Sequential Teamwork in Competitive Environments: Theory and Evidence from Swimming Data

Sequential Teamwork in Competitive Environments: Theory and Evidence from Swimming Data

Michael Neugart* and Matteo G. Richiardi†

March 2013

Abstract

Many tasks require the input by more than one person very often with members of the team contributing sequentially. However, team production is plagued by disincentive problems. We investigate individual incentives to team production with sequential contributions and competing teams. We show that earlier contributors free-ride on team members contributing later on. We test our predictions on sports data using an athlete's performance in the individual race as a natural control for his relay performance. Our empirical findings strongly support the theoretical claims.

Keywords: team production, contest, intergroup competition, sequential contribution, free-riding
JEL-Classification: C70, D20, D70, H40

*Technical University of Darmstadt, Department of Law and Economics, Marktplatz 15, D-64283 Darmstadt, Germany; e-mail: neugart@vwl.tu-darmstadt.de

†University of Turin, Department of Economics and Statistics, lungo Dora Siena 100A, 10153 Torino, Italy; Collegio Carlo Alberto and LABORatorio Revelli, via Real Collegio 30, 10024 Moncalieri (Torino), Italy; e-mail: matteo.richiardi@unito.it

1 Introduction

It is widely acknowledged that team production is plagued by disincentive problems because individuals free-ride on the contributions of other team members (Alchian and Demsetz, 1972; Holmstrom, 1982). Yet, little seems to be known on the structure of such disincentives when there is between-team competition and individuals' contributions build on previous work done by the other team members. In this paper we provide both a theoretical and an empirical investigation of the issue.

Sequential contributions to a team are indeed quite common in real-world production processes. For instance, in the era of globalization work is done around the clock, with computer programmers, planners and other workers producing output that can easily be sent electronically around the globe; at the end of the day they hand over their work to colleagues located in a country where the sun is about to rise. The problems involved are the same that occur when drafting a document: a bad draft requires more effort by the people working on it at later stages in order to achieve a certain level of quality. Another example is the training of students at universities. Colleagues who do a proper job teaching first year courses prepare students well for courses to be taught by other faculty in the second year. One may also think of mail delivery. In many countries, parcels or registered letters have to be delivered personally by the postman. It is a common perception that postmen sometimes do not even bother to ring the bell and just leave a notice on the door saying that the mail can be retrieved at some mail delivery center. Thus, they save on their time while increasing the workload at the center. Finally, in production processes organized along assembly lines a poor job by some worker at an early stage can increase the workload of subsequent workers in order to achieve the target quality. In all these examples it may be difficult or very expensive to track the responsibility for a poor final outcome down to the contributions of all the team members involved.

It is most natural to identify a team with an individual firm. In this interpretation, competition comes directly from market forces. However, having multiple teams performing the same task in a competitive environment might also be a strategy of internal organization within the firm, specifically targeted at reducing free-riding of employees. Whatever the interpretation, it is interesting to know *a priori* whether and to what extent there are incentives to free ride in a competitive environment, and whether such incentives change along the production process.

To this aim we develop a model where members of competing teams contribute sequentially to win a commonly shared prize. The main finding is that free-riding remains pervasive, with earlier contributors exploiting team members contributing later on.

For testing our predictions we turn to swimming data. This has several advantages. Typically, it is difficult to measure the performance of workers and their individual contribution to a team in standard work situations. Moreover, from a researcher's perspective it is usually infeasible to construct a convincing counterfactual that would allow to draw causal inference from the observations in an environment of an operating firm. In order to overcome these problems economists have increasingly turned to sports data recognizing that these environments provide a number of natural experiments which allow for the testing of the effects of incentives on labour market behavior (Ehrenberg and Bognanno, 1990; Kahn, 2000; Szymanski, 2003). Our comprehensive data-set covers swimming competitions from all over the world during the years 1972 to 2009 with a total of more than 300,000 observations. It seems to be particularly suited for our purposes as it allows us to construct a counterfactual for each individual's performance by comparing times swam in individual races with the same swimmer's performance in a relay at the same event typically taking place within a time period smaller than 2 days. This solves a major identification problem, that arises because the team members starting order in a relay depends on individual ability, with better swimmers generally placed first or last. Moreover, the richness of the data allows us to exclude a series of other potentially confounding factors. The evidence supports the prediction that even competitive environments are characterized by substantial free-riding, with a marked first-mover advantage.

We proceed in the following way. After reporting on the related literature in the next section, in section 3 we set-up the theoretical model. In section 4 we provide some background on the rules and main characteristics of swimming competitions, describe the data that we use and present our empirical testing strategy. In section 5 we present and discuss our empirical results. Section 6 offers our conclusions.

2 Related literature

Our paper builds a bridge between two separate strands of the literature. The first one looks at contributions to public goods.Varian (1994) argued that with sequential choice the free-riding problem is exacerbated with respect to the simultaneous contributions mechanism, and that there exists a first mover advantage with early contributors contributing less. While this contribution is akin to ours in the sense that agents make sequential choices, there is no competition between teams.

Experimental evidence by Andreoni et al. (2002) specifically tailored to test the predictions of Varian (1994) confirms the first mover behavior but also shows that the difference between simultaneous and sequential play vanishes to the end of the experiment. More recent experiments by Gaechter et al. (2009) support the prediction that the overall contribution is lower in sequential contributions but do not find evidence for the predicted first mover advantage. In another, earlier experimental study Erev and Rapoport (1990) compared sequential and simultaneous contributions to public goods showing that simultaneous choice is significantly less effective in solving the dilemma. However, we are not aware of an empirical assessment of free-riding with sequential contributions in natural work environments. Moreover, none of the studies reviewed above considers the effect of competition between groups on within-group performance.

This is the focus of the second strand of the literature we are relating to. Models of between-group competitions with simultaneous contributions have been developed in the vast literature on contests, which goes back to the seminal contribution of Tullock (1980).[1] Again, empirical analyses of between-group competition have mainly involved laboratory experiments. For instance, Bornstein et al. (1990) compared the performance of groups in a social dilemma situation under two conditions: one in which groups were not facing competition and another in which groups were competing for an additional reward. They found that between-group competition significantly increased the contributions of the simultaneously acting team members, a finding that was replicated by Erev et al. (1993) in a different work environment where subjects picked oranges, as well as by Gunnthorsdottir and Rapoport (2006). However, all these experiments involved simultaneous contributions. In summary, surprisingly, little seems to be known on whether competition between teams eliminates free-riding as team members contribute sequentially.

Our contribution aims at filling this gap. In our theoretical model we follow the approach developed by Tullock, and extend it to the case when team members contribute sequentially. We are then able to test our implications in a setting where one would *a priori* expect very little free-riding. The fact that swimmers in the first legs of a relay slow down with respect to their potential performance strongly suggests that they free-ride on swimmers in subsequent legs, who appear to exert close to full effort. That this happens to be true even in the Olympic finals speaks of the ubiquity of free-riding.

We are not the first to look at swimming competitions as a test-bed. Existing studies mostly involve experimental work stemming from the area of social psychology with mixed results on the existence of free riding. One of the earliest contributions (Sorrentino and Sheppard, 1978) classified swimmers on the basis of a questionnaire in approval-oriented ones and others threatened by rejections. They found that the approval-oriented types had faster swimming times in relays. Williams et al. (1989) devised an experiment in which swimmers in the treatment group knew their swimming time would be announced at the end

[1]See for instance Katz et al. (1990), Ursprung (1990), and Gradstein (1993). Two recent surveys are Corchòn (2007) and Konrad (2009).

of the competition, while swimmers in the control group knew their swimming time would remain concealed. Relay swimming times were faster when scores were identifiable. In a similar setting, Everett et al. (1992) measured social cohesion in teams, again by means of a questionnaire. They looked at the difference between relay and individual competition swimming times. Social cohesion was found to affect free-riding for female but not for male swimmers, while the announcement of performances bore no effects. Miles and Greenberg (1993), too, compared performances in relays with individual competitions and analyzed the effect of punishment threats. They found that the threat of punishment attenuated free-riding. In a more recent contribution, Hüffmeier and Hertel (2011) found that later swimmers in a relay tend to perform better which they interpret as a self-motivation effect found in an accompanying field study on local sports club swimmers who were asked about their perceived indispensability in relays.

Differently from these contributions, we draw inference from swimmers who performed in a natural setting. Moreover, the statistical power of our dataset is much stronger, involving more than 300,000 observations rather than a handful of observations (often less than 30) as in many experimental studies.

3 A model for inter-group competition with intra-group choices

3.1 The model

We model competition between groups along the lines of a Tullock-contest. Tullock (1980) employed a contest success function (CSF) where the probability of winning the competition is equal to the ratio of own effort to global effort, the sum of efforts of all contenders.[2]

There are two teams $t = A, B$ competing for a prize S. Each team has two players, denoted with $n = 1, 2$. Inputs of players are substitutes. All players are of homogeneous ability. The prize S has equal value to each of the team members.

Team members exert effort e at cost $c(e)$. For simplicity, we employ a linear cost function $c(e) = ke$, with $k > 0$. The contributions of team members to the overall team output enter additively. In order to reflect the sequential nature of the game, however, first members of teams A and B choose their level of effort first. Last players make their choice on the basis of first players' outcome. Thus, members 1 are Stackelberg leaders *vis-á-vis* members 2 in their teams. We denote

[2]Tullock's idea was to compare rent seeking activity –group contribution in our setting– to the purchase of lottery tickets: the higher the number of tickets, the more likely to win the lottery. Skaperdas (1996) provided an axiomatic foundation for the Tullock CSF, while more recently Jia (2008) offered a distribution based justification for its ratio form.

with e_{A1} effort spent by the team member moving first in team A, and e_{A2} as effort spent by the team member who moves second in team A. The notation for team B is chosen accordingly.

The fact that effort is not deterministically transformed into performance, and that there are stochastic elements in the competition is taken into consideration by CSFs. In our setting, however, uncertainty is partly resolved after the first players finish their task. Hence, second players face a different informational content. We model this by introducing an additional random term

$$\epsilon = \begin{cases} \varepsilon \\ -\varepsilon \end{cases} \tag{1}$$

having two support points with equal probability, and actual realizations restricted to $0 < \varepsilon < S/(4k)$. Realization of this random variable takes place after first players made their choices and before second players choose their effort; following Hirshleifer's micro-foundation of the Tullock contest function (Hirshleifer, 1989) it contributes to the odds of winning of each team according to:

$$\frac{p_A}{p_B} = \frac{e_{A1} + \epsilon + e_{A2}}{e_{B1} - \epsilon + e_{B2}}. \tag{2}$$

This is to say that the advantage unfolding for team A as its first player finished the task is to the disadvantage of team B, and vice versa.[3]

As $p_A + p_B = 1$ we can write team A's and B's respective probabilities of success in this either-or competition as

$$p_A = \frac{e_{A1} + \epsilon + e_{A2}}{e_{A1} + e_{A2} + e_{B1} + e_{B2}}, \tag{3a}$$

$$p_B = \frac{e_{B1} - \epsilon + e_{B2}}{e_{A1} + e_{A2} + e_{B1} + e_{B2}}. \tag{3b}$$

Expected payoffs follow as

$$EV_{tn} = p_t S - k e_{tn} \quad t = \{A, B\},\ n = \{1, 2\} \tag{4}$$

where each player of a team t weighs the prize S with the team's probability of winning and subtracts from that his individual effort costs.

[3] In the lottery analogy, nature is allowed to buy some tickets and give them to one of the teams, thus increasing its chances of winning and decreasing those of the opponent correspondingly.

3.2 Effort choices

The game is solved by backward induction. Second players take the level of effort of first players as given, know the realization of ϵ, and choose their own effort simultaneously. First players make their choices taking into account the later realization of ϵ, and the reaction of the second players.

Proposition 1 ***[First mover advantage]*** *In equilibrium, the optimal level of effort provided by first players is $e^*_{A1} = e^*_{B1} = 0$ and is lower than the average level of effort provided by second players, which is $e^*_{A2} = e^*_{B2} = S/(4k)$.*

Proof Any positive level of effort put by last players must solve the following first order conditions (s.o.c. being satisfied):

$$(e_{B1} + e_{B2})S + \epsilon S = kD^2 \tag{5a}$$

$$(e_{A1} + e_{A2})S - \epsilon S = kD^2, \tag{5b}$$

with $D = e_{A1} + e_{A2} + e_{B1} + e_{B2}$. Optimal choices follow as

$$e^*_{A2} = \max\left(0, -e_{A1} + \frac{S}{4k} - \epsilon\right) \tag{6a}$$

$$e^*_{B2} = \max\left(0, -e_{B1} + \frac{S}{4k} + \epsilon\right). \tag{6b}$$

These are the reaction functions of second players to effort choices of the first players. Corner solutions with zero effort level by one or both second players may arise. Due to substitutability of efforts within teams, second players decrease their optimal efforts as first players put in more effort. Efforts of second players are non-decreasing in the valuation of the prize and non-increasing in the marginal effort costs k. A positive realization of the random term ϵ, implying that team A has a lead over team B, induces the second player of team A to decrease his level of effort (if not already in the corner solution), while the second player of team B puts in more effort.

We now turn to the optimal choices of first players. These two players also make their decisions simultaneously. They take into account the reaction of the subsequent players and form expectations on the realization of ϵ. Accordingly, their payoff functions follow as

$$EV_{A1} = E[\frac{e_{A1} + \epsilon + e^*_{A2}}{e_{A1} + e^*_{A2} + e_{B1} + e^*_{B2}} S - ke_{A1}] \tag{7a}$$

$$EV_{B1} = E[\frac{e_{B1} - \epsilon + e^*_{B2}}{e_{A1} + e^*_{A2} + e_{B1} + e^*_{B2}} S - ke_{B1}]. \tag{7b}$$

In the Appendix we show that the condition $0 < \varepsilon < S/(4k)$ rules out the possibility of multiple equilibria, and guarantees an interior solution for second players. In such a case, inserting the reaction functions (6) in (7) and simplifying gives

$$E[V_{A1}] = S/2 - ke_{A1} \tag{8a}$$

$$E[V_{B1}] = S/2 - ke_{B1}. \tag{8b}$$

Obviously, we get corner solutions and first players choose effort

$$e^*_{A1} = e^*_{B1} = 0. \tag{9}$$

Inserting back $e^*_{A1} = e^*_{B1} = 0$ into the reactions functions (6) yields

$$e^*_{A2} = \frac{S}{4k} - \epsilon > 0 \tag{10a}$$

$$e^*_{B2} = \frac{S}{4k} + \epsilon > 0. \tag{10b}$$

□

These results stem from the substitutability of within team members' efforts. The team member moving first knows that an increase in his effort is leading to a decrease in the effort of the team member moving second. Thus higher effort on his side is not reflected by a larger chance of winning the competition, though he would still have to carry the burden of a higher cost of effort. Interestingly, what deters the first player from contributing is the reaction of the second player, who will respond one-by-one reducing his contribution. Under this perspective, the first player is simply restraining from a useless (but costly) action. It is the second player's self interest that eventually turns against himself.

That we get a corner solution where first players reduce effort to zero is due to the assumption of linear effort costs. Applying a convex function in numerical simulations yields positive efforts for first players, which, however, remain below the level for second players.[4] Given the inherently difficult tractability of contest models with quadratic cost functions, we opted for a presentation of a model with linear cost functions.

[4]For instance, assuming a quadratic cost function $c(e) = e^2$, a random term $\epsilon = 0.1$ and a prize $S = 2$, we get $e_{i1} \approx 0.2$ for first players, while e_2 is approximately equal to 0.37 for the second player of the leading team, and to 0.45 for the second player of the lagging team.

4 Empirical strategy

We use swimming data on relays and individual competitions to test the main implications of our model. Relays come very close to the setting we focus on. In a relay, there is sequential contribution to a team with one swimmer performing after the other, and teams competing against each other. Moreover, the task is simple, the rules are clear to everybody and performance is exactly measured.

4.1 Swimming competitions

Swimming competitions entail four competitive styles –backstroke, breaststroke, butterfly and freestyle– at varying distances (e.g. 100 meters, 200 meters, etc.) typically in 25 or 50 meter pools.[5] Relays are a group of swimmers who either all swim freestyle or each swim one different style in the order of backstroke, breaststroke, butterfly and freestyle (medley relay). Except for some specific (usually minor) events, relay teams, according to FINA rules, consist of four swimmers. Unless specified by the Promoter's conditions the nomination of team members and the relay swimming order must be made before the competition. Any relay team member may compete in a race only once.

4.2 Data description

Our data-set was kindly provided by GeoLogix AG, a Suisse company which gets the data directly from the European Swimming Federation (LEN) and other participating federations. In total we have 302,576 observations of performances of individual swimmers at about 7,000 events which took place worldwide between 1972 and 2009. The data comprises athletes who took part in the same event and for the same style, both in the individual competition and in the relay.

The events included in our sample are major events such as the Olympic Games, World Championships, European Championships, Pan Pacific Games, the Commonwealth Games or Universiades, and other events, like national championships (see table 1).

As for the *personal characteristics* of the swimmers, we have information on age, gender, nationality, and FINA points. Age is between 6 and 109 years with a mean of 17.8 (the median is 16). Gender composition of the sample is more or less equally split. The FINA Point Scoring assigns point values to swimming performances. Points are assigned at every competition, by comparing a swimmer's performance with a base time that is recalculated every year, taking the average of

[5] Rules for these swimming competitions, both national and international, are set by the Federation Internationale de Natation (FINA) (`www.fina.org`).

the top ten of the All Time World Rankings. More points go along with better performance. In the sample, FINA points are related to the individual competitions and vary between 5 and 1,181 with a mean of 502.8 (the median is 506). Michael Phelps had 1,063 in the year of the Olympic Games in Beijing.

Next we have information on the *event* (event name, location and beginning and ending day), the *competition* (style, distance, date of attendance and round – heats, preliminary, semifinals, or finals) with the day of the competition allowing to some extent to control for the sequence of the individual and the relay race, and finally *performances*, which include the time in the individual and the relay competition, the total relay time, as well as the starting order in the relay and the final placement both for the relay and the individual competition.

For our analysis we classify event importance according to three criteria. The first one (*major1*) includes Olympic games, world championships, European championships, pan Pacific games, Commonwealth games and Universiades (3,785 observations at 31 events). The second one (*major2*) adds also national championships and major regional championships (140,570 observations at 2,308 events).[6] The third one (*major750*) considers only events for which the average FINA points of the participants we observe in our sample is above 750 (17,785 observations at 236 events): this corresponds to the top 8% of the distribution of FINA points. All *major1* events are also *major2* events. On the other hand, the *major750* classification has some overlapping with the other two. Of all *major1* observations, more than 90% are also classified as *major750*, while of all non-*major1* events, less than 5% are classified as *major750*.[7] Overall, it seems safe enough to consider the following ranking of events, in increasing order of importance: all events, *major2*, *major750*, *major1*.

4.3 Bringing the theory to the data

Our model predicts that earlier swimmers slack off more than later swimmers which should be reflected in a sequence of decreasing swimming times of relay swimmers. However, taking this result straightforwardly to the data is not feasible. First swimmers in a relay competition start upon hearing the starting signal (a "flat start") while the following swimmers start after the previous swimmer touched the

[6] Event identification for the first two criteria is performed by textual search of the relevant strings (e.g. "world", "European", *etc.*) and variations in the meeting name, in all the relevant languages.

[7] The *major1* events that are not classified as *major750*, are the XX Olympic Games (for which we have only one observation about a swimmer who happened to be below 750 FINA points), Canadian World Championship Trials (170 observations), the Canadian Pan Pacific Trials (144 observations) and the Junior Pan Pacific Championships (14 observations).
The non-*major1* events that are classified as *major750* are mainly national championships.

wall of the pool. Hence, all relay swimmers starting after the first ones can see their preceding team-mate approaching and fine-tune their start (what is called a "flying start"): they enjoy a reaction time advantage. Therefore, we are not able to disentangle the reaction time effect from a potentially slacking off of later swimmers.

However, as all swimmers but the first ones benefit from the same reaction time advantage, we may employ those swimmers to test free-riding on later contributors to the team. But again, one has to be cautious to straightforwardly go for such a test. If swimmers are not allocated randomly to slots, say faster swimmers are systematically swimming at later positions in the relay, we would not be able to infer that there is free-riding from the observation that later swimmers are faster in a relay. And, indeed, we do find that the starting order is not random. In order to overcome this problem, we use information on each relay swimmers' performance in the individual competition of the same event to control for his ability: our dependent variable is, therefore, the *relative difference in swimming time between the relay and the individual competition*. In other words, each swimmer acts as his own control.

Note that our empirical test does not require any assumption on how athletes evaluate the prize to be won in the individual competition, as long as that evaluation is common to all swimmers in the team. If one is also willing to accept that individuals value relay and individual competition prizes equally –and there are good reasons to believe that this is a reasonable assumption, which we will discuss in the Conclusions– then our data allows for another test of free-riding by employing first swimmers only. Given the assumption of equal valuation of the prize, the relay and the individual competition have the same structure of incentives and costs. Hence, the performance in the individual competition offers a natural counterfactual to evaluate the optimal effort, absent any free-riding concern. It follows that we can directly compare swimming times in the relay with performances in the individual competition for first swimmers as there is no reaction time issue involved. If free-riding prevails we should find slower swimming times in relays than in individual races for first swimmers.

Finally, one can exploit the data on fourth swimmers in order to learn more on the reaction time advantage relative to free-riding. The fourth swimmers are the last swimmers in the relay and, according to our model, bear the bulk of the burden of the competition. Consensus estimates among experts and practitioners quantify the reaction time advantage to about .6 secs (.7 secs for a good flat start[8], and .1 secs for a good flying start[9]). Finding a similar advantage for swimming times

[8]See http://www.swimmingscience.net/2012/06/reaction-time.html

[9]See http://forums.usms.org/showthread.php?16261-FINA-Relay-Take-Off-Rule-w-Automated-Equip

in relays with respect to the individual competition would suggest that the last swimmers perform about the same in the relay and in the individual competition. Again, if we are willing to assume equal valuation of the prize between the relay and the individual competition, we could conclude that free-riding vanishes at the end of the race.[10]

To recapitulate, in order to get rid of composition issues we look at the relative difference between the relay and the individual competition. We interpret a *diminishing* difference as we move from the second to subsequent swimmers as evidence in favor of our main proposition, which states that early swimmers enjoy a first-mover advantage. The swimming time in the individual competition is used only as a control here. Moreover, we interpret a *positive* difference for first swimmers as a measure of free-riding in our context, conditional on an equal valuation of the prizes in the relay and in the individual competition. Turning to last swimmers only, finding an absolute difference in swimming times between the relay and the individual competition of about .5-.7 seconds. would suggest that last swimmers perform in relays as well as in individual competitions. With equal valuation of the prizes, this would point to free-riding vanishing toward the end of the competition.

5 Empirical evidence

5.1 Starting order is not random

As a preliminary analysis, we check whether the assignment by team managers of swimmers to particular slots depending on their ability is a concern. First we focus on the subset of observations for which we have data on all swimmers in a team (5.8% of the sample, or 17,532 observations).[11] Table 2 reports the average FINA points of these swimmers by starting order. First and last swimmers are on average better than second and third swimmers. Second, we ran OLS regressions of FINA points on the characteristics of the competition, the characteristics of the swimmer and the starting order, first restrincting only to teams for which we have information on all players (table 3, column 1) and then considering the whole sam-

[10]Extending our model to individual competitions indeed suggests that the optimal effort should be the same as for last swimmers in the relay. In an individual competition, only two players compete to win the prize S; the game has only one stage, and there is no revelation of information during the play. The odds of winning are $p_A/p_B = e_A/e_B$, and expected payoffs are $V_t = p_tS - ke_t$, $t = \{A, B\}$. Equating marginal benefits to marginal costs, we obtain $e^* = S/4k$, which is equal to the average effort for last swimmers in the relay, over all possible realizations of ϵ.

[11]Our data, as already stressed, contains only swimmers that swam both the individual and the relay competition.

ple of swimmers (column 2). The evidence suggests that stronger swimmers are placed either first or last, while weaker swimmers are in the middle. These findings are in line with common knowledge among team managers and athletes, and motivates our choice of the relative difference between the relay and individual performance as the outcome variable in the analyses that follow.

5.2 Descriptive evidence on free-riding

The first column of table 4 reports average swimming time in the individual competition. The additional columns show the relative difference in swimming times between the relay and the individual competition for different starting orders in the relay. First swimmers are, on average, slower in relays with respect to their performance in the individual competition. The difference amounts to .21%, i.e. 12/100 of a second in absolute terms.

Testing the relative difference for the first swimmers against the null of there being no difference in performance yields a highly significant p-value. This result is robust against splitting the sample along gender or age. It is furthermore valid for swimmers with higher or lower FINA points than the median swimmer. It also holds over all styles if we focus on the sign of the difference in swimming times and in 6 out of 7 subgroups for the various styles in terms of significance.

There is some evidence that the difference in performance decreases with the importance of the competition, although it becomes non significant only for *major1* events.

There are no indications that training or the use of illegal substances targeted to a specific competition (individual or relay) might disturb our results. In 87% of the observations individual and relay competitions are within 1 day of separation which implies that training efforts influence individual and relay competitions equally. With illegal substances targeting longer term goals such as the building up of red blood cells basically the same logic applies as with legal training methods.

One might also be concerned that fatigue decreases performance in later events. However, the distribution of days of separation between individual and relay competitions is quite symmetric. In any case, we also split the sample along the timing of competitions to check whether it makes a difference if the individual race took place before or after the relay at the particular swimming event for which we compare the swimming times. It is still true that relay performances are weaker than individual performances.

As we have already noted, direct comparison between the relay and the individual competition is possible only for first swimmers as subsequent swimmers enjoy an advantage in terms of reaction time. This explains why the time difference turns negative for the second to the fourth swimmers in the relay (columns

3-5). More importantly, however, last swimmers in relays seem to be relatively faster (with respect to their own individual performance) than swimmers starting 2nd or 3rd. Again, this is what we would have expected from our theory.

Summarizing these findings, faster relative swimming times for later swimmers in the relays provide evidence in support of our main proposition. Furthermore, if one is willing to accept the assumption of equally valuable prizes in the individual and the relay competition, the evidence presented for first swimmers can be interpreted as a measure of free-riding.

The fact that first swimmers under-perform in relays with respect to individual competitions, and that relative swimming times in relays decrease at the end of the race raises the question whether fourth swimmers still under-perform, over-perform or perform about the same in relays and in individual competitions. Table 5 reports the estimated intercepts in a linear regression model of the (absolute) difference between relay and individual competition swimming times, controlling for age, gender, style and length of the race, estimated on fourth swimmers only with different sample selections. For each subsample, the intercept is the estimated advantage in relays over individual competitions for the reference category (100m Freestyle, male swimmers aged 15-30), and can be attributed to the reaction time advantage coming from the flying start and a possible free-riding effect. Considering all events it seems that fourth swimmers are still slower in relays than in individual competitions (their advantage is lower than the presumed reaction time advantage). For more competitive events we find an absolute difference of comparable size to the reaction time advantage (.5-.6 secs). This suggests that –at least in more competitive events– fourth swimmers swim the relay about as fast as the individual competition. If one is willing to assume that the prize is equally evaluated in the two circumstances, one might conclude that fourth swimmers exert approximately full effort.

In the remaining part of the section we elaborate on the findings of table 4 by means of a multivariate analysis, which allows us to increase the number of controls. The multivariate analysis is run for first swimmers and for second to fourth swimmers, separately.

5.3 Regression analysis for first swimmers

The analysis on first swimmers allows to identify how much first players reduce their effort in relays with respect to their individual performance under the assumption of equal valuation of the prizes. We run both an OLS and a fixed effects model, the latter to take into account unobserved heterogeneity. The controls include gender, age[12], style, event importance and schedule (whether the individual competition is on a day before the relay, on the same day, or on a day after). Standard errors are clustered at the individual level.[13]

Table 6 reports the results. The coefficient in the OLS regression of the *swimmer1* indicator gives the average of the relative difference in swimming performance for the reference category. It implies that first swimmers slow down by 0.1 percentage points (ppt). In the fixed-effects model we get a point estimate of 0.2 ppt, the coefficient of the *swimmer1* indicator being the average of all individual effects. For both specifications the p-values for the coefficient on the first swimmer are smaller than .001. As for the controls, we find that both young and old swimmers perform relatively worse in relays, while the gap in relative performance is reduced in major events.

These results are robust to various variations of the underlying sample. For instance, one could object that free-riding depends on competitive pressure. While we already control for the type of events with a set of dummies, we may further restrict the sample to competitions with closer outcomes. So we estimate the same models focusing on freestyle competitions with swimmers aged 15-30 swimming in finals.[14] We do this on all events, and then separately on major events only, using all three definitions. The results are reported in table 7: still swimmers are slower in relays as compared to their individual performances in a range of 0.15-0.3 ppt.

The filter on *major2* events is our preferred sample, as it achieves a good balance between selecting highly competitive events and not throwing away too much data. We therefore use it as a basis for further robustness checks. First, we select swimmers who finished in the first four positions both in the relay and in the individual competition. This should shield our results from the influence of uncompetitive teams and uncompetitive swimmers. The results (first two columns of table 8) still go through, with even higher estimated coefficients (0.3-0.4 ppt). We then check more thoroughly whether the timing of the competitions matters,

[12] We use three age groups rather than a continuous age variable, in order for the coefficients of the starting order indicators to show the effects for the reference group (swimmers aged 15-30), rather than for swimmers of a specific age; the consequential reduction in explanatory power –as measured by R^2– is very small.

[13] For fixed effect estimation, we use the `xtreg` procedure in Stata, which does not adjust the degrees of freedom for the number of fixed effects swept away in the within-group transformation. This approach is appropriate when the standard errors are nested within clusters (meaning all the observations for any given group are in the same cluster), as in our case, see Wooldridge (2010, ch. 20) and Gormley and Matsa (2012). The R^2, on the other hand, is obtained with the `areg` procedure, which takes into account the contribution of the individual effects to the overall fitness of the model, as unobserved heterogeneity is part of the explanation of the outcome.

[14] Further restricting the sample to a specific length –50m, 100m or 200m– does not alter the results in any significant way, both here and in what follows.

by estimating our models separately when the individual competition is on a day before, on the same day, or on a day after the relay (columns 3-8). The differences turn out to be very small. Indeed, if one lesson can be learnt from all the regressions, it is that the coefficient for first swimmers is remarkably stable: we find robust evidence of free-riding for first swimmers, and we quantify it in the order of 0.2-0.3 ppt.

5.4 Regression analysis for swimmers 2-4

The analysis on swimmers 2-4 allows to test for free-riding without assuming equal valuation of prizes. As suggested by our main proposition we should observe earlier swimmers free-riding on later swimmers, and due to the reaction time advantage, we expect the coefficients for the indicator variables to be negative. If our main proposition holds true, we would then observe that the indicator variables change in size with higher orders becoming more negative (later swimmers exert on average more effort).

We estimate on swimmers 2-4 the same models as on first swimmers. Table 9 reports the results of OLS and fixed-effect regressions on the full sample. The coefficients of the order indicators are negative due to the reaction time advantage. More importantly, as predicted by our theoretical model, their absolute value is increasing from -0.4 to -0.7 for the OLS model. Also, there seems to be a big jump from the third to the fourth swimmer, coherently with our intuition that it is the last swimmer who bears most of the burden of the competition. Controlling for fixed-effects reduces the coefficients of the order indicators, implying that a part of the relative difference in relay and individual swimming times is actually explained by individual idiosyncratic characteristics, possibly reaction times.[15] What matters, however, for our test is that the coefficients of the starting order indicators are still decreasing. Moreover, the fact that the differences between the starting order coefficients remain roughly constant and are statistically different from zero speaks for our free-riding interpretation. For what concerns the control variables, again, both junior and senior swimmers appear to slack off more than prime age swimmers, and the same pattern as with first swimmers is found with respect to major events.

This very robust result that free-riding wanes at the end of the competition and that unobserved heterogeneity is not responsible for this persists all our sample selections. As with first swimmers, we first filter on individuals aged 15-30 swimming a freestyle final. We then run our OLS and fixed-effect regressions first on

[15]Note that the near-zero coefficient for 2nd and 3rd swimmers does not mean that there is no free-riding. Their swimming times in the relay and in the individual competition are about the same, but in the relay they enjoy the reaction time advantage.

all events (table 10, columns 1 and 2), and then separately on major events only, using all three definitions (columns 3-8). The coefficients of the order indicators tend, if anything, to become more negative, but the pattern we look for does not change: last swimmers are relatively faster, and the amount by which they are relatively faster is not affected by the inclusion of fixed-effects. The only difference with respect to our analysis on the full sample is that the coefficients of 2nd and 3rd swimmers are no longer significantly different. This is in line with what our model suggests: the burden of the competition lies on the shoulders (and arms and legs) of last swimmers.

Finally, we perform the additional robustness checks on our preferred setting (15-30 age group, freestyle finals, *major2* events). We first estimate our two specifications on those individuals who finished in the first four positions both in the relay and in the individual competition (table 11, columns 1-2), then perform separate estimations depending on the timing of the competitions (columns 3-8). Our results still hold.

5.5 Regression analysis for swimmers 1-4

In order to fully exploit our fixed-effects estimation strategy to control for time-invariant pro-relay unobserved factors in later swimmers (motivation, for instance), we also perform a joint analysis of all swimmers, in our preferred setting (15-30 age group, freestyle finals, *major2* events). This is reported in table 12. The OLS coefficients of the order indicators are 0.26 (first swimmers), -0.44 (second and third swimmers alike) and -0.66 (fourth swimmers), respectively. With fixed-effects, the coefficients change slightly in size, but the difference among them remains approximately constant, with the coefficient of the 1st swimmer being positive and the coefficient of the last swimmer being smaller than those of 2nd and 3rd swimmers.

5.6 Size of the effects

For establishing a benchmark against which to evaluate the size of free-riding, we calculate the average lag in swimming time between individuals belonging to teams that finished in nth place in relays and individuals belonging to teams that finished in n-1th place, for $n > 1$. On average this lag is between 1% and 2% for those who finished in the first 3 positions, and between 0.3% and 1% for those who finished in the first 10 positions (with little differences between males and females, and between different event importance, see table 13).

These numbers can then be compared with the size of free-riding for first players as shown in tables 6 and 7, which is in the range 0.2-0.3 ppt: if a first swimmer did not engage in free-riding, keeping constant the behavior of his opponents, he would recoup between 10% and 100% of his lag, a non-negligible amount.

6 Summary and conclusions

In this paper we developed a simple model of sequential contributions to a team when teams compete against each other. We find that team members contributing earlier to a team's common task contribute less than the team members contributing later. The mechanism underlying the free-riding in teams is substitutability of inputs between team members to a Tullock contest. At the margin a team member contributing earlier refrains to increase costly efforts as he can foresee that the following team members will reduce their input.

Our theoretical claims find considerable empirical support. Drawing on a unique data set of more than 300,000 observation from swimming competitions from all over the world during the last four decades we find evidence for free-riding and the pattern of efforts over the course of sequential contributions to a team as suggested by our model.

The basic idea employed for empirical testing was to compare for a given event the swimming performance of the same individual athletes for individual and relay competitions. By definition no free-riding occurs in an individual competition which is why swimmers should exert full effort at these occasions. Taking their performance in the individual race as a control we find that on average first swimmers swim slower in relays. Moreover, controlling for reaction times and confounding factors including individual effects we find that free-riding diminishes as we move from the second, to the third and finally to the fourth swimmers in relays. These estimated time differences occur to be of meaningful size.

That free-riding is still prevalent as we use each swimmer's time in the individual competition as a control for individual ability and as we additionally run a fixed-effect model may also be interpreted as supporting evidence for our modeling choice in the first part of the paper. We assumed that all swimmers have the same ability when deriving our proposition. If heterogenous ability mattered, then we should have found that free-riding vanished as we run fixed-effects models on top of controlling for individual ability with swimming times of individual competitions. As it did not, we feel quite confident with the (simplifying) modeling choices made. Although heterogeneous ability is certainly given among the swimmers of the teams, it seems that team managers cannot solve the free-riding problem by allocating swimmers to particular slots.

Our attribution of the lower performance of the first swimmer in the relay –but not the pattern of free-riding for swimmers two to four– with respect to individual competitions depends on the assumption that the prize S is equally valuable in relays and individual competitions. This assumption may need some discussion for those who believe that individual competition might be more valuable as the honors do not have to be shared.[16] One may argue that for a vast majority of athletes what actually matters is that they are medalists. The difference in utility from winning it in the individual competition rather than in the relay may be of minor importance. Indirect evidence for this comes from the observation that sport federations generally award the same monetary prizes to athletes irrespective of whether they have won a medal in an individual competition or as a member of a relay team. The *Deutsche Sportförderung*, the major German funding body for top-level sports, adopts this incentives scheme, for instance. Similarly, in Italy individual race and relay medalists including those who only swam in the prelims receive the same prize money. This is true not only for the Olympic games, but also for the world and European championships (though, obviously, the prize for more important events is bigger). Turning to endorsement deals that may follow from winning a competition, it occurs that very few athletes are having considerable incomes coming from this source. Quite the contrary: cases have been reported that medal winners are trying to sell their medals for cash.[17] What seems to shape public perception, however, are the very few superstars that are actually able to monetize their success. At the Olympic games in Beijing, Michael Phelps was offered a contract by his sponsor which would award him one million dollars in case he won at least seven gold medals.[18] But actually, with such an incentive scheme, there is no difference between winning an individual race or a relay. In the end, he won eight gold medals. While certainly not representative for a discussion on prize money, it is, however, another example that the assumption of equal valuation of prizes in individual and relay competitions cannot be easily dismissed. A similar circumstance occurred to UK athletes at the 2012 Olympic games in London. To them, no prize money was awarded, but the Royal Mail promised to issue a special postage stamp for any British gold medal winner.[19] Finally, for many athletes winning a medal in the relay is the only possibility to become a medalist at all. Accordingly, we would expect them to show particular effort in relays. That we, nevertheless, find free-riding speaks for the strength and ubiquity of our findings.

[16]The other side of 'two in distress make sorrow less'.

[17]See `www.news.de`, "Von Gold zu Geld, "August 18th, 2010.

[18]See `www.spiegel.de`, "Es geht um die Wurst", August 2nd, 2012.

[19]Same source as above.

Acknowledgements

We would like to thank GeoLogix AG for providing us with the data and in particular Christian Kaufmann for his kind and knowledgable support through all stages of this project. We also thank Giuseppe Bertola, Henry Ohlsson, Alois Prinz, Matteo Rizzoli, Alberto Zazzaro and two referees for their helpful suggestions. We also profited from the many remarks made by participants of the 37th Annual Conference of the Eastern Economic Association in New York, the Conference on Tournaments, Contests and Relative Performance Evaluation in Raleigh, the conference of the Verein für Socialpolitik in Frankfurt, the 11th journées Louis-André Gérard-Varet in Marseille, the 53rd Annual Meeting of the Italian Economic Association in Matera, the 11th Brucchi Luchino Labour Economic Workshop in Trento and research seminar participants at the University of Hamburg, the Johannes-Gutenberg University of Mainz, the Carlo Alberto College in Moncalieri and the University of Salento.

Appendix: Complete proof of Proposition 1

Distinguishing between interior and corner solutions, the optimal choices of second players (their reaction functions) can be written as

$$
\begin{aligned}
e^*_{A2} &= \max[0, -e_{A1} + S/(4k) - \varepsilon] \\
e^*_{B2} &= \max[0, -e_{B1} + S/(4k) + \varepsilon]
\end{aligned} \tag{11}
$$

if team A has the lead, and

$$
\begin{aligned}
e^*_{A2} &= \max[0, -e_{A1} + S/(4k) + \varepsilon] \\
e^*_{B2} &= \max[0, -e_{B1} + S/(4k) - \varepsilon]
\end{aligned} \tag{12}
$$

if team B has the lead. First players do not know the realization of ϵ and form expectations. The expected payoffs for the first player of team A are

$$
\begin{aligned}
EV_{A1} = {} & \frac{1}{2}\frac{e_{A1} + \varepsilon + \max[0, -e_{A1} + S/(4k) - \varepsilon]}{e_{A1} + \max[0, -e_{A1} + S/(4k) - \varepsilon] + e_{B1} + \max[0, -e_{B1} + S/(4k) + \varepsilon]} S + \\
& + \frac{1}{2}\frac{e_{A1} - \varepsilon + \max[0, -e_{A1} + S/(4k) + \varepsilon]}{e_{A1} + \max[0, -e_{A1} + S/(4k) + \varepsilon] + e_{B1} + \max[0, -e_{B1} + S/(4k) - \varepsilon]} S \\
& - k e_{A1},
\end{aligned} \tag{13}
$$

where the first line relates to the case that team A has the lead, the second line to the case that team B has the lead, and the last line are the effort costs. Analogously, the expected payoffs for the first player of team B are

$$
\begin{aligned}
EV_{B1} = {} & \frac{1}{2}\frac{e_{B1} - \varepsilon + \max[0, -e_{B1} + S/(4k) + \varepsilon]}{e_{A1} + \max[0, -e_{A1} + S/(4k) - \varepsilon] + e_{B1} + \max[0, -e_{B1} + S/(4k) + \varepsilon]} S + \\
& + \frac{1}{2}\frac{e_{B1} + \varepsilon + \max[0, -e_{B1} + S/(4k) - \varepsilon]}{e_{A1} + \max[0, -e_{A1} + S/(4k) + \varepsilon] + e_{B1} + \max[0, -e_{B1} + S/(4k) - \varepsilon]} S \\
& - k e_{B1}.
\end{aligned} \tag{14}
$$

Resolving the $\max$ functions gives us nine different cases that we have to consider when deriving the reactions functions of first players of team A and B (figure 6). The cases are defined in terms of e_{A1} and e_{B1}. It is therefore possible, in principle, to have multiple equilibria, with the reaction functions intersecting at different values of e_{A1} and e_{B1} in different cases. However, for an equilibrium to be valid, it must lie within the relevant range of values of e_{A1} and e_{B1} which define the case being considered. We will proceed by showing that this happens only for Case 1, given the assumption $\varepsilon < S/(4k)$. The equilibrium identified in Case 1 is therefore unique.

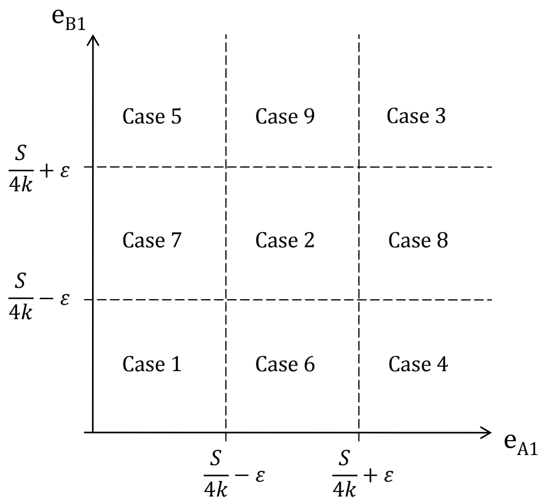

Figure 1: Case distinction for the calculation of the reaction functions for first players

- Case 1: $e_{A1}, e_{B1} < S/(4k) - \varepsilon$. The expected payoffs for players A1 and B1 write

$$\begin{aligned} EV_{A1} &= \frac{1}{2}S - ke_{A1} \\ EV_{B1} &= \frac{1}{2}S - ke_{B1}, \end{aligned} \tag{15}$$

respectively. Zero effort choices of both first players constitute the Nash equilibrium.

- Case 2: $e_{A1}, e_{B1} \in [S/(4k) - \varepsilon, S/(4k) + \varepsilon)$. The expected payoffs for players A1 and B1 write

$$\begin{aligned} EV_{A1} &= \frac{1}{2}\frac{e_{A1} + \varepsilon}{e_{A1} + S/(4k) + \varepsilon}S + \frac{1}{2}\frac{S/(4k)}{S/(4k) + \varepsilon + e_{B1}}S - ke_{A1} \\ EV_{B1} &= \frac{1}{2}\frac{S/(4k)}{e_{A1} + S/(4k) + \varepsilon}S + +\frac{1}{2}\frac{e_{B1} + \varepsilon}{S/(4k) + \varepsilon + e_{B1}}S - ke_{B1}, \end{aligned} \tag{16}$$

respectively. The first order conditions follow as

$$\begin{aligned} &\frac{1}{2}\frac{S/(4k)}{(e_{A1} + S/(4k) + \varepsilon)^2}S = k \qquad\qquad (playerA1) \\ &\frac{1}{2}\frac{S/(4k)}{(S/(4k) + \varepsilon + e_{B1})^2}S = k. \qquad\qquad (playerB1) \end{aligned} \tag{17}$$

Solving yields

$$\begin{aligned} e^*_{A1} &= -\left(\frac{S}{4k} + \varepsilon\right) \pm \sqrt{2}\frac{S}{4k} \\ e^*_{B1} &= -\left(\frac{S}{4k} + \varepsilon\right) \pm \sqrt{2}\frac{S}{4k}. \end{aligned} \tag{18}$$

We can rule out the negative solution. The positive solution violates the lower bound of Case 2 as $e^*_{A1} = e^*_{B1} = \frac{S}{4k}(\sqrt{2} - 1) - \varepsilon) < S/(4k) - \varepsilon$. No Nash equilibrium exists for this case.

- Case 3: $e_{A1}, e_{B1} \geq S/(4k) + \varepsilon$. The expected payoffs for players A1 and B1 write

$$\begin{aligned} EV_{A1} &= \frac{e_{A1}}{e_{A1} + e_{B1}}S - ke_{A1} \\ EV_{B1} &= \frac{e_{B1}}{e_{A1} + e_{B1}}S - ke_{B1}, \end{aligned} \tag{19}$$

respectively. First order conditions follow as

$$\frac{(e_{A1}+e_{B1})-e_{A1}}{(e_{A1}+e_{B1})^2}S=k \qquad (player A1) \tag{20}$$
$$\frac{(e_{A1}+e_{B1})-e_{B1}}{(e_{A1}+e_{B1})^2}S=k. \qquad (player B1)$$

The solution to these first order conditions is $e_{A1}=e_{B2}=S/(4k)$ which is outside of the restriction of Case 3. No Nash equilibrium exists.

- Case 4: $e_{A1}\geq S/(4k)+\varepsilon$ and $e_{B1}<S/(4k)-\varepsilon$. The expected payoffs for players A1 and B1 write

$$EV_{A1}=\frac{1}{2}\frac{e_{A1}+\varepsilon}{e_{A1}+S/(4k)+\varepsilon}S+\frac{1}{2}\frac{e_{A1}-\varepsilon}{e_{A1}+S/(4k)-\varepsilon}S-ke_{A1} \tag{21}$$
$$EV_{B1}=\frac{1}{2}\frac{S/(4k)}{e_{A1}+S/(4k)+\varepsilon}S+\frac{1}{2}\frac{S/(4k)}{e_{A1}+S/(4k)-\varepsilon}S-ke_{B1}.$$

  Player B1 will choose zero effort given any choice of player A1. The reaction function of player A1 follows from the first order condition

$$\frac{1}{2}\frac{S/(4k)}{(e_{A1}+S/(4k)+\varepsilon)^2}S+\frac{1}{2}\frac{S/(4k)}{(e_{A1}+S/(4k)-\varepsilon)^2}S=k. \tag{22}$$

  As the left hand side is decreasing in e_{A1} there will be no intersection of reactions function if for the smallest possible realization of e_{A1} the left hand side is smaller than k:

$$\frac{1}{2}\frac{S/(4k)}{(S/(4k)+\varepsilon+S/(4k)+\varepsilon)^2}S+\frac{1}{2}\frac{S/(4k)}{(S/(4k)+\varepsilon+S/(4k)-\varepsilon)^2}S<k. \tag{23}$$

  The inequality is fulfilled for any $\varepsilon>0$, implying that there is no Nash equilibrium for Case 4.

- Case 5: $e_{A1}<S/(4k)-\varepsilon$ and $e_{B1}\geq S/(4k)+\varepsilon$. No Nash equilibrium exists. The result follows from the same reasoning as in Case 4.

- Case 6: $e_{A1}\in[S/(4k)-\varepsilon,S/(4k)+\varepsilon)$ and $e_{B1}<S/(4k)-\varepsilon$. The expected payoffs for players A1 and B1 write

$$EV_{A1}=\frac{1}{2}\frac{e_{A1}+\varepsilon}{e_{A1}+S/(4k)+\varepsilon}S+\frac{1}{4}S-ke_{A1} \tag{24}$$
$$EV_{B1}=\frac{1}{2}\frac{S/(4k)}{e_{A1}+S/(4k)+\varepsilon}S+\frac{1}{4}S-ke_{B1},$$

respectively. Player B1 chooses zero effort. Player A1's choice follows from the first order condition

$$\frac{1}{2}\frac{S/(4k)}{(e_{A1}+S/(4k)+\varepsilon)^2}S=k \tag{25}$$

which solves as:

$$e_{A1}=-(\frac{S}{4k}+\varepsilon)\pm\sqrt{2}\frac{S}{4k}. \tag{26}$$

The negative solution can be excluded. The positive solution fulfills $e^*_{A1}=\frac{S}{4k}(\sqrt{2}-1)-\varepsilon<S/(4k)-\varepsilon$ which violates the lower bound of Case 6. No Nash equilibrium exists.

- Case 7: $e_{A1}<S/(4k)-\varepsilon$ and $e_{B1}\in[e_{B1}\geq S/(4k)-\varepsilon,S/(4k)+\varepsilon)$. No Nash equilibrium exists. The result follows from the same reasoning as in Case 6.

- Case 8: $e_{A1}\geq S/(4k)+\varepsilon$ and $e_{B1}\in[S/(4k)-\varepsilon,S/(4k)+\varepsilon)$. The expected payoffs for players A1 and B1 write

$$\begin{aligned} EV_{A1} &= \frac{1}{2}\frac{e_{A1}+\varepsilon}{e_{A1}+S/(4k)+\varepsilon}S+\frac{1}{2}\frac{e_{A1}-\varepsilon}{e_{A1}+e_{B1}}S-ke_{A1} \\ EV_{B1} &= \frac{1}{2}\frac{S/(4k)}{e_{A1}+S/(4k)+\varepsilon}S+\frac{1}{2}\frac{e_{B1}+\varepsilon}{e_{A1}+e_{B1}}S-ke_{B1}, \end{aligned} \tag{27}$$

respectively. Effort choice of player B1 follows from the first order condition

$$\frac{1}{2}\frac{(e_{A1}+e_{B1})-(e_{B1}+\varepsilon)}{(e_{A1}+e_{B1})^2}S=k \tag{28}$$

as

$$e_{B1}=\frac{-2e_{A1}\pm\sqrt{2\frac{e_{A1}-\varepsilon}{k}S}}{2}. \tag{29}$$

For the positive solution the slope of the reaction function is

$$\frac{de_{B1}}{de_{A1}}=-1+\frac{S}{2k}\frac{1}{\sqrt{2\frac{e_{A1}-\varepsilon}{k}S}}. \tag{30}$$

Note that the second part of the right hand side is decreasing in e_{A1}. If the slope is negative at the smallest possible choice of e_{A1} it would be decreasing everywhere within Case 8:

$$\frac{de_{B1}}{de_{A1}}|_{e_{A1}=S/(4k)+\varepsilon}=-1+\frac{\sqrt{2}}{2}<0. \tag{31}$$

Thus, the reaction function is sloping downwards everywhere. Furthermore, we have

$$e_{B1}|_{e_{A1}=S/(4k)+\varepsilon} = \frac{1}{\sqrt{2}}\frac{S}{2k} < S/(2k) \tag{32}$$

so that the reaction function of player B1 is outside of the area of Case 8. No Nash equilibrium exists.

- Case 9: $e_{A1} \in [S/(4k) - \varepsilon, S/(4k) + \varepsilon)$ and $e_{B1} \geq S/(4k) + \varepsilon$. No Nash equilibrium exists. The result follows from the same reasoning as in Case 8.

This completes the proof that the only Nash equilibrium with $\varepsilon < S/(4k)$ entails zero effort for first players.

References

ALCHIAN, A. AND H. DEMSETZ (1972): "Production, information costs, and economic organization," American Economic Review, 62, 777–795.

ANDREONI, J., P. BROWN, AND L. VESTERLUND (2002): "What makes an allocation fair? Some experimental evidence," Games and Economic Behavior, 40, 1–24.

BORNSTEIN, G., I. EREV, AND O. ROSEN (1990): "Intergroup competition as a structural solution for social dilemma," Social Behavior, 5, 247–260.

CORCHÒN, L. C. (2007): "The theory of contests: a survey," Review of Economic Design, 11, 69–100.

EHRENBERG, R. AND M. BOGNANNO (1990): "Do tournaments have incentive effects?" Journal of Political Economy, 98, 1307–1324.

EREV, I., G. BORNSTEIN, AND R. GALILI (1993): "Constructive intergroup competition as a solution to the free rider problem: a field experiment," Journal of Experimental Social Psychology, 29, 463–478.

EREV, I. AND A. RAPOPORT (1990): "Provision of step-level public goods: The sequential contribution mechanism," The Journal of Conflict Resolution, 34, 401–425.

EVERETT, J., R. SMITH, AND K. WILLIAMS (1992): "Effects of team cohesion and identifiability on social loafing in relay swimming performance," International Journal of Sport Psychology, 23, 311–324.

GAECHTER, S., D. NOSENZO, E. RENNER, AND M. SEFTON (2009): "Sequential versus simultaneous contributions to public goods: experimental evidence," CESifo Working Paper Series 2602, CESifo.

GORMLEY, T. A. AND D. A. MATSA (2012): "Common Errors: How to (and Not to) Control for Unobserved Heterogeneity," Available at SSRN: `http://ssrn.com/abstract=2023868` or `http://dx.doi.org/10.2139/ssrn.2023868`.

GRADSTEIN, M. (1993): "Rent seeking and the provision of public goods," Economic Journal, 103, 1236–1243.

GUNNTHORSDOTTIR, A. AND A. RAPOPORT (2006): "Embedding social dilemmas in intergroup competition reduces free-riding," Organizational Behavior and Human Decision Processes, 101, 184–199.

HIRSHLEIFER, J. (1989): "Conflict and rent-seeking success functions: ratio vs. difference models of relative success," Public Choice, 63, 101–112.

HOLMSTROM, B. (1982): "Moral harzard in teams," The Bell Journal of Economics, 13, 324–340.

HÜFFMEIER, J. AND G. HERTEL (2011): "When the whole is more than the sum of its parts: group motivation gains in the wild," Journal of Experimental Social Psychology, 47, 455–459.

JIA, H. (2008): "A stochastic derivation of the ratio form of contest success functions," Public Choice, 135, 125–130.

KAHN, L. M. (2000): "The sports business as a labor market laboratory," Journal of Economic Perspectives, 14, 75–94.

KATZ, E., S. NITZAN, AND J. ROSENBERG (1990): "Rent-seeking for pure public goods," Public Choice, 65, 49–60.

KONRAD, K. (2009): Strategy and Dynamics in Contests, New York: Oxford University Press.

MILES, J. AND J. GREENBERG (1993): "Using punishment threats to attenuate social loafing effects among swimmers," Organizational Behavior and Human Decision Processes, 56, 246–265.

SKAPERDAS, S. (1996): "Contest success functions," Economic Theory, 7, 283–290.

SORRENTINO, R. AND B. SHEPPARD (1978): "Effects of affiliation-related motives on swimmers in individual versus group competition: a field experiment," Journal of Personality and Social Psychology, 36, 704–714.

SZYMANSKI, S. (2003): "The assessment: the economics of sport," Oxford Review of Economic Policy, 19, 467–477.

TULLOCK, G. (1980): "Efficient rent-seeking," in Towards a theory of a rent-seeking society, ed. by J. M. Buchanan, R. Tollison, and G. Tullock, College Station: Texas A&M University Press, pp97.

URSPRUNG, H. W. (1990): "Public goods, rent dissipation, and candidate competition," Economics and Politics, 2, 115–132.

VARIAN, H. R. (1994): "Sequential contributions to public goods," Journal of Public Economics, 53, 165–186.

WILLIAMS, K., S. NIDA, L. BACA, AND B. LATANÉ (1989): "Social loafing and swimming: effects of identifiability on individual and relay performance of intercollegiate swimmers," Basic and Applied Social Psychology, 10, 73–81.

WOOLDRIDGE, J. M. (2010): Econometric Analysis of Cross Section and Panel Data, Cambridge, MA: The MIT Press.

Table 1: Descriptive statistics: distribution of the sample

Number of events	6,973	Schedule	
Overall no. of observations	302,576	individual first[(b)]	79,564
(of which)	(144,130)	relay first[(c)]	81,671
Olympic games	660	same day	141,341
World championships	1,531	Round (individual comp.)	
European championships	716	timed finals (default)	136,917
National championships[(a)]	136,785	finals	75,290
Pan Pacific games	390	semi-finals	2,620
Commonwealth games	184	preliminaries	87,489
Universiades	304	others[(d)]	260
Nationalities	142	Round (relay)	
Gender		timed finals (default)	193,881
male	150,243	finals	94,280
female	152,333	preliminaries	13,847
Age: mean [min-max]	17.8 [6-109]	others[(d)]	568
FINA points: mean [min-max]	502.8 [5-1181]	Order (relay)	
Style		1st	64,481
50m Breaststroke	20,092	2nd	86,841
50m Fly	17,599	3rd	78,358
50m Freestyle	99,081	4th	72,896
100m Breaststroke	28,588		
100m Fly	25,096		
100m Freestyle	91,079		
200m Freestyle	21,041		

(a) Includes major regional championships.

(b) Day of individual competition before day of relay

(c) Day of individual competition after day of relay

(d) Swim-off after semi-finals, swim-off after preliminaries

Table 2: Average FINA points by starting order

Starting order	FINA points	std.dev.
1	486.8	160.5
2	450.3	161.5
3	441.9	157.6
4	474.1	161.1

Table 3: Predicting FINA points based on starting order (OLS regression)

Dependent variable	FINA points			
	complete teams		all swimmers	
age <15	-152.3	***	-175.2	***
age >30	-125.4	***	-161.1	***
major1[(a)]	-6.4		20.6	***
major750[(b)]	223.1	***	206.0	***
major2[(c)]	25.4	***	32.1	***
female	42.0	***	25.6	***
same day	-36.3	***	-39.1	***
relay first	-11.6	***	-13.7	***
50m Breaststroke	-86.1	***	-31.9	***
100m Breaststroke	-109.0	***	6.6	***
50m Fly	-110.0	***	-31.5	***
100m Fly			-3.8	***
50m Freestyle	-75.9	***	-67.0	***
200m Freestyle	97.1	***	52.9	***
swimmer 1	588.8	***	596.6	***
swimmer 2	556.7	***	573.6	***
swimmer 3	549.9	***	561.3	***
swimmer 4	578.3	***	592.8	***
obs.	17,532		302,576	
R2	0.95		0.95	
test sw.1 = sw.2	168.3	***	661.0	***
test sw.2 = sw.3	8.12	***	195.0	***
test sw.3 = sw.4	135.2	***	1,276.0	***

(a) Olympic, Pan Pacific and Commonwealth Games, World and European Championships, Universiades.

(b) *major1* events plus national and major regional championships.

(c) Average FINA points equal or above 750.

*** $p < .01$, ** $p < .05$, * $p < .10$.

Table 4: Descriptive statistics: individual and relay swimming time

	Swimming times					
	Individual	Relative difference b/w relay and individual				
		1st swimmers		2nd to 4th swimmers		
	(secs.)	(%)				
	I	$\frac{R_1-I}{I}$		$\frac{R_2-I}{I}$	$\frac{R_3-I}{I}$	$\frac{R_4-I}{I}$
Overall	55.97	.21	***	-.33	-.34	-.77
(std.dev.)	(27.15)	(1.34)		(1.48)	(1.58)	(1.64)
Gender						
male	53.80	.20	***	-.42	-.42	-.84
female	58.12	.23	***	-.25	-.26	-.70
Age (yrs)						
< 15	55.03	.36	***	-.03	-.02	-.48
$15 - 30$	58.56	.11	***	-.48	-.51	-.94
> 30	38.10	.31	***	-.58	-.52	-1.00
Ability of swimmer$^{(a)}$						
$\leq$ median	52.63	.26	***	-.19	-.17	-.61
$>$ median	59.31	.17	***	-.48	-.53	-.93
Style						
50m Breaststroke	39.67	.44	***	-.51	-.12	-.25
50m Fly	33.58	.62	***	-.10	-.48	-.32
50m Freestyle	32.05	.32	***	-.70	-.71	-1.15
100m Breaststroke	78.30	.32	***	-.05	-.05	-.22
100m Fly	67.89	.05		-.53	-.04	-.47
100m Freestyle	62.87	.14	***	-.31	-.32	-.46
200m Freestyle	128.54	.07	***	-.03	-.08	-.04
Event importance						
major1 events$^{(b)}$	71.06	.03		-.54	-.57	-.91
major750 events$^{(c)}$	63.12	.04	***	-.67	-.75	-1.12
major2 events$^{(d)}$	60.89	.13	***	-.41	-.42	-.84
others	55.66	.22	***	-.33	-.34	-.75
Schedule of competitions						
individual first$^{(e)}$	54.93	.19	***	-.39	-.41	-.99
relay first$^{(f)}$	60.00	.14	***	-.49	-.50	-.77
same day	54.23	.27	***	-.22	-.21	-.66

I - individual competition swimming time.

$R_1, \cdots, R_4$ - relay swimming time, starting order $1, \cdots, 4$.

T-test, $H_0 : \frac{R_1-I}{I} = 0$; *** $p < .01$

(a) As measured by FINA points.

(b) Olympic, Pan Pacific and Commonwealth Games, World and European Championships, Universiades.

(c) Average FINA points equal or above 750.

(d) Major1 events plus national and major regional championships.

(e) Day of individual competition before day of relay.

(f) Day of individual competition after day of relay.

Table 5: Estimated absolute difference in swimming times between the relay and the individual competition, fourth swimmers.

Event type	$R_4 - I$	std.dev.	obs.
Olympic games	-0.66	0.06	167
World championships	-0.54	0.05	357
European championships	-0.66	0.06	174
major1	-0.59	0.03	906
major750	-0.55	0.01	3,870
major2	-0.38	0.01	33,236
all events	-0.34	0.01	72,896

I - individual competition swimming time.

R_4 - relay swimming time, fourth swimmers.

The table reports the OLS estimated intercepts of a linear regression model of $R_4 - I$, controlling for age, gender, style and length of the competition, estimated on different subsamples. The reference category is 100m Freestyle, male swimmers aged 15-30. Standard errors clustered at the individual level.

Table 6: Regression results, first swimmers, full sample

Dependent variable	$(R-I)/I$			
	OLS		FE	
age <15	0.19	***		
age >30	0.11	***		
major1[(a)]	0.04		-0.07	
major750[(b)]	-0.03	*	-0.13	***
major2[(c)]	-0.07	***	-0.08	***
50m Breastroke	0.19	***	0.33	**
100m Breastroke	0.20	***	0.27	*
50m Fly	0.36	***	0.24	
100m Fly	-0.03		-0.06	
50m Freestyle	0.12	***	0.04	**
200m Freestyle	-0.02	**	-0.04	**
female	-0.01			
same day	0.04	***	0.07	***
relay first	-0.01		0.00	
swimmer 1	**0.10**	***	**0.20**	***
obs.		64,481		
swimmers		36,796		
R2	0.01		0.67	

(a) Olympic, Pan Pacific and Commonwealth Games, World and European Championships, Universiades.
(b) Average FINA points equal or above 750.
(c) Major1 events plus national and major regional championships.
Reference category: Male swimmers, 100m freestyle, individual competition on a day prior to the relay.
Standard errors clustered at the individual level.
*** $p < .01$, ** $p < .05$, * $p < .10$.

Table 7: Regression results, first swimmers, filtered sample (15-30 age group, freestyle finals)

Dependent variable	$(R-I)/I$															
	Filter: 15-30 age group, freestyle finals															
	all events				major2 events$^{(c)}$				major750 events$^{(b)}$				major1 events$^{(a)}$			
	OLS		FE		OLS		FE		OLS		FE		OLS		FE	
major1$^{(a)}$	0.05		-0.02													
major750$^{(b)}$	-0.01		-0.14	***												
major2$^{(c)}$	0.00		-0.02													
50m Breastroke																
100m Breastroke																
50m Fly																
100m Fly																
50m Freestyle	0.05	***	-0.01		-0.04		-0.12	***	0.07		-0.04					
200m Freestyle	-0.09	***	-0.08	***	-0.11	***	-0.12	***	-0.06		-0.01		0.11		0.17	
female	0.06	***			0.07	***			0.16	***			-0.16	*		
same day	-0.06	***	-0.06	*	-0.02		-0.00		-0.08		-0.16		-0.11		-0.40	***
relay first	-0.01		0.00		-0.03		-0.05		-0.05		-0.00		-0.07		0.41	
swimmer 1	**0.23**	***	**0.29**	***	**0.24**	***	**0.30**	***	**0.18**	***	**0.27**	***	**0.28**	***	**0.15**	**
obs.	23,150				10,925				1,305				138			
swimmers	13,694				6,874				760				114			
R2	0.00		0.67		0.00		0.71		0.01		0.57		0.06		0.86	

(a) Olympic, Pan Pacific and Commonwealth Games, World and European Championships, Universiades.

(b) Average FINA points equal or above 750.

(c) Major1 events plus national and major regional championships.

Reference category: Male swimmers, 100m freestyle, individual competition on a day prior to the relay.

Standard errors clustered at the individual level.

*** $p < .01$, ** $p < .05$, * $p < .10$.

Table 8: Regression results, first swimmers, filtered sample, major events

Dependent variable	$(R-I)/I$							
	Filter: 15-30 age group, freestyle finals, major2 events[(a)]							
	first 4[(b)]		individual first		same day		relay first	
	OLS	FE	OLS	FE	OLS	FE	OLS	FE
50m Freestyle	0.03	-0.02	0.03	0.21	-0.04	-0.09	-0.08	-0.28 ***
200m Freestyle	-0.14 ***	-0.25 ***	-0.10 **	-0.11	-0.14 ***	-0.20 ***	-0.09 ***	-0.05
female	0.10 **		0.18 ***		0.06 *		0.04	
same day	0.04	0.03						
relay first	-0.01	-0.00						
swimmer 1	**0.35** ***	**0.42** ***	**0.17** ***	**0.20** ***	**0.23** ***	**0.29** ***	**0.23** ***	**0.27** ***
obs.	2,833		2,148		4,565		4,212	
swimmers	1,912		1,718		3,566		2,991	
R2	0.00	0.76	0.01	0.87	0.00	0.84	0.00	0.78

(a) Olympic, Pan Pacific and Commonwealth Games, World and European Championships, Universiades, national and major regional championships.
(b) Swimmers finishing in the first four positions both in the individual and in the relay competition. Reference category: Male swimmers, 100m freestyle, individual competition on a day prior to the relay.
Standard errors clustered at the individual level.
*** $p < .01$, ** $p < .05$, * $p < .10$.

Table 9: Regression results, swimmers 2-4, full sample

Dependent variable	$(R-I)/I$			
	OLS		FE	
age <15	0.56	***		
age >30	0.36	***		
major1[(a)]	0.09	***	-0.04	
major750[(b)]	-0.33	***	-0.22	***
major2[(c)]	-0.17	***	-0.14	***
50m Breastroke	-0.41	***	-0.68	***
100m Breastroke	0.20	***	0.18	***
50m Fly	-0.34	***	-0.55	***
100m Fly	0.27	***	0.30	***
50m Freestyle	-0.75	***	-1.08	***
200m Freestyle	0.48	***	0.55	***
female	0.09	***		
same day	0.18	***	0.08	***
relay first	-0.04	***	-0.07	***
swimmer 2	**-0.42**	***	**0.04**	***
swimmer 3	**-0.44**	***	**-0.03**	***
swimmer 4	**-0.69**	***	**-0.26**	***
obs.		238,095		
swimmers		90,696		
R2	0.17		0.54	
test sw.2 = sw.3	5.0	**	32.3	***
test sw.3 = sw.4	600.0	***	351.0	***

(a) Olympic, Pan Pacific and Commonwealth Games, World and European Championships, Universiades.
(b) Average FINA points equal or above 750.
(c) Major1 events plus national and major regional championships.
Reference category: Male swimmers, 100m Freestyle, individual competition on a day prior to the relay.
Standard errors clustered at the individual level.
*** $p < .01$, ** $p < .05$, * $p < .10$.

Table 10: Regression results, swimmers 2-4, filtered sample (15-30 age group, freestyle finals)

Dependent variable	$(R-I)/I$															
	Filter: 15-30 age group, freestyle finals															
	all events				major2 events[c]				major750 events[b]				major1 events[a]			
	OLS		FE		OLS		FE		OLS		FE		OLS		FE	
major1[a]	-0.29	***	-0.32	***												
major750[b]	-0.35	***	-0.18	***												
major2[c]	-0.15	***	-0.09	***												
50m Freestyle	-0.90	***	-1.24	***	-1.02	***	-1.37	***	-1.33	***	-1.29	***				
200m Freestyle	0.42	***	0.61	***	0.42	***	0.61	***	0.72	***	0.78	***	1.04	***	1.06	***
female	0.19	***			0.20	***			0.32	***			0.41	***		
same day	0.15	***	0.02		0.13	***	-0.02		0.08		0.00		-0.11		-0.68	***
relay first	0.05	***	-0.02		0.03		-0.12	***	-0.08		-0.14	*	-0.10		-0.35	*
swimmer 2	**-0.31**	***	**-0.05**	***	**-0.46**	***	**-0.18**	***	**-0.81**	***	**-0.68**	***	**-1.41**	***	**-0.96**	***
swimmer 3	**-0.33**	***	**-0.11**	***	**-0.46**	***	**-0.21**	***	**-0.88**	***	**-0.72**	***	**-1.41**	***	**-1.04**	***
swimmer 4	**-0.54**	***	**-0.30**	***	**-0.66**	***	**-0.39**	***	**-1.04**	***	**-0.84**	***	**-1.51**	***	**-1.18**	***
obs.	45,033				22,035				2,063				320			
swimmers	23,598				12,452				1,200				207			
R2	0.24		0.68		0.28		0.73		0.50		0.77		0.41		0.83	
test sw.2 = sw.3	1.08		3.69	*	0.00		0.99		0.90		0.11		0.00		0.06	
test sw.3 = sw.4	141.4	***	63.3	***	80.6	***	29.5	***	6.58	**	1.21		0.50		0.22	

(a) Olympic, Pan Pacific and Commonwealth Games, World and European Championships, Universiades.
(b) Average FINA points equal or above 750.
(c) Major1 events plus national and major regional championships.
Reference category: Male swimmers, 100m freestyle, individual competition on a day prior to the relay.
Standard errors clustered at the individual level.
*** $p < .01$, ** $p < .05$, * $p < .10$.

Table 11: Regression results, swimmers 2-4, filtered sample, major events

Dependent variable	$(R-I)/I$															
	Filter: 15-30 age group, freestyle finals, major2 events[(a)]															
	first 4[(b)]				individual first				same day				relay first			
	OLS		FE		OLS		FE		OLS		FE		OLS		FE	
50m Freestyle	-1.08	***	-1.39	***	-1.23	***	-1.43	***	-0.85	***	-1.13	***	-1.14	***	-1.55	***
200m Freestyle	0.43	***	0.65	***	0.43	***	0.74	***	0.38	***	0.64	***	0.43	***	0.69	***
female	0.21	***			0.35	***			0.18	***			0.17	***		
same day	0.15	***	0.03													
relay first	-0.05		-0.18	**												
swimmer 2	**-0.30**	***	**-0.11**	*	**-0.44**	***	**-0.26**	**	**-0.30**	***	**-0.20**	***	**-0.47**	***	**-0.39**	***
swimmer 3	**-0.29**	***	**-0.09**		**-0.43**	***	**-0.27**	**	**-0.33**	***	**-0.24**	***	**-0.45**	***	**-0.34**	***
swimmer 4	**-0.57**	***	**-0.32**	***	**-0.70**	***	**-0.51**	***	**-0.61**	***	**-0.46**	***	**-0.53**	***	**-0.42**	***
obs.	5,386				4,079				9,807				8,149			
swimmers	3,258				3,212				6,885				5,534			
R2	0.27		0.75		0.36		0.88		0.25		0.82		0.30		0.80	
test sw.2 = sw.3	0.01		0.13		0.03		0.01		0.77		0.40		0.31		0.65	
test sw.3 = sw.4	33.3	***	9.75	***	24.3	***	3.91	*	60.0	***	11.8	***	4.86	**	1.46	

(a) Olympic, Pan Pacific and Commonwealth Games, World and European Championships, Universiades, national and major regional championships.

(b) Swimmers finishing in the first four positions both in the individual and in the relay competition. Reference category: Male swimmers, 100m freestyle, individual competition on a day prior to the relay.

Standard errors clustered at the individual level.

*** $p < .01$, ** $p < .05$, * $p < .10$.

Table 12: Regression results, swimmers 1-4, filtered sample, major events

Dependent variable	$(R-I)/I$			
Filter: 15-30 age group, freestyle finals, major2 events[a]				
	OLS		FE	
50m Freestyle	-0.70	***	-0.93	***
200m Freestyle	0.18	***	0.33	***
female	0.16	***		
same day	0.07	***	0.01	
relay first	0.01		-0.09	***
swimmer 1	**0.27**	***	**0.44**	***
swimmer 2	**-0.69**	***	**-0.70**	***
swimmer 3	**-0.71**	***	**-0.76**	***
swimmer 4	**-0.93**	***	**-0.96**	***
obs.		32,960		
swimmers		16,264		
R2	0.15		0.66	
test sw.2 = sw.3	0.18		2.88	*
test sw.3 = sw.4	94.6	***	42.0	***

(a) Olympic, Pan Pacific and Commonwealth Games, World and European Championships, Universiades, national and major regional championships.
Reference category: Male swimmers, 100m freestyle, individual competition on a day prior to the relay.
Standard errors clustered at the individual level.
*** $p < .01$, ** $p < .05$, * $p < .10$.

Table 13: Average lag over a preceding team

	all	filtered
	first 3 positions	
Males		
50m freestyle	1.7%	1.6%
100m freestyle	1.4%	1.5%
200m freestyle	0.9%	1.0%
Females		
50m freestyle	1.7%	2.0%
100m freestyle	1.7%	1.9%
200m freestyle	1.1%	1.1%
	first 10 positions	
Males		
50m freestyle	0.5%	0.9%
100m freestyle	0.6%	0.7%
200m freestyle	0.3%	0.5%
Females		
50m freestyle	0.8%	1.0%
100m freestyle	0.7%	0.9%
200m freestyle	0.3%	1.0%

Filter: 15-30 age group, freestyle finals, major2 events