Do Setbacks Delay the Participation in Repeated Competitions? Evidence from a Natural Experiment with Amateur Tennis Players*

Simon Haenni†

June 29, 2016

Abstract

Many important life goals require repeated confrontation with competitors. Losing in such competitions may discourage individuals and make them postpone the next stage of the competition and thereby harm future prospects. This study shows new evidence from a large natural experiment with amateur tennis players on how competition outcomes causally affect the time to the next tournament. The results suggest that individuals take on average 10% longer to compete again after losing than after winning. The comprehensive data-set allows to identify individual rankings and predicted competition outcomes as reference points, suggesting a complementary role of status-quo and expectation-based reference points.

Keywords: competition, reference points, field data

JEL Classificattion: D03, D81, L83

* *Acknowledgements*: I am grateful to Adrian Bruhin, Lorenz Goette, Katharina Erhardt, and Alon Dolev for helpful comments and support.

† *Affiliation*: University of Lausanne, Faculty of Business and Economics (HEC Lausanne), 1015 Lausanne, Switzerland. E-mail: simon.haenni.ch@gmail.com

1 Introduction

"Ever tried. Ever failed. No matter. Try again. Fail again. Fail better."

— Samuel Beckett, *Worstward Ho*

Repeatedly participating in competitions is essential to achieve many life goals: students choose to repeatedly pass competitive tests and apply to selective schools in order to obtain valuable degrees to continue their education or to enter the job market (e.g. Hoekstra 2009), workers routinely compete in tournaments to get a new job, a promotion, or a pay raise (e.g. Prendergast 1999), and when dating, people looking for partners have to leave competitors behind in order to find their best match (e.g. Fisman et al. 2006; Hitsch, Hortaçsu and Ariely 2010). Competition tries to separate the skilled from the unskilled, but even the best lose from time to time. Yet losing is linked to negative emotions (Ding et al. 2005) and could potentially discourage individuals and delay their participation in future competitions. An important example is long-term unemployment. It has been commonly argued that one of the main reasons for long-term unemployment is the job seekers' frustration from rejected applications and the resulting reduction in job search effort (e.g. Pissarides, Layard and Hellwig 1986; Budd, Levine and Smith 1988; Jackman and Layard 1991). Even though the relationship between bad performance in competitions and the ensuing lack of competitiveness is relevant in various fields in economics, we yet need a clean identification of this causal effect.

This paper studies how success and failure in a repeated competition setting causally affect the waiting time between individual competitions. Establishing this causal relationship is complex for two reasons. First, one needs an appropriate framework to measure the waiting time between competitions and second, one needs a convincing empirical setup to identify the competition outcome as the main cause. Since competitiveness is widely regarded as stable personality trait (Roberts et al. 2009), long-run changes in the propensity to compete are hard to measure in laboratory experiments and the external validity of short-run responses in artificial competition situations might be limited. Purely observational studies on the other hand do not allow for a clean identification of causal effects. Hence one needs a natural experiment with real-life competition data to properly address this research question.

I show new evidence from over 60,000 amateur tennis players who competed in 1.4 million matches in single tennis tournaments between 2007 and 2014. Using an amateur population has two distinct advantages. First, the sample reflects the general population as it covers all socioeconomic and demographic groups (Lamprecht, Fischer and Stamm 2015). Second, individuals participate voluntarily in competitions and do not rely on monetary benefits. As individuals are completely free to sign up for as many tournaments as they want, the time to the next tournament participation is strongly influenced by the individuals' motivation to compete.

A crucial property of this specific tournament series is the random generation of tournament draws. This creates exogenously assigned pairs of opponents in every match and thereby provides a natural experiment that can be exploited with an IV strategy. In the tournaments I analyze, each individual is listed in a national ranking which is updated semi-annually based on the past year's performance. The exogenous difference between two opponents in the national ranking provides a valid instrument for the competition outcome that does not affect the motivation to compete through any unobserved channels.

The first result from a linear IV regression suggests that individuals postpone the next competition by 10 days when losing, compared to winning. This corresponds to an 11% increase from a baseline time to the next competition of 93 days. Hence losing increases the time until the next competition even though staying away from competitions is harmful for the individual ranking. But the psychological effects seem to outweigh ranking-incentives.

One mechanism that is compatible with this behavior are reference-dependent preferences with respect to the competition outcome. Reference points separate gains from losses and allow for the valuation of outcomes relative to those reference points. Even though a vast theoretical and empirical literature has evolved in this field, it remains unclear what exactly constitutes such a reference point. (Kahneman and Tversky 1979; Kőszegi and Rabin 2006; Bordalo, Gennaioli and Shleifer 2012).

The comprehensive dataset allows to investigate this question. A first psychologically plausible mechanism is that losing against a worse ranked opponent hurts more than losing against a better ranked opponent. A second mechanism are expectations, whereas unexpected outcomes have a stronger impact than anticipated outcomes. I provide new evidence by looking at two aspects of reference dependence: individuals' positions in the national ranking and predicted ex-ante winning probabilities. The exogenous assignment of opponents with different individual rankings allows for a regression discontinuity design, separating opponents ranked slightly better from opponents ranked slightly worse than a focal individual. The results suggest that the relative ranking order between opponents does not matter when individuals win. But individuals stay absent about 10 days longer after losing against an opponent ranked slightly worse compared to an opponent ranked slightly better than themselves. I then construct a measure of surprise for victories and defeats by predicting the probability of winning, using a large set of variables that are observable by all individuals. These variables are able to correctly predict the outcome of the competition 75% of the time. The deviation of the actual from the predicted outcome is an accurate proxy for the surprise effect. The results from local polynomial regressions indicate that individuals react more sensitively to surprising than to anticipated victories and defeats.

The results of this paper are compatible with the predictions from prospect theory (Kahneman and Tversky 1979), where individuals use their position in the national ranking as status quo reference point. Expectation-based reference points in line with Kőszegi and Rabin (2006) are compatible with observed behavior against opponents further away from an individual in the ranking, but fail to explain the strong regression discontinuity in the immediate neighborhood of individuals' own rankings. If individuals formed proper expectations about the match outcome there should be no such discontinuous jump in the waiting time until the next competition. There are very few papers that carefully distinguish between expectation-based and status quo reference points. Particularly with field data this is a very challenging task. Song (2015) conducted lab experiments specifically designed to this question and concludes that both concepts matter equally. Even though the setup is completely different, the conclusions from my field evidence are remarkably similar.

This paper directly contributes to different strands of the literature. So far, very little is known about the impact of actual defeats on the motivation to compete again in the future. Buser (forthcoming) shows evidence from a lab experiment and finds that losers in a two-person tournament subsequently set more challenging and risky targets but fail to reach them and consequently earn less than winners of the same tournament. This is an interesting finding on how losing affects behavior in competitions but the experimental setup does not allow individuals to postpone future participation in competitions.

A broader literature looks at the effects of losing or staying behind targets on effort and performance. Mas (2006) shows that police performance declines after representing unions lose in final-offer arbitration over salary demands. He further shows that reference points, namely expectations, play an important role: staying behind expected pay rises is especially harmful to future work efforts. Similarly, Ockenfels, Sliwka and Werner (2014) find that the performance of managers decreases if their bonuses fall behind pre-assigned bonus targets. An important new contribution to this literature is the different recipient of the behavioral reaction. While police officers and managers retaliate against their employer, their colleagues, or the general public, I show that individuals also harm themselves after losing in a competition. Another related paper looks at effects of losing in the lab. Gill and Prowse (2014) find that subjects reduce effort after losing - partly due to chance - in a real-effort competition. Another strand of this literature looks at situations where individuals can avoid the loss. Multiple papers find that individuals are willing to provide additional effort or cost to avoid losing (Pope and Schweitzer 2011; Bartling, Brandes and Schunk 2015; Allen et al. forthcoming). On the contrary, Gill and Prowse (2012) find that subjects in the lab reduce effort when they expect to lose in order to avoid the feeling of disappointment. All these papers look at some type of expectation-based reference points but do not try to distinguish them from status quo reference points. This is another important contribution of my work to this field.

More generally, there is also a strand of the literature that looks at different psychological effects of losing in competitions. Card and Dahl (2011) look at domestic violence and find that unexpected losses of local professional American football teams lead to a 10% increase in the rate of at-home violence by male football fans against women. Also using data from National Football League games, Cornil and Chandon (2013) show that football fans eat less healthy after their team has lost. Moreover these studies differ from the current paper in the sense that individuals do not take part in competitions themselves in but rather suffer from vicarious losses by the team they support.

The paper is organized as follows. Section 2 describes the data set. Section 3 looks at the causal effect of losing in competitions on the waiting time until the next competition. Section 4 shows how reference points influence the individuals' perception of competition outcomes. Finally, section 5 concludes.

2 Data Set

2.1 Origin and Structure of the Data

This study uses a large panel data set provided by the Swiss Tennis association covering every single match played by 78,420 amateur tennis players with a ranking in Switzerland between January 2007 and December 2014.[1] Although the sample is not representative, it contains almost one percent of the country's population. Statistics from the federal sports office show that tennis is popular among all age groups, regions and nationalities. Men and people with high incomes are particularly likely to play tennis (Lamprecht, Fischer and Stamm 2015).

Individuals repeatedly participate in single tournaments and in a yearly team-season. Over the observed period of 8 years this results in 2,502,744 observations. Based on the results from these competitions, players obtain a nation-wide, gender-separated ranking on a semi-annual basis.

2.2 Determination of the National Ranking

The ranking is determined in the following fashion. Starting with the discounted ranking value from the previous period, victories have a positive impact while defeats have a negative impact on the ranking. Importantly, not competing also leads to a ranking decrease. Furthermore, the rankings of the opponents matter. Victories against strong opponents have a larger positive impact on the ranking than victories against weak opponents. Naturally, the opposite is true for defeats. Losing against weak opponents hurts the ranking more than losing against strong opponents. Additionally, there is an explicit incentive to compete in the form of a participation bonus that is added for every match and only depends on the opponent's ranking and not the match outcome.

[1] (Semi-) Professional tennis players are removed from the analysis. The 150 best men and 75 best women have a separate ranking and regularly compete in international tournaments. Including them does not change the results.

Table 1: Numeric example of ranking determination for a men ranked 6'000

Opponent's ranking	Match outcome	Next period ranking
6'600	Victory	8'700
	Defeat	13'200
6'000	Victory	7'850
	Defeat	12'900
5'400	Victory	7'000
	Defeat	12'600
No participation in competitions		12'300

Table 1 illustratively shows an example of the next period's ranking calculation for a men currently ranked 6'000 if he plays one single match against an opponent ranked 10% worse, the same, and 10% better than himself, while all other player's rankings are assumed to stay constant. As reference, the last line reports the new ranking in case the individual stays absent from competitions all together. Not competing sharply reduces the ranking. Already one victory has a strong positive effect on the ranking. The effect is considerably larger if the individual wins against a strong opponent than when he wins against a weaker opponent. The effect of losing is very small. While defeats against strong opponents only hurt the ranking marginally, effects against weaker opponents have a somewhat larger effect. The positive effects from winning outweigh the negative effects from losing by far.

There is little room for strategic behavior. Targeting a specific ranking is not possible because the ranking calculations are repeated five times and the baseline values applied in the calculation rounds are not visible to the individuals. Furthermore, in order to optimize the ranking it is beneficial to play as often as possible while choosing the optimal tournament difficulty [2]. The exact

[2]Unless we assume that an individual always expects to lose with an extremely high probability.

mechanism of the ranking calculation is explained in more detail in appendix A.1.

To facilitate the organization of tournaments and team competitions, players are assigned to a ranking category based on their ranking. As an example, figure 1 shows the distribution of rankings into the nine different ranking categories in spring 2015 for women and men. These categories are filled from top to bottom. Generally every subsequent category contains twice as many players as the previous one. The distribution hence mimics the shape of a pyramid. The two lowest categories are not filled completely due to varying numbers of active players.

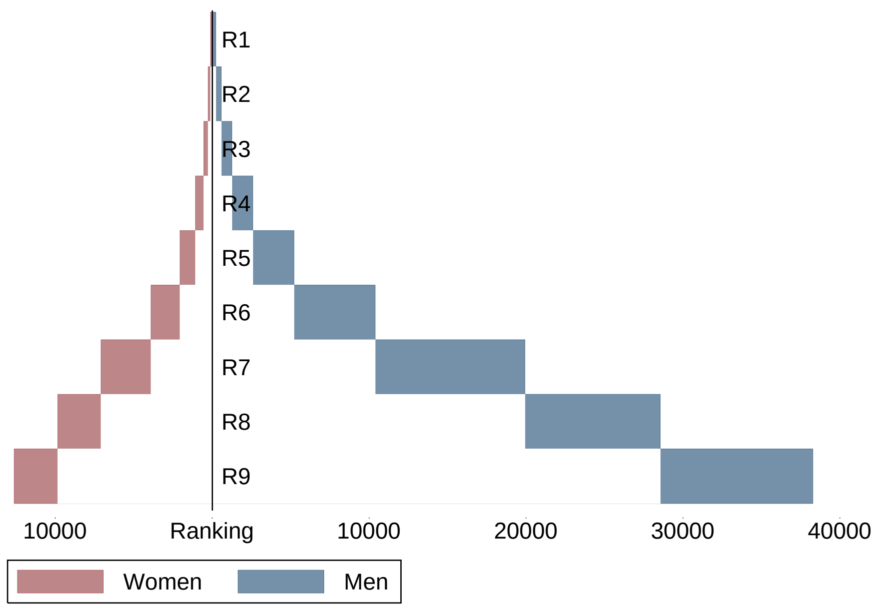


Figure 1: The distribution of rankings into ranking categories for women and men in spring 2015. The ranking categories are filled from top to bottom, starting with the strongest player. A subsequent category generally contains twice as many players as the previous one. For the lowest two categories this rule is not strictly applied due to varying numbers of active players.

2.3 Participation in Single Tournaments

In this paper, I am interested in individuals' motivation to compete. Team competitions with fixed season schedules are not feasible for such an analysis. Single tournaments on the other hand provide a great opportunity to look at individual-level motivation to compete. Consequently, for the rest of the paper I restrict my attention to the performance in the last match of every player in single tournaments.

There is an abundance of these single tournaments. Under the supervision of the Swiss Tennis association 5,522 different organizers carry out tournaments for different ranking categories from once a year up to several times a month. All tournaments are announced through a centralized website and individuals can participate as often as they like. Generally, there are several tournaments close-by every weekend for players of any ranking category. Many tournaments combine multiple ranking categories into one draw. Individuals therefore have the choice to compete in more or less difficult events. To make sure that I do not capture such self-selection effects I only consider fixtures between individuals of the same ranking category who had exactly the same selection of tournaments available to choose from.

Participation in tournaments is completely voluntary and participation costs are low, just covering costs of court rental and provided equipment. Small natural prices are usually provided by sponsors. Subscribing and unsubscribing is possible until three days before the start of a tournament without any consequences in terms of ranking and monetary costs. Hence, the time to the next competition is heavily influenced by the individual motivation to compete.[3] Figure 2 shows the distribution of the time to the next competition. The mode of the distribution is one week. This indicates that many observations are generated by individuals who play frequently. But the distribution is very wide spread.

[3]The date of play of any match is reported as the last day of the tournament.

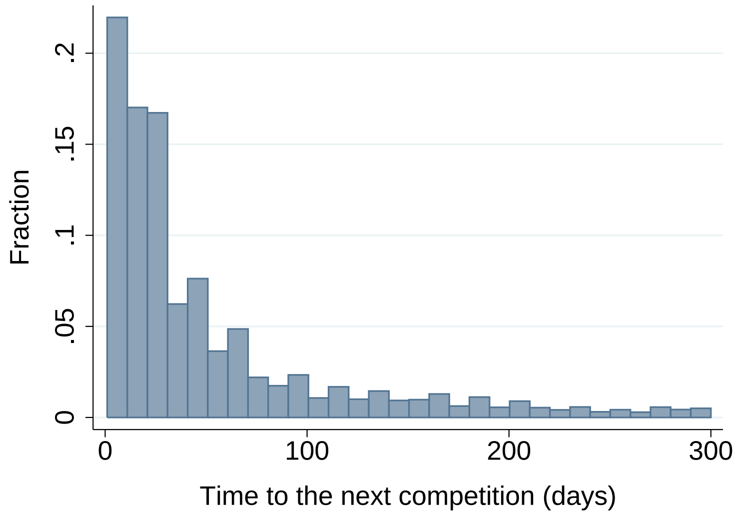


Figure 2: Distribution of the time to the next competition. Note that the distribution is cut at 300 days covering 90% of the observations to achieve a more informative representation.

Once the sign-up window for a tournament is closed, the organizer uses a computer software provided by the Swiss Tennis association to create the tournament draw by randomly assigning individuals to positions in the draw. About the best 20% of the players are seeded what indicates that they cannot face each other in the first round of the tournament.[4] Tournaments are generally carried out as a single-elimination tournament, meaning that the winner continues to the next round while the loser drops out.[5] The exogenous opponent-matching is a crucial property of the data set because it helps identifying causal effects. This matter is discussed in more detail in section 3.2.

Another crucial variable in the data set is the tournament outcome. It is measured for each player by the last round performance at that tournament. In order to link a match outcome to the time until the next competition, it is not feasible to consider observations from previous rounds. By construction an individual keeps competing in a given tournament until she loses. However,

[4]Omitting seeded players does not affect the results.
[5]I only use completed matches in the analyses.

Table 2: Descriptive statistics of individual observations

Variable	Mean	Std. dev.
Time to the next competition (days)	92.85	162.52
Lose	0.80	0.40
Age	29.82	18.15
Male	0.77	0.42
Number of individuals	36,761	
Number of observations	222,831	

the empirical analysis will account for the previous course of the tournament by absorbing tournament round fixed effects.

Table 2 reports the descriptive statistics of the final data set that covers 222,831 observations from 36,761 individuals. The distribution of the time to the next competition is wide spread with an average of 93 days. In 80% of the observations individuals lost in their last match of the tournament while in 20% of the observations individuals ended the tournament with a victory, what usually indicates that they won the tournament.[6] The age ranges from 4 to 88 years with an average of 30 years. The gender-ratio of the observations is proportional to the share of women in the sample and equal to 77%. This indicates that women play as many matches as men conditional on being a ranked tennis player.

[6]Some outdoor tournaments cannot be finished due to bad weather conditions. In this case there are several individuals who win their last match of the tournament.

3 The Causal Effect of Losing in a Competition

This section analyses the causal effect of losing in a competition on the waiting time until the next competition. I first show some graphical evidence, afterwards I describe the IV approach and the econometric analysis used to identify the causal effect, before I eventually discuss the regression results.

3.1 Graphical Evidence

I first provide some graphical evidence of the association between losing in a competition and the waiting time until the next competition. The graphical evidence is purely correlational and ignores potential endogeneity problems. Furthermore the confidence intervals do not take into account that individuals are observed multiple times. Hence, it is for illustration only.

Figure 3 shows the relationship between the outcome in the final of a com-

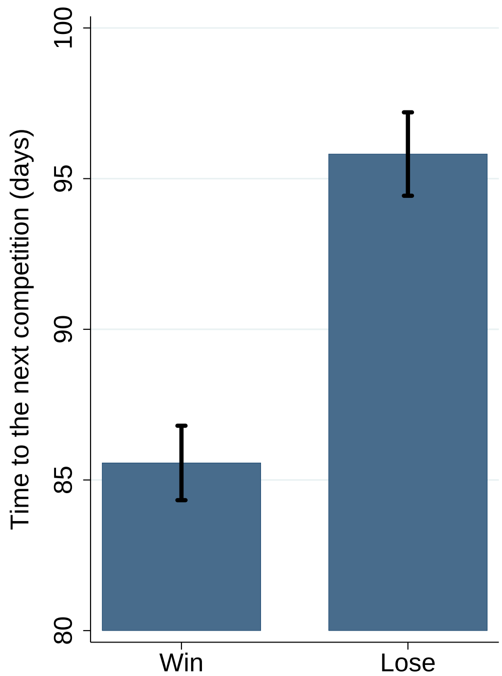


Figure 3: Time to the next competition as a function of the match outcome in the final of a tournament (with 95% confidence intervals).

petition and the number of days individuals wait before competing in the next tournament. Individuals wait on average 86 days after having won, while they wait 10 days longer after having lost. The confidence intervals are sufficiently far apart to make this difference statistically significant.

This simple graphical evidence supports the hypothesis that the waiting time is affected by victories and defeats, but cannot rule out reverse causality or unobserved channels that might jointly affect the match outcome and the time to the next competition. One such channel could be the lack of spare time to play tennis. This affects the probability of winning (through the lack of practice) and the frequency of participation in competitions at the same time. If this lack of spare time is persistent over time, individual fixed effects would take care of it. But if a temporary shock leads to less frequent practicing and competing, like temporary stress at work or a temporary medical condition, it is hard to properly account for it with fixed effects. Because there are many more potential channels of that kind and it is impossible to account for all of them, a good IV strategy is crucial to establish a plausible causal relationship.

3.2 Exogenous Assignment of Opponents as an Instrument

A property of single tennis tournaments is the exogenous assignment of opponents that provides a natural experiment. As described in section 2.3 the draw of every tournament is created randomly by a computer software such that individuals never have control over whom they play against in any round. Hence it seems natural to exploit this exogenous variation in the data for an IV strategy.

The semi-annual national ranking of players (see section 2.2) describes their current strength very well. In a exogenously drawn encounter between two players, the better ranked player clearly wins more often than his weaker opponent. The ranking difference between two opponents is thus strongly correlated with the probability to win.

There are some natural requirements for a measure intended to appropriately

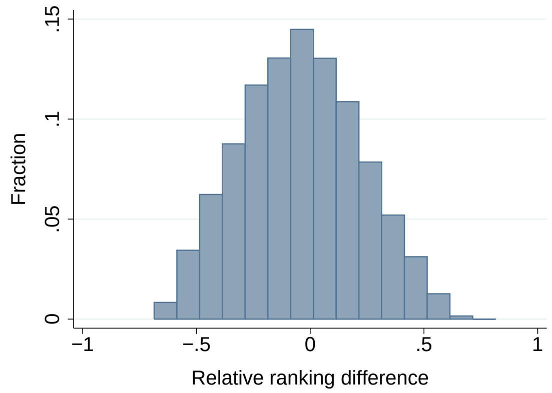


Figure 4: Distribution of the relative ranking difference

proxy the winning probability. First, it should have explanatory power for a large range of values. Obviously, a ranking difference of 1,000 positions between the two players ranked 1 and 1,001 is not comparable to the same ranking difference of 1,000 positions between two players ranked 10,000 and 11,000. Second, the measure should be symmetrical, i.e. the winning probability of a player should be equal to the losing probability of his opponent. Consequently I consider the relative ranking difference[7] defined as

$$R_{it} = \frac{\text{Opponent's Ranking}_{-it} - \text{Individual's Ranking}_{it}}{0.5 \times (\text{Opponent's Ranking}_{-it} + \text{Individual's Ranking}_{it})}.$$

Figure 4 shows the distribution of the relative ranking difference. Positive values indicate that an individual is ranked better than her opponent while negative numbers state that the opponent is ranked better than the individual herself. Two properties are noteworthy. First, most fixtures are played between individuals with similar rankings. Second, the distribution is slightly left skewed. This is due to better ranked players reaching the later tournament rounds more often. Thus they appear more often as opponents in the data set than weaker ranked players.

[7]Using the absolute ranking difference instead does not change the results.

The relative ranking difference is the natural candidate as an instrument for the match outcome in an IV specification. An instrument needs to satisfy two conditions: it has to be strong and fulfill the exclusion restriction. While the strength of the instrument can be easily assessed (section 3.4 shows a first-stage F-test greater than 6,000), showing the exogeneity of an instrument is generally a more delicate matter.

The exclusion restriction of the instrument requires that, conditional on an individual's ranking, the relative ranking difference between two opponents of a given ranking category does not affect an individual's waiting time until the next competition directly, but only through the outcome of the match. This assumption is a priori plausible, because the assignment of opponents in a tournament is exogenous. While ultimately the exclusion restriction remains untestable it is still possible to rule out systematic correlations with the individual's baseline motivation to compete. If the instrument is indeed exogenous, the lagged motivation to compete, i.e. the time since the last competition, should not be affected by the ranking difference between the opponents of the current match. If there was such a correlation with the baseline motivation level, the instrument would likely be invalid.

Table 3 shows that, conditional on a player's ranking, there is no such correlation between the time since the last competition and the current ranking difference between the opponents. The point estimate of the relative ranking difference is very small and insignificant. Being assigned against an opponent ranked 10% worse is related to an increase in the time since the last competition by 0.05 days. This confirms that the assignment of opponents is indeed exogenous and not directly correlated with players' baseline motivation to compete.

The proposed instrument has been shown to be strong and exogenous. The next section describes how the instrument can be used in an econometric analysis to identify the causal effect of losing in a competition on the waiting time until the next competition.

Table 3: Correlation between instrument and time since the last competition

Dependent variable: Break since last competition	
	OLS regression
Relative ranking difference	0.489
	(1.183)
Ranking	0.00193***
	(5.42e-05)
Constant	65.21***
	(0.677)
Observations	210,327
R-squared	0.016

Individual cluster robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

3.3 Econometric Analysis

I now discuss the econometric analysis I use to estimate the effect of losing in a competition on the waiting time until the next competition.

I first lay out a standard OLS model that treats all variables as exogenous. This model is later used as reference for the IV estimates. The regression equation (ignoring fixed effects for notational simplicity) is given by

$$Y_{it} = \gamma_0 + \gamma_1 L_{it} + \gamma_2' X_{it} + \gamma_3' X_{-it} + \varepsilon_{it}. \tag{3.1}$$

The dependent variable Y_{it} captures the time to the next competition for each individual i at time t. The explanatory variables consist of a constant, the binary indicator L_{it} that states whether or not an individual i loses his match in time t, as well as individual (mostly time-varying) characteristics of the individual (X_{it}) and her opponent (X_{-it}).

To determine the causal effect of losing on the waiting time until the next competition I estimate an IV regression model by two stage least squares (2SLS), using R_{it} (i.e. the relative ranking difference between the two opponents) as an instrument for the match outcome. The first stage regression equation is given by

$$L_{it} = \delta_0 + \delta_1 R_{it} + \delta_2' X_{it} + \delta_3' X_{-it} + \varepsilon_{it}' \rightarrow \text{predict } \hat{L}_{it}. \tag{3.2}$$

The binary indicator L_{it} is regressed on a constant, the instrument R_{it}, and all other regressors from equation 3.1. The predicted values $\hat{L}_{it}$ are used as regressor in the second stage of the 2SLS. The second stage regression equation is given by

$$Y_{it} = \eta_0 + \eta_1 \hat{L}_{it} + \eta_2' X_{it} + \eta_3' X_{-it} + \varepsilon_{it}''. \tag{3.3}$$

The coefficient η_1 measures the causal effect of losing in a competition on the time to the next competition.

3.4 Baseline Regression Results

This section shows the results for the OLS and the IV model. Table 4 reports the regression results. The first column states the estimated coefficients from the OLS model and the second column shows the second-stage coefficients of the IV model. The first stage of the IV model is reported in appendix table 8. All models absorb the same fixed effects and include all control variables. Individual fixed effects capture all permanent characteristics of the individuals, while organizer fixed effects account for tournament and location specific characteristics like size or prestige of a tournament. Controlling for different tournament rounds captures the course of the event prior to the current competition. Finally, month and year fixed effects control for reoccurring seasonal effects and pick up year-specific shocks. Standard errors are clustered at the individual level to account for serial correlations.

I first look at the causal effect of losing on the waiting time to the next competition.[8] Losing increases the break after the competition by 10.32 days ($p<0.01$)

[8] One type of heterogeneous effect that is potentially interesting to look at is repeated vs. initial defeats. As the analysis reveals no evidence for such a difference, it is reported in appendix section A.3.

Table 4: Baseline results

Dependent variable: time to the next competition (days)		
	OLS regression	2. stage of IV regression
Lose	7.395***	10.32***
	(0.692)	(3.514)
Ranking	-0.00114***	-0.00106***
	(0.000269)	(0.000285)
Ranking^2	3.34e-08**	3.05e-08**
	(1.41e-08)	(1.45e-08)
Ranking^3	-2.85e-13	-2.50e-13
	(2.08e-13)	(2.11e-13)
Age	1.420*	1.424*
	(0.860)	(0.860)
Opponent's age	0.0496	0.0551
	(0.0381)	(0.0387)
First stage F-test		6,007.14
Observations	212,025	212,025
R-squared	0.534	0.534

All regressions absorb individual, organizer, tournament round, year, and month fixed effects. Individual cluster robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

in the IV-specification. This is roughly a 10% increase from the baseline level of 93 days. Considering the OLS coefficient, there is evidence for underestimation of the effect when ignoring potential endogeneity. But the difference is moderate and the 95% confidence intervals of the two estimated coefficients overlap. Even though the instrument is very strong (first stage F-test of 6,007) the IV model is less precise than the OLS model. This is a general property of the 2SLS estimator (e.g. Wooldridge 2010).

I know turn to the estimated coefficient of the ranking. The current individual ranking is an important control variable as the instrument is exogenous conditional on this variable. The estimated relationship with the time to the next competition is statistically significant. The estimated first-order term is meaningful while second and third order terms of the polynomial have a negligible impact only. In combination these coefficients indicate that individuals ranked 1,000 positions better on average compete again 1 day later. This correlational effect likely captures individuals playing very frequently when they are young and poorly ranked [9].

Age also has a strongly positive (though only marginally significant) relationship with the time to the next competition. For every year an individual grows older she takes 1.4 days longer between competitions.

[9] Including the interaction between ranking and age renders the ranking coefficient insignificant.

4 Reference Points Influence the Perception of Competition Outcomes

The last section showed that the waiting time until the next competition is causally affected by competition outcomes. One mechanism that can explain this behavior are reference-dependent preferences. Compared to a neutral reference-point victories are in the gains domain and increase an individual's motivation to compete while defeats are in the loss domain and decrease the motivation to compete. In this section I now look at the reference point formation in more detail by showing evidence for two types of reference points. The comprehensive dataset allows to construct two measures that mimic status quo and expectation-based reference points. First, I look at the influence of individual rankings. I find that not every loss is equally demotivating but that the ranking order between two opponents strongly influences the perception of a loss. Afterwards I estimate a second model that considers individuals' ex-ante winning probabilities. The results indicate that individuals are more sensitive to defeats and victories when they come as a surprise.

4.1 Ranking as Reference Point

To explore the influence of individual rankings as reference point, I exploit a quasi-experimental feature of the data that allows for a clean identification by applying a regression discontinuity (RD) design.

Within a small in the national ranking, individuals can be considered as being similarly strong. The expected match outcome and the consequences on individual rankings change gradually and not discontinuously when competing against opponents ranked slightly above or ranked slightly below an individual's own ranking. Also the quality of a match and other unobserved characteristics are likely to be similar for individuals within a narrow range of each other.

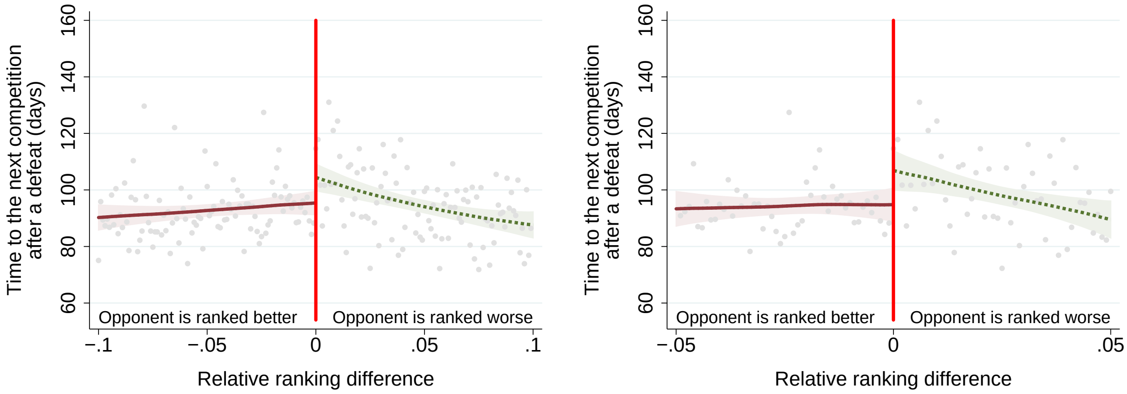
An individual's own ranking separates opponents ranked better from opponents ranked worse than the individual herself. If this cutoff strongly affects the time until the next competition, this indicates that individuals use their own rankings as a reference point when evaluating victories and defeats in competitions.

Graphical Evidence

Figures 5 and 6 show graphical evidence for the role of rankings as reference points for defeats and victories respectively. I first consider defeats. Panel a) in figure 5 shows a clear regression discontinuity at the individuals' own ranking using a bandwidth of 10%. When an individual loses against an opponent ranked slightly worse than herself, she waits about 9 days longer to compete again compared to when she loses against an opponent ranked slightly better than herself. Panel b) shows an even bigger effect of almost 12 days if we consider a smaller bandwidth of only 5%. From a baseline time to the next competition of about 100 days, this is an increase of 9-12%.

I now turn to the role of the individuals' rankings when they win. Figure 6 shows no clear discontinuity, neither in panel a) with a 10% bandwidth, nor in panel b) with a 5% bandwidth. The 95% confidence intervals are wide in both panels and I thus do not reject the null hypothesis that there is no regression discontinuity at the individuals' own ranking.

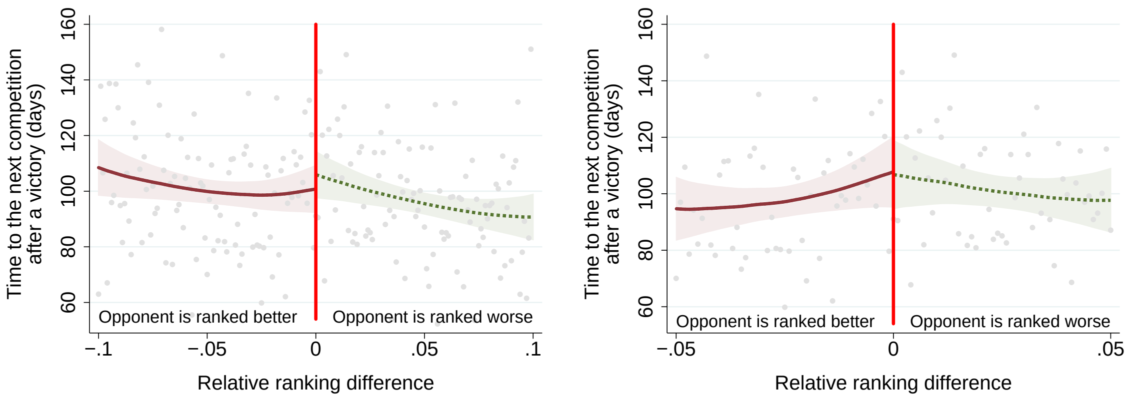
The graphical evidence indicates that there is a strong effect own individuals' ranking when they lose but not when they win. To establish a clean causal relationship in a regression discontinuity framework it is crucial to discuss some sensitive assumptions related to such a design. In the next sections I first look at the validity of the proposed regression discontinuity design, later discuss the econometric method, and finally show that the results are insensitive to various bandwidth choices.



a) 10% bandwidth

b) 5% bandwidth

Figure 5: Defeats - regression discontinuity of the time to the next competition at the individual's own ranking. An individual who loses against an opponent who is ranked slightly worse than herself stays away from competitions about 9-12 days longer compared to an individual who loses against an opponent ranked slightly better than herself.



a) 10% bandwidth

b) 5% bandwidth

Figure 6: Victories - no regression discontinuity of the time to the next competition at the individual's own ranking. When an individual wins it does not matter if the opponent is ranked slightly better or slightly worse than herself.

Sharp Regression Discontinuity Framework

The design applied in the graphical analysis conceptually corresponds to a sharp regression discontinuity design as proposed by Thistlethwaite and Campbell (1960) (see Imbens and Lemieux (2008) and Lee and Lemieux (2010) for practitioner's guides).

The individuals' ranking is a sharp cutoff that separates opponents into two groups - the ones who are relatively better and the ones who are relatively worse than the individual herself. Hence, the treatment variable is the binary variable,

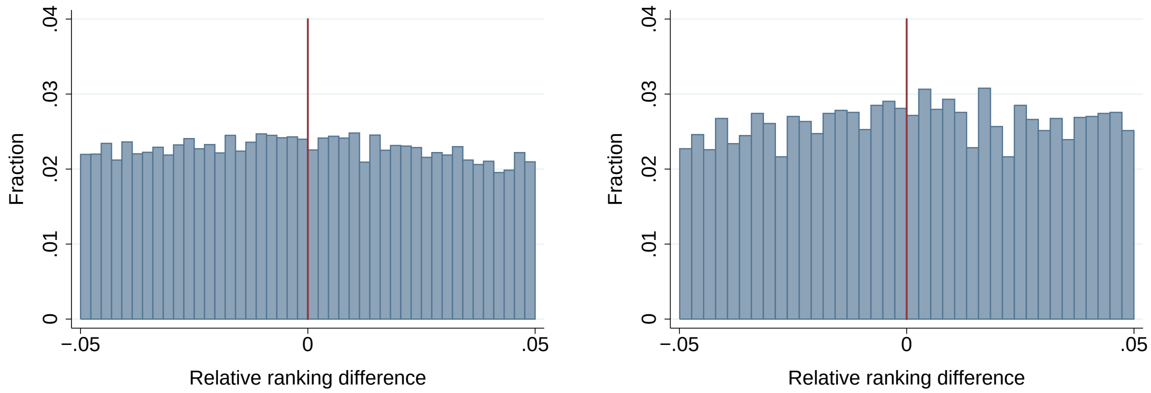
$$D_{it} = 1\{R_{it} \geq 0\},$$

that distinguishes these two groups.

The crucial requirement for a valid RD-design is the inability of individuals to precisely manipulate the treatment variable (Lee and Lemieux 2010). In this setting, this is ensured by the exogenous assignment of opponents. Individuals have no possibility to effectively choose if they are better or worse than their opponent. One way to somewhat manipulate the cutoff would be to fail to report to matches against opponents with a very similar ranking.[10] Figure 7 shows the density of the relative ranking difference on both sides of the cutoff. Panel a) shows defeats while panel b) shows victories. In neither case there is evidence that individuals manipulate the cutoff by missing competitions if they are ranked very close to their opponent. Hence, the local randomization assumption is satisfied.

It follows directly from the local randomization that on both sides of the cutoff opponents are comparable in all characteristics except for the treatment variable (Lee and Lemieux 2010). These characteristics include observable variables such as gender and age, but also unobservable variables like the relative

[10] I show that the frequency of forfeits does not change discontinuously at the cutoff in appendix table 10.



a) Defeats b) Victories

Figure 7: Density of relative ranking difference near the cutoff. Panel a) shows the distribution for defeats and panel b) shows the distribution of victories. There is no evidence that individuals manipulate the cutoff.

strength (i.e. the probability to win) or the baseline motivation to compete. Ultimately, it is not possible to rule out every unobserved channel. Choosing narrow bandwidths and showing that observed covariates do not jump at the cutoff make the RD-design credible. Appendix table 10 reports estimated discontinuities for various covariates but finds no significant jumps for any bandwidth.

Estimation

The sharp regression discontinuity design is routinely estimated by a nonparametric local linear regression (Lee and Lemieux 2010). In this setup the regression equation is given by

$$Y_{it} = \alpha_l + \tau D_{it} + \beta_l(R_{it} - c) + (\beta_r - \beta_l)D_{it}(R_{it} - c) + \varepsilon_{it}, \tag{4.1}$$

where $c - h \leq R_{it} \leq h - c$, while $c = 0$, and h is the bandwidth.

Choosing the bandwidth is not trivial as Wald estimates in RD-settings are very sensitive to the bandwidth choice (Lee and Lemieux 2010). The starting point for bandwidth selection has traditionally been the Silverman (1986) rule of thumb (ROT) with some undersmoothing. ROT bandwidth choices are generally suspected to yield too large bandwidths in RD frameworks. Consequently Calonico, Cattaneo and Titiunik (2014) (CCT) propose a bias-corrected approach that yields smaller bandwidth choices and consistent standard errors.

In the next section I will show the results for the two narrow bandwidths used in the graphical analysis as well as for the optimal bandwidth chosen by CCT. Later, I show robustness checks, indicating that the results are robust to a large range of different bandwidth choices.

RD Results

I now show the estimation results of the local linear regressions. Table 5 reports the results for defeats while table 6 reports the results for victories. In both tables, columns 1 and 2 apply the two specifications with narrow bandwidths used in the graphical analyses. Column 3 shows the robust, bias-corrected CCT estimates. Standard errors are clustered at the individual level in all specifications but column 3 that shows robust standard errors.

First I look at defeats. The first column in table 5 shows the local Wald estimate using strictest bandwidth of 5%. The estimated discontinuity is 12 days. This implies that an individual takes a break between competitions that is significantly longer when she loses against an opponent ranked just worse than herself compared to when she loses against an opponent ranked just better than herself. This result is robust to choosing a larger bandwidth as reported in column 2 and 3. While a larger bandwidth includes more observations on both sides of the cutoff and thereby increases the efficiency of the estimation, it also increases the potential bias. Hence, results from specifications with smaller bandwidth are more credible as they follow the idea of a local estimation more strictly.

I now turn to victories. In line with the graphical evidence there is no effect in any specification. Individuals take the same break between competitions no matter if they win against relatively stronger or weaker opponents.

Table 5: Ranking influences evaluation of defeats - regression discontinuity in time to the next competition at relative ranking difference of 0

	(1)	(2)	(3)
	5%	10%	CCT
Wald estimator (τ)	12.059***	9.028***	9.528***
	(4.658)	(3.295)	(3.467)
Considered observations	25,228	48,115	55,269
Bandwidth loc. poly.	0.050	0.100	0.117
Bandwidth bias correction			0.206

Results from local linear regressions using a triangular kernel. The 3 columns show different bandwidth choices in line with common selection criteria. (Individual cluster) robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Ranking does not influence evaluation of victories - regression discontinuity in time to the next competition at relative ranking difference of 0

	(1)	(2)	(3)
	5%	10%	CCT
Wald estimator (τ)	-1.000	5.182	5.685
	(9.213)	(6.459)	(6.7646)
Considered observations	7,439	13,871	16,231
Bandwidth loc. poly.	0.050	0.100	0.121
Bandwidth bias correction			0.203

Results from local linear regressions using a triangular kernel. The 3 columns show different bandwidth choices in line with common selection criteria. (Individual cluster) robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Robustness Checks

I show two types of robustness checks. The first analysis shows that Wald estimates from the regression discontinuity are stable over a large range of bandwidths. The second analysis shows that there are no regression discontinuities at other values of the treatment variable, where we would not expect to find such a discontinuity.

The last section already reported regression results with four different bandwidth choices. Figure 8 additionally plots the Wald estimates against a large range of bandwidths. The smallest bandwidth I consider is 1% and the largest bandwidth is 30%. The graph shows that Wald estimates are significantly positive for any bandwidth greater or equal than 2%. The more local the regression, the larger is the treatment effect. When we increase the bandwidth, the estimated treatment effects get somewhat smaller. All considered estimates are between 5 and 15 days and can thus be considered robust.

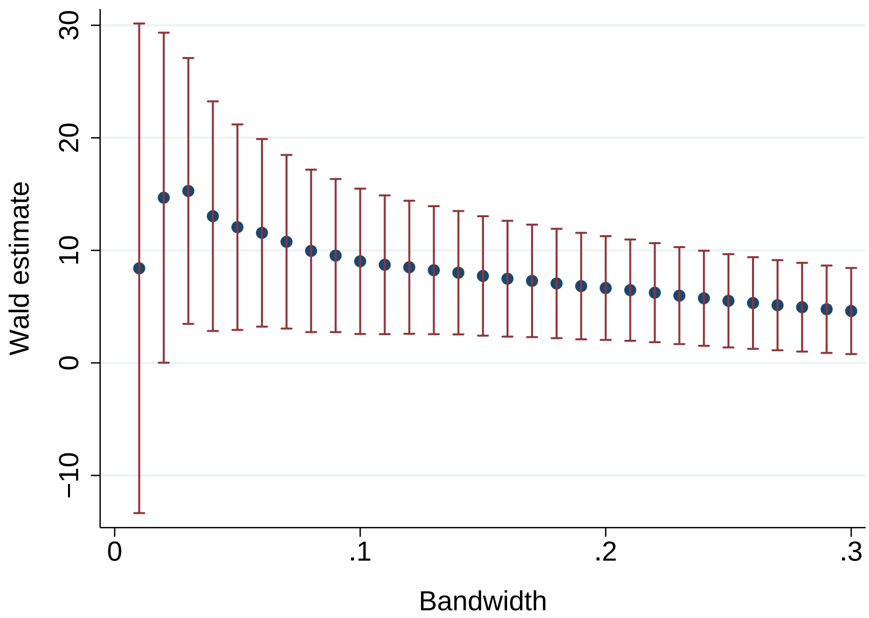


Figure 8: Defeats - Sensitivity of Wald estimates to the bandwidth choice. Wald estimates along with 95% confidence intervals.

Table 7: No jump in the time to the next competition at non-discontinuity points, when losing

	(1)	(2)	(3)	(4)
	-10%	-5%	+5%	+10%
Wald estimator	-1.579	1.527	3.787	-2.449
	(4.548)	(4.640)	(4.833)	(4.711)
Considered observations	24,087	25,083	23,032	19,961
Bandwidth loc. poly.	0.050	0.050	0.050	0.050

Results from local linear regressions using a triangular kernel. The four columns report different non-discontinuity points in the neighborhood of the discontinuity cut-off. Individual cluster robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

The second robustness check involves testing for discontinuities in the treatment variable at points where we do not expect to find such a discontinuity. If we were to find such regression discontinuities this would be a warning sign as we could not be sure that the regression discontinuity described in the last section is not an outcome of chance. Table 7 reports Wald estimates at four non-discontinuity cutoffs. The considered values range from a relative ranking difference of -10% to a relative ranking difference of +10%. All four Wald estimates are close to zero and estimated with little precision. Hence, there is no evidence of discontinuities at non-discontinuity points.

In summary, the results from this section indicate that individuals use their own ranking as a reference point when evaluating defeats. Losing against opponents who are a little stronger than themselves is less hurtful than losing against a supposedly weaker opponent. Individuals seem to have the desire to defend their ranking against opponents just behind them in the national ranking. This mechanism does not exist for victories. It seems that winning always has the same impact, no matter if individuals win against relatively stronger or weaker opponents.

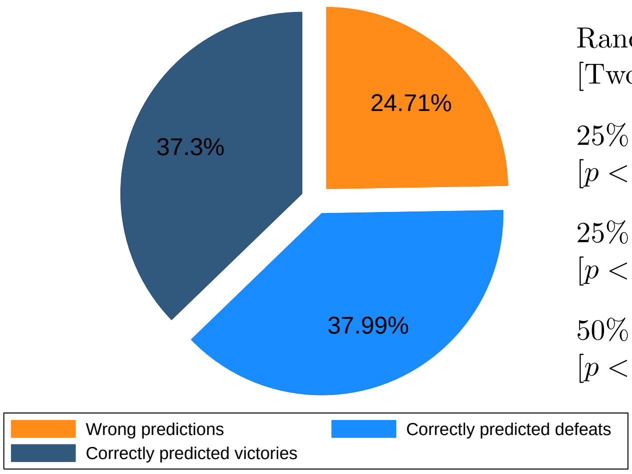
4.2 Ex-ante Winning Probability as Reference Point

While the last section showed evidence for reference-dependent behavior when an individual is faced with an opponent who is immediately behind her in the national ranking, I now consider the full range of opponents and show how surprising victories and defeats influence the waiting time until the next competition. Therefore, I model and predict the winning probability of each individual for every match so that we can see how they assess outcomes relative to those ex ante winning probabilities.

Method

First, I estimate winning probabilities for every match. There are many factors that decide who wins in a competition. Some of them are unobservable, like the daily shape, preferences for certain weather or surface conditions, or simply preferences with respect to the opponent's style of play. Yet many other factors are observable or can be proxied for. I use these observable factors, that are also available to the individuals before every competition, to form accurate predictions about the match outcomes. Namely, I include the relative ranking difference between opponents (including up to third order polynomial terms), individual characteristics of both opponents, like the ranking (including polynomials), age, the number of matches played in the current season, the win/lose-ratio , and the number of consecutive wins. It is possible to use all observations from single tournaments to estimate the correlation of these variables with the match outcome. Subsequently I predict the winning probability of each individual for every match.

These predictions are very accurate in foretelling who wins the match, as can be seen in figure 9. To assess the prediction accuracy we can generate a binary variable defining predicted victories and defeats by the most likely outcome. If we compare the predicted with the observed competition outcomes we see that



Random predictions as benchmark [Two-sample t test]:

25% correctly predicted victories [$p < 0.001$]

25% correctly predicted defeats [$p < 0.001$]

50% wrong predictions [$p < 0.001$]

Figure 9: Prediction accuracy for competition outcomes. Actual predictions are reported in the left panel while the right panel shows random predictions as benchmark.

75% of the observations are indeed predicted correctly. Victories and defeats are predicted equally well. All predictions are significantly better than random predictions that are reported for comparison in the right panel of figure 9.

We can now use the (continuous) predicted winning probability to calculate the deviation of the actual from the predicted outcome for every individual in every competition. This variable proxies for the level of surprise. If an individual is likely to win but loses this is equally surprising as if an individual, likely to lose, in fact wins.

In the last step I look at the correlation between the deviation from the predicted outcome and the time until the next competition. To allow for a flexible relationship between these two variables, I use a nonparametric approach in this last step. The next section describes this in more detail.

Econometric Analysis

The first stage of the two stage estimation procedure described in the last section is a Logit regression of the winning probability on the relative ranking difference and various individual characteristics of the individual and her opponent, as well as a constant. The first stage regression equation is given by

$$P(win)_{it} = \zeta_0 + \zeta_1' R_{it} + \zeta_2' X_{it} + \zeta_3' X_{-it} + \epsilon_{it}^v \rightarrow \text{predict } \hat{P}_{it}. \tag{4.2}$$

The fitted values ($\hat{P}_{it}$) obtained from this first stage regression are then subtracted from the competition outcome variable in the following way:

$$S_{it} = \begin{cases} W_{it} - \hat{P}_{it} & \text{if } W_{it} = 1 \\ L_{it} - (1 - \hat{P}_{it}) & \text{if } L_{it} = 1 \end{cases}$$

If an individual wins, the level of surprise S_{it} is given by the difference between 1 and the predicted probability to win $\hat{P}_{it}$. If an individual loses, S_{it} is calculated by the difference between 1 and the predicted probability to lose $(1 - \hat{P}_{it})$.

In the second stage the time to the next competition Y_{it} is regressed on S_{it} using a local polynomial regression to allow for flexibility. The regression equation is given by

$$Y_{it} = f(S_{it}) + \epsilon_{it}, \tag{4.3}$$

using the Epanechnikov kernel smoother. The optimal bandwidth is calculated by Silverman's rule of thumb (Silverman 1986) [11]. Individual cluster robust standard errors from this two-stage estimator are obtained from 1,000 bootstrap replications (Horowitz 2001).

[11] Cross validation yields a very similar bandwidth in the original sample (less than 5%-points difference) but is computationally infeasible for bootstrapping standard errors. The estimated slopes are robust to using the cross validated bandwidth.

Results

The first stage works very well in accurately predicting winning probabilities. The pseudo-R^2 is 28%. The most important variables are the relative ranking difference and the individuals' win/lose-ratio in the current season. The first stage with the full set of estimated coefficients is reported in appendix table 11.

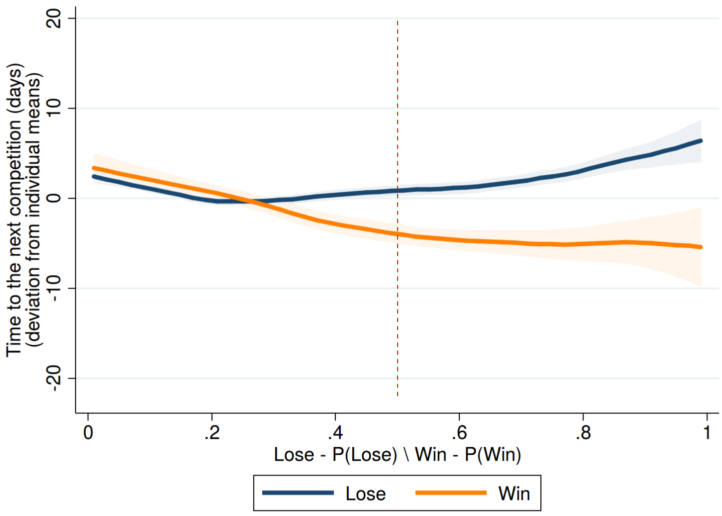
The results from the second stage are reported in a graphical way. A robustness check using OLS as second stage yields qualitatively the same results and is reported in appendix table 12. Figure 10 shows the estimated functions along with the bootstrapped 95% confidence intervals. Panel a) reports the results from a local mean regression while panel b) shows results from a local linear regression. The results indicate that falling short of the predicted outcome (i.e. losing when the predicted probability to win was greater than 0) is related with a longer time to the next competition. The stronger the deviation from the prediction, the stronger is this effect. On the contrary, over-performing (winning while the probability to win was smaller than 1) is linked with a shorter time to the next competition. Again, the stronger the deviation from the prediction, the stronger is this effect.

In line with the results from section 3.4 we see that individuals on average take a longer break between competitions when they lose compared to when they win. For outcomes close to the predicted values (left half of the graphs) this effect is only very weak. For surprising outcomes (right half of the graphs) the difference between losing and winning is very pronounced. These findings indicate that expectations might play an important role when individuals evaluate their performance in competitions. Losing seems to hurt more when the outcome was unexpected while winning motivates more when the victory came as a surprise.

One might be concerned that the analysis merely catches individuals in transition, i.e. individuals getting better when they play more often and individuals getting worse when they play less frequently. The control variables in the first stage, namely the number of matches, the win/lose ratio and the number of consecutive wins in the current season account for this potentially confounding effect by adjusting the predicted winning probability.

The next section integrates the results of this paper into the discussion about reference point formation.

a) Local mean regression



b) Local linear regression

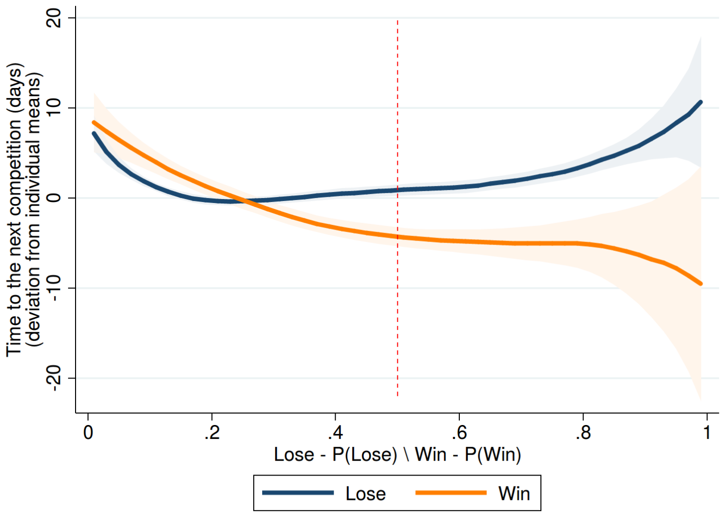


Figure 10: Time to next competition as a function of deviation from predicted outcome. Fit from local polynomial regression along with 95% confidence bounds based on 1,000 bootstrap replications.

4.3 Compatibility of results with different concepts of reference points

There are two main distinctions with respect to reference points that have received a lot of attention in the competition literature (Mas 2006; Pope and Schweitzer 2011; Ockenfels, Sliwka and Werner 2014; Bartling, Brandes and Schunk 2015; Allen et al. forthcoming): status quo and expectation-based reference points.

Status quo reference points

Kahneman and Tversky (1979) originally proposed the status quo as reference point. In the context of this paper the value of the uncertain competition outcome is measured by the time break until the next competition and the status quo reference point of an individual is neither losing nor winning. Victories should thereby be seen as the gains domain while defeats are located in the loss domain. The baseline results discussed in section 3 are compatible with this view. The individuals' motivation to compete is lower after they have lost then after they have won in a competition.

More explicitly, the regression discontinuity analysis in section 4.1 identified individuals' current rankings as status quo reference point. Individuals seem to feel entitled to their current ranking and willing to defend it against their immediate rivals [12] (Gill et al. 2015). This implies that individuals weight defeats by the relative strength of opponents. In that sense an individual has more to lose when competing against a relatively weaker compared to a relatively stronger opponent. The RD-results from section 4.1 support this point of view for the loss domain. Individuals are less motivated to compete again after they have lost against relatively weaker opponents than when they have lost against relatively stronger opponents. The fact that this result does not hold in the gains domain is

[12] This is a purely psychological reasoning as the actual ranking is not computed that way.

compatible with Kahneman and Tversky (1979) who generally predict a steeper value function for losses than for gains.

Expectation-based reference points

Alternatively, Kahneman and Tversky (1979) explicitly allow for expectations about uncertain outcomes as reference point and Kőszegi and Rabin (2006) formalize this idea. This concept is a natural candidate for the context of this paper. Let us conjecture that individuals form expectations about the outcome before every competition. If they perform better than expected motivation should increase, while the opposite is true if they perform worse than expected. Defeats can never have a positive motivational effect because at most the defeat was completely expected, i.e. it has no effect. Normally defeats decrease the motivation to compete. The opposite is true for victories. They almost always have a positive impact on the motivation to compete, unless they were completely anticipated. In this case they have no effect on the individuals' motivation.

In the close neighborhood of an individual's own ranking expectation-based reference points are not applicable as discontinuities are not compatible with the formation of expectations. If individuals formed proper expectations about the match outcome they should not be influenced by small ranking differences. But if we consider the whole range of opponents, including opponents further away from the individuals' own ranking, we see that individuals indeed behave as if they considered expectations about the competition outcomes as reference point. Positive surprises are related to a larger motivation to compete while negative surprises have the opposite effect.

In conclusion the results hint to the coexistence of status quo and expectation-based reference point. Neither of the concepts can independently explain the results but they complement each other. Even though the setups are completely different, the conclusions are remarkably similar to lab evidence by Song (2015).

5 Conclusion

This paper uses a large natural experiment with amateur tennis players to identify the causal effect of losing in competitions on the waiting time until the next competition. The random assignment of players in tournament draws generates an exogenous ranking difference between opponents that is used as instrument for the competition outcome. I find that losing in a competition increases the time break between competitions by 11%, on average. The comprehensive data set allows to investigate the reference point formation in this context. In the narrow neighborhood of an individual's ranking, losing against relatively weaker opponents has a stronger demotivational effect than losing against relatively stronger opponents. This is evidence that individuals use their own position in the national ranking as reference point. By predicting ex-ante winning probabilities I finally find that surprising victories and defeats, measured as deviations from predicted values, weigh heavier than more likely outcomes. This is evidence that away from the local ranking reference point, expectations indeed matter.

One important open question is the precise psychological mechanism behind the demotivational effect of losing in competitions. The results are compatible with individuals disliking disappointment and staying away from competition to avoid losing again. Alternatively, individuals might decide to take a break from competition to avoid being reminded of the past defeat. Another issue that remains open for further research is identifying the interaction mechanisms of status quo and expectation based reference points using field data. The results of this paper suggest that both concepts are important. But how exactly the two concepts work together remains an open question.

Independent of the precise psychological mechanism the results suggest that disappointing competition outcomes can create time inconsistencies that prevent individuals from competing at an optimal rate. This finding has wide implications. Consider the job market where individuals might shy away from repeated competition by not applying regularly to new jobs or promotions after having experienced a setback. They might in turn fail to get a wage raise, promotion, or a new job. In particular, disappointment from losing in competitions might help to explain why long-term unemployed have such a hard time finding a job.

The findings from this paper have several policy implications. Individuals should be encouraged to prearrange a competition schedule they stick to, independent of the latest competition outcome. Furthermore it might be optimal for individuals to aim high in competitions in order to contain disappointment after a potential setback.

A Appendix

A.1 Calculation of the individual ranking

Individuals are ranked according to their ranking value. The ranking value consists of two parts, the competition value C and the participation bonus P. The competition value is given by

$$C = \frac{1}{2}\left\{\ln\left(\sum_{i=1}^{v} e^{c_i} + e^{c_0}\right) - \ln\left(\sum_{j=1}^{d} e^{-c_j} + e^{-c_0}\right)\right\},$$

where v is the number of victories, d is the number of defeats, c_0 is last period's competition value multiplied with 0.75, c_i are the competition values of defeated opponents and c_j are the competition values of lost opponents. Victories have a positive impact while defeats have a negative impact but the positive impact of a victory is always larger than the negative impact of a defeat against the same opponent. Not playing at all leads to a decrease of the competition value by 25%.

To the competition value a participation bonus

$$P = \frac{1}{6}\left\{\ln\left(\sum_{i=1}^{v} e^{c_i} + e^{c_0}\right) + \ln\left(\sum_{j=1}^{d} e^{-c_j} + e^{-c_0}\right)\right\}$$

is added that only depends on the number of matches and the competition values of the opponents but is independent of the match outcome. It always creates an incentive to participate more often.

A.2 The Causal Effect of Losing in a Competition

Table 8: First stage regression of baseline results

Binary dependent variable: Lose	
	OLS regression
Relative ranking difference	-0.282***
	(0.00364)
Ranking	-3.44e-05***
	(1.07e-06)
Ranking^2	1.25e-09***
	(5.37e-11)
Ranking^3	-1.51e-14
	(7.56e-16)
Age	-0.000808
	(0.00306)
Opponent's age	-0.00185***
	(0.000110)
Observations	212,025
R-squared	0.353

The regression absorbs individual, organizer, tournament round, year, and month fixed effects. Individual cluster robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

A.3 Repeated Defeats

A potentially interesting question to look at is whether the effect of losing in competitions on the time until the next competition is different for repeated defeats. Different mechanisms could give rise to such heterogeneities. While repeated defeats might feel more disappointing, it is also possible that individuals get used to losing (Peterson, Maier and Seligman 1993).

This question can be analyzed by including the binary indicator L_{it}^{t-1} whether or not the individual lost the competition one period before and the interaction between the two instances of L_{it}^{t-1} and L_{it}.

$$Y_{it} = \lambda_0 + \lambda_1[L_{it} \mid L_{it}^{t-1} = 1] + \lambda_2[L_{it} \mid L_{it}^{t-1} = 0] + \lambda_3 L_{it}^{t-1} + \lambda_4' X_{it} + \lambda_5' X_{-it} + \varepsilon_{it}'''. \tag{A.1}$$

To determine the causal effect of losing on the waiting time until the next competition, I again estimate an IV regression model by 2SLS, using the interaction between R_{it} and the two instances of L_{it}^{t-1} as instruments.

There is no evidence that the prior competition outcome influences the perception of the current loss. The full set of results is reported in table 9.

Table 9: Repeated defeats

Dependent variable: time to the next competition (days)		
	OLS regression	2. stage of IV regression
Lose | Lost previous competition	6.635***	9.537**
	(0.738)	(3.924)
Lose | Won previous competition	6.681***	8.582
	(1.266)	(5.906)
Lost previous competition	1.316	0.570
	(1.230)	(5.032)
Ranking	-0.000896***	-0.000823***
	(0.000261)	(0.000277)
Ranking^2	1.69e-08	1.42e-08
	(1.39e-08)	(1.43e-08)
Ranking^3	-7.82e-14	-4.65e-14
	(2.06e-13)	(2.10e-13)
Age	0.906	0.906
	(0.826)	(0.826)
Opponent's age	0.0399	0.0448
	(0.0369)	(0.0374)
First stage F-test		2471.07
$\beta_1 = \beta_2$ (p-value)	0.974	0.889
Observations	201,080	201,080
R-squared	0.527	0.527

All regressions absorb individual, organizer, tournament round, year, and month fixed effects. Individual cluster robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

A.4 Ranking as reference point

Table 10: No jump of covariates when losing[13] at relative ranking difference of 0

	(1)	(2)	(3)
	5%	10%	CCT
Wald estimators:			
Days before the last competition	1.039	0.129	0.0449
	(3.794)	(2.669)	(2.814)
Likelihood to lose	-0.000918	0.00528	0.00776
	(0.0103)	(0.00736)	(0.00763)
Frequency of forfeits	-0.00166	-0.00222	-0.00255
	(0.00286)	(0.00208)	(0.00222)
Ranking	-43.84	-221.7	-205.7
	(322.7)	(223.7)	(229.2)
Age	-0.0440	-0.324	-0.357
	(0.523)	(0.375)	(0.382)
Male	0.00161	0.00119	0.00110
	(0.0118)	(0.00851)	(0.00862)
Bandwidth loc. poly.	0.050	0.100	0.117
Bandwidth bias correction			0.206

Results from local linear regressions using a triangular kernel. The 3 columns show different bandwidth choices in line with common selection criteria. (Individual cluster) robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

[13] Coefficient for likelihood to lose and frequency of forfeits are not conditioned on losing.

A.5 Winning probability as reference point

Table 11: First stage regression of surprise

Binary dependent variable: Win	Logit regression
Relative ranking difference	1.971***
	(0.00630)
Relative ranking difference^2	-0.00823*
	(0.00478)
Relative ranking difference^3	0.0216***
	(0.00524)
Individual's characteristics:	
Ranking	2.69e-06**
	(1.24e-06)
Ranking^2	-4.41e-10***
	(7.14e-11)
Ranking^3	1.12e-14
	(1.08e-15)
Age	-0.00992***
	(0.000202)
No. of matches in current season	-0.00136***
	(0.000196)
Win/lose-ratio in current season	-1.550***
	(0.0107)
No. of consecutive wins	0.0532***
	(0.00122)
Opponents characteristics:	
Age	0.00964***
	(0.000202)
No. of matches in current season	0.000852***
	(0.000196)
Win/lose-ratio in current season	1.504***
	(0.0108)
No. of consecutive wins	-0.0249***
	(0.00121)
Constant	-0.0354***
	(0.0117)
Observations	1,425,048
Pseudo R-squared	0.280

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 12: Second stage from deviations from predictions model using OLS

Dependent variable: time to the next competition (days)	
	OLS Regression
$S_{it} \times W_{it}$	-8.913***
	(1.299)
$S_{it} \times L_{it}$	7.350***
	(0.768)
Ranking	-0.00164***
	(0.000173)
Ranking^2	3.19e-08***
	(8.24e-09)
Ranking^3	-1.63e-13
	1.14e-13
Age	0.940
	(0.588)
Opponent's age	0.0153
	(0.0248)
Observations	610,560
R-squared	0.492

The regression absorbs individual, organizer, tournament round, year, and month fixed effects. Individual cluster robust standard errors (based on 1,000 bootstrap replications) in parentheses. *** p<0.01, ** p<0.05, * p<0.1

References

Allen, Eric J, Patricia M Dechow, Devin G Pope and George Wu. forthcoming. "Reference-dependent preferences: Evidence from marathon runners." *Management Science* .

Bartling, Björn, Leif Brandes and Daniel Schunk. 2015. "Expectations as reference points: Field evidence from professional soccer." *Management Science* 61(11):2646–2661.

Bordalo, Pedro, Nicola Gennaioli and Andrei Shleifer. 2012. "Salience Theory of Choice Under Risk." *The Quarterly Journal of Economics* p. qjs018.

Budd, Alan, Paul Levine and Peter Smith. 1988. "Unemployment, vacancies and the long-term unemployed." *The Economic Journal* 98(393):1071–1091.

Buser, Thomas. forthcoming. "The impact of losing in a competition on the willingness to seek further challenges." *Management Science* .

Calonico, Sebastian, Matias D Cattaneo and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82(6):2295–2326.

Card, David and Gordon B Dahl. 2011. "Family violence and football: The effect of unexpected emotional cues on violent behavior." *The Quarterly Journal of Economics* 126(1):103–143.

Cornil, Yann and Pierre Chandon. 2013. "From fan to fat? Vicarious losing increases unhealthy eating, but self-affirmation is an effective remedy." *Psychological science* 24(10):1936–1946.

Ding, Min, Jehoshua Eliashberg, Joel Huber and Ritesh Saini. 2005. "Emotional bidders - An analytical and experimental examination of consumers' behavior in a priceline-like reverse auction." *Management Science* 51(3):352–364.

Fisman, Raymond, Sheena S Iyengar, Emir Kamenica and Itamar Simonson. 2006. "Gender differences in mate selection: Evidence from a speed dating experiment." *The Quarterly Journal of Economics* pp. 673–697.

Gill, David and Victoria Prowse. 2012. "A structural analysis of disappointment aversion in a real effort competition." *The American economic review* pp. 469–503.

Gill, David and Victoria Prowse. 2014. "Gender differences and dynamics in competition: The role of luck." *Quantitative Economics* 5(2):351–376.

Gill, David, Zdenka Kissová, Jaesun Lee and Victoria Prowse. 2015. "First-place loving and last-place loathing: How rank in the distribution of performance affects effort provision." *Available at SSRN 2641875* .

Hitsch, Günter J, Ali Hortaçsu and Dan Ariely. 2010. "What makes you click?Mate preferences in online dating." *Quantitative marketing and Economics* 8(4):393–427.

Hoekstra, Mark. 2009. "The effect of attending the flagship state university on earnings: A discontinuity-based approach." *The Review of Economics and Statistics* 91(4):717–724.

Horowitz, Joel L. 2001. "The bootstrap." *Handbook of econometrics* 5:3159–3228.

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

Jackman, Richard and Richard Layard. 1991. "Does long-term unemployment reduce a person's chance of a job? A time-series test." *Economica* pp. 93–106.

Kahneman, Daniel and Amos Tversky. 1979. "Prospect theory: An analysis of decision under risk." *Econometrica* pp. 263–291.

Kőszegi, Botond and Matthew Rabin. 2006. "A model of reference-dependent preferences." *The Quarterly Journal of Economics* pp. 1133–1165.

Lamprecht, Markus, Adrian Fischer and Hanspeter Stamm. 2015. *Sport Schweiz 2014: Factsheets Sportarten*. Magglingen: Bundesamt für Sport BASPO.

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

Mas, Alexandre. 2006. "Pay, Reference Points, and Police Performance." *The Quarterly Journal of Economics* 121(3):783–821.

Ockenfels, Axel, Dirk Sliwka and Peter Werner. 2014. "Bonus payments and reference point violations." *Management Science* 61(7):1496–1513.

Peterson, Christopher, Steven F Maier and Martin EP Seligman. 1993. *Learned helplessness: A theory for the age of personal control*. Oxford University Press, USA.

Pissarides, Christopher, Richard Layard and Martin Hellwig. 1986. "Unemployment and Vacancies in Britain." *Economic Policy* 1(3):499–541.

Pope, Devin G and Maurice E Schweitzer. 2011. "Is Tiger Woods loss averse? Persistent bias in the face of experience, competition, and high stakes." *The American Economic Review* 101(1):129–157.

Prendergast, Canice. 1999. "The provision of incentives in firms." *Journal of economic literature* 37(1):7–63.

Roberts, BW, JJ Jackson, JV Fayard, G Edmonds and J Meints. 2009. Conscientiousness. In *Handbook of Individual Differences in Social Behavior*, ed. M Leary and R Hoyle. New York: Guilford pp. 369–381.

Silverman, Bernard W. 1986. *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.

Song, Changcheng. 2015. "An Experiment on Reference Points and Expectations." *Available at SSRN 2580852* .

Thistlethwaite, Donald L and Donald T Campbell. 1960. "Regression-discontinuity analysis: An alternative to the ex post facto experiment." *Journal of Educational Psychology* 51(6):309.

Wooldridge, Jeffrey M. 2010. *Econometric analysis of cross section and panel data.* MIT press.