The Superstar Effect in Gymnastics

Leah Meissner
Seton Hall University, Former Student

Aishwarya Rai
Seton Hall University, Former Student

Kurt W. Rotthoff[1]
Seton Hall University, Professor

Fall 2020

Abstract
Tournament structures provide a unique opportunity to assess the factors that affect an agent's decision-making. In this paper, we seek to determine the peer-effects of Simone Biles' "superstar" presence on her fellow competitors, both before and after her rise to dominance in 2013. Specifically, with our unique data, we can utilize both meet-level and athlete-level fixed effects on the athletes' performances. We find that the gymnasts attempted more difficult routines in Biles' weakest events. There is also evidence of support for the idea that risk taking changes when a superstar is present, showing how athletes change their approach to big events changes with and without a superstar. Lastly, we find that the standard for perfection in the sport has changed over time, which could be related to forms of judging bias.

Keywords: Superstar Effect, Judging Bias, Risk Strategies, Tournaments

JEL: D4, D81, Z2

---

[1] Kurt W Rotthoff: Seton Hall University's Stillman School of Business, 400 South Orange Ave, JH 674, South Orange, NJ 07920. He can be reached at Kurt.Rotthoff@shu.edu or Rotthoff@gmail.com. Thank you to the participants at the 2018 Southern Economic Association's annual meetings. Any mistakes are our own.

I. Introduction

There have been many recent studies pertaining to the use of tournaments in various settings as a way of compensating employees and maximizing productivity. Pioneer literature on this topic can be traced back to Lazear and Rosen (1981), which found that the tournament structures involving rank-order payments for contestants have the capability to induce greater worker effort than other forms of compensation schemes. The study of tournaments has since been extended to investigate effort, compensation, and risk in multiple areas.

In order to measure the impact on employees' effort and risk taking responses to tournaments, it is useful to understand the role of tournament structures on performance in individual sports. Specifically, the related literature has focused on professional golf. Mainly analyzing the superstar effect of Tiger Woods (Brown, 2011 and McFall and Rotthoff, 2020) and peer effects from putting in Hickman and Metz (2018). The findings from these studies provide insight into how the structure of tournaments impacts optimizing or hampering a player's execution on the golf course. This, in turn, helps us understand how these incentives impact the workplace, which is often far more difficult to measure properly. This allows us to generalize these results to see how peer effects and superstars impact employee incentives. While Brown claimed that having a superstar present in a tournament lead to less effort by participants, McFall and Rotthoff show evidence that it is not a lack of effort, but rather that those involved in a tournament with a superstar are increasing their risk levels (as their only chance to beat the superstar).

The ability to further understand a superstar's impact on effort and risk in a tournament (and thus the effectiveness of implementing tournament structures in the corporate setting), we expand these studies by utilizing a unique dataset of elite level gymnastics. Specifically, we focus on the rise to complete dominance of Simone Biles. We use data from elite level gymnastics competitions from 2011-2016 to investigate a possible superstar effect on individual performance due to the presence of Simone Biles.

To properly analyze these, we use data from both before and after Simone Biles became a dominant force on the competitive gymnastics scene to detect any possible trends or changes in the odds of a penalty and both the difficulty and execution scores of the other gymnasts. Prior to the World Artistic Gymnastics Championships in 2013, Simone was a powerful, junior-level gymnast, but inconsistent. She often fell on one (or more) events, although the events she hit – she hit well. At the 2013 World Championships, everything clicked (late in 2013). That was the first meet she completed without falter, and from that point on she was able to control her power and become the most dominant athlete the sport has ever seen.

Once the results of the superstar effect in tournaments is more clear, then the impact of superstars in tournaments in the workplace will also become more clear, clarifying our understanding of workplace behavior peer-effects amongst employees and helping inform managerial decision-making.

II. Gymnastics

In women's gymnastics, all-around meets feature four distinct events: vault, balance beam, uneven bars, and a floor routine. For each event, gymnasts receive a score for the difficulty of the routine they choose and another score based on their execution of that routine, which then are added together to get the individual's overall score for that particular performance (this separation of execution and difficulty occurred in 2006, after a controversy in the 2004 Olympics. See Morgan and Rotthoff, 2014, for a discussion of this change). Hence, each gymnast receives eight scores, a difficulty and execution score for each of the four events, which are summed up to find the gymnast's total score for the meet.

Additionally, gymnasts can also lose points for various neutral errors such as falling or going out of bounds. Thus, each event requires calculated decisions on which level of difficulty to attempt based on the individual's strengths and weaknesses. A gymnast may attempt a more difficult routine in hopes of earning a higher difficulty score for, say, the floor exercise if she feels confident in her abilities for that event or if she feels the risk is worth it. This translates into a measure of "riskiness" per routine, as higher difficulty correlates to more inherent risk. This provides a unique environment to study not only the role of tournament structure in gymnastics but also whether or not there is evidence of a peer effect on competitors, as found in both Brown (2011) and McFall and Rotthoff (2020).

Simone Biles began her reign as the dominant force on the international competitive gymnastics scene in 2013. First, Biles took home the gold medal in the all-around competition at the P&G Championships, followed by an impressive first-place finish at World Artistic Gymnastic Championships in Antwerp, Belgium, winning the Worlds by almost 0.900 points (for context, in the prior year's Olympics, the gap between first and third was 0.675). Before the World Championships that year, Biles had performed well in many competitions, but her success was inconsistent, and she was prone to making errors. But, after the 2013 season, Biles began to outperform her peers – and in a big way.

Gymnastics scores are typically very precise, as sometimes gymnasts can be separated by 0.001 of a point; for example, in the 2016 Olympics, Jessica Brizeida Lopez Arocha (Venezuela) placed 7th with a final all-around score of 57.966, while Asuka Teramoto (Japan) placed 8th with a final all-around score of 57.965. Hence, part of what has made Biles so extraordinary is not just the fact that she has won so many competitions since her acceleration to the top, but she also has a lengthy track record of winning over other gymnasts by significant margins, sometimes even 2.000 or more points.[2] In the 2019 World Artistic Gymnastics Championships, Biles' all-around score was 58.999, placing her at first, while Tang Xijing scored 56.899, placing her at second (with second through fifth all scoring in the 56s). This kind of dominance is virtually unprecedented in gymnastics, making Biles a prime candidate for analyzing the possibility of a superstar effect. She brings an unmatched level of execution to her sport, much like Tiger Woods to golf, which uniquely lends itself to studying if other gymnasts are positively or negatively impacted by her presence at a meet.

Since gymnastics scoring is so sensitive to minor changes in performance, each individual's decisions regarding the difficulty of their routines play a critical role in determining their overall score. Thus, the athlete's decision to take on a more conservative (risky) routine will be measurable. The more conservative (risky) approach means the athlete is more (less) likely to easily land that particular routine, as each athlete has many different skills that can be thrown during a given meet. The harder the routine attempted, the higher the cost (more likely not to hit the routine) of attempting that skill. Given the athletes' elite skill sets, these decisions are tailored on a case by case, meet by meet basis.[3] Thus, the athlete analyzes the situation they are in and can change their routine to match that particular situation – such as if there is a superstar present.

[2] Historically victories are won by fractions of a point. In the 2008 Olympics, U.S. gymnasts scored first and second. Nastia Liukin won with a score of 63.325, Shawn Johnson finished second with a score of 62.725, with China's Yang Yilin finishing third with a score of 62.650.

[3] It is important to note that the athletes' going to a given competition know who will be at the meet, as their countries have to register their athletes for the meet beforehand. The countries have to submit a list to the federation months in advance, but the country can change the list at any point up to the final submission. The U.S. is notorious for naming their official teams very late in the process, which is still (typically) at least a month before the event. Thus, barring an injury, they know who they will be competing with before an event begins.

But, given this is a competition, competitors also must respond to other gymnasts' choices. Brown (2011) finds that competitors (golfers) are in fact impacted by a superstar's presence at a tournament and will adjust their strategy accordingly. Brown's results indicate that this effect is negative, which translates to a decreased effort from the other golfers when Tiger Woods is present at a particular tournament. On the other hand, McFall and Rotthoff (2020) assess risk-taking changes on par-five holes and find that golfers are more likely to adopt riskier strategies at tournaments where Tiger Woods is present. They agree with Brown's general finding that golfers do worse with the superstar present but find that this worse performance is caused by golfers' taking excessive risks. Analogously, competitive gymnasts must take into account their ideal level of risk, as well as the amount of risk a superstar like Simone Biles is incorporating into her routines on any given day to optimize their performance (or give them a chance to compete with the superstar that is present). Thus, there are clearly high stakes involved when choosing whether to attempt a more challenging (or easier) routine for a particular event, as it could result in gaining (or losing) crucial points.

The insights gained from assessing difficulty scores in gymnastics provide benefits in other settings. They provide a way to model and understand the decision-making processes people face daily. Individual risk-taking is inherently a function of judgments about our own capabilities and a response to those around us' perceived abilities. Each decision one faces as to what level of risk to take on over a lifetime, in both personal and professional settings, incorporates one's judgment heuristics, which in turn shapes effort levels and can have high monetary consequences based on whether the individual has either a poor or exemplary performance. Thus, it is clear that workers and athletes alike adjust their effort and strategies based on the structure of compensation and the effort of those around them.

Risks matter outside of sports as well. Chevalier and Ellison (1997) look at how investment funds adjust their risk-taking strategies to increase performance, while Aseff and Santos (2005) and Core and Guay (2001) look at the effectiveness of stock options to compensate employees in the firm. They all show that employees respond to present and future financial incentives by taking on more risk as the end of a tournament draws near. This is comparable to gymnasts attempting more difficult (risky) routines and helps us understand, in general, how employees handle the impact of superstars. More specifically, how they adjust their riskiness with the presence of superstars in the firm or industry.

III. Data

For this paper, we use data from the U.S.A Gymnastics website for women's artistic gymnastics competitions, including results from nearly 50 all-around meets that took place domestically and abroad between 2011 and 2016. This includes meets with and without Simone Biles and include when she was competing as a high-level junior and then at the elite level. Our data set is meant to reflect a relatively balanced ratio of meets in which Simone was and was not present both prior to and after her rise to dominance in the gymnastics world in 2013.

Each of the competitions has data for each athlete's performance on four separate events: vault, uneven parallel bars, balance beam, and the floor routine. There is a score breakdown for each of these events for the execution and the difficulty of that routine. The difficulty score is composed of three unique aspects. These criteria are: the number of different stunts attempted by the gymnast, whether or not the competitor implemented more advanced skills that are connected to one another (a connection value), and a compositional requirement based on primary elements that should be incorporated into every routine for a specific apparatus. Furthermore, the execution score is a value of how well the gymnast performed her chosen stunts during the event, which is scored starting at 10 points, with score reductions for mistakes in the overall technique.

Thus, the total score for an individual event (e.g., floor) is the sum of the difficulty score and the execution score minus any further point deductions for neutral errors, such as stepping out of bounds. In turn, these four individual event scores are summed together for an athlete's overall score in a given competition. Accordingly, this scoring method is beneficial for analyzing changes in the overall scores of gymnasts and changes by event. The difficulty measure associated with the four different event scores allows us to gain insight into potential variations in the level of riskiness of skills attempted by the other athletes, prompted by Simone's presence or absence at a meet.[4] Hence, if competitive gymnasts tried a greater number of more difficult stunts and/or incorporated more intricate elements into their performances over time, this could be interpreted as a peer effect induced by Simone, leading to athletes taking on more risk than would be the case otherwise.

IV. Methodology and Results

To analyze if other athletes were taking on more risk when Simone was present is to look at the probability that a penalty is given during a specific event,

$$\text{Penalty in Event } (0, 1) = \alpha + \beta \text{ (Simone)} + \varepsilon$$

In addition to the probability of the penalty, we can see if the dummy variable on Simone being present at a given meet is statistically significant. This is particularly important if it is statistically significant before and after she was a superstar. We use the data from 2011 to 2016, which we split for years before her superstardom (2011-2013) and during her dominance (2014-2016). To

[4] Difficulty score is determined during judging by the A-Panel (Two-judge panel), based on the *Code of Points* (points correspond to a skill); they then reach a consensus about the difficulty score.

help control for variations over time we also include the time fixed effects and cluster the standard errors at the meet level.

We exclude Simone from the data (as we are testing her impact on others), and since we do not have the order data for these competitions, we cannot match the order in which the athletes competed in in all these competitions either (thus, we cannot test for the rank order bias studied in Rotthoff, 2015). In Table 1, we first analyze the probability of an athlete receiving a penalty before and after Biles' rise to dominance with meet level and time fixed effects (and standard errors clustered at the meet level). When looking at the probability of penalties, there are negative and significant effects both before and after her dominance. With a significantly lower probability of penalty on the vault after her rise to dominance on the vault, bars, and floor (her two best events are the vault and floor), the decrease in penalties on the beam is only less after her dominance in that specific event.

[Table 1]

It is also clear in both the difficulty and execution scores that her being at a meet after her dominance caused other athletes to get lower difficulty and execution scores on her two best events, vault and floor, while getting better scores on her worst events, bars and beam. In the 2016 Olympics, Simone Biles won the all-around gold, but also the individual gold in vault and floor; she got the bronze medal on the beam and did not medal on the bars (historically her worst event). This shows some form of superstar effects where athletes are not only changing how they perform with and without her, but also differently after her rise to dominance.

[Table 2]

To increase accuracy, we also have the ability to add athlete level fixed effects in Table 2. This measure lets us look at how an individual athlete changes their performance statistics when Simone is present during all the years before and after her rise to dominance. Looking at the probability of a penalty, we find no effect on the odds of an athlete receiving a penalty before or after her dominance. However, when looking at the impact on the athletes' difficulty and execution scores, we find that other athletes were trying significantly less difficult routines (and receiving significantly lower execution scores) on the vault after her dominance. While athletes were trying less difficult floor routines before her rise (and simultaneously earning lower execution scores), their performance scores improved after her rise.

[Table 3]

Our last measure adds an interaction term of the Simone impact interacted with the biggest meets in our dataset (World Cup competitions and Olympic Games and trials). This interaction will show if the changes that we observe are different at the biggest competitions relative to the smaller competitions (but all are big enough to get the most talented gymnasts to attend the meet). Before her superstardom, the odds of a penalty were lower in Simone's best events (vault and floor), but those decrease in odds of penalty go away after she becomes a superstar – aligning precisely with the findings in McFall and Rotthoff (2020), that the odds of risk are changing when a superstar is present relative to what they would have done in the biggest meets without a superstar. These athletes would have done routines that they could do with fewer penalties in the biggest competitions without a superstar, but they continue to do routines that induce penalties with her present.

When looking at the difficulty and execution scores, we find that after her dominance other athletes did less difficult (and lowered their execution scores) routines on the vault and floor (her best events) when she was present, but they did less difficult routines at the biggest meets. In contrast, there was no effect of her presence, in big or non-big meets, before her rise to superstardom.

V. Conclusion

Simone Biles has come to dominate the sport of gymnastics in a way never seen before. McFall and Rotthoff (2020) also find a superstar effect utilizing golf data, finding that players took on additional risks when Tiger Woods was present (during his domination, but not after). When looking at competitions in which superstar Simone competed, we generally find that other gymnasts tended to do simpler routines but performed them worse – shown by the drop in the difficulty and execution scores after Simone Biles became a superstar presence at the end of 2013.

There are two possible outcomes here: (1) The other athletes performed less difficult vaults and did worse on them, or (2) the expectations of what constitutes the highest scores in the sport have changed. We think that it is more likely the latter of these explanations and the general drop in scores for difficulty and execution may be more attributable to a judging bias (or more specifically, the judges redefining the standard of what is needed for a given score).

This leaves the explanation of the judge's perception of what constitutes a "perfect score" for future research. It is plausible that after Simone Biles' rise to dominance in 2013, the concept of "perfection" in gymnastics was readdressed to account for this new level of performance. For instance, Nadia Comaneci earned a perfect 10 in the 1976 Montreal Olympics on the uneven bars. Assessing her performance relative to today's athletes and the standards they are held to, would not be near a perfect 10.

We also find very similar results to McFall and Rotthoff (2020) when interacting the superstar effect with the biggest meets in our data. Athletes responded differently in the biggest competitions before her dominance. However, after her dominance, they continued to act the same at the biggest competitions. This could be driven by the fact that they used to try things in the less-big competitions, but only do what they were good at in the biggest competitions. However, after superstar Simone was present, they continued to try these routines into the biggest meets because it was their only chance to beat her. This change after her dominance, relative to before, confirm the findings in the McFall and Rotthoff paper – that risk levels change in the presence of superstars.

Understanding how people respond to incentives in tournaments has impacts in many employment areas outside of sports. This study also adds to what needs to be considered in those structures. Specifically, if there is a superstar, how the other employees will shift their focus from the specific task the superstar is good at to other tasks – which may, or may not, be the desired outcome of the firm.

If, as proposed, a judging bias exists in such tournament settings, it is also worthwhile considering how much of the superstar effect is determined by the judges, not just the athletes (workers). In a workplace, this may translate into how employees' performances are judged based on a superstar-peer's performance, not so much their work. If you are in a department with a superstar researcher, are you held to a different standard (even at the same school with the same teaching load) then you would have been if they were not present?

Works Cited

Aseff, J. G., & Santos, M. S. (2005) "Stock Options and Managerial Optimal Contracts" *Economic Theory*, Volume 26, No. 4, 813–837.

Brown, J. (2011). "Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars" *Journal of Political Economy*, Volume 119, No. 5, 982–1013.

Chevalier, J., & Ellison, G. (1997) "Risk Taking by Mutual Funds as a Response to Incentives" *Journal of Political Economy*, Volume 105, No. 6, 1167–1200.

Core, J., & Guay, W. (2001) "Stock Option Plans for Non-Executive Employees" *Journal of Financial Economics*, Volume 61, No. 2, 253–287.

Groothuis, Peter A., Kurt W. Rotthoff, and Mark Strazicich (2017) "Structural Breaks in the Game: The Case of Major League Baseball" *Journal of Sports Economics* Volume 18, Issue 6, pp. 622-637

Hickman, D. C., Kerr, C., & Metz, N. (2019). "Rank and Performance in Dynamic Tournaments: Evidence from the PGA Tour" *Journal of Sports Economics*, Volume 20, No. 4, 509–534.

Lazear, E. P., & Rosen, S. (1981). "Rank-order Tournaments as Optimum Labor Contracts" *Journal of Political Economy*, Volume 89, No. 5, 841–864.

McFall, T., & Rotthoff, K. W. (2020) "Risk Taking Dynamics in Tournaments: Evidence from Professional Golf" Journal of Economic Behavior and Organization Vol. 129, November, 378-394

McFall, Todd, Charles Knoeber, and Walter Therman (2009) "Contests, Grand Prizes, and the Hot Hand," *Journal of Sports Economics*, Volume 10, No.3, 236-255

Morgan, H. N., & Rotthoff, K. W. (2014) “The Harder the Task, the Higher the Score: Findings of a Difficulty Bias” *Economic Inquiry*, Volume 52, No. 3, 1014–1026.

Rotthoff, K. W. (2015) “(Not Finding a) Sequential Order Bias in Elite Level Gymnastics: Sequential Order Bias” *Southern Economic Journal*, Volume 81, No. 3, 724–741.

Table 1: Estimation of Simone impact with Meet Level and Year fixed effect. (clustered by meet)

VARIABLES	Vault	Bars	Beam	Floor	
		Panel A: Probability of Penalty			
Simone	-0.35736***	-0.04389***	-0.17398***	-0.41307***	All Data
	(0.000)	(0.000)	(0.000)	(0.000)	
N	804	99	520	918	
Simone	-0.05677***	0.00595***	-0.17571***	-0.12491***	2011-2013
	(0.000)	(0.000)	(0.000)	(0.000)	
N	426	52	333	502	
Simone	-0.42704***	-0.05848***	-0.04349***	-0.46549***	2014-2016
	(0.000)	(0.000)	(0.000)	(0.000)	
N	378	47	187	416	
		Panel B: Difficulty Score			
Simone	0.09172***	0.14566***	-0.01690***	0.29419***	All Data
	(0.000)	(0.000)	(0.000)	(0.000)	
N	671	454	444	663	
Simone	0.07338***	-0.04290***	-0.05567***	0.12998***	2011-2013
	(0.000)	(0.000)	(0.000)	(0.000)	
N	334	246	213	347	
Simone	-0.13845***	0.05591***	-0.01656***	-0.19435***	2014-2016
	(0.000)	(0.000)	(0.000)	(0.000)	
N	337	208	231	316	
		Panel C: Execution Score			
Simone	0.34643***	0.14566***	-0.01690***	0.29419***	All Data
	(0.000)	(0.000)	(0.000)	(0.000)	
N	691	454	444	663	
Simone	0.07045***	-0.04290***	-0.05567***	0.12998***	2011-2013
	(0.000)	(0.000)	(0.000)	(0.000)	
N	354	246	213	347	
Simone	-0.13845***	0.05591***	-0.01656***	-0.19435***	2014-2016
	(0.000)	(0.000)	(0.000)	(0.000)	
N	337	208	231	316	

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2: Estimation of Simone impact with Meet Leve, Athlete Level, and Year fixed effect. (clustered by meet).

VARIABLES	Vault	Bars	Beam	Floor	
		Panel A: Probability of Penalty			
Simone	0.01046	38.51449	-0.93372**	-0.01590	All
	(0.325)	(2.827e+08)	(0.465)	(0.262)	Data
N (IDs)	430 (58)	15 (3)	236 (32)	551 (84)	
Simone	-0.09166	39.37500	-0.70899	-0.08420	2011-
	(0.477)	(4.347e+08)	(0.564)	(0.395)	2013
N (IDs)	192 (32)	7 (1)	127 (25)	268 (52)	
Simone	-0.38479	-28.39131	-1.00767	0.24710	2014-
	(0.467)	(0.000)	(1.056)	(0.395)	2016
N (IDs)	159 (27)	2 (1)	54 (8)	222 (42)	
		Panel B: Difficulty Score			
Simone	-1.59147***	-2.08740*	-3.07780***	-1.05478**	All
	(0.602)	(1.177)	(1.170)	(0.457)	Data
N (IDs)	287 (33)	135 (23)	167 (17)	354 (40)	
Simone	-1.10855	0.16257	-18.86062	-1.60293**	2011-
	(0.933)	(1.215)	(3,983.389)	(0.802)	2013
N (IDs)	104 (17)	49 (11)	73 (10)	131 (22)	
Simone	-2.28512**	-18.81645	-18.59461	-1.27414	2014-
	(1.093)	(5,523.819)	(3,474.016)	(0.819)	2016
N (IDs)	118 (18)	56 (11)	45 (6)	136 (17)	
		Panel C: Execution Score			
Simone	-1.59924***	-2.03774*	-3.07780***	-1.05478**	All
	(0.604)	(1.167)	(1.170)	(0.457)	Data
N (IDs)	293 (34)	132 (22)	167 (17)	354 (40)	
Simone	-1.15394	0.16251	-18.86062	-1.60293**	2011-
	(0.939)	(1.215)	(3,983.389)	(0.802)	2013
N (IDs)	113 (19)	46 (10)	73 (10)	131 (22)	
Simone	-2.28512**	-18.81645	-18.59461	-1.27414	2014-
	(1.093)	(5,523.819)	(3,474.016)	(0.819)	2016
N (IDs)	118 (18)	56 (11)	45 (6)	136 (17)	

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Interacting Big Meets with Simone, estimation of Simone impact with Meet Leve, Athlete Level, and Year fixed effect. (clustered by meet).

VARIABLES	Vault	Bars	Beam	Floor	
			Panel A: Probability of Penalty		
Simone	0.32868	-4,957.80346	-0.50312	0.10454	All
	(0.385)	(0.000)	(0.525)	(0.314)	Data
Big Meet x	-0.48916	11,533.93004	-0.81215	-0.18509	
Simone	(0.331)	(0.000)	(0.529)	(0.269)	
N (IDs)	430 (58)	15 (3)	236 (32)	551 (84)	
Simone	0.67455	-0.00000	-0.40092	0.40390	2011-
	(0.540)	(0.000)	(0.668)	(0.452)	2013
Big Meet x	-1.46997***	59.06250	-0.53054	-0.84724**	
Simone	(0.519)	(0.000)	(0.667)	(0.400)	
N (IDs)	192 (32)	7 (1)	127 (25)	268 (52)	
Simone	-0.89168	-18.92754	-0.83998	-0.06409	2014-
	(0.670)	(0.000)	(1.045)	(0.500)	2016
Big Meet x	0.64093	-18.92754	-0.63142	0.42084	
Simone	(0.583)	(0.000)	(0.970)	(0.406)	
N (IDs)	159 (27)	2 (1)	54 (8)	222 (42)	
			Panel B: Difficulty Score		
Simone	-2.33183***	-4.20205***	-4.04219***	-1.96726***	All
	(0.643)	(1.540)	(1.317)	(0.507)	Data
Big Meet x	1.48460***	3.04165***	1.38742**	1.77372***	
Simone	(0.405)	(1.078)	(0.610)	(0.385)	
N (IDs)	287 (33)	135 (23)	167 (17)	354 (40)	
Simone	-1.48407	-0.62162	-35.49329	-2.23868***	2011-
	(0.938)	(1.217)	(4,816.262)	(0.864)	2013
Big Meet x	1.14747	31.70192	17.55225	1.36166**	
Simone	(0.713)	(3067879.635)	(3,421.383)	(0.634)	
N (IDs)	104 (17)	49 (11)	73 (10)	131 (22)	
Simone	-3.24772***	-38.25798	-19.17167	-2.62229***	2014-
	(1.160)	(8,895.197)	(3,757.725)	(0.905)	2016
Big Meet x	1.70792***	19.24863	1.49254	2.26138***	
Simone	(0.539)	(4,367.072)	(0.916)	(0.568)	
N (IDs)	118 (18)	56 (11)	45 (6)	136 (17)	
			Panel C: Execution Score		
Simone	-2.34760***	-4.18877***	-4.04219***	-1.96726***	All
	(0.645)	(1.543)	(1.317)	(0.507)	Data
Big Meet x	1.48287***	3.03382***	1.38742**	1.77372***	
Simone	(0.406)	(1.078)	(0.610)	(0.385)	
N (IDs)	293 (34)	132 (22)	167 (17)	354 (40)	
Simone	-1.52640	-0.62151	-35.49329	-2.23868***	2011-

	(0.946)	(1.217)	(4,816.262)	(0.864)	2013
Big Meet x	1.14172	31.78377	17.55225	1.36166**	
Simone	(0.714)	(3197398.220)	(3,421.383)	(0.634)	
N (IDs)	113 (19)	46 (10)	73 (10)	131 (22)	
Simone	-3.24772***	-38.25798	-19.17167	-2.62229***	2014-2016
	(1.160)	(8,895.197)	(3,757.725)	(0.905)	
Big Meet x	1.70792***	19.24863	1.49254	2.26138***	
Simone	(0.539)	(4,367.072)	(0.916)	(0.568)	
N (IDs)	118 (18)	56 (11)	45 (6)	136 (17)	

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1