How success breeds success

Ambroise Descamps, Changxia Ke, and Lionel Page*

School of Economics and Finance, Queensland University of Technology, and QuBE

Abstract

We study whether and how success increases the chance of subsequent success using a real-effort laboratory experiment. We identify the causal effect of winning in a simple dynamic contest (best-of-three) using the random component of a stochastic contest success function that determines the winner of each round. We find a positive effect of an initial success on subsequent performance. Replacing either the first round or the last round of the contest with a die selecting the winner at random, we disentangle two competing explanations of the positive effect: strategic thinking and psychological effect of winning. Our results clearly support the existence of a psychological effect of winning. On the contrary, we do not find evidence that strategic thinking can explain the effect of winning. Varying the amount of feedback provided in contest, we find that the psychological effect is likely driven by improved self-confidence after experiencing a success. We suggest that contest models need to venture beyond the framework of games with complete information to explain behaviour in dynamic contests.

*Descamps: ambroise.descamps@qut.edu.au; Ke: changxia.ke@qut.edu.au; Page: lionel.page@qut.edu.au.
For detailed and constructive comments we gratefully thank Ralph Bayer, Jeff Borland, Patrick DeJarnette, Ming He, Ningbo Huang, Boon Han Koh, Steffen Huck, Joseph Tao-yi Wang, Tom Wilkening, Jingjing Zhang, Jun Zhang. We also benefited from comments from seminar and conference participants at the 2016 Economics and Biology of Contest Conference (QUT, Brisbane), 2016 Australia New Zealand Workshop on Experimental Economics (UQ, Brisbane), GATE Lyon Saint-Etienne and the 2018 Asia Pacific Economics Science Association conference (Brisbane).

1 Introduction

One success could be critical in triggering a string of subsequent successes. Anecdotes abound of high achievers (in business, sports or academia) describing how one critical success paved the way for what they became (Robertson, 2012). However, whether success breeds success is debated and even if it is true, different explanations could be driving it. Empirical studies have found mixed results (Ferrall and Smith Jr, 1999; Tong and Leung, 2002; Fu et al., 2015a) though recent evidence tends to support the existence of a positive effect of winning on later performances (Malueg and Yates, 2010; Mago et al., 2013; Van de Rijt et al., 2014; Miller and Sanjurjo, 2014; Gauriot and Page, 2015; Cohen-Zada et al., 2017).

In this paper, we investigate this question experimentally using variations of a simple dynamic (best-of-three) contest where players compete on a real-effort task. The advantage of our design is twofold. First, using a stochastic contest success function, we cleanly identify the causal effect of an early success from the random component of the success function that determines the winner in each round. Second, we toggle on and off different features of the contest in a way that selectively eliminates one possible explanation at a time. Our results are striking. We find clear evidence for a positive effect of winning. However, there is no evidence of the strategic effect suggested by standard economic models. The psychological momentum effect we find is likely driven by improved self-confidence after experiencing an initial success.

In many instances, not everyone can be a winner. It is a truism in social competitions such as job applications, internal promotions, political races, competing exams, grant applications, lawsuits, sports, and R&D races. Even in less open competitions, whenever everyone strives to succeed, success is typically measured by how people compare to others (e.g. fame). Understanding what drives success is key to identify where socio-economic inequalities arise. To a large extent, existing inequalities are driven by prior differences in economic and human capital. However, even among people with similar advantages, substantial income differences can arise. People having similar starting points may see their chances of economic success diverge as differences in initial success lead to different opportunities. An easy explanation is that assets such as wealth, recognition, social networks grow after an initial success and give an advantage to the initial winner later on. Another possible factor is that success can act on people's mindset and motivation in a way which fosters later successes. Which one plays a more important role has long been debated, mainly because identifying the causal effect of a success on subsequent performance is challenging, due to an obvious endogeneity problem. Early and later successes are influenced by individual traits which are imperfectly observed. Early success tends therefore to be associated with later successes as it signals unobserved traits associated with high performance.

Even if one could cleanly identify the causal effect of being successful, disentangling the competing explanations are nontrivial. Broadly, two different types of explanations co-exist. First, Harris and Vickers (1987) have shown in their seminal paper, that rational players should not play competitions taking place over time as a series of standalone sub-contests. In particular, they identify the existence of a "strategic momentum effect" which arises from the asymmetry of incentives between past winners and losers. This result has later been replicated across a range of contests (Ferrall and Smith Jr, 1999; Konrad and Kovenock, 2006; Klumpp and Polborn, 2006; Konrad and Kovenock, 2009; Malueg and Yates, 2010; de Roos and Sarafidis, 2015; Krumer et al., 2017; Gauriot and Page, 2015). Whenever an early success gets you closer to a final prize, the incentives to perform again are higher than if you did not succeed. On the contrary, an initial failure means that a player may have first to catch up in order to secure future successes. Catching up is costly (involving effort and resources) without bringing rewards in itself, therefore, it decreases the incentives of the laggard. A strategic effect can arise from winning because, by backward induction, the players' *future* expected rewards from the competition are asymmetric as a function of their initial success. This asymmetry of incentives between early winners and losers can lead to differences in motivation to expend effort and resources and therefore in later chances of successes.

Beside the strategic effect of winning, there is a competing behavioural explanation whereby psychological factors influence the competing agents after an initial success. In this view, *past* performance has a direct causal effect, for instance influencing self-confidence. The idea of "psychological momentum" is a robust element of folk psychology with success being perceived as increasing later performance (Markman and Guenther, 2007). There is a substantial literature in psychology supporting the existence of this effect (Iso-Ahola and Mobily, 1980; Taylor and Demick, 1994; Markman and Guenther, 2007). However the empirical evidence on such a phenomenon is still debated (Bar-Eli et al., 2006).

These two approaches propose radically different frameworks to explain the effect of an early success. The mechanism underlying economic models is clear. However, their ecological validity is limited by their restrictive assumptions. Since Harris and Vickers (1987)'s study, most models of dynamic contests assume complete information and homogeneous players which imperfectly capture complex real world competitions. In comparison, the behavioural approach provides appealing intuitions, but it lacks a clear theoretical framework which articulates underlying mechanisms.

Our experimental design aims to study the existence and origin of an effect of winning on later performance. Doing so requires two things: cleanly identifying the causal effect of winning and disentangling between its different possible explanations. We solve the identification problem using random variations in winning generated in a controlled laboratory experiment. We use the workhorse design of a best-of-three game as a simplified dynamic contest (Konrad and Kovenock, 2009; Mago et al., 2013; Dechenaux et al., 2015; Fu et al., 2015a), and we implement each round as a real effort contest with a Tullock contest success function (Tullock, 1980). We then estimate the causal effect of winning by matching winners and losers with identical winning probabilities. Conditional on these winning probabilities, the success in one round is entirely exogenous. We find a clear positive effect of winning (vs losing) early in the contest on the performance in the following round effort task. We observe this effect both at the intensive margin (productivity) and at the extensive margin (time worked).

Having established the existence of a causal effect of winning, we disentangle the two explanations with modified versions of the best-of-three contest that turn on and off the driving factors of each explanation. We use the fact that strategic and psychological explanations are predicted to arise from different sources of information: information about the *future* (strategic effect) and information about the *past* (psychological effect). We replace either the first round or the last round of the best-of-three contest with strategically neutral rounds (where the winner is decided by the throw of a dice) to eliminate this information. Replacing the last round with a neutral round, we eliminate the cost of effort in the last round which drives the asymmetric incentives and strategic momentum by backward induction. Similarly, replacing the first round of the contest with a neutral round, we eliminate the source of a psychological momentum as the initial success is not driven by players' performances and winning itself contains no information about the players' relative strength. Surprisingly, we find no evidence of a strategic effect of winning, but instead find clear evidence of a psychological effect.

We further investigate the mechanisms behind the psychological effect with an additional treatment: By giving more information to players about their performance, we reduce the marginal informational content of a win. We find that the effect of winning stops being significant when players do not learn anything about their own performance from winning early. The psychological effect of winning seems therefore driven by the gain in self-confidence a player enjoys after having experienced an initial success. This last result suggests that understanding behaviour in competition may require a departure from models with complete information which leave out the role of players' beliefs in their relative strength.

The remainder of the paper is as follows: Section 2 describes the simplified dynamic contest we are considering as well as our experimental design; Section 3 presents our identification strategy; Section 4 describes our results; Section 5 discusses the origins of the effect we observe and Section 6 concludes.

2 A simplified dynamic contest experiment

We set up a dynamic contest experiment using a best-of-three contest. We present here a description of the strategic features of this game. Extended discussions and equilibrium analysis of best-of-n contests can be found in Konrad and Kovenock (2009) and Fu et al. (2015b).

2.1 Best-of-three contest: conceptual framework

Consider a game of complete information with players who are payoff maximising and homogenous in terms of ability. They compete over (up to) three rounds. The player winning two rounds wins a prize V. We denote e_{it} the effort exerted by player i ($i \in \{A; B\}$) in round t ($t \in \{1, 2, 3\}$), and $c(e_{it}) > 0$ the associated cost.

The winner of a round is determined according to a contest success function, which assigns a probability of success, depending on a player and his opponent's efforts. Let $p_i(e_{it}; e_{-it})$ be this function. Figure 1 represents the structure of such a contest.[1]

[1] The numbers given in brackets in Figure 1 are the score of player A versus the score of player B at each point in time. Starting from a score of 0:0 at the beginning of the contest, a player can increase his score by one after winning a round. The game ends with one of the four potential outcome (namely, 2:0, 2:1, 1:2, or 0:2.)

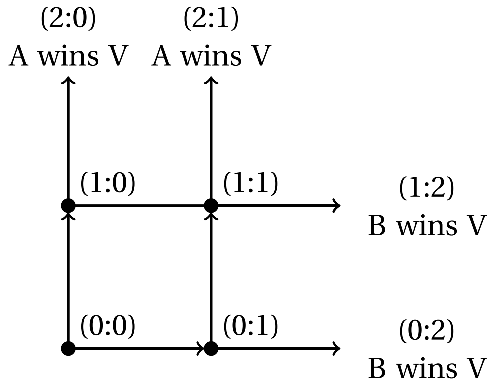


Figure 1: Representation of a best-of-three game.

The equilibrium strategy of this game is found by backward induction. In round 3, the players face symmetric incentives. Whenever there is a symmetric equilibrium to the game (which is the case for standard contest functions) both players will have the same expected equilibrium level of effort e_3^* which induces an expected level of effort cost c_3^*.[2] By symmetry, they have the same expected chance of winning the last round. In equilibrium, the expected payoff of player i when reaching the third round is therefore $v_3 = p(e_{i3}, e_{-i3})V - c(e_{i3}) = V/2 - c_3^*$.

In round 2, one of the two players has already won the first round. Without loss of generality, let's assume that it is Player A. In case of success in round 2, A gets a value of V. If A is not successful he gets v_3, the expected payoff when entering the third round. In comparison, B gets a value of v_3 in case of success in the second round and 0 otherwise. The two players' incentives to exert effort only depend on the difference in expected payoffs between winning and losing in round 2 (i.e., the effective prize spread). For A, the incentive is $\delta_A = V - v_3 = V/2 + c_3^*$ and for B it is $\delta_B = v_3 = V/2 - c_3^*$. The incentive to exert effort is greater for the leading contestant (A), than it is for the lagging one (B): $\delta_A > V/2 > \delta_B$.

This asymmetry in incentives generates a strategic effect of winning whereby the player winning the first round expends more effort in the second round than the player who lost. Such an asymmetry has been found in a wide range of dynamic contests leading there to a similar strategic effect of winning (Konrad, 2009)

In comparison, the psychological effect is not supported by a single and well defined conceptual framework (Taylor and Demick, 1994; Markman and Guen-

[2] If the equilibrium is in mixed strategy, this expected level of effort represents the expectation of the equilibrium mixture over the different levels of possible effort.

ther, 2007). It relies on the idea that winning, in itself, can affect behaviour and future performances. Cohen-Zada et al. (2017) describe this idea of psychological momentum in contests as "the tendency for an outcome to be followed by a similar outcome not caused by any strategic incentive of the players." Descriptions of this effect typically involve an increase in efficacy (the ability to transform effort in a good performance), motivation and or aggressiveness. This lack of clear theoretical underpinning and the associated uncertainty about the precise mechanisms underlying such an effect has been stressed by economists investigating the reality of such an effect (Mago et al., 2013; Gauriot and Page, 2015; Cohen-Zada et al., 2017).

2.2 Real effort competition

To implement the best-of-three contest, in each round, players compete on an real-effort task. The task is inspired by Huck et al. (2015). Participants observe on their computer screen a block of 20 characters (numbers, lower and upper cases letters), and have to enter it backwards in a text box.[3] Each time a block is correctly entered, a new one appears on the screen. In each round, players have 10 minutes to enter as many blocks as possible. The higher the number of blocks they complete correctly, the more likely they are to win that round.

Huck et al. (2015)'s results suggest that this task is unpleasant enough for participants to adjust their effort when the reward associated with entering an extra block varies. In addition, we follow Gächter et al. (2016) and add to this subjective opportunity cost of effort a monetary opportunity cost.[4] Participants are endowed with $3 per round, and have to pay half a cent for each second they spend completing the task.[5] At any time they have the opportunity to hit a "stop" button to stop working on the task. They can then keep whatever is left from their endowment if they decide not to spend all the endowment on the effort task.[6]

[3] To avoid confusion, characters that are too similar were not included, such as capital "o" and zeros. Participants need to give the correct answer before a new block is displayed.

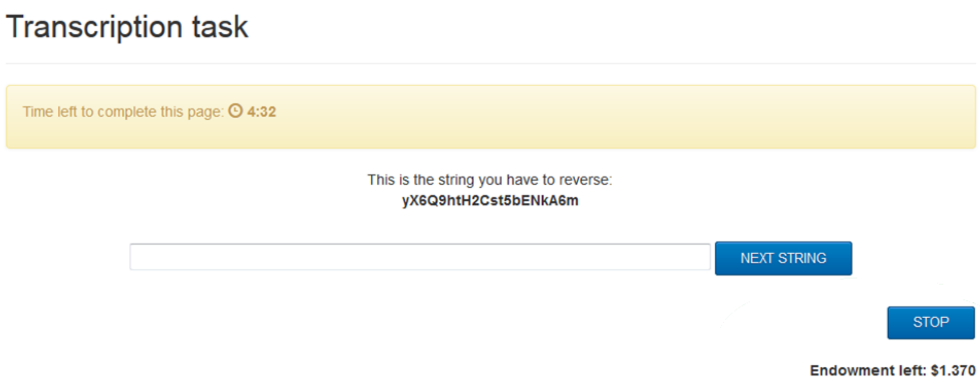
[4] Real effort tasks have been criticized on the grounds that there is not much opportunity cost to them (Erkal et al., 2016). It has been suggested that once participants are is in the lab, they have nothing to do but complete the tasks and they try as hard as possible at all time, regardless of incentives (Araujo et al., 2016). Our design addresses this problem by choosing an effort task which has proven to be painful enough for participants to react to incentives and by creating an additional monetary opportunity cost.

[5] All amounts are in Australian dollars. If a winner occurs after two rounds, both players keep the full amount of endowment for the last round since no one need to work on the effort task.

[6] The experimental instructions in Appendix D provide the exact phrasing of the explanation of this procedure to the participants. The effectiveness of the "stop" button as a way to induce

Figure 2 shows a screen-shot of the effort task in our experiment.

Figure 2: A screen-shot of the effort task faced by participants.



2.3 Experimental design and testable hypotheses

In order to disentangle the strategic and psychological effects of winning, we design three variations of the experimental best-of-three contest.

The first one is our baseline treatment (called *Baseline*). It is a standard best-of-three contest. The timing of *Baseline* is depicted in the top left corner of Figure 3. Before starting the contest, the participants enter a piece-rate round, without the "stop" button. This piece-rate round lasts 10 minutes and each correct block is rewarded with 20cts. The participants are then informed they will play a contest.

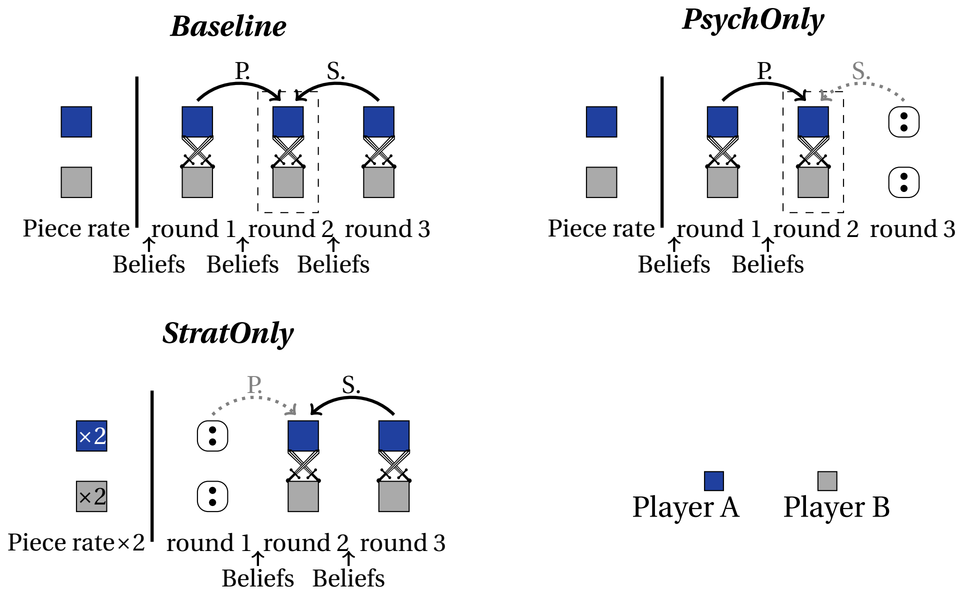
At the beginning of the contest, participants are told that they will be placed in pairs of similar ability (the first ranked at the piece rate round with the second, and so on). It is the only information they receive about their opponent.[7] The purpose of the piece-rate round is to pair the participants by ability in the contest and for it to be common knowledge. Our experimental design follows therefore closely the set up of a game of complete information with homogeneous players as described in Section 2.1.

Players' probability to win a round increases with their effort and decreases with the effort of their opponent. This probability is determined by a stochastic

opportunity cost is supported by the study of Erkal et al. (2016).

[7]The information about the contest and the pairing is given only after the piece-rate round to ensure that participants do not play strategically in the piece rate round. We follow here the approach of Fu et al. (2015a).

Figure 3: Representation of the different experimental treatments. The"P" stands for psychological effect, and "S" stands for strategic effect."Beliefs" with an arrow represents the point in time when participants' beliefs about their winning probabilities are elicited.



contest function which links players efforts to a winning probability. We use the Tullock contest function, the most widely used in the literature (Dechenaux et al., 2015). With such a function, the probability to win for player i conditional on his own and his opponent's efforts in the first round (e_{i1} and e_{j1} respectively) is:

$$p_{i1} = \frac{e_{i1}}{e_{i1} + e_{j1}} \tag{1}$$

Before each round, participants are also asked their beliefs about their chance to win the next round. We use here a simple non incentivised question: "According to you, how likely are you to win the next round (in %)". It has been shown that simply asking participants their subjective beliefs works well as an elicitation method (Hollard et al., 2016). After each round, participants are informed of the winner of that round. Neither the actual number of blocks completed nor the winning probability of each player is revealed. In the real world, information is often limited to outcomes with actual levels of effort being unobserved. Our experimental set up reflects such a situation. Furthermore, in a game of complete information with homogeneous players, the information on the performance of the opponent only signals his/her level of effort and this information has no strategic value in equilibrium. We change this informational structure in a complementary treatment discussed in Section 5.

Let's define the causal effect of winning as:

Definition 1 ***Effect of winning.*** *There is a causal effect of winning on performance in a dynamic contest if a player displays a greater performance after a win than in the counter-factual situation where he/she does not win.*

The *Baseline* treatment is designed to test whether winning has any effect on later performance. By using a stochastic contest function in each round we can use a matching approach to estimate whether players with a similar chance of winning the first round have a higher or lower performance in the second round after winning or losing the first round (see Section 3).

Hypothesis 1 ***Positive effect of winning.*** *A positive effect of winning will be observed in the Baseline treatment.*

This hypothesis reflects the widely held idea that winning can have a postive effect on performance. If an effect is observed in the baseline treatment, it is impossible to tell whether it is psychological or strategic. Indeed, both effects point in the same direction, namely the first-round winner performs better in the second round relative to his/her opponent. To disentangle the strategic and psychological effects we use a key difference in the sources of these effects. The strategic effect arises from backward-induction, and therefore from expectations about the *future* evolution of the game. Instead, the psychological momentum assumes a direct effect of *past* performance on future motivation and performance. We use this difference to identify which effect is prevalent by alternatively toggling on and off the information players can derive from the past and the future of the game.

Our second treatment (*PsychOnly*) is designed to identify the psychological effect by turning the strategic effect off. As mentioned in Section 2.1, the prospect of effort cost expended in the third round is the cause of the strategic effect. If no effort is to be exerted in the last round, the expected value of going to the last round is equal to $V/2$ for both players. As a consequence, the round-1 winner's incentive in round 2 (i.e., $V - V/2$) is the same as round-1 loser's incentive in round 2 (i.e., $V/2 - 0$). We eliminate effort in round 3 by using a random device, to select the outcome.

The players are randomly assigned to either odd or even numbers, and a computer simulated dice determines the winner. The top right corner of Figure 3 shows the timing of the *PsychOnly* treatment. In this scenario, if a player is more likely to win round 2 after winning in round 1, it cannot be attributed to standard strategic considerations. Note, that the game structures are different across the two conditions with players expecting to play two more rounds after a win in *Baseline* and only one more round in *PsychOnly*. Therefore, we do not make any prediction on a psychological effect being necessarily equal in the *Baseline* and in *PsychOnly*. But if a psychological effect of winning exists, we would expect to observe an effect of winning in *PsychOnly*.

Hypothesis 2 ***Psychological effect.*** *There is a psychological effect of winning. Therefore, we will observe an effect of winning in PsychOnly even though the strategic effect is turned off.*

Conversely, the third treatment (*StratOnly*) is designed to isolate the strategic effect. We deactivate the psychological effect by randomly assigning a winner in round 1, thus randomly affecting the history of the game. This is achieved using the same dice procedure as in *PsychOnly*. As a consequence, the first round does not provide successful players with the experience of winning as a consequence of their performance. Most discussions about a psychological effect of winning assume that this effect arises from past performance itself. This possible mechanism is turned off in *PsychOnly*. One cannot exclude that even an explicitly random success unrelated to any performance could have a psychological effect. In that sense our treatment does not fully turn off all possibilities of a residual psychological effect but rather its most commonly assumed versions. Observing a causal effect of winning in the treatment *StratOnly* would reveal that this effect cannot be explained as a commonly assumed version of a psychological effect.

Hypothesis 3 ***Strategic effect.*** *There is a strategic effect of winning. Therefore, we will observe an effect of winning in StratOnly even though the commonly assumed source of psychological effect is turned off.*

In *StratOnly* participants do not exert effort in the round 1. To avoid differences in fatigue in the second round relative to the other treatments, participants play two piece-rate rounds before the contest. The bottom left panel of Figure 3 displays the timing of this treatment.

2.4 Data collection

The experimental sessions were conducted in a large Australian University between March and April 2017. We ran 11 sessions, and each was around 75 minutes long, including instructions and payments. The same experimenter ran all the sessions. The experiment was programmed in oTree (Chen et al., 2016). Participants were recruited from various faculties using ORSEE (Greiner, 2015). In total, 146 students took part in the experiment (mean age 21, 54% males, N=50 in *Baseline*, N=48 in *PsychOnly* and N=46 in *StratOnly*).[8] The average payment is composed of a $5 show-up fee, and a variable earning of $17, ranging from $3 to $34.5.

Before starting the experiment the participants received written instructions about the piece-rate round(s). After reading the instructions, the participants were given a short presentation of the task using projected slides to ensure common knowledge. After the piece-rate round(s), participants were given a separate set of instructions for the best-of-three contest to read, followed by a short presentation. A few control questions were then displayed on their screens, testing their understanding of the game. At the end of the experiment, participants were asked a few demographic questions. The full list of questions is available in Appendix. Overall, there is no significant difference in demographics across treatments (see Table 8 in Appendix).

3 Identification strategy

We aim to estimate the effect of winning the first round on the performance in the second round. For simplicity, and in line with the game theoretic model, we use “effort" as the main driver of performance in the discussion about the identification strategy. This framework naturally generalises to situations where performance can be affected by other factors, such as the effectiveness of the effort (what psychologists call “efficacy"). To model the effort in the second round (e_{i2}) let's consider the following equation:

$$e_{i2} = \alpha + \beta win_{i1} + \delta_2 + c_i + \varepsilon_{i2} \tag{2}$$

[8] In *PsychOnly*, one pair is dropped from analysis as one participant declared in the post experiment survey not having understood the rules.

The variable win_{i1} in (2) is a dummy taking the value 1 if individual i won in round 1, and 0 otherwise. The intercept δ_2 is a round specific element which accounts for learning or exhaustion as the participants move through the contest. The term c_i is an individual effect which accounts for heterogeneity, such as individual differences in ability. Finally, ε_{i2}, is a round and individual specific disturbance which captures residual variations in effort in a given round for a given individual.

Equation (2) is, de facto, a dynamic panel data model. As a consequence, usual estimation procedures will deliver biased estimates. It is for instance easy to see that estimating (2) by OLS will suffer from an endogeneity problem. The individual effect (c_i) impacts effort in the first round (e_{i1}), which in turn affects individual i's winning odds (win_{i1}). Individuals who exert more effort than their opponent in each round will be more likely to win round 1 and round 2. It creates a spurious correlation between the outcome of the two rounds.

A seemingly intuitive way to solve this problem is to estimate a fixed effects regression for equation (2). However, the fixed effects estimation will suffer from the Nickell bias arising from the fact that the fixed effects absorb part of the noises ε_{i2} (Nickell, 1981). It creates an attenuation bias which can be very large when the panel dimension is short as in our case.

Another intuitive solution is to take the first difference of equation (2):

$$\Delta e_i = \beta win_{i1} + \Delta\delta + \Delta\varepsilon_i \tag{3}$$

with $\Delta e_i = e_{i2} - e_{i1}$, $\Delta\delta = \delta_2 - \delta_1$, $\Delta\varepsilon_i = \varepsilon_{i2} - \varepsilon_{i1}$. The individual heterogeneity c_i is netted out of the estimation, solving the endogeneity problem of equation (2). However, a different endogeneity problem appears. Effort in the first round (e_{i1}) affects the likelihood of winning round 1, but is itself affected by unobserved individual heterogeneity in round 1 (ε_{i1}). Hence, the exogeneity assumption is violated. We can expect in particular a negative bias in $\widehat{\beta}$ due to a *regression towards the mean.* Random variations of ε_{i1} in round 1 are positively correlated with the winning probability. As a consequence there is a negative correlation between $\Delta\varepsilon_i$ and win_{i1}. A win in round 1 partially signals a likely high ε_{i1}, meaning that the effort in round 2 is not likely to be as high in round 2 due to a lower ε_{i2}.

One solution, proposed by Gill and Prowse (2014), is to use as instrumental variable the effort of a contestant's opponent. The opponent's effort is not correlated with the contestant's effort choice and directly affects his winning chances.

However, in our experiment participants are paired with each others by ability in order to closely match the hypothesis of homogeneous ability from the game theoretic model. Effort levels are therefore highly correlated within a pair, thereby making the instrument invalid.[9]

We propose here a novel approach. Since we use a stochastic contest success function, the performances of both participants do not determine the result of the round but the winning probability of each player. We therefore use this probability to match winners and losers with similar winning probabilities. This approach is similar to a propensity score matching procedure (Todd, 2010), but contrary to most propensity score matching applications we *perfectly* know the probability of an observation being in one or the other conditions.

In order to recover the causal effect of winning, we want to estimate the (counterfactual) potential change in performance a player would experience after a win or a loss. Using the Rubin (1974) framework, let's denote Δe_i^1 and Δe_i^0 the potential outcomes in terms of change of effort for player i if, respectively, the player wins in round 1 ($win_{i1} = 1$) or not ($win_{i1} = 0$). Due to the endogeneity problem, players who won would likely have had a different performance if they had lost than players who actually lost. The potential changes in performance are therefore not independent of the observed win/loss outcome. But given that we know the exact winning probability determined by the performance of the player, conditional on this probability the win/loss outcome is purely random (i.e. unrelated to the player's characteristics). As a consequence, the conditional independence assumption holds:

$$(\Delta e_i^1, \Delta e_i^0) \perp\!\!\!\perp win_{i1} | p_i \tag{4}$$

Conditioning on winning probability (using the matching approach) we can therefore identify the causal effect of winning. To do so we match and compare winners and losers who have similar ex-ante winning probabilities. We implement a local linear regression matching which compares each winner to a weighted average of losers with similar probabilities (Heckman et al., 1998). More weights are given to counterfactual observations with closer matching probabil-

[9]Let's consider player i and his opponent j. The performance $e_{j,1}$ of the opponent j in the first round of the contest is likely correlated with his performance in the piece rate round $e_{j,0}$ (due to unobserved heterogeneity c_j). Due to the pairing, this performance is itself correlated with the performance of the player i in the piece rate round, $e_{i,0}$. And, therefore, it is also correlated with the performance $e_{i,1}$ of the player i in the first round of the contest (due to unobserved heterogeneity c_i).

ity. As shown by Fan (1992), local linear regression performs strictly better than local weighted averaging like kernel regression.

Let's consider a game where n participants compete in pairs in a given round. Let $\mathscr{M}_i$ denote the matching neighbourhood of observation i, which includes all observations j that had a different outcome win_{j1} (win/loss) in the first round and were located within a bandwidth h in regard to their winning probability:

$$\mathscr{M}_i = \Big\{ j \in \{1,..,n\} : \ ||p_i - p_j|| < h \cap win_{j1} \neq win_{i1} \Big\} \tag{5}$$

We estimate the following regression in a given matching neighbourhood of observation i:

$$\min_{a_i, b_i} \sum_{j \in \mathscr{M}_i} \left(\Delta e_j - a_i - b_i \times (p_i - p_j) \right)^2 K\left(\frac{p_j - p_i}{h} \right)$$

Where a_i and b_i are the parameters of the local linear regression and K a kernel weighting function with a bandwitdh h (see Fan, 1992). The prediction of the above regression is a synthetic counter-factual to observation i. Let $\Delta \widehat{e_i}^1$ and $\Delta \widehat{e_i}^0$ be these estimated counterfactuals after a win and a loss, respectively. We can compute the individual effect ($\widehat{\beta}_i$) of winning as:

$$\widehat{\beta}_i = \begin{cases} \Delta e_i - \Delta \widehat{e_i}^0 & \text{if } win_{i1} = 1 \\ \Delta \widehat{e_i}^1 - \Delta e_i & \text{if } win_{i1} = 0 \end{cases}$$

And the average treatment effects ($\widehat{\beta}$), henceforth denoted as ATE, is:

$$\widehat{\beta} = \frac{1}{n} \sum_{i=1}^{N} \widehat{\beta}_i$$

For our estimation procedure to be valid, we need observations to lie in an area where there exists potential counter-factuals (Smith and Todd, 2005). That is, we need to impose a common support condition, whereby each observation can be compared with at least one counter-factual that had a similar probability to win. To do so we only consider the set of observations where the empirical distributions of the winning odds of the winners and losers overlap:

$$S_p = \Big\{ p_i \ : \ \widehat{f}(p_i | win_{i1} = 0) > 0 \ \cap \ \widehat{f}(p_i | win_{i1} = 1) > 1 \Big\}, \tag{6}$$

where $\widehat{f}(p_i)$ is the empirical distribution of the winning probabilities.

We estimate the effect's standard error by bootstrap clustered at the session level as suggested by Fréchette (2012). While bootstrapping fails for nearest neighbour matching, it provides reliable standard errors in the local linear regression case (Abadie and Imbens, 2008).

In the *StratOnly* treatment, where participants are randomly allocated to a win or loss outcome in round 1 with a 50% chance, we simply use ordinary least squares (OLS) to measure the effect of winning on effort.

4 Results

We first provide an overview of the data in terms of summary statistics. It gives an opportunity to discuss how simple measures of an effect of winning can be biased by the endogeneity issue discussed in Section 3. We then present our results from the matching procedure which solve this issue. We present robustness checks in Appendix.

4.1 Summary Statistics

Table 1 shows the number of blocks completed, overall and by treatment. On average, participants completed 20 blocks over the 10 minutes. Our randomisation was successful with no significant difference in average performance across treatments. The spread of performances is characterised by a standard deviation of around 5.5 (see Table 6 in Appendix for detailed summary statistics).

After the piece rate, participants are paired by performance (the first participant with the second one, and so on). The resulting within-pair differences of piece rate performance is fairly small. The interquartile range of the difference in performance within a pair is 1.5. This difference is even smaller when looking at the sample of participants who end up being on the common support in terms of winning probability at the end of the first round (our matching estimation is restricted to this sample). For these participants, the average difference within a pair is 0.9.

In round 1 the contest starts and participants have to pay half a cent per second of time spent on performing the effort task. We observe that in the two treatments with a first round, *Baseline* and *PsychOnly*, the average number of blocks completed is respectively 19.96 and 19.64. These averages are very close to piece rate performances, except that participants spent less than full 10 minutes on the task (8m32s and 8m53s respectively) given the choice of “stop" button (see Table

		Number of blocks				Kruskal-Wallis test
		Overall	*Baseline*	*PsychOnly*	*StratOnly*	
	Mean	20.38	20.84	19.92	20.35	$p = 0.561$
Piece-rate	(sd)	(5.466)	(5.377)	(6.253)	(4.799)	
	N	144	50	48	46	
	Mean	19.81	19.96	19.64	-	$p = 0.349$
Round 1	(sd)	(7.893)	(8.174)	(7.673)	-	
	N	98	50	48	46	
	Mean	19.51	20.3	19.54	18.63	$p = 0.304$
Round 2	(sd)	(7.502)	(8.229)	(6.925)	(7.316)	
	N	144	50	48	46	
	Mean	22.39	21.63	-	23.14	
Round 3	(sd)	(6.314)	(5.332)	-	(7.213)	$p = 0.335$
	N	44	22	26	22	

Table 1: Comparison of effort indicators between treatments. Only the outcome of the first piece-rate round is presented for the *StratOnly* treatment in order to have a comparable situation to the other treatments. Note that, by design, the overall number of observations is only the sum of observations in *Baseline* and *PsychOnly* in round 1 and only the sum of observations in *Baseline* and *StratOnly* in round 3.

6 in Appendix).

In round 2, we want to assess whether participants' performance are influenced by the result of round 1 as described in Section 2.3. Comparing directly the performance in round 2 of the winners and losers from round 1 does not give clear results. Figure 4 displays the differences in performance as a function of the round 1 result. When looking at absolute performance (left panels), winners in round 1 seem to complete more blocks in round 2. This comparison is similar to estimating equation (2). The observed difference cannot be interpreted as a causal effect of winning given the endogeneity problem: winners in round 1 may just happen to be better at the task. While such a concern should be alleviated with our pairing of contestants, residual differences can exist. When looking at the changes in performance between round 1 and round 2 (right panels), we do not observe any clear difference.[10] This comparison is similar to estimating equation (3). As discussed, another endogeneity problem exists here. Participants who happened to perform unusually well in round 1 were more likely to win then. In round 2, these participants are more likely to be back to a normal

[10]The right panels do not include the *StratOnly* treatment, because participants do not to make effort in a first round there.

(lower) level of performance. Such a selection pattern can create a regression towards the mean, biasing downward the estimate of the effect of winning.

A few pairs of participants enter the round 3 after the initial loser won the round 2. They represent a selected subsample and any differences can be driven by the selection process. If participants react differently to winning/losing, pairs reaching the third round may have different behavioural traits than in the original sample. We therefore do not include the round 3 performances in our analyses.

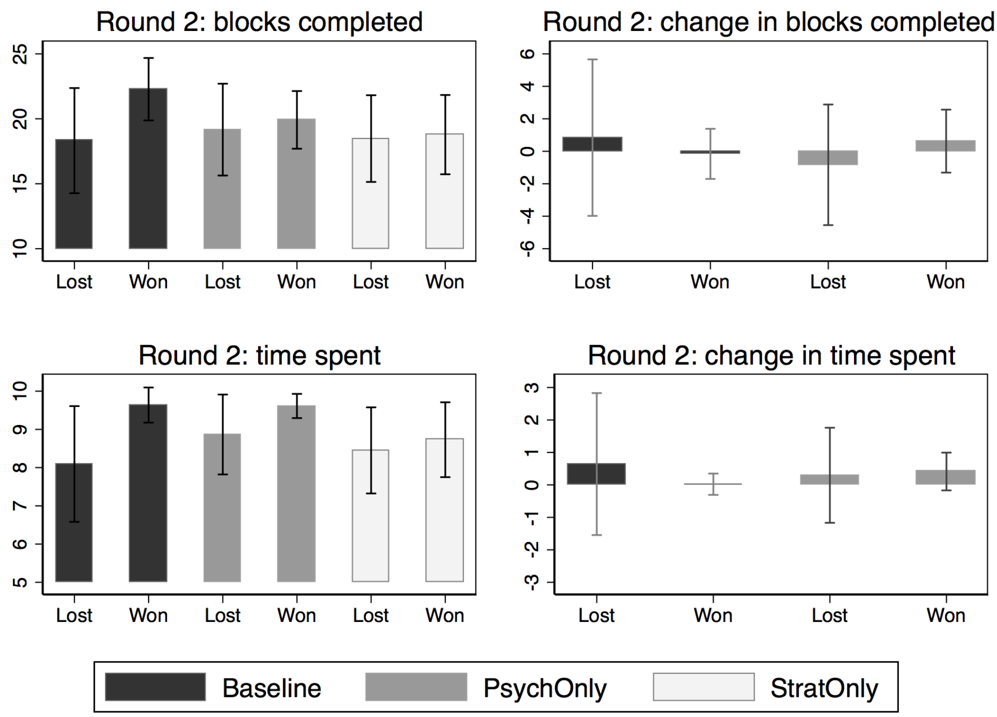


Figure 4: Number of blocks completed and the time spent on it by treatment and round-1 outcome

4.2 Matching results

Table 2 shows the ATE ($\widehat{\beta}$) of winning on effort variables. For treatments *Baseline* and *PsychOnly*, we use the matching approach described in Section 3 with a bandwidth of $h = 0.025$ for the local linear regression (we show that our results are robust to other choices of bandwidth in Section B.4). It means that a round 1 winner with a 50% chance of winning is compared to round 1 losers with a minimum probability of winning of 47.5% and a maximum one of 52.5%. For the

StratOnly treatment, winning is fully random, so we use a standard OLS regression to analyze the effect of winning.

In *Baseline*, we observe a positive causal effect of winning on effort across all measures. At the extensive margin, we find that winners spent on average more time in round 2 than losers (difference of 1.96min, $p = 0.073$). At the intensive margin, winners display a higher productivity. On average they complete 0.65 more blocks per minute ($p = 0.022$). These two joint effects combine into a significantly larger performance for winners who complete 4.46 blocks more in round 2 ($p = 0.039$). It means that participants with similar probability of winning in round 1, will diverge in their variation of performance in the second round based on their success in round 1.

Result 1 (Positive effect of winning) *We observe a positive effect of winning on performance in the Baseline treatment.*

Such an effect can be driven by a strategic effect on the decision to expend higher effort, a psychological momentum, or both. Taken on its own, the result from the *Baseline* treatment does not allow us to disentangle the two possible effects. For instance the difference in productivity could either reflect a difference in effort or a difference in efficacy whereby round 1 losers take more time to complete blocks as they lose some confidence following the round 1 result and become more cautious and hence less productive.

In order to identify the source of the effect, we now turn to the *PsychOnly* and *StratOnly* treatments. In *PsychOnly*, we find that winning has a strong impact on the number of blocks completed and productivity. The average performance of round 1 winners is greater by an average of 2.93 blocks in round 2 compared to the situation where they would not have won ($p = 0.020$). The productivity of winner is also higher by 0.26 tasks per minute ($p = 0.006$). Winners spent on average half a minute more on the task, however, the difference is not statistically significant ($p = 0.239$).

As the round 3 winner is randomly determined in *PsychOnly*, there can't be any strategic effect. The presence of a positive effect in *PsychOnly* points to the existence of a psychological effect of winning.

Result 2 (Positive Psychological effect) *We observe a positive effect of winning on performance in the PsychOnly treatment.*

In *StratOnly*, we do not observe any effect in round 2 of having won in round 1.[11] Randomly assigned round 1 winners do not spend significantly more time in the second round (p = 0.446), they do not become more productive (p = 0.952) and they do not complete more blocks in round 2 (p = 0.859).

Result 3 (No Strategic effect) *We do not observe an effect of winning when we turn off the source of psychological effect (even if a strategic effect is still predicted to be present).*

Table 2: Effect of winning round 1 on effort in round 2. This table displays the regression results of the impact of winning on various effort indicators. For *Baseline* and *PsychOnly* the LLR matching procedure is used, and in *StratOnly* an OLS regression is employed. The bandwidth for the LLR is set to 0.025. Standard errors are clustered at the session level, and constructed by bootstrap (2000 replications). Significance levels are denoted as *** p<0.01, ** p<0.05, *p<0.1.

Treatment		Blocks completed	Time spent	Productivity
Baseline	$\widehat{\beta}$	4.46**	1.96*	0.65**
	se	(2.156)	(1.091)	(0.285)
	p-value	$p = 0.039$	$p = 0.073$	$p = 0.022$
PsychOnly	$\widehat{\beta}$	2.93**	0.52	0.26***
	se	(1.264)	(0.441)	(0.094)
	p-value	$p = 0.020$	$p = 0.239$	$p = 0.006$
StratOnly	$\widehat{\beta}$	0.30	0.28	-0.01
	se	(1.715)	(0.366)	(0.152)
	p-value	$p = 0.859$	$p = 0.446$	$p = 0.952$

The joint results of our three experimental conditions provide two clear conclusions. First, winning has indeed a positive causal effect on players' performance. Second, this effect persists when the source of strategic effect is turned off and it disappears when the source of the psychological effect is turned off. Therefore, we do not observe evidence of a strategic effect while at the same time we find evidence supporting a psychological effect.[12]

[11] We look here directly at the effort level in round 2. The absence of effort in round 1 does not make it possible to look at a difference of effort between round 2 and 1. Note that we do not use the change in effort between the second piece-rate and round 2 as it might capture heterogeneity in the willingness to use the "stop" button, or in behaviour in a competitive environment rather than actual effect of winning on effort. Using such a difference gives similar results though.

[12] Note that as we do not find a strategic effect, one could be tempted to expect for the effect in

5 Investigating the psychological effect

Our experimental results point to the existence of a positive effect of winning on later performance driven by a psychological effect. These results naturally open new questions about the mechanisms underlying this psychological effect. In psychology, there is no unified conceptual framework describing the mechanisms behind such an effect. And in economics, the term psychological momentum has been used by default to describe behavioural effects unrelated to rational strategies in contest.

The most frequent interpretation of the psychological momentum is that it comes from an increase of players' self-confidence after experiencing a success. The idea that self-confidence can enhance performance idea was formalised in economics by Compte and Postlewaite (2004) who coined the term *confidence-enhanced performance*. In their model, the information on past successes raise the self-confidence of agents in their ability to succeed again. This increase in self-confidence raises their ability to perform and their willingness to take risks in order to succeed.

We designed a new treatment, to investigate whether the *confidence-enhanced performance hypothesis* is a possible explanation for the effect of winning. Starting from the design of the *PsychOnly* treatment, we removed all the informational content of a win. We call this treatment *PsychOnlyInfo*. In this treatment, after each round of the best-of-three contest, every participants receive information about the number of blocks completed, the winning odds, the time spent in the round and productivity, of both paired participants.

If the effect of winning in *PsychOnly* comes from the informational content of a win, the psychological effect should disappear in *PsychOnlyInfo*. Once participants know their relative performance and their winning odds, winning or losing itself does not bring any new information on their relative strength and it should therefore have no impact on their self-confidence and performance.

Hypothesis 4 ***Effect of a win without information value*** *Winning should not have a causal effect on performance in PsychOnlyInfo given that it has no informational value to raise self-confidence.*

Baseline and *PsychOnly* to be equal. However, as indicated in Section 2.3, the *Baseline* and the *PsychOnly* treatments have different game structures and we therefore can't make a prediction about the equality of a psychological effect in these two conditions. The coefficients we find are different (though not significantly, see Table 12 in Appendix) across the two conditions.

Second, if the effect of winning comes from the impact of information about past performance then the direct provision of information itself should have a positive effect on performance. We can here use the fact that participants get different level of information about their round 1 performance in *PsychOnly* and *PsychOnlyInfo* to investigate the effect of information. In *PsychOnly*, players get an imperfect signal of their relative strength (whether they won or not). In comparison, in *PsychOnlyInfo* they get detailed information of how well they did relative to their opponent. Furthermore, in each pair the strong players always get a positive information in *PsychOnlyInfo* but they get a positive information only when they win in *PsychOnly*. As a consequence, one would expect the best players from each pair to increase their effort more in *PsychOnlyInfo* than in *PsychOnly* as they get better information about their strength.

Hypothesis 5 ***Effect of information in itself*** *Relatively strong players should have a higher performance in round 2 in PsychOnlyInfo compared to PsychOnly.*

These two predictions are indeed observed in *PsychOnlyInfo*. First, the causal effect of winning stops being significant in *PsychOnlyInfo*. The results of our matching approach in this treatment are displayed in Table 3. We do not observe a significant effect of winning on any of the effort variables in round 2. Winners did not spend significantly more time than losers in round 2 ($p = 0.188$), they did not complete more blocks per minute ($p = 0.424$, Wilcoxon signed-rank test, henceforth denoted as WSR) and as a consequence they did not increase their number of blocks completed relative to losers ($p = 0.376$).

Second, players who are relatively strong in their pair indeed react to information in *PsychOnlyInfo* more than in *PsychOnly*. Table 4 shows the regression results of the performance of a player in round 2 as a function of the difference in performance relative to the other player in round 1. We observe that in the *PsychOnly* condition where winning the round 1 is the only information, the relative performance of the player in round 1 has a positive, but insignificant, effect on measures of performance in round 2. While winning has an effect on beliefs in itself (see Table 5), many weak players win by chance and many strong players lose by chance so the player's strength does not fully determine the signal received (win/loss). On the contrary, in the *PsychOnlyInfo*, players receive an accurate information about their relative strength, whether they won or lost. We observe that the interaction term between relative strength and the dummy for the *PsychOnlyInfo* is large and clearly significant for number of blocks performed and the time spent on the task. It indicates that strong players increased their effort more in the *PsychOnlyInfo* treatment relative to strong players in *PsychOnly*.

Table 3: Effect of winning round 1 on measures of performance in round 2 for *PsychOnlyInfo*. This table displays the OLS regression results of the impact of winning on various effort indicators in *PsychOnlyInfo*. The bandwidth for the LLR is set to 0.025. The common support for this treatment includes 36 observations, and covers participants with winning odds between 37% and 63%. Standard errors are clustered at the pair level and constructed by bootstrap (2000 replications). Significance levels are denoted as *** p<0.01, ** p<0.05, *p<0.1.

Treatment		Blocks completed	Time spent	Productivity
PsychOnlyInfo	$\widehat{\beta}$	1.46	1.15	-0.06
	se	(1.654)	(0.872)	(0.079)
	p-value	$p = 0.376$	$p = 0.188$	$p = 0.424$

Table 4: Effect of information on effort in round 2, depending on differences within a pair. The table displays the regression results of the impact of receiving information about relative performance in round 1 on measures of effort. Info is a dummy variable that takes the value 0 in *PsychOnly* and 1 in *PsychOnlyInfo*.

	Blocks completed	Time spent	Productivity
Info	-1.97	-0.98	0.07
(se)	(1.656)	(0.347)	(0.078)
Rel perf R1	0.07	0.02	-0.00
(se)	(0.075)	0.026	(0.005)
Info × Rel perf R1	0.34*	0.11**	0.01
(se)	(0.182)	(0.052)	(0.010)
Constant	19.54***	9.24***	2.07***
(se)	(0.779)	(0.222)	(0.052)
N	102	102	102

Another sign that the players' self-confidence drives the positive effect of winning should be that players' beliefs in their chance of winning are positively influenced by a win in *Baseline* and *PsychOnly* but not in *PsychOnlyInfo* where win-

Table 5: Effect of winning round 1 on elicited confidence in round 2. This table displays the regression results of the impact of winning on elicited confidence. The LLR matching procedure is used. The bandwidth for the LLR is set to 0.025. We use WSR tests to test for significance. Standard errors are clustered at the pair level and constructed by bootstrap (2000 replications). Significance levels are denoted as *** p<0.01, ** p<0.05, *p<0.1.

Treatment	*Baseline*	*PsychOnly*	*PsychOnlyInfo*
$\widehat{\beta}$	18.58***	14.76***	7.23
se	(2.544)	(3.967)	(4.661)
p-value	$p < 0.001$	$p < 0.001$	$p = 0.121$

ning does not have informational value in itself. It is indeed what we find. Using our matching approach we can estimate the causal effect of winning on players' belief in their winning chances in the next round. Table 5 shows the results of this estimation. In all three treatments, winning leads participants to become more confident. In *Baseline* and *PsychOnly* the observed effect is of similar magnitude (18.58 and 14.76 respectively), and significantly different from zero ($p < 0.001$ in both cases). However, the effect is not significantly different from zero ($p = 0.121$) in *PsychOnlyInfo*.

These results on belief updating and the effect of information in itself in *PsychOnlyInfo* suggests that the effect we observed in *PsychOnly* was driven by the effect of information on past performance on self-confidence.[13]

Result 4 (Confidence-enhanced performance) *The causal effect of winning seems associated with the informational content of a win about the players performance which increases their self-confidence and as a consequence their later performance.*

6 Conclusion

We studied experimentally the existence and origin of a causal effect of winning. To do so we designed a series of simplified dynamic contests with randomly generated variations in early success. We find a clear positive causal effect

[13] To be clear, the absence of evidence is not evidence of absence and therefore the fact that we do not find a significant effect of winning in *PsychOnlyInfo* does not prove that there is no effect of winning in this treatment. We simply interpret this result as being in support of the predictions form Hypotheses 4 and 5.

of winning. Participants who won early in the contest had relatively higher performance in a subsequent period both at the extensive margin (time spent) and at the intensive margin (quality of performance per unit of time). This effect is causal and not due to unobserved differences between early winners and losers.

By turning on and off the mechanisms considered to be underlying the causal effect of winning we are able to disentangle two possible explanations of this effect: strategic thinking and psychological factors. Surprisingly, we do not find support for a role of strategic thinking. Instead, our results support the psychological explanation: winning has a direct effect on motivation, confidence and/or competitiveness. To understand further the mechanism of this psychological effect we designed an additional treatment allowing us to remove all the informational content from a win. By giving all the relevant information to the participants about their and their opponent's performance we ensure that winning or losing does not bring any additional information to the players about their strength relative to their opponent. We find that doing so eliminates the positive effect of winning.

Our results pave the way to new research on competitions as games of incomplete information. While we did not find evidence of the strategic effect predicted in a complete information framework with symmetric players, we find evidence compatible with the "confidence-enhanced performance hypothesis", the idea that the self confidence (which can be boosted by a positive performance) can foster a rise in later performance. Our study suggests that self-confidence plays a critical role in people's performance. Understanding how self-confidence is shaped by past successes and in turn shapes future successes can play an important role in understanding how identical people can end up having very different success path. Indeed, even with an initially even playing field, differences in success at some point can play a critical role in future success and therefore contribute to the rise of inequalities.

References

ABADIE, A. AND G. W. IMBENS (2008): "On the failure of the bootstrap for matching estimators," *Econometrica*, 76, 1537–1557.

ARAUJO, F. A., E. CARBONE, L. CONELL-PRICE, M. W. DUNIETZ, A. JAROSZEWICZ, R. LANDSMAN, D. LAMÉ, L. VESTERLUND, S. W. WANG, AND A. J. WILSON (2016): "The slider task: an example of restricted inference on incentive effects," *Journal of the Economic Science Association*, 2, 1–12.

BAR-ELI, M., S. AVUGOS, AND M. RAAB (2006): "Twenty years of "hot hand" research: Review and critique," *Psychology of Sport and Exercise*, 7, 525–553.

CHEN, D. L., M. SCHONGER, AND C. WICKENS (2016): "oTree—An open-source platform for laboratory, online, and field experiments," *Journal of Behavioral and Experimental Finance*, 9, 88–97.

COHEN-ZADA, D., A. KRUMER, ET AL. (2017): "Psychological momentum and gender," *Journal of Economic Behavior & Organization*, 135, 66–81.

COMPTE, O. AND A. POSTLEWAITE (2004): "Confidence-enhanced performance," *The American Economic Review*, 94, 1536–1557.

DE ROOS, N. AND Y. SARAFIDIS (2015): "Momentum in Dynamic Contests," *Available at SSRN 2696539.*

DECHENAUX, E., D. KOVENOCK, AND R. SHEREMETA (2015): "A survey of experimental research on contests, all-pay auctions and tournaments," *Experimental Economics*, 18, 609–669.

ERKAL, N., L. GANGADHARAN, AND B. H. KOH (2016): "Monetary and Non-Monetary Incentives in Real-Effort Tournaments," *Experimental economics.*

FAN, J. (1992): "Design-adaptive nonparametric regression," *Journal of the American statistical Association*, 87, 998–1004.

FERRALL, C. AND A. A. SMITH JR (1999): "A sequential game model of sports championship series: theory and estimation," *The Review of Economics and Statistics*, 81, 704–719.

FRÉCHETTE, G. R. (2012): "Session-effects in the laboratory," *Experimental Economics*, 15, 485–498.

FU, Q., C. KE, AND F. TAN (2015a): “"Success breeds success" or "Pride goes before a fall"?: Teams and individuals in multi-contest tournaments,” *Games and Economic Behavior*, 94, 57 – 79.

FU, Q., J. LU, AND Y. PAN (2015b): “Team contests with multiple pairwise battles,” *The American Economic Review*, 105, 2120–2140.

GÄCHTER, S., L. HUANG, AND M. SEFTON (2016): “Combining “real effort” with induced effort costs: the ball-catching task,” *Experimental economics*, 19, 687.

GAURIOT, R. AND L. PAGE (2015): “Does success breed success? A quasi-experiment on strategic momentum in dynamic contests,” Tech. rep., QUT Business School.

GILL, D. AND V. PROWSE (2014): “Gender differences and dynamics in competition: The role of luck,” *Quantitative Economics*, 5, 351–376.

GREINER, B. (2015): “Subject pool recruitment procedures: organizing experiments with ORSEE,” *Journal of the Economic Science Association*, 1, 114–125.

HÄRDLE, W. K., M. MÜLLER, S. SPERLICH, AND A. WERWATZ (2012): *Nonparametric and semiparametric models*, Springer Science & Business Media.

HARRIS, C. AND J. VICKERS (1987): “Racing with uncertainty,” *The Review of Economic Studies*, 54, 1–21.

HECKMAN, J. J., H. ICHIMURA, AND P. TODD (1998): “Matching As An Econometric Evaluation Estimator,” *The Review of Economic Studies*, 65, 261–294.

HECKMAN, J. J., H. ICHIMURA, AND P. E. TODD (1997): “Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme,” *The Review of Economic Studies*, 64, 605–654.

HOLLARD, G., S. MASSONI, AND J.-C. VERGNAUD (2016): “In search of good probability assessors: an experimental comparison of elicitation rules for confidence judgments,” *Theory and Decision*, 80, 363–387.

HUCK, S., N. SZECH, AND L. M. WENNER (2015): “More effort with less pay: On information avoidance, belief design and performance,” .

ISO-AHOLA, S. E. AND K. MOBILY (1980): "Psychological momentum: A phenomenon and an empirical (unobtrusive) validation of its influence in a competitive sport tournament," *Psychological Reports*, 46, 391–401.

KLUMPP, T. AND M. K. POLBORN (2006): "Primaries and the New Hampshire effect," *Journal of Public Economics*, 90, 1073–1114.

KONRAD, K. A. (2009): "Strategy and dynamics in contests," *OUP Catalogue*.

KONRAD, K. A. AND D. KOVENOCK (2006): "Equilibrium and Efficiency in the Tug-of-War," Tech. rep., SFB/TR 15 Discussion Paper.

——— (2009): "Multi-battle contests," *Games and Economic Behavior*, 66, 256 – 274.

KRUMER, A., R. MEGIDISH, AND A. SELA (2017): "First-mover advantage in round-robin tournaments," *Social Choice and Welfare*, 48, 633–658.

MAGO, S. D., R. M. SHEREMETA, AND A. YATES (2013): "Best-of-three contest experiments: Strategic versus psychological momentum," *International Journal of Industrial Organization*, 31, 287 – 296.

MALUEG, D. A. AND A. J. YATES (2010): "Testing contest theory: evidence from best-of-three tennis matches," *The Review of Economics and Statistics*, 92, 689–692.

MARKMAN, K. D. AND C. L. GUENTHER (2007): "Psychological momentum: Intuitive physics and naive beliefs," *Personality and Social Psychology Bulletin*, 33, 800–812.

MILLER, J. B. AND A. SANJURJO (2014): "A cold shower for the hot hand fallacy," .

NICKELL, S. (1981): "Biases in dynamic models with fixed effects," *Econometrica: Journal of the Econometric Society*, 1417–1426.

ROBERTSON, I. H. (2012): *The winner effect: The neuroscience of success and failure*, Macmillan.

RUBIN, D. B. (1974): "Estimating causal effects of treatments in randomized and nonrandomized studies." *Journal of educational Psychology*, 66, 688.

SMITH, J. A. AND P. E. TODD (2005): "Does matching overcome LaLonde's critique of nonexperimental estimators?" *Journal of Econometrics*, 125, 305 – 353.

TAYLOR, J. AND A. DEMICK (1994): "A multidimensional model of momentum in sports," *Journal of Applied Sport Psychology*, 6, 51–70.

TODD, P. E. (2010): *Matching Estimators*, London: Palgrave Macmillan UK, 108–121.

TONG, K.-K. AND K. LEUNG (2002): "Tournament as a motivational strategy: Extension to dynamic situations with uncertain duration," *Journal of Economic Psychology*, 23, 399–420.

TULLOCK, G. (1980): "Efficient rent seeking," in *Towards a theory of the rent-seeking society*, ed. by J. Buchanan, T. RD, and T. G, Texas AM University Press, 97–112.

VAN DE RIJT, A., S. M. KANG, M. RESTIVO, AND A. PATIL (2014): "Field experiments of success-breeds-success dynamics," *Proceedings of the National Academy of Sciences*, 111, 6934–6939.

A Summary statistics

A.1 Detailed summary statistics on the first three treatments

Table 6: Summary statistics on performance and effort provision, conditional on the outcome in round 1. The measures of performance and effort are the number of blocks completed, whether the participant decided to stop in the round (binary variable) and the average time spent in the round.

	Baseline			PsychOnly			StratOnly		
	# blocks	*Stop*	*Time spent*	*# blocks*	*Stop*	*Time spent*	*# blocks*	*Stop*	*Time spent*
Piece-rate 1									
Mean	20.84	-	-	19.92	-	-	20.35	-	-
(sd)	(5.377)	-	-	(6.253)	-	-	(4.799)	-	-
Piece-rate 2									
Mean	-	-	-	-	-	-	21.13	-	-
(sd)	-	-	-	-	-	-	(5.188)	-	-
Contest - round 1		*N=50*			*N=48*			*N=46*	
Mean	19.96	48%	8mins32s	19.64	43.75%	8mins53s			
(sd)	(8.174)	(0.505)	(2.662)	(7.673)	(0.501)	(2.074)			
Contest - round 2		*N=50*			*N=48*			*N=46*	
R1 Winners - Mean	22.28	36%	9mins38s	19.91	54%	9mins36s	18.78	43%	8mins42s
(sd)	(5.820)	(0.490)	(1.113)	(5.258)	(0.509)	(0.750)	(7.058)	(0.507)	(2.265)
R1 Winners - Mean change	-0.16	-4%	1.2s	0.63	12.5%	18s	-	-	-
(sd)	(3.738)	(0.200)	(0.795)	(4.600)	(0.680)	(1.376)	-	-	-
R1 Losers - Mean	18.32	44%	8mins4s	19.16	46%	8mins52s	18.48	48%	8mins27s
(sd)	(9.810)	(0.507)	(3.668)	(8.370)	(0.509)	(2.475.51)	(7.722)	(0.511)	(2.604)
R1 Losers - Mean change	0.84	-12%	38s	-0.83	0%	24s	-	-	-
(sd)	(11.679)	(0.726)	(5.300)	(8.800)	(0.722)	(3.468)	-	-	-
Contest - round 3		*N=22*			*N=24*			*N=22*	
N									
R1 Winners - Mean	23	9%	9mins58s				22.45	64%	8mins55s
(sd)	(5.099)	(0.302)	(0.060)				(7.005)	(0.505)	(1.907)
R1 Winners - Mean change	2.45	-0.18	17s				2.18	9%	1s
(sd)	(3.671)	(0.603)	(1.176)				(4.557)	(0.539)	(1.165)
R1 Losers - Mean	20.27	36%	9mins47s				23.82	45%	9mins16s
(sd)	(5.442)	(0.505)	(0.876)				(7.692)	(0.522)	(1.984)
R1 Losers - Mean change	-0.55	0%	-1s				1	9%	-3s
(sd)	(3.110)	(0.775)	(1.176)				(3.975)	(0.701)	(0.307)

A.2 Summary statistics of the additional treatment *PsychOnlyInfo*

Table 7 provides summary statistics on the number of blocks completed and the time spent in this new treatment. We find that participants in a pair are similar in the piece-rate round, as we cannot reject the hypothesis that they had the same performance in the piece rate. In round 2, the observed behaviour is very close to that of round 1: losers spent less time competing and completed less blocks.

Table 7: Summary statistics on effort provision ***PsychOnlyInfo***, conditional on outcome of round 1. Significance levels for WSR tests are denoted as *** p<0.01, ** p<0.05, *p<0.1. P-values are for MW tests, testing whether the value of the variable is the same for winners and losers.

	PsychOnlyInfo		
	Winners (1)	*Losers* (2)	Δ [(1) - (2)]
Piece-rate 1	*N=54*		
Mean # blocks	20.3	18.59	1.70
(se)	(1.27)	(1.25)	(1.78)
Contest - round 1	*N=54*		
Mean # blocks	19.59	15.52	
(se)	(1.62)	(1.65)	
Use of button	59%	59%	
Mean time spent	8min44s	7min37s	
(se)	(0.38)	(0.61)	
Contest - round 2	*N=54*		
Mean # blocks	19.52	15.63	
(se)	(1.47)	(1.74)	
Use of button	59%	59%	
Mean time spent	8min53s	7min37s	
(se)	(0.34)	(0.65)	
Contest - round 3	*N=22*		

B Robustness checks

B.1 Balance tests across treatments

	Overall	*Baseline*	*PychOnly*	*StratOnly*	Kruskal-Wallis test
% Males	54%	60%	46%	57%	$p = 0.347$
(sd)	(0.500)	(0.495)	(0.504)	(0.501)	
Mean age	21.07	21.7	20.58	20.91	$p = 0.978$
(sd)	(4.738)	(5.59)	(3.506)	(4.871)	
% Enjoyed task	73%	76%	73%	70%	$p = 0.779$
(sd)	(0.446)	(0.431)	(0.449)	(0.465)	
% Effort pays off	69%	66%	67%	74%	$p = 0.658$
(sd)	(0.465)	(0.479)	(0.476)	(0.444)	
Mean risk assessment	6.01	5.68	6.40	5.98	$p = 0.318$
(sd)	(2.052)	(2.084)	(2.111)	(1.926)	

Table 8: Comparison of demographics between treatments

B.2 Common support

It is standard to check the size of the common support (set of observations where the matching scores overlap) when using a propensity score matching. In our case, our matching strategy is facilitated by the fact that the Tullock function produces winning probabilities concentrated around 50%. Therefore most winners in our sample can be matched with losers with a similar winning probability and vice-versa.

The empirical distribution of the round 1 winning probabilities is represented in Figure 5. In *Baseline*, the common support includes participants having a chance to win the first round between 41% and 59% ($N = 32$). In *PsychOnly*, the common support includes pairs whose propensity score range between 27% and 73% ($N = 46$). Detailed summary statistics over the common support are presented in Table 9.

Figure 5: Distribution of winning probability in each treatment conditions

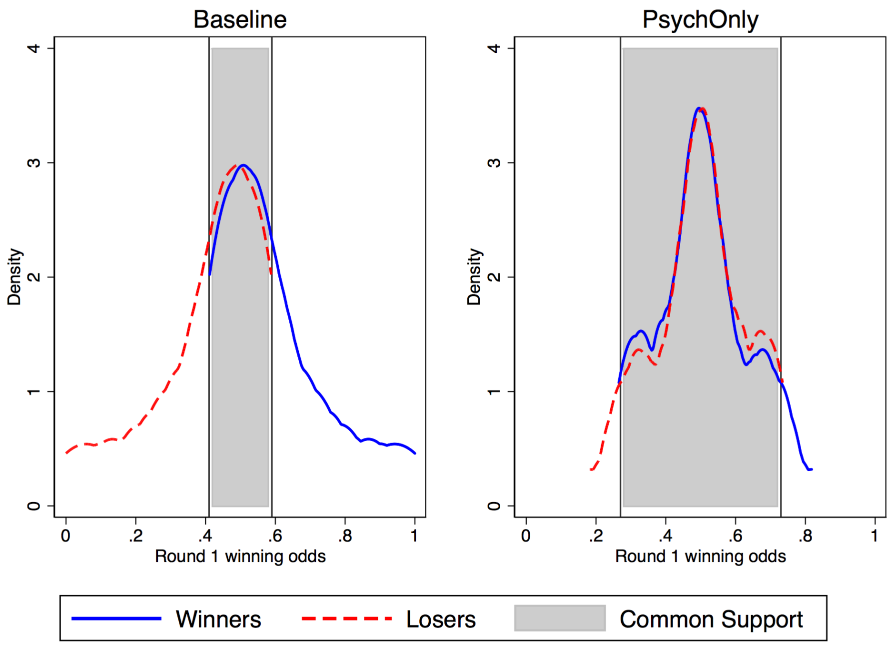


Table 9: Summary statistics over the common support.

		Change in # blocks		Change in time spent		Change in productivity	
		Winners	Losers	Winners	Losers	Winners	Losers
Baseline	Mean	-0.25	-4.63	-0.066	-2	0	-0.58
	se	(0.955)	(2.427)	(0.232)	(1.035)	(0.075)	(0.261)
	MW test	$p = 0.279$		$p = 0.480$		$p = 0.151$	
PsychOnly	Mean	0.86	-1.65	0.43	0.02	-0.02	-0.24
	se	(0.946)	(1.669)	(0.292)	(0.681)	(0.133)	(0.060)
	MW test	$p = 0.473$		$p = 0.432$		$p = 0.176$	

Note: p-values are for Mann-Whitney tests testing whether the effort measure is from the same distribution for winners and losers.

As the common supports differ in size, we checked that it does not affect our results. We ran our matching approach in *PsychOnly* using only observations having a winning probability between 41% and 59% (N=26), like for the common support of the *Baseline*. The results of this regression are displayed in Table 10.

The magnitude of the effects stay the same (while not significant due to small number of observations).

Table 10: Effect of winning round 1 on effort in round 2 for *PsychOnly* with the same common support as *Baseline*

Treatment		Blocks completed	Time spent	Productivity
PsychOnly	$\widehat{\beta}$	2.41	0.51	0.15
	se	(2.407)	(0.838)	(0.116)
	p-value	$p = 0.317$	$p = 0.546$	$p = 0.198$

Note: This table displays the regression results of the impact of winning on various effort indicators in *PsychOnly* while considering the same common support as *Baseline*. The bandwidth for the LLR is set to 0.025. Standard errors are clustered at the pair level and constructed by bootstrap (2000 replications). Significance levels for WSR tests are denoted as *** p<0.01, ** p<0.05, *p<0.1.

B.3 Details of the local linear regression matching

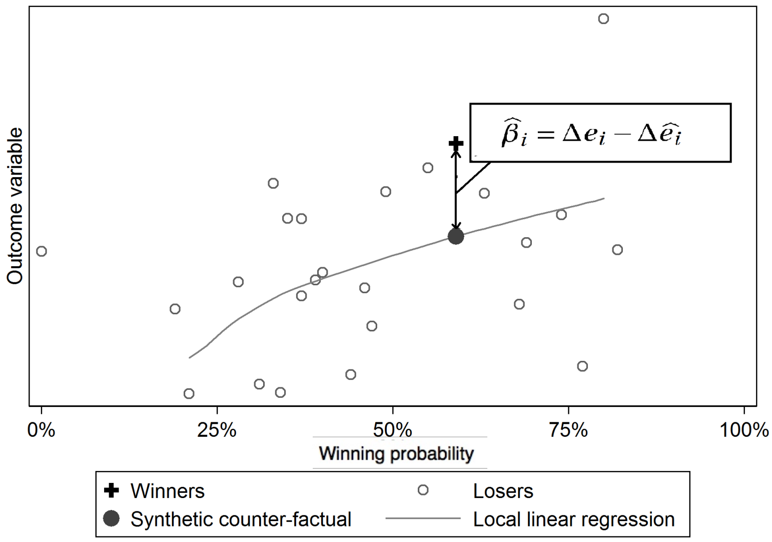
We build the composite control group using a local linear regression matching (LLR), which is depicted on Figure 6. It has the advantage of using more, and more precise, information than the standard nearest neighbour matching (Heckman et al., 1997; Todd, 2010). Indeed, k nearest neighbour matching is a situation where a unit weight is given to k observations. On the other hand, the LLR procedure compares a treated observation with a weighted combination of all control observations in a given neighbourhood in the propensity score space (Fan, 1992). More precisely, we estimate the following regression in a given neighbourhood of observation i:

$$\min_{a_i, b_i} \sum_{j \in \mathcal{M}_i} \Big(\Delta e_j - a_i - b_i \times (p_i - p_j)\Big)^2 K\Big(\frac{p_j - p_i}{h}\Big)$$

Where $K()$ is a kernel weighting function, and a_i and b_i are the parameters of the weighted ordinary least squares (OLS) regression. The prediction of the above estimation at propensity score p_i is the constructed counter-factual to observation i ($\Delta \widehat{e_i}$) .

For our estimation procedure to be valid, we need that every treated observation lies in an area where there exist potential counter-factuals (Smith and Todd, 2005). That is, we need to impose the common support condition, whereby each

Figure 6: Representation of the construction of the synthetic control group. This figure represents how the synthetic control group is built. It represents how the counterfactual value of a variable of interest Y is estimated for a winning player, using observerations from losing players.



treated observation can be compared with at least one untreated that had a similar probability to win (and vice versa). The estimation therefore relies only on observations where the distributions of the winning probabilities of the treated and controls overlap.

B.4 Bandwidth selection

We relax the assumption of a bandwidth of 2.5% for the local linear regression to check the robustness of our results. First, we use a leave-one-out cross-validation method (Härdle et al., 2012) to determine the bandwidth minimising the Asymptotic Mean Integrated Squared Errors (AMISE). This bandwidth is referred as the "optimal bandwidth" in typical applications. It is however not necessarily optimal in a matching estimation. The identification strategy requires observations to be matched with very close observations in order to ensure that they are similar. The best bandwidth in a matching approach may therefore be smaller than the one minimising the AMISE. We therefore only use this different bandwidth as an objective benchmark, different from our initial choice of a small bandwidth.

The leave-one-out cross-validation method consists in estimating the AMISE of the estimator by running the model on the whole sample minus one observation and compare the model prediction for this observation with the actual value of the variable studied. By successively leaving out each observation in the sample once, one can estimate an error for each observation. The average of these errors provides an estimate of the AMISE of the model given its bandwidth. The "optimal bandwidth" is the one that minimizes the mean square error of the predictions. The results of this procedure are displayed in Table 11.

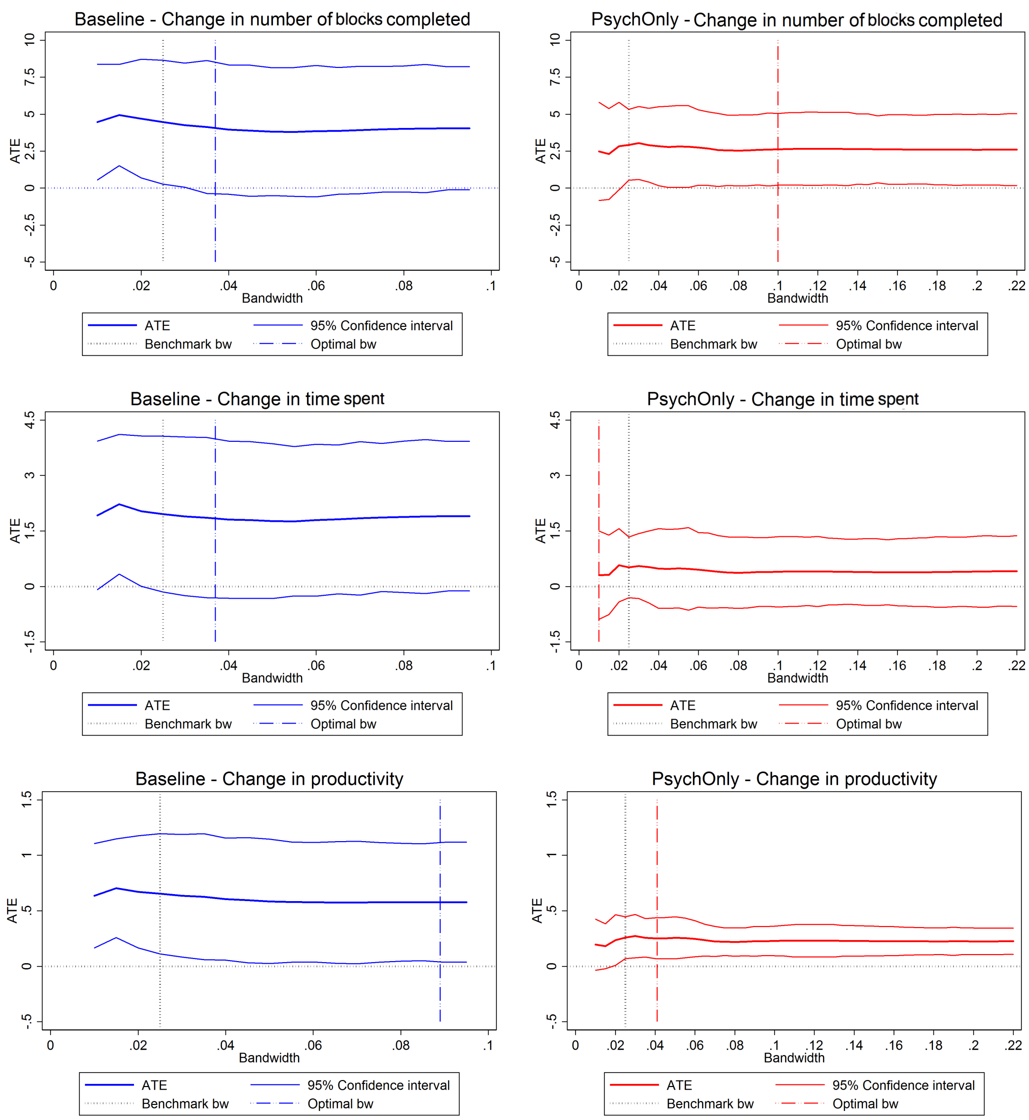
Table 11: Results of our cross validation procedure. This table displays the regression results of the impact of winning on various effort indicators, at the optimally chosen bandwidth. Standard errors are clustered at the session level and constructed by bootstrap (2000 replications). Significance levels for WSR tests are denoted as *** p<0.01, ** p<0.05, *p<0.1.

		Blocks completed	Time spent	Productivity
Baseline	Optimal bw	0.037	0.037	0.089
	Estimated effect	4.06*	1.84*	0.58**
PsychOnly	Optimal bw	0.10	0.01	0.041
	Estimated effect	2.82**	0.3	0.25***

In *Baseline*, we find that this bandwidth varies from 0.01 in to 0.09, depending on the outcome variable considered. Given that the common support covers the range [41,59], a bandwidth of 0.09 means that the whole common support is considered. The average synthetic counter-factual is therefore the outcome of a weighted OLS regression. Our main results still hold for these optimal bandwidths: winning leads to more blocks completed ($p = 0.075$), more time spent in the second round ($p = 0.096$) and an increase in productivity ($p = 0.037$). In *PsychOnly* our findings also hold at the optimal bandwidth: winning has a significant positive impact on the number of blocks completed ($p = 0.040$) and productivity ($p = 0.008$).

Second, we assess the sensitivity of our results to bandwidth selection by looking at how our results vary for all the possible bandwidths. We start by estimating the model with a bandwidth of 0.01, and progressively increase it up to the point where all observations on the common support are included. The estimated effects are displayed in Figure 7. Our findings hold regardless of the chosen bandwidth. In *Baseline*, the positive impact of winning on productivity is significantly different from zero for all the bandwidths considered. We also find that the effect of winning on time spent on the task and number of blocks completed is weakly significantly different from zero for all cases we considered. In

Figure 7: Impact of varying bandwidth on effort measures



PsychOnly, the effect of winning on the number of blocks completed and productivity is significantly different from zero for 90% of the bandwidths. Overall, varying the bandwidth for the construction of the synthetic control group seems to have little impact on our results.

C Comparison of the effects in the *Baseline* and *PsychOnly* treatments

Table 12: Comparison of treatment effects between *Baseline* and *PsychOnly*. This table compares the relative size of the average treatment effects. The standard errors for the ratio of expenditures were computed by bootstrapping (5 000 replications), and clustered at the session level. Significance levels are denoted as *** p<0.01, ** p<0.05, *p<0.1.

	Blocks completed	Time spent	Productivity
$\beta_{Baseline} - \beta_{PsychOnly}$	1.53	1.44	0.39
(se)	(2.834)	(1.284)	(0.325)
$H_0 : \beta_{Baseline} - \beta_{PsychOnly} = 0$	$p = 0.589$	$p = 0.263$	$p = 0.223$

D Instructions

Experimental Instructions

Welcome to our experiment! You will receive AUD 5 for showing up on time. Please read these instructions carefully and completely. Properly understanding the instructions will help you to make better decisions and, hence, to earn more money. If you read these instructions carefully and perform well in the experiment, you can earn a significant amount of money (which will be paid out to you in cash at the end of the experiment).

Please keep in mind that you are not allowed to communicate with other participants during the experiment. You are not allowed to use your mobile phone at any time either. If you do not obey these rules you will be asked to leave the laboratory and will not be paid. Whenever you have a question, please raise your hand; an experimenter will come to assist you.

This experiment consists of two parts. You will receive separate instructions for each part. Your final payment will be the sum of your earnings in both parts, plus your show-up fee.

Part 1

In this experiment you will be asked to reverse strings of characters that will appear on your screen. Each string is randomly generated and has 20 characters.

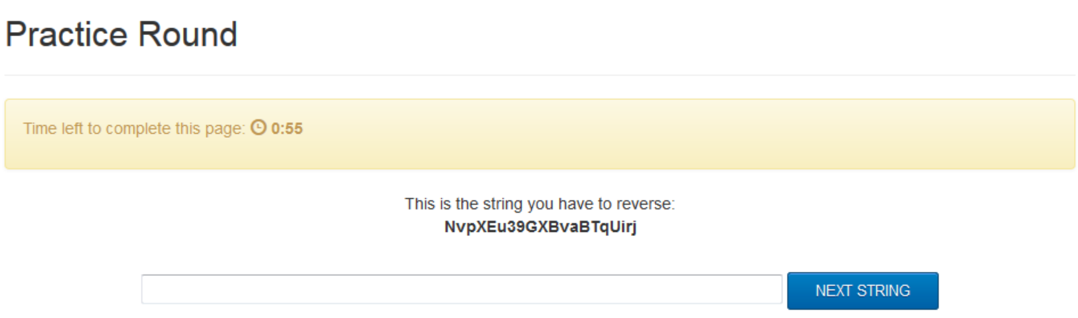
For example, if you see:

NvpXEu39GXBvaBTqUirj

You have to enter:

jriUqTBavBXG93uEXpvN

You will be doing this task for 10minutes. For each string you correctly reverse you will get 20cts. A typical screen that you will be seeing is as follows:



Your payoff will be computed as follows:

Number of tasks completed correctly × 20cts

For instance, if you completed 8 tasks, you will earn: 8 × \$0.2 = \$1.6.

To make sure that you understand the task and know how to work, you will have the opportunity to practice for 1 minute without payment.

Part 2

Matching

In this part, the computer will rank all the participants in this session according to the number of tasks completed in Part 1. Then it will match the two participants with the closest ranks into a pair.
For instance, a participant with rank 1 will be paired with a participant with rank 2, and a participant with rank 3 will be paired with the one in rank 4, etc.
The ranking and pairing assignments remain anonymous throughout the entire experiment. You will not be informed about the identity of the participant in this room you have been paired with. You will also not be able to learn your actual rank or the rank of anyone else. All you should keep in mind is that your opponent has the closest rank to you (based on performance records from Part 1).

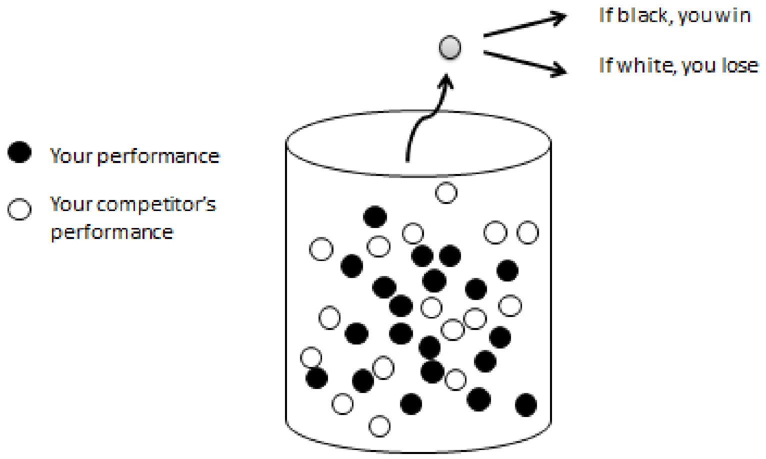
Competition

In this part, the two participants in a pair will compete against each other. The competition consists of maximum three independent rounds. The one who wins two out of the three rounds wins the competition and will be awarded a prize of AUD 20.

The competition in the first two rounds is organized as follows: both you and your competitor will work simultaneously and independently on the same 10-minute task as in Part 1.You will **NOT** learn how many tasks you or your opponent has completed during the competition. Nonetheless, the more tasks you complete, the more likely you are to win in each round.

The winner of round 1 and round 2 is determined according to the following procedure:

Imagine that you are facing an urn. Each time you complete a task correctly, you earn one black ball that will go into the urn. Each time your competitor complete a task correctly, he/she earn a white ball that goes into the urn. At the end of the round, the computer randomly picks a ball from the urn. If it is black you win that round, if it is white your competitor wins that round. This procedure is depicted in the following figure.



This means that your probability to win a round depends on both your performance and your competitor's performance in that round. Specifically, it is computed as:

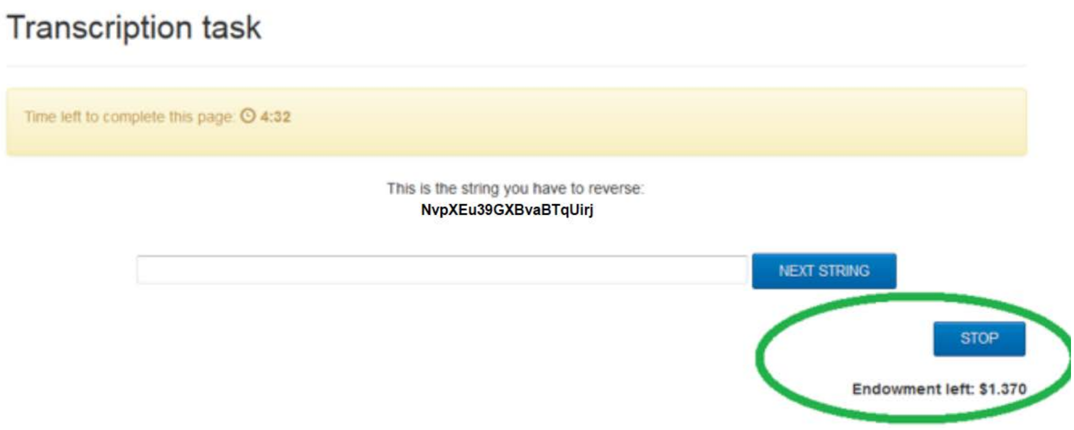
$$Your\ Probability\ to\ win = \frac{total\ number\ of\ tasks\ completed\ correctly\ by\ you}{total\ number\ of\ tasks\ completed\ by\ you\ and\ your\ competitor}$$

For example, if you and your competitor complete equal number of tasks (including if you both solve 0 tasks) you have 50% chance of winning. If you solve more (or less) tasks than your competitor, you have more than (or less than) 50% chance to win.

Time is money

Spending time to work on the tasks in this competition is costly. In the first two rounds, you are endowed with AUD 3 for each round. As soon as you enter a round, your endowment starts depleting. You pay AUD 0.005 (i.e., half a cent) per second for the time you spend working on the tasks. You are free to stop working at any point during the competition though.

Once you decide that you have worked enough for the competition, you could simply click the "STOP" button displayed at the right-bottom of the screen (see the screen shot below).



All the tasks you completed correctly before hitting the Stop button will be accounted for, when the winner is determined (i.e., the black balls you have earned stay in the urn).

For instance, if you stop as soon as the round starts, you will receive the full endowment AUD 3 for that round, but you will lose that round for sure if your competitor completes at least one task correctly; If you work for 5 minutes (300 seconds) and then press the Stop button, your initial endowment will be reduced by: AUD 1.50 (=300 x AUD 0.005) and the remaining endowment will be included in your final payment; At the same time, your chance of winning that round is determined by the total number of tasks you have completed correctly within that 5 minutes, together with the total number of tasks your competitor has completed.

Please note that once the “stop” button is hit, you cannot come back to the task anymore.

Tie breaking rule

Remember that you will not see the number of correct tasks completed by you and your competitor during each round. This is also true when the round is finished. However, you will be informed who the winner is at the end of each round. After two rounds of competition, if you have won (or lost) both rounds, you have won (or lost) the prize. There is no need to compete in the third round. However, if you and your competitor each have won one round, you will enter a third round to break the tie and the winner of the third round receives the prize.

In the third round, the computer will first randomly assign you and your opponent to odd numbers and even numbers respectively. And then the computer will throw a rolling die (with numbers 1, 2, 3, 4, 5, 6 on each side of the dice). If the number showing up on the die is an odd number (either 1, or 3, or 5), then the one who has been assigned to odd numbers will win round 3; if the number showing up on the dice is an even number (either 2, or 4, or 6), then the one who has been assigned to even numbers will win round 3. The winner in this round will be awarded the prize.

After the competition

After everyone has finished the competition, you will need to finish a simple exit questionnaire. All the information you provide in this questionnaire, as well as your performance data in the experiment will only be used for statistical analysis and will be kept anonymous and strictly confidential.

Once the questionnaire is done, your total payment from this experiment will be calculated as the following:

If you win:
Total payment=show-up fee (AUD 5) + earnings from Part 1 + prize (AUD 20) + total endowment you kept

If you lose:
Total payment=show-up fee (AUD 5) + earnings from Part 1 +total endowment you kept

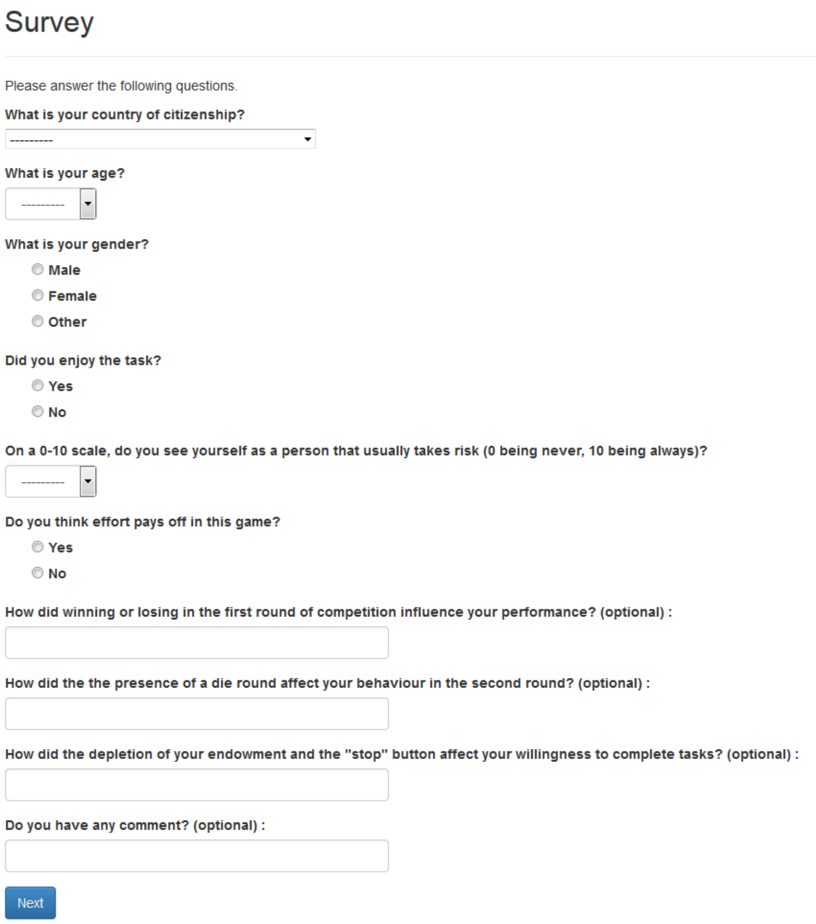
At the end of this session, your performance and your competitor's performance in each round, as well as your payoff will be displayed to you. See the following screenshot as an example.

The participation fee is $5.
Your total payoffs are: $22.73

Round number	Your performance	Your competitor's performance	Number of wins	Time left on counter (in s.)	Payoff
Part 1	5	-	-	-	$1.00
Part 2-round1	5	5	1	0.0	$0.00
Part 2-round2	5	3	1	346.0	$1.73
Part 2-round3	-	-	2	-	$20.00

D.1 Demographic questions

Figure 8: Screenshot of the end of experiment survey questions



Survey

Please answer the following questions.

What is your country of citizenship?

What is your age?

What is your gender?

- Male
- Female
- Other

Did you enjoy the task?

- Yes
- No

On a 0-10 scale, do you see yourself as a person that usually takes risk (0 being never, 10 being always)?

Do you think effort pays off in this game?

- Yes
- No

How did winning or losing in the first round of competition influence your performance? (optional) :

How did the the presence of a die round affect your behaviour in the second round? (optional) :

How did the depletion of your endowment and the "stop" button affect your willingness to complete tasks? (optional) :

Do you have any comment? (optional) :

Next