV	No	No	No	No
Covariates	No	No	No	No

Notes: Standard errors are clustered on the strokes * tournament level (in parentheses). */**/*** significant at the 10 /5/1 percent level. RV=running variable.

4.3 Additional results: high and low stakes

In Section 4.1 I found that marginal winners, in the treatment tournament, outperform marginal losers with respect to the performance in the outcome tournament. In this section I investigate how sensitive this performance difference is to the magnitude of the initial relative success/failure (i.e. the total prize money in the treatment tournament) and to the stakes in the outcome tournament (i.e. the total prize money in the outcome tournament).[9] The exercises are performed separately for men and women and the sensitivity of the effect sizes to the prize money is then compared. The magnitude of the initial relative success/failure is a binary variable that takes on the value 1 if the total prize money in the treatment tournament was above the median of the prize money in the treatment tournaments within the particular tour and season and 0 otherwise. The stakes in the outcome tournament are also coded as a binary variable where 1 indicates that the prize money in the outcome tournament was above the median of the prize money in the outcome tournaments within the relevant combination of tour and season.

I also group the players into four categories according to the values of the aforementioned two variables (i.e. zero-zero, zero-one, one-zero and one-one). Doing so, I can in a simple way investigate how the success/failure effect is affected by the stakes in the outcome tournament holding the magnitude of the initial success/failure constant and vice versa.[10]

My findings are presented in Table 5, where for ease of presentation, I only use making the cut in the outcome tournament as the outcome variable. In Panel A, I focus on players who participated in a treatment tournament with below-median prize money. Thus, marginal winners and losers in this sample experienced relatively small successes and failures. I then investigate how the performance difference between these players in the outcome tournament is affected by the size of the prize money in the outcome tournament. In panel B, I do the corresponding exercise for players that participated in a treatment tournament with above-median prize money. In panels C and D, I keep the prize money in the outcome tournaments fixed and vary the prize money in the treatment tournaments.

Azmat et al. (forthcoming) and Gill and Prowse (2014) have previously conducted similar investigations in other settings. Azmat et al. (forthcoming) study potential gender differences in the reaction to changed stakes. They studied Spanish high school students and found that female students tend to choke under pressure in the sense that

[9] Note that I only have data on the total prize money in the tournaments, not the prize money for the individual players. Thus, I use variation in prize money between tournaments and not between players within tournaments.

[10] This was also done in Rosenqvist and Skans (2015) using a slightly different empirical approach. They found that the success/failure effect for male golfers on the European Tour was entirely driven by high stakes outcome tournaments.

the gender gap in test results (to the advantage of female students) is smaller in high stakes tests than in low stakes tests. Intuitively, positive recollections of previous performances should be particularly important in situations where the probability of choking under pressure is relatively high (i.e. high stakes situations). Thus, the effect of making the previous cut on current performance should generally be higher in high stakes outcome tournaments, and if the results in Azmat et al. (forthcoming) are relevant for adult competitive women as well, this pattern should be particularly pronounced for female golfers, since they are suggested to be more likely to choke under pressure. This reasoning implies that the estimates in column (2) of panels A and B should generally be higher than the corresponding estimates in column (1) and that the difference in column (3) should be higher for women than for men. The estimates in columns (1–2) of Panels A and B are only statistically significant on one occasion (see the men in panel B), but the fact that high stakes outcome tournaments always produce greater point estimates than low stakes outcome tournaments strengthens the notion that earlier successes (which are assumed to build confidence) are most valuable in high stakes environments when players are likely to be under pressure. The success/failure effect for women is, however, not more sensitive to the prize money in the outcome tournament than the effect is for men; instead, if anything, the estimates suggest that men are more confidence-dependent than women in high stakes situations since they seem to be very sensitive to the outcome of previous performances in exactly those cases. It should, however, be noted that the gender difference in the effect sensitivity to the prize money in the outcome tournaments is statistically insignificant in both Panel A and Panel B (see the difference-in-differences estimates at the bottom of the respective panels).

Gill and Prowse (2014) studied male and female university students in a laboratory setting. They found that men react to the size of an initial success/failure in such a way that, conditional on losing, their subsequent effort is only negatively affected if the loss was considerable. Conditional on winning, subsequent effort was not affected by the size of the win. For my setting, this would suggest that the performance difference between marginal winners and marginal losers among men should be highest after a treatment tournament with above-median prize money. This implies that the estimates for men in column (2) of Panels C and D should be higher than the corresponding estimates in column (1). The differences between the estimates are, however, very small and go in opposite directions in Panels C and D, which suggests that the success/failure effect for male golfers is insensitive to the prize money in the treatment tournament. For women, Gill and Prowse (2014) find that conditional on losing, the subsequent effort is not affected by the size of the loss. Conditional on winning, however, subsequent effort decreases in the prize money. Thus, if the results in Gill and Prowse (2014) hold true in a wider context, the performance difference between marginal winners and marginal losers among female golfers should be at its maximum after a treatment tournament with below-median prize money. This implies that the estimates for women in column (1) of panels C and D should be higher than the corresponding estimates in column (2). Looking at the point estimates, this is true in both cases, although the differences fail to exhibit statistical significance (see column [3]). Still, the rather surprising result from Gill and Prowse (2014) about small previous successes being more beneficial for women's current performance than large ones is tentatively confirmed by my results, which calls for more research on the potential mechanisms behind this peculiar result. Comparing men and women with respect to the effect sensitivity to the prize money in the treatment tournaments, we see that according to the point estimates, women are more sensitive (see the difference-in-differences estimates at the bottom of panels C and D). But since the data are cut relatively thin in this exercise, the difference-in-differences estimates are not statistically significant.

Overall, the most important findings from this exercise are that women benefit from relatively small rather than large previous successes and that both men and women (especially men) are particularly dependent on positive recollections of previous performances when competing in high stakes situations.

Table 5. The importance of a previous success in situations with high and low stakes

Column:	(1)	(2)	(3)
Panel A. Prize money in treatment tournament low			
Prize money in outcome:	Low	High	Difference (High-Low)
Men: Making the cut	0.0085	0.0457*	0.0372
	(0.0150)	(0.0255)	(0.0296)
Women: Making the cut	0.0296	0.0474	0.0177
	(0.0202)	(0.0313)	(0.0372)
Difference-in-differences (Women-Men):	-0.0195 (0.0476)		
Panel B. Prize money in treatment tournament high			
Prize money in outcome:	Low	High	Difference (High-Low)
Men: Making the cut	0.0008	0.0526***	0.0518*
	(0.0242)	(0.0178)	(0.0301)
Women: Making the cut	0.0054	0.0100	0.0047
	(0.0320)	(0.0278)	(0.0424)
Difference-in-differences (Women-Men):	-0.0471 (0.0520)		
Panel C. Prize money in outcome tournament low			
Prize money in treatment:	Low	High	Difference (High-Low)
Men: Making the cut	0.0085	0.0008	-0.0077
	(0.0150)	(0.0242)	(0.0285)
Women: Making the cut	0.0296	0.0054	-0.0243
	(0.0202)	(0.0320)	(0.0379)
Difference-in-differences (Women-Men):	-0.0166 (0.0474)		
Panel D. Prize money in outcome tournament high			
Prize money in treatment:	Low	High	Difference (High-Low)
Men: Making the cut	0.0457*	0.0526***	0.0069
	(0.0255)	(0.0178)	(0.0311)
Women: Making the cut	0.0474	0.0100	-0.0373
	(0.0313)	(0.0278)	(0.0419)
Difference-in-differences (Women-Men):	-0.0442 (0.0522)		

Notes: Standard errors are clustered on the strokes * tournament level (in parentheses). */**/*** significant at the 10 /5/1 percent level.

4.4 Additional results: good and bad days

For about 80% of the observations I have the individual scores for the two first rounds of the outcome tournaments.[11] This makes it possible to investigate whether a previous success is more beneficial on relatively good or bad days, or if the effect is constant. For each observation with non-missing data, I calculate the best day and the worst day and then examine how a previous success affects the results on the respective days

[11] The detailed information exists for 100% of the tournaments on the PGA Tour, about 85% of the tournaments on the LPGA Tour and about 50% of the tournaments on the European Tour. For the other tournaments I only have the aggregate score over the two rounds.

conditional on the RD model.[12] Table 6 shows the results. The pattern of the results suggests that a previous success is particularly important on a relatively bad day when the players are struggling on the course. The success/failure effect is roughly twice as big on a bad day compared with a good day and the effect is only statistically significant on bad days (see column [1]). The same general pattern is apparent for both women and men (columns [2–3]). Arguably, these results strengthen the notion that confidence is the main factor behind the positive success, effect since players are more likely to start doubting their ability on relatively bad days. But with a fresh memory of success, these negative thoughts might be easier to keep at bay, making the recently successful players less likely to post very bad results. Effectively, this suggests that confidence reduces variance in performance.

Table 6. The effect of a previous success/failure on good and bad days

Column:	(1)	(2)	(3)	(4)
Sample:	All	Men	Women	All
Estimate:	Main	Main	Main	Women-Men
Panel A.				
Outcome:	Number of strokes on a good day			
Making the cut	-0.0425	-0.0541	-0.0197	0.0344
	(0.0403)	(0.0495)	(0.0685)	(0.0845)
Observations	42,160	27,520	14,640	42,160
Mean of dep. var.	70.2238	69.8351	70.9546	70.2238
Panel B.				
Outcome:	Number of strokes on a bad day			
Making the cut	-0.0931**	-0.0983*	-0.0833	0.0150
	(0.0442)	(0.0549)	(0.0730)	(0.0914)
Observations	42,160	27,520	14,640	42,160
Mean of dep. var.	73.3259	72.9582	74.0170	73.3259
Sample window	[-5,6]	[-5,6]	[-5,6]	[-5,6]
Linear RV	Yes	Yes	Yes	Yes
By treatment RV	Yes	Yes	Yes	Yes
Quadratic RV	No	No	No	No
By tournament RV	No	No	No	No
Covariates	No	No	No	No

Notes: Standard errors are clustered on the strokes * tournament level (in parentheses). */**/*** significant at the 10/5/1 percent level. RV=running variable.

5 Conclusion

In experiments, women have been found to decrease their performance following a setback, while men appear to be unaffected (Gill and Prowse [2014]). It has also been

[12] About 10% of the observations have identical results on the two rounds and consequently, their worst day score is identical to their best day score.

suggested that this gender difference in dealing with failures might partly explain the presence of a glass ceiling for women on the labor market, based on the logic that early career failures leave deeper scars on women than on men. However, this result has not been replicated among men and women that have actually chosen to enter competition-intensive work environments (see Jetter and Walker [2015] and Banko et al. [2016], who have investigated the behavior of professional male and female tennis players). Instead, these studies have found that competitive men and women are equally sensitive to previous results. But in all field studies on this issue so far, identification has relied on selection-on-observables strategies, which leaves uncertainty regarding the robustness of the results.

In this paper, I contribute to the literature by providing quasi-experimental evidence from a field setting very well suited for identifying causal effects of a previous success/failure on current performance. I use data from about 200 all-female and 450 all-male golf tournaments. These tournaments involve a stringent qualification rule that can be used to study the effect of a previous success, relative to a failure, on current performance holding ability constant (this empirical strategy was previously used in Rosenqvist and Skans [2015] but only for male golfers). Halfway through professional golf tournaments, the half of the players with the highest scores is eliminated from the tournament (players outside top 70 are eliminated and the field size is typically around 140 players). The other players continue the tournament and earn at least some prize money in the end. Players that are just barely eliminated (marginal losers) and players that are just barely allowed to complete the tournament and earn prize money (marginal winners) performed almost equally well, but will arguably experience the performance differently in terms of success or failure. Using an RD design, I estimate the performance difference in the tournament the following week between marginal winners and marginal losers.

The analysis reveals two main findings. First, marginal winners generally outperform marginal losers in the subsequent tournament. Marginal winners have roughly 0.16 fewer shots (significant on the 5 percent level) after 36 holes in the outcome tournament and they are 2.5 percentage points more likely to make the cut (significant on the 1 percent level). This result shows that the finding in Rosenqvist and Skans (2015) about the existence of substantial causal success/failure effects also holds true in a wider setting where women are included. Second, men and women exhibit virtually identical results, suggesting that top-performing women can tackle failures just as well as top-performing men. Thus, if the behavior of these professional male and female athletes is similar to the behavior of competitive men and women in the rest of the society, it seems unlikely that women are unable to reach top-positions in firms because they are worse than men at dealing with failures. Instead, it seems likely that women's difficulties in reaching top-positions in firms and other organizations are caused by external barriers rather than internal barriers, which calls for more research on the structure of these external barriers and how to penetrate them. However, it can of course not be ruled out that internal barriers other than sensitivity to failures partially can explain women's difficulties in reaching top positions.

The analysis has also produced three additional results. First, female golfers benefit more from relatively small previous successes than from larger ones, which replicates the finding in Gill and Prowse (2014). Second, both men and women (especially men) are particularly sensitive to recent results when competing in high stakes environments. This result strongly suggests that confidence is the main factor behind the success/failure effect for both women and men, since confidence is arguably crucial when players are under intense pressure. Third, for both men and women, higher confidence (from a previous success) tends to help the players by improving their lowest ability level rather than their highest, effectively reducing between-day variance in performance.

The first additional result is particularly interesting since it could imply that women are more easily satisfied than men, i.e. when women have reached a major success they might not have the same hunger as men for further progress. The result could also be explained by equity concerns. Women might to a greater extent than men feel that they have had their share of prosperity after a major success and that it is unfair to advance further at the expense of others. In a more regular labor market context this could mean that women, after being promoted, do not strive as much for further promotion as men do which potentially, at least partly, can explain women's underrepresentation in absolute top positions on the labor market.

References

Albrecht, J., Björklund, A. and Vroman, S. (2003), “Is There a Glass Ceiling in Sweden?”, *Journal of Labor Economics*, 21(1), 145–177

Arulampalam, W., Booth, A. L. and Bryan, M. L. (2007), “Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution”, *Industrial & Labor Relations Review*, 60(2), 163–186.

Azmat, G., Calsamiglia, C. and Iriberri, N. (forthcoming), “Gender differences in response to big stakes”, *Journal of European Economic Association*.

Banko, L., Leeds, E. M. and Leeds, M. A. (forthcoming), “Gender Differences in Response to Setbacks: Evidence from Professional Tennis”, *Social Science Quarterly*.

Barber, B. M. and Odean, T. (2001), “Boys will be boys: Gender, overconfidence, and common stock investment”, *Quarterly Journal of Economics*, 116(1), 261–292.

Bertrand, M. (2009), “CEOs”, *Annual Review of Economics*, 1(1), 121–150.

Bertrand, M. and Hallock, K. F. (2001), “The gender gap in top corporate jobs”, *Industrial & Labor Relations Review*, 55(1), 3–21.

Dohmen, T. and Falk, A. (2011), “Performance pay and multidimensional sorting: Productivity, preferences, and gender”, *American Economic Review*, 101(2), 556–590.

Ehrenberg, R. G. and Bognanno, M. L. (1990a), “The Incentive Effects of Tournaments Revisited: Evidence from the European PGA Tour”, *Industrial and Labor Relations Review*, 43(3), 74S–88S.

Ehrenberg, R. G. and Bognanno, M. L. (1990b), “Do Tournaments Have Incentive Effects?”, *Journal of Political Economy*, 98(6), 1307–1324.

Gill, D. and Prowse, V. L. (2014), “Gender differences and dynamics in competition: The role of luck”, *Quantitative Economics*, 5 (2), 351–376.

Guryan, J., Kroft, K. and Notowidigdo, M. J. (2009), “Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments”, *American Economic Journal: Applied Economics*, 1(4), 34–68.

Hensvik, L. (2014), “Manager impartiality: Worker-firm matching and the gender wage gap”, *Industrial & Labor Relations Review,* 67(2), 395–421.

Jetter, M. and Walker, J. K. (2015), “Game, set, and match: Do women and men perform differently in competitive situations?”, *Journal of Economic Behavior & Organization*, 119, 96–108.

Lanfranchi, J. and Narcy, M. (2015), “Female Overrepresentation in Public and Nonprofit Sector Jobs: Evidence From a French National Survey”, *Nonprofit and Voluntary Sector Quarterly*, 44(1), 47–74.

Lee, D. and Card, D. (2008), “Regression discontinuity inference with specification error”, *Journal of Econometrics*, 142(2), 655–674.

Lee, D. and Lemieux, T. (2010), “Regression Discontinuity Designs in Economics”, *Journal of Economic Literature*, 48(2), 281–355.

Lundeberg, M. A., Fox, P. W. and Punćcohaŕ, J. (1994), “Highly confident but wrong: gender differences and similarities in confidence judgments”, *Journal of Educational Psychology*, 86(1), 114–121.

Melton, M. and Zorn, T. S. (2000), “An Empirical Test of Tournament Theory: The Senior PGA Tour”, *Managerial Finance*, 26(7), 16–32.

Niederle, M. and Vesterlund, L. (2011), “Gender and competition”, *Annual Review of Economics,* 3(1), 601–630.

Niederle, M. and Vesterlund, L. (2007), “Do women shy away from competition? Do men compete too much?”, *Quarterly Journal of Economics*, 122(3), 1067–1101.

Orszag, J. (1994), “A new look at incentive effects and golf tournaments”, *Economics Letters*, 46(1), 77–88.

Pope, D. G. and Schweitzer, M. E. (2011), “Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes”, *American Economic Review,* 101(1), 129–157.

Rosenqvist, O. and Skans, O. N. (2015), “Confidence enhanced performance?–The causal effects of success on future performance in professional golf tournaments”, *Journal of Economic Behavior & Organization,* 117, 281–295.

Wolfers, J. (2006), “Diagnosing discrimination: Stock returns and CEO gender”, *Journal of the European Economic Association*, 4(2–3), 531–541.

Wozniak, D. (2012), “Gender differences in a market with relative performance feedback: Professional tennis players”, *Journal of Economic Behavior & Organization*, 83(1), 158–171.

Appendix

Table A 1. Robustness checks

Column:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Strokes after 36 holes in the outcome tournament - Men							
Making the cut	-0.1087	-0.1051	-0.0230	-0.1624	-0.1604**	-0.1513*	-0.1647**
	(0.0855)	(0.0940)	(0.1101)	(0.1433)	(0.0775)	(0.0791)	(0.0749)
Observations	32,532	27,938	22,410	15,783	36,119	36,119	36,119
Mean of dep.	143.0738	143.0718	143.0944	143.0561	143.0593	143.0593	143.0593
Panel B. Strokes after 36 holes in the outcome tournament - Women							
Making the cut	-0.1315	-0.0976	-0.2866*	-0.3061	-0.1885	-0.1131	-0.1544
	(0.1291)	(0.1423)	(0.1643)	(0.2197)	(0.1169)	(0.1134)	(0.1140)
Observations	15,193	12,932	10,228	7,181	17,014	17,014	17,014
Mean of dep.	145.6794	145.6854	145.7111	145.7179	145.6649	145.6649	145.6649
Panel C. Making the cut in the outcome tournament - Men							
Making the cut	0.0188*	0.0194*	0.0076	0.0213	0.0248***	0.0248***	0.0268***
	(0.0104)	(0.0116)	(0.0134)	(0.0177)	(0.0093)	(0.0095)	(0.0091)
Observations	32,532	27,938	22,410	15,783	36,119	36,119	36,119
Mean of dep.	0.5471	0.5475	0.5434	0.5452	0.5487	0.5487	0.5487
Panel D. Making the cut in the outcome tournament - Women							
Making the cut	0.0207	0.0075	0.0292	0.0212	0.0260**	0.0206	0.0271**
	(0.0142)	(0.0159)	(0.0186)	(0.0246)	(0.0128)	(0.0130)	(0.0126)
Observations	15,193	12,932	10,228	7,181	17,014	17,014	17,014
Mean of dep.	0.5604	0.5568	0.5548	0.5523	0.5612	0.5612	0.5612
Sample window	[-4,5]	[-3,4]	[-2,3]	[-1,2]	[-5,6]	[-5,6]	[-5,6]
Linear RV	Yes	Yes	Yes	Yes	Yes	Yes	Yes
By treatment RV	Yes	Yes	Yes	Yes	No	Yes	Yes
Quadratic RV	No	No	No	No	Yes	No	No
By tournament RV	No	No	No	No	No	No	Yes
Covariates	No	No	No	No	No	Yes	No

Notes: Standard errors are clustered on the strokes * tournament level (in parentheses). */**/*** significant at the 10 /5/1 percent level. RV=running variable.

Publication series published by IFAU – latest issues

Rapporter/Reports

2016:1 Engdahl Mattias and Anders Forslund ”En förlorad generation? Om ungas etablering på arbetsmarknaden”

2016:2 Bastani Spencer, Ylva Moberg and Håkan Selin ”Hur känslig är gifta kvinnors sysselsättning för förändringar i skatte- och bidragssystemet?”

2016:3 Lundin Martin, Oskar Nordström Skans and Pär Zetterberg ”Kåren och karriären: Student-politiken som språngbräda”

2016:4 Brommesson Douglas, Gissur Erlingsson, Johan Karlsson Schaffer, Jörgen Ödalen and Mattias Fogelgren ”Att möta den högre utbildningens utmaningar”

2016:5 Egebark Johan ”Effekter av skatter på ungas egenföretagande”

2016:6 Mannelqivst Ruth, Berndt Karlsson and Bengt Järvholm ”Arbete och arbetsmarknad i sjuk-försäkringen”

2016:7 Rosenqvist Olof ”Rösträtt och ungdomars kunskap om politik”

2016:8 Lindgren Karl-Oskar, Sven Oskarsson and Mikael Persson ”Leder bättre tillgång till utbildning till ökat politiskt deltagande?”

2016:9 Moberg Ylva ”Är lesbiska föräldrar mer jämställda?”

2016:10 Hinnerich Björn T. and Jonas Vlachos “Skillnader i resultat mellan gymnasieelever i fristående och kommunala skolor”

2016:11 Engdahl Mattias ”Invandringens arbetsmarknadseffekter: lärdomar från den internationella litteraturen och svenska resultat”

2016:12 Hagen Johannes ”Hälsoeffekter av senarelagd pensionering”

Working papers

2016:1 Bastani Spencer, Ylva Moberg and Håkan Selin “Estimating participation responses using transfer program reform”

2016:2 Lundin Martin, Oskar Nordström Skans and Pär Zetterberg ”Leadership experiences, labor market entry and early career trajectories”

2016:3 van den Berg Gerard J., Lena Janys, Enno Mammen and Jens P. Nielsen ”A general semi-parametric approach to inference with marker-dependent hazard rate models”

2016:4 Egebark Johan “Effects of taxes on youth self-employment and income”

2016:5 Holmlund Helena “Education and equality of opportunity: what have we learned from educational reforms?”

2016:6 Rosenqvist Olof “Rising to the occasion? Youth political knowledge and the voting age”

2016:7 Lindgren Karl-Oskar, Sven Oskarsson and Mikael Persson “How does access to education influence political candidacy? Lessons from school openings in Sweden”

2016:8 Moberg Ylva “Does the gender composition in couples matter for the division of labor after childbirth?”

2016:9 Hinnerich Björn T. and Jonas Vlachos “The impact of upper-secondary voucher school attendance on student achievement. Swedish evidence using external and internal evaluations”

2016:10 Farbmacher Helmut, Raphael Guber and Johan Vikström “Increasing the credibility of the twin birth instrument”

2016:11 Hagen Johannes “What are the health effects of postponing retirement? An instrumental variable approach”

2016:12 de Luna Xavier, Philip Fowler and Per Johansson “Proxy variables and nonparametric identification of causal effects”

2016:13 Björklund Maria, Mikael Carlsson and Oskar Nordström Skans “Fixed wage contracts and monetary non-neutrality”

2016:14 Rosenqvist Olof “Is there a gender difference in the ability to deal with failures? Evidence from professional golf tournaments”

Dissertation series

2016:1 Hagen Johannes “Essays on pensions, retirement and tax evasion”