

**Raising the Bar: Causal evidence on gender differences
in risk-taking from a natural experiment**

by

René BÖHEIM

Mario LACKNER

Wilhelm WAGNER

Working Paper No. 2001

January 2020

Johannes Kepler University of Linz
Department of Economics
Altenberger Strasse 69
A-4040 Linz - Auhof, Austria
www.econ.jku.at

Corresponding author: mario.lackner@jku.at

Raising the Bar: Causal evidence on gender differences in risk-taking from a natural experiment*

René Böheim^{a,b}, Mario Lackner^{a,c}, and Wilhelm Wagner^a

^aJohannes Kepler University Linz (JKU)

^bAustrian Institute of Economic Research; IZA; CESifo

^cChristian Doppler Labor, Aging and the Labor Market

January 21, 2020

Abstract

We analyze data from top-tier professional athletes and find that female and male athletes differ in the timing and in the extent of their reactions to a change of the rules which increased the risk of failure. Male athletes increased risk-taking in the more risky environment immediately after the changes. Female athletes, however, increased risk-taking two years after the rule change. Over time, female athletes reverted to pre-reform risk-taking levels and male athletes' continued to make more risky decisions in the new environment. We attribute our findings to gender differences in competitiveness and risk preferences.

JEL Classification: J16, J44

Keywords: Competitiveness, risk-taking, gender differences

*Corresponding author: Mario Lackner, Department of Economics, Altenberger Str. 69, 4040 Linz; Tel.:+43/732/2468-7390; Email: Mario.Lackner@jku.at. We are grateful for funding by the Austrian National Bank, project number 16242. Thanks to Alexander Ahammer, Dominik Grübl, Simone HÄd'ckl, Rudolf Winter-Ebmer, Christine Zulehner, and participants of the NOeG 2017 and ESPE 2017 conferences for valuable comments.

1 Introduction

Standard economic theory explains gender differences by different preferences, distinct confidence in abilities, specialization or (statistical) discrimination. One potential explanation for gender differences in wages or representation in top-level jobs are differences in competitiveness and risk preferences. Most evidence on gender differences in competitiveness and risk-taking is derived from laboratory experiments ([Eckel and Grossman, 2008](#); [Croson and Gneezy, 2009](#); [Niederle and Vesterlund, 2011](#)), where the selection of participants into different payment schemes is used as a measure for competitiveness. These results show male participants as more competitive than female ones ([Niederle and Vesterlund, 2007](#); [Buser et al., 2014](#)). It is difficult to study differences in competitiveness in the field as field data typically lack information on the set of available options that people have, the associated risk-levels or their costs or benefits.¹ However, evidence from the field is important as results from experimental studies might be driven by an experiment's design ([Filippin and Crosetto, 2016](#)).

Data from sport competitions allow to study competitiveness and risk-taking decisions as sports have clear incentive structures, rewards are linked to performance, and performance is precisely monitored ([Böheim et al., 2017](#)). We use data from 91 major sport events as the Olympic Games or the World Championships, 1988 to 2012, to analyze the effect of a more risky environment on the risk-taking of female and male athletes. We use the reform of the pole vault regulations in 2003 to identify the causal effect of the more risky environment on risk-taking. The reform made it more difficult to clear any height with a non-optimal attempt. We follow [Böheim and Lackner \(2015\)](#) and interpret passing a height as a risky strategy. The causal effect is estimated using a difference-in-differences estimation where we compare the risk-taking of pole vault jumpers to the risk-taking of high jumpers who were not subject to a reform.

¹Data from financial markets, e.g. portfolio choices or pension fund strategies, provide an opportunity to study risk-taking ([Barber and Odean, 2001](#); [Dwyer et al., 2002](#); [Säve-Söderbergh, 2012](#)).

The reactions of athletes to the reform provide valuable information on risk-taking preference and competitiveness. The empirical evidence on gender differences in risk-taking is unambiguous. [Gerdes and Gränsmark \(2010\)](#), for example, analyze professional chess competitions and find that female players are more risk averse than male players. Using data from professional basketball, [Böheim et al. \(2016\)](#) find a significant gender gap in risk-taking in crucial game situations. [Böheim and Lackner \(2015\)](#) find that female athletes take less risk than male athletes in jumping competitions. Results from sports data on the gender gap in competitiveness, however, is mixed. [Frick \(2010\)](#) analyze long distance running competitions and find female contests to be less competitive. [Booth and Yamamura \(2018\)](#) confirm a gap in competitiveness and performance in speedboat races when women compete in mix-gender environments. [Pikos and Straub \(2019\)](#) find no gender gap in competitiveness.

In 2003, the International Association of Athletics Federation (IAAF) made several significant changes to the pole vault regulations. The main objective of the reform was to make competitions more attractive to watch. The new rules made the sport more difficult and changed the benefits of risk-taking. [Lobinger et al. \(2010\)](#) provide a qualitative analysis of twelve German pole vault athletes' risk preferences and their subjective evaluation of the reform. Overall, the athletes were concerned that the reform would lower their performance as the new rules made success more difficult. However, [Lobinger et al. \(2010\)](#) did not find a relation between an athlete's subjective assessment of risk and their objective performance. We do indeed find that the reform made the pole vault more difficult and the share of failed attempts increased significantly after the reform.

We estimate that, as a reaction to the reform, both male and female athletes increased risk-taking. The effect of the reform on male athletes is large and of similar magnitude in the 9 post-reform years. For female athletes, however, we find that risk-taking increased only two years after the reform and reverted to pre-reform levels five years after the reform. We estimate that increased risk-taking after the reform has no direct effect on the chances

of failing an attempt. This suggests that our results are not due to men’s overconfidence but are due to a significant gender gap in risk-taking or competitiveness.

2 Background

The pole vault and the high jump are two Olympic events with a long history and tradition. In high jump competitions, athletes have to jump over a bar at a certain height. After a short inrun, the athlete has to push herself by the use of one leg only over the bar. The pole vault is a similar sport with almost identical rules where athletes use a pole to propel themselves over the bar. Both sports require similar physical attributes and share a similar set of regulations and settings.²

An athlete has a maximum of three attempts to clear a height. If she jumps over the bar, the athlete advances to the next height. If the bar does not remain on the finger pads or the inrun time is exceeded, the attempt is counted as a failed attempt. Three consecutive fails lead to the elimination of the athlete. The athlete who clears the highest height wins. In case of a tie, the athlete with fewer attempts at the last height wins. If this does not break the tie, there will be a head-to-head jump-off competition. In case there are more than two competitors, this is repeated until there is an obvious ranking.

Jumping over the bar requires practice, talent, and energy, which are limited at the date of the competition. Each attempt depletes the available energy and athletes may pass a height to preserve effort or to reduce the risk from injury. In this case, the athlete sits out for one height. Athletes may pass multiple heights in a row, but they may not attempt lower heights once they declared a pass. Athletes may also pass heights strategically in order to have a low number of attempts, which could determine a tie-break.

Passing provides an athlete with recovery time, but for the final ranking in a competition it is the highest cleared height which counts. Each pass, however, invariably makes the next attempt more difficult since difficulty increases monotonously with height, a

²[Böheim and Lackner \(2015\)](#) provide additional details on competition formats.

higher height is never easier to clear than a lower height. We follow [Böheim and Lackner \(2015\)](#) and interpret passing a height as a risky strategy.

In 2003, the International Association of Athletics Federation (IAAF) introduced significant changes in the rules for pole vault competitions. The new rules specify shorter finger pads on which the bar is positioned, from 75mm to 55mm, and “Hungarian” bars, which have an oval rather than a square profile. The changes intended to improve safety standards through larger jumping mats and to increase the spectators’ value of the sport through a tightened procedure and a shorter inrun time ([Lobinger et al., 2010](#)).

The changes made by the 2003 reform significantly changed the consequences of touching the bar, making dislodgment more likely. Consequently, it is more difficult to clear a certain height with a badly executed attempt. As a consequence of the reform, we should see a larger fraction of attempts to result in failures. In addition, we also should see a smaller share of passes after the reform as passing any height is more difficult.

For the individual athlete, fewer overall attempts are more valuable after the reform. This is due to the increase in the relative number of fails caused by the more difficult environment. Consequently, competitive athletes will increase risk-taking in more risky situations, i.e., pass more often, as they will consider this to help them gain an advantage over their competitors. In contrast, less competitive athletes have a greater incentive to attempt the current height to avoid dropping out of the contest without clearing the current or the next, more difficult, height. This has the potential costs of additional failed attempts and spent energy.

Athletes who are more confident will pass more often than athletes who are less confident. As summarized by [Niederle and Vesterlund \(2011\)](#), women are typically found to be less confident than men. Consequently, following the existing evidence on gender differences, we expect to see any gender gap in risk-taking widen after the reform. This is consistent with [Lobinger et al. \(2010\)](#) who find that male pole vault competitors did in fact welcome the reform while female athletes were sceptical.

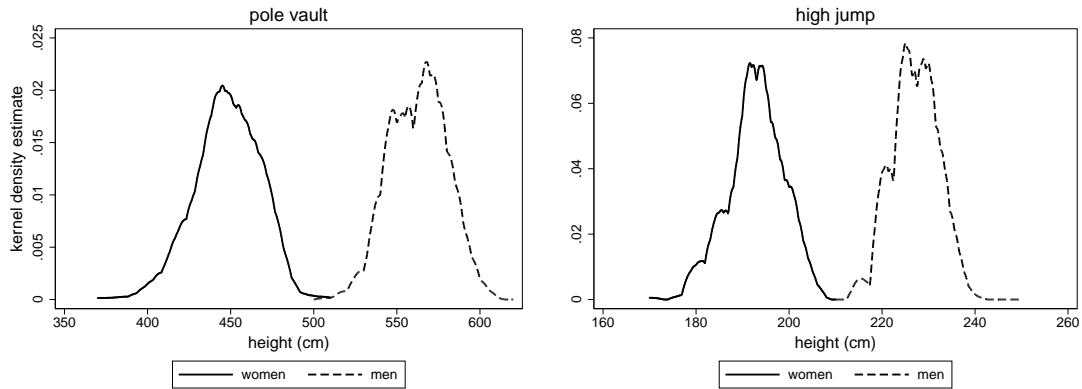
The trade-off between risk-taking and the difficulty of jumps has potential consequences for our analysis. Athletes who are more willing to increase their risk-taking to lower the number of attempts might have a lower share of failed attempts after the reform. In contrast, athletes who are less likely to increase risk-taking might be observed as having a greater share of failed attempts due to their overall greater number of attempts.

3 Data and Estimation Strategy

We use data from jumping competitions to identify the causal effect of the 2003 reform on risk-taking behavior. The data include 91 contests for women and 94 contests for men, collected from final rounds of major sport events as the Olympic Games, the World Championships (in- and outdoors), the European Championships (in- and outdoors), and IAAF Diamond League competitions, where both high jump and pole vault competitions were held. The data cover the time from 1997 through 2012 and contain information on 278 female and 330 male athletes. We analyze only attempts from the first to tenth height of competitions. The sample consist of 15,579 attempts which are either coded as a successful clearing of a height, a failed attempt or a pass.

Figure 1 shows the distributions of cleared heights for female and male athletes. The average cleared height (standard deviation) for female high jumpers is 194.2 cm (5.7) and 451.8 cm (18.8) for female pole vault athletes. Male high jumpers reach on average 228.0 cm (4.8) and pole vault athletes 566.7 cm (16.7). Male athletes perform on average better than female athletes due to differences in physical strength. Female and male athletes fail heights in similar patterns during high jump competitions. In pole vault competitions, male athletes are more likely to fail earlier than women. (Figure A. 1 in the Appendix presents survival graphs.) Below, we focus on the stage of the progression rather than the actual height (e.g., in centimeters).

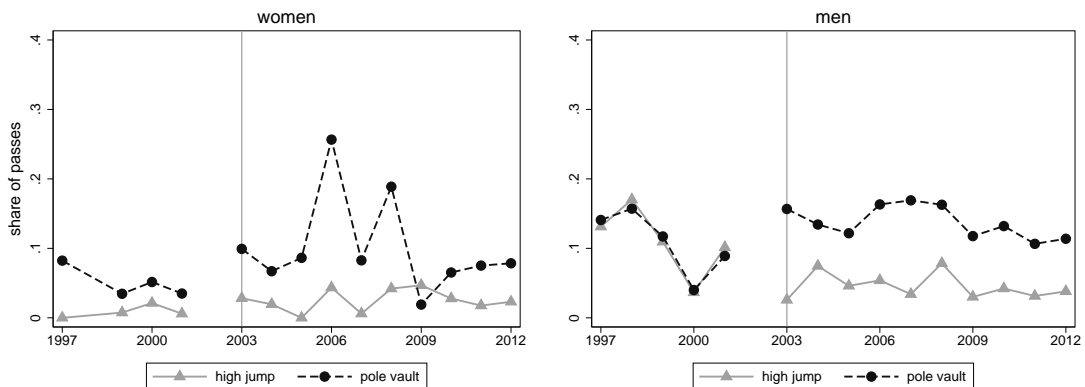
Figure 1: Kernel density estimates for recorded maximum height, by gender and discipline.



Note: Kernel density estimates for performances of female and male athletes. Only each athlete's best height by competition is taken into account. x-axis presents the height in cm (scales differ).

About half of the attempts or passes, 49.3 percent, are from female athletes. About 12.8 percent of all observations are pre-reform. Of all attempts, about 5.6 percent are passes (a total of 879 observations). First descriptive evidence for our data is presented in Figure 2. The graph shows that, for each gender, risk-taking levels in high jump and pole vault were almost identical before the reform. Consequently, we are confident that the crucial common-trend assumption is fulfilled. Risk taking in the pole vault increased after the reform. However, this increase was only temporary for female athletes, while male athletes appear to have permanently increased risk taking. Table A.1 presents more detailed descriptive statistics.

Figure 2: Risk-taking before and after the 2003 reform, by gender.



Note: share of passes (risky decisions) over all attempts for female and male athletes. Women: $N = 6,206$, men: $N = 6,412$.

We estimate the following difference-in-difference equation:

$$risk_{i,e,s} = \alpha_0 + \alpha_1 PV_{i,e} + \alpha_2 T_{i,e} + \alpha_3 (T_{i,e} * PV_{i,e}) + \mathbf{X}_i' \beta + e_{i,e,s}, \quad (1)$$

where the dependent variable, *risk*, is a binary indicator which is 1 if athlete *i* decided to pass at height *r* in competition *e*. It is 0, if the athlete decided to attempt stage *s* (of the height progression). *PV* is an indicator variable equal to one if the athlete *i* competed in the pole vault at event *e* and zero if she/he competed in the high jump. *T* is an indicator which is equal to one if the attempt happened in 2003 or later and 0 for pre-reform events. The coefficient of interest is α_3 as it measures the causal average treatment effect (ATE) of the 2003 reform on the risk-taking of pole vaulters.

\mathbf{X}_i is a vector of control variables, controlling for heterogeneity between athletes and events. Note that under the assumption of a quasi-experimental situation the inclusion of additional covariates does not change the estimate of the causal effect but serves to obtain more precisely estimated effects by potentially explaining more variation in the dependent variable. Standard errors are clustered on (individuals \times events) to correct for heteroscedasticity of unknown form.

Control variables include binary variables indicating the competition's stage during which the attempt was recorded. For example, we include 2 binary indicators to distinguish between the early and the later stages of a competition to take into account differential effects of the reform on risk taking at easier and more difficult stages. We also include indicators to account for effects of different types of competitions, because athletes might be motivated differently at relative more prestigious contests. We control for heterogeneity in the failed attempts using an indicator of past attempts as an athlete who passes after a successful attempt (at a lower height) might differ from an athlete who passes after two failed attempts.

4 Results

Table 1 presents the results for estimating equation (1) using different samples, stratified by gender and discipline.³ Using a full set of controls, we estimate an average treatment effect of an about 6.4 percentage points greater incidence of risk-taking due to the reform. Note that the mean of the dependent variable is about 7 percent, i.e., this is a large effect. The results of the specifications with and without controls do not differ qualitatively. If we estimate the equation separately for female and male athletes, we estimate an ATE of 3.5 percentage points for female and of 8.8 percentage points for male athletes. The difference between the point estimates for women and men is significant at the 10% level.

Table 1: Estimated effect of the reform on risk-taking.

	Pooled	Female	Male	Pooled	Female	Male
Treatment Effect ^a	0.0617*** (0.0128)	0.0287** (0.0127)	0.0894*** (0.0213)	0.0636*** (0.0121)	0.0347*** (0.0124)	0.0876*** (0.0199)
Difference ^b $b_{men} - b_{women}$			0.0607***			0.0529***
<i>N</i>	12,618	6,206	6,412	12,618	6,206	6,412
<i>R</i> ²	0.024	0.022	0.021	0.035	0.035	0.032
<i>Female</i>	x			x		
<i>Competition type</i>				x	x	x
<i>Stage</i> ^c				x	x	x
<i>2 failed attempts</i> ^d				x	x	x
Mean of dep. var.	0.0697	0.0514	0.0873	0.0697	0.0514	0.0873

Notes: The table presents the estimated treatment effects of the rule change on male and female pole vaulters. Standard errors are clustered on (individual athlete × competition). *, ** and *** indicate statistical significance at the 10, 5, and 1 percent level.

^a Coefficient α_3 of the interaction term between the post-reform period and treatment group indicators.

^b Difference in estimated coefficients tested with Fisher's permutation test for differences in estimates for different groups.

^c Binary variables indicating height progressions 5-7, and 8-10 respectively. Heights 1-4 are the reference group.

^d Binary variable equal to 1 if observed attempt is a potential drop-out with two failed attempts at the same height level.

To analyze the dynamic effects over time for different periods after the reform, we use an alternative specification of equation (1), where we replace the post-treatment indicator with several indicators to distinguish between immediate, short-term, and mid-term effects. We use different indicators for the years 2003 and 2004 (T_1), 2005 and 2006 (T_2),

³We cannot identify a pass as a risky decision before an athlete has either cleared or failed an attempt. We therefore cannot use the 2,961 attempts which were the first observed attempts.

2007–2009 (T_3), and 2010–2012 (T_4). The estimates are presented in Table 2. The results for this specification for the pooled sample suggest that athletes changed their risk-taking behavior immediately after the reform. This effect was still present in the period 2010–2012. When we split the sample by the gender of the athletes, we obtain a different pattern for men and women. Female athletes did not initially (T_1) increase risk taking. However, for period T_2 , we estimate a sizeable and significant increase of 12.1 percentage points. For period T_3 , this effect is estimated at 4.2 percentage points, and we do not obtain a significant effect for the years 2010–2012 (T_4). Female athletes reacted strongly to the reform, but with a delay and only temporarily. Over time, their risk-taking reverted to pre-reform levels.

In contrast, male athletes increased their risk-taking immediately after the reform (in T_1) and this increase is found in all periods after the reform. Their average reaction to the reform is stronger and more permanent than that of female athletes.

Table 2: Estimated treatment effects for different post-treatment periods.

	All	Female	Male
T_1	0.0612*** (0.0189)	0.0247 (0.0220)	0.0941*** (0.0299)
T_2	0.1157*** (0.0223)	0.1214*** (0.0290)	0.0935*** (0.0309)
T_3	0.0727*** (0.0176)	0.0415* (0.0228)	0.1022*** (0.0264)
T_4	0.0542*** (0.0125)	0.0192 (0.0128)	0.0816*** (0.0206)
<i>Female</i>	x		
<i>Competition type</i>	x	x	x
<i>Stage</i>	x	x	x
<i>2 failed attempts</i>	x	x	x
N	12,618	6,206	6,412
R^2	0.037	0.042	0.033

Notes: The table presents estimated treatment effects for the post-treatment period for the analyzed groups. Standard errors are clustered on (individual athlete \times competition). *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level. Coefficients $T_1 - T_4$ are the coefficients for the interaction terms between different post-reform periods and treatment indicators. T_1 indicates the years 2003 and 2004, T_2 corresponds to 2005 and 2006, T_3 corresponds to 2007–2009, and T_4 corresponds to 2010–2012. Full set of controls use as for cols. 4-6 of Table 1

Performance. In the first part of our empirical analysis, we presented evidence in favor of an increase in risk-taking due to the 2003 reform. However, we know little about the effects of the reform — and the resulting increase in risk-taking — on overall success. The empirical literature on gender differences in risk-taking or competitiveness typically raises interest in the consequences of gender differences in preferences or beliefs. For example, [Barber and Odean \(2001\)](#) stresses that excessive risk-taking by men harms their overall performance. Also, [Niederle and Vesterlund \(2007\)](#) concludes that men compete too much. One problem of studies that use field data to study this relationship is the difficulty of establishing a link between risk-taking decisions and subsequent success. Our setting provides an ideal opportunity to analyze the consequences of risk-taking on success.

We demonstrate this by an analysis of the athletes' performances, as we can distinguish between failed and successful attempts. [Figure A. 2](#) in the Appendix illustrates the evolution of all failed attempts over time for high jump and pole vault competitions and both genders. The focus is a potential change in the failure rate after the 2003 reform. Over time, men and women had similar failure rates.

To further investigate performance, we estimate a modified version of equation (1) where we use a binary dependent variable which has a value of 1 if the observed attempt is a fail, 0 if it is a success. We include an interaction term of the treatment effect and a dummy variable indicating attempts directly following the risky strategy of a pass:

$$\begin{aligned} fail_{i,e,s} = & \alpha_0 + \alpha_1 PV_{i,e} + \alpha_2 T_{i,e} + \alpha_3 (T_{i,e} * PV_{i,e}) + \\ & + \alpha_4 after\ risk_{i,e} + \alpha_5 (T_{i,e} * PV_{i,e}) * after\ risk_{i,e} + \mathbf{X}_i' \beta + e_{i,e,s}, \end{aligned} \quad (2)$$

Results are presented in [Table 3](#). We estimate a general increase in the probability of a failed attempt after the reform in the pooled sample by 5.5 percentage points. If we stratify the analysis by gender, we find that female athletes had a greater likelihood of a fail after the reform, while it did not increase for male athletes. This is consistent with

an interpretation that female athletes, who did not increase risk-taking after the reform, experienced more fails. In contrast, male athletes, who did increase risk-taking, saw no increase in the relative share of failed attempts.

We estimate that men in general seem to perform worse on attempts directly following a risky pass, which can be interpreted as excessive risk-taking or overconfidence (Moore and Healy, 2008; Grieco and Hogarth, 2009; Fellner and Krügel, 2012; Krawczyk, 2012). This is not the case for female athletes.

The coefficient on the interaction between the treatment variable ($T_{i,e} * PV_{i,e}$) and a binary indicator ($after\ risk_{i,e}$) equal to 1 if the attempt happened after a risky pass (0 else) is insignificant. This suggests that neither female nor male athletes experience any performance decrements after changes in risk-taking due to the reform. Thus, male athletes, who increase risk taking as a consequence of the reform, were not overconfident. (If athletes were overconfident, this coefficient would be positive.)

Table 3: Effect of the reform on performance.

	Pooled	Women	Men
after risky	0.064** (0.025)	-0.038 (0.049)	0.105*** (0.029)
Treatment effect	0.055** (0.022)	0.065** (0.033)	0.040 (0.028)
Treatment effect × after risk	0.020 (0.032)	0.054 (0.060)	0.020 (0.037)
<i>Female</i>	x		
<i>Competition type</i>	x	x	x
<i>Stage</i>	x	x	x
<i>2 failed attempts</i>	x	x	x
Mean of dep. var.	0.5576	0.5316	0.5836
<i>N</i>	14,700	7,366	7,334
<i>R</i> ²	0.101	0.107	0.094

Note: The table presents treatment effects for the post treatment periods (see text for explanation) for the analyzed groups - including an additional treatment effect for post-risk taking attempts. Dependent variable is equal to 1 if observed attempt is a fail, 0 else. Only actual attempts were included, all passes were excluded from the sample. All attempts at the beginning of competitions where passing cannot unambiguously be identified as a strategy strategy were also included. Standard errors are clustered for individual. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. Full set of controls use as for cols. 4-6 of Table 1

5 Conclusion

We analyze the reactions of top-level athletes in high-stakes competitions to an exogenous shock in the incentive structure. A rule change for the pole vault in 2003 made clearing any height more difficult. We interpret a pass as a risky strategy and estimate the reactions of male and female athletes to this change. As the pole vault and the high jump are almost identical in terms of formats and the set of possible strategies, we use attempts in height jump competitions as control observations for attempts in pole vault competitions.

The 2003 reform influenced the incentives for choosing risky strategies. In particular, the probability of fails increased, which increased the incentive to reduce risk-taking. In contrast, the incentive to lower the number of attempts increased, making risk-taking more valuable. This is due to the fact that a slightly flawed attempt will be more likely to result in a fail after the reform. Consequently, athletes feel a more pronounced trade-off between the greater difficulty of an attempt and the incentive to lower the number of attempts by passing on certain heights.

Using a difference-in-differences approach, we find that male professional athletes changed their risk-taking behavior. After the reform, male pole vault athletes increased risk-taking and continue to take more risk in the following years. We estimate that the probability of a fail after a pass did not change for male athletes after the reform. This suggests that although male athletes overall increased risk-taking, their risk-taking is not due to male overconfidence.

In contrast, female athletes increased risk-taking after the reform only for a short period, starting about two years after the reform. After this temporary increase, risk-taking returned to pre-reform levels. We estimate that the reform resulted in a greater probability of a failed attempt for female athletes. We interpret this last result as the consequence of female athletes not avoiding the more difficult attempts by passing more often. The probability of a failed attempt did not change for male athletes who did increase risk-taking after the reform.

Our results have strong implications for a multitude of different settings. In a corporate environment, we might expect that male employees will permanently increase risk-taking if associated incentive structures for risk-taking are changed. In an investment firm, for example, male traders could permanently increase risk-taking if their compensation is linked to their performance relative to co-workers. According to our results, we do not expect such a change from female traders.

An unobserved factor which could be important for the athletes's decisions is the gender of the coach. Coaches could have a crucial influence on the risk-taking decision. As most coaches in the pole vault are men, the estimated effect for female athletes is an upper bound. Controlling for the coach's sex should lead to lower estimates if risk and the coaches's sex are positively correlated.

References

- Barber, Brad M. and Terrance Odean**, “Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment,” *The Quarterly Journal of Economics*, 2001, *116* (1), 261–292.
- Böheim, René and Mario Lackner**, “Gender and Risk Taking: Evidence from Jumping Competitions,” *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2015, *178* (4), 883–902.
- , **Christoph Freudenthaler**, and **Mario Lackner**, “Gender Differences in Risk-Taking: Evidence from Professional Basketball,” *IZA discussion paper 10011*, 2016.
- , **Dominik Grübl**, and **Mario Lackner**, “Gender Differences in Competitiveness,” *ifo DICE Report*, 2017, *15* (2), 13–17.
- Booth, Alison and Eiji Yamamura**, “Performance in mixed-sex and single-sex competitions: What we can learn from speedboat races in Japan,” *Review of Economics and Statistics*, 2018, *100* (4), 581–593.
- Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek**, “Gender, competitiveness, and career choices,” *The Quarterly Journal of Economics*, 2014, *129* (3), 1409–1447.
- Croson, Rachel and Uri Gneezy**, “Gender Differences in Preferences,” *Journal of Economic Literature*, 2009, *47* (2), 448–474.
- Dwyer, Peggy D., James H. Gilkeson, and John A. List**, “Gender Differences in Revealed Risk Taking: Evidence from Mutual Fund Investors,” *Economics Letters*, 2002, *76* (2), 151 – 158.
- Eckel, Catherine C. and Philip J. Grossman**, “Men, Women and Risk Aversion: Experimental Evidence,” *Handbook of Experimental Economics Results*, 2008, *1*, 1061–1073.
- Fellner, Gerlinde and Sebastian Krügel**, “Judgmental overconfidence: Three measures, one bias?,” *Journal of Economic Psychology*, 2012, *33* (1), 142 – 154.
- Filippin, Antonio and Paolo Crosetto**, “A Reconsideration of Gender Differences in Risk Attitudes,” *Management Science*, 2016, *62* (11), 3138–3160.
- Frick, Bernd**, “Gender Differences in Competitiveness: Empirical Evidence from Professional Distance Running,” *Labour Economics*, 2010, *18* (3), 389–398.

- Gerdes, Christer and Patrik Gränsmark**, “Strategic Behavior across Gender: A Comparison of Female and Male Expert Chess Players,” *Labour Economics*, 2010.
- Grieco, Daniela and Robin M. Hogarth**, “Overconfidence in absolute and relative performance: The regression hypothesis and Bayesian updating,” *Journal of Economic Psychology*, 2009, *30* (5), 756 – 771.
- Krawczyk, Michał**, “Incentives and timing in relative performance judgments: A field experiment,” *Journal of Economic Psychology*, 2012, *33* (6), 1240 – 1246.
- Lobinger, Babet, Tanja Hohmann, and Andreas Nicklisch**, “Analyse subjektiver und objektiver Auswirkungen von Regeländerungen im Stabhochsprung [Analysis of the subjective and objective effects of rule changes in pole-vaulting],” *Zeitschrift für Sportpsychologie*, 2010, *17* (1), 12–20.
- Moore, Don A. and Paul J. Healy**, “The Trouble with Overconfidence,” *Psychological review*, 2008, *115* (2), 502.
- Niederle, Muriel and Lise Vesterlund**, “Do women shy away from competition? Do men compete too much?,” *The Quarterly Journal of Economics*, 2007, *122* (3), 1067–1101.
- and –, “Gender and Competition,” *Annual Review of Economics*, 9 2011, *3*, 601–630.
- Pikos, Anna Katharina and Alexander Straub**, “Mind the Absent Gap: Gender-Specific Competitive Behavior in Nonprofessional Sports,” *Journal of Sports Economics*, 2019, *forthcoming*.
- Säve-Söderbergh, Jenny**, “Self-Directed Pensions: Gender, Risk, and Portfolio Choices,” *The Scandinavian Journal of Economics*, 2012, *114* (3), 705–728.

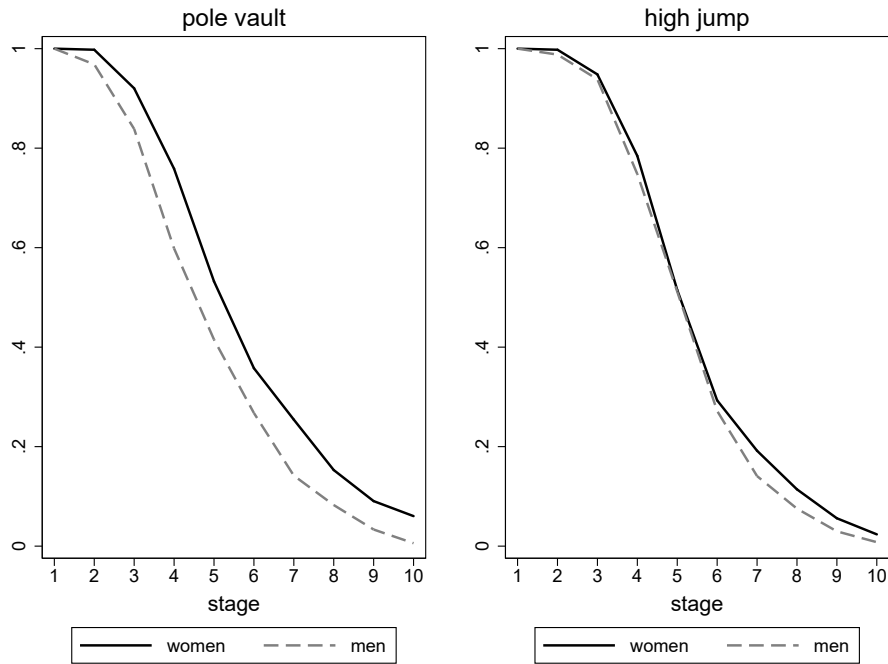
A Appendix

Table A.1: Observed actions by sport, gender, and period.

	High Jump			Pole Vault		
	<i>Pre-reform</i>	<i>Post-reform</i>	<i>Difference</i>	<i>Pre-reform</i>	<i>Post-reform</i>	<i>Difference</i>
Women						
Pass	0.7	2.5	1.8	4.4	9.1	4.7
Clear	37.8	38.3	0.5	37.6	30.3	-7.3
Fail	61.5	59.2	-2.3	58.0	60.6	2.6
<i>N</i>	553	2,592	-	585	2,447	-
Men						
Pass	11.3	4.2	-7.1	11.0	12.8	1.8
Clear	27.1	35.5	8.4	27.6	24.6	-3.0
Fail	60.6	60.3	-0.3	61.4	62.6	1.2
<i>N</i>	611	2,811	-	544	2,446	-

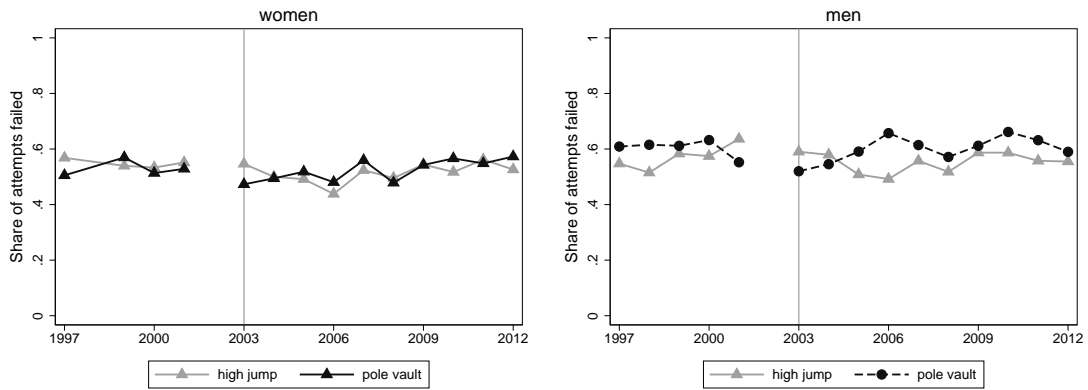
Note: Percentage shares of all observed actions by gender before and after the reform. *Clear* indicates a successful attempt. A *fail* denotes any unsuccessful attempt. A *pass* denotes a forgone attempt. Observations correspond to all observations used for Tables 1 and 2.

Figure A. 1: Survival graphs by gender and discipline.



Note: Survival graphs of high jump and pole vault competition, plotting the share of athletes remaining in the competition after each height. Stratified into competition formats for specialized female athletes, specialized male athletes, and multi-competition male athletes.

Figure A. 2: Fails after passes before and after the 2003 reform.



Notes: Share of fails among all attempts, conditional on following a pass, $N = 14,700$. The circles indicate high jump events, the triangles indicate pole vault events.