American Economic Review 2012, 102(1): 469–503
http://dx.doi.org/10.1257/aer.102.1.469

A Structural Analysis of Disappointment Aversion in a Real Effort Competition†

By David Gill and Victoria Prowse*

We develop a novel computerized real effort task, based on moving sliders across a screen, to test experimentally whether agents are disappointment averse when they compete in a real effort sequential-move tournament. We predict that a disappointment averse agent, who is loss averse around her endogenous choice-acclimating expectations-based reference point, responds negatively to her rival's effort. We find significant evidence for this discouragement effect, and use the Method of Simulated Moments to estimate the strength of disappointment aversion on average and the heterogeneity in disappointment aversion across the population. (*JEL* C91, D12, D81, D84)

Disappointment at doing worse than expected can be a powerful emotion. This emotion may be particularly intense when the disappointed agent exerted effort in competing for a prize, thus raising her expectation of winning. Furthermore, a rational agent who anticipates possible disappointment will optimize taking into account the expected disappointment arising from her choice.

In this paper we use a laboratory experiment to test whether agents are disappointment averse when they compete in a real effort tournament. In particular, we test whether our subjects are loss averse around reference points given by endogenous expectations that adjust to both an agent's own effort choice and that of her rival. Pairs of subjects complete a novel computerized real effort task, called the "slider task," which involves moving sliders across a screen. The *First Mover* completes the task, followed by the *Second Mover*, who observes the First Mover's effort before choosing how hard to work.[1] A money prize is awarded to one of the pair members based on the pair's relative work efforts and some element of chance, which we control. After

* Gill: Department of Economics, Oxford University, Manor Road, OX13UQ, UK (e-mail: david.gill@economics.ox.ac.uk); Prowse: Department of Economics, Cornell University, Ithaca, NY 14853 (e-mail: prowse@cornell.edu). We thank the University of Southampton, the George Webb Medley Fund, and the Zvi Meitar/Vice-Chancellor Oxford University Social Sciences Research Prize for financial support, the Nuffield Centre for Experimental Social Sciences and the Economic Science Laboratory at the University of Arizona for hosting our experiments, the Oxford Supercomputing Centre for computing facilities, and the following for helpful comments and discussion: three anonymous referees, Johannes Abeler, Charles Bellemare, Daniel Benjamin, John Bluedorn, Gary Charness, Miguel Costa-Gomas, Vincent P. Crawford, Gordon B. Dahl, Martin Dufwenberg, Jim Engle-Warnick, Shachar Kariv, Rosario Macera, Juanjuan Meng, Meg Meyer, Matthew Rabin, Charles Sprenger, Michael Vlassopoulos, John Wooders, and seminar participants at Aberdeen, Alicante, Arizona, Arizona State, Nottingham, Nuffield CESS, Southampton, UCLA, ESA meetings in Copenhagen, Innsbruck and Tucson, the Maastricht Symposium, the SITE Workshop on Psychology and Economics, the Stony Brook Workshop on Behavioral Game Theory and the Warwick Mathematics Institute Workshop on Game Theory.

† To view additional materials, visit the article page at http://dx.doi.org/10.1257/aer.102.1.469.

[1] We use a sequential tournament to give clean identification, rather than because most competitive situations involve sequential effort choices.

each repetition, the subjects are re-paired. We impose probabilities of winning the prize which are linear in the difference in the agents' efforts, so the marginal impact of a Second Mover's effort on her probability of winning does not depend on the effort of the First Mover she is paired with. Therefore, if agents care only about money and their cost of effort, the Second Mover's work effort should not depend on the effort of the First Mover. However, as predicted by our model of disappointment aversion, the experimental data show a *discouragement effect*: the Second Mover shies away from working hard when she observes that the First Mover has worked hard, and tends to work relatively hard when she observes that her competitor has put in low effort. Thus, First and Second Movers' efforts are strategic substitutes.

Our primary contribution is empirical. First, we offer evidence consistent with disappointment aversion from a reduced form linear random effects panel regression. More substantively, we exploit the richness of our experimental dataset to estimate the parameters of a structural model of disappointment aversion using the Method of Simulated Moments. This allows us to estimate the strength of disappointment aversion on average and the extent of heterogeneity in disappointment aversion across the population. Goodness of fit analysis shows that the estimated model fits our data well.

Together with random variation in the monetary prize across pairs of subjects, the design of our slider task generates sufficient variation in behavior to enable us to estimate the structural parameters of our model of disappointment aversion. In particular, the slider task gives a finely gradated measure of performance over a short time scale. As the task takes only two minutes to complete, we can collect repeated observations of the same Second Movers facing different prizes and First Mover efforts, while the fineness of the performance measure allows us to observe accurately how Second Movers respond to different prizes and First Mover efforts. The resulting panel data permit precise quantification of the distribution of the cost of effort and the strength of disappointment aversion across agents in the population.

The formal model that we test is a natural extension of disappointment aversion to situations in which agents compete. Models of disappointment aversion (e.g., Bell 1985; Loomes and Sugden 1986; Delquié and Cillo 2006; and Kőszegi and Rabin 2006, 2007) build on the idea that agents are sensitive to deviations from what they expected to receive; in particular, agents are loss averse around their expected payoff, so losses relative to this expectation are more painful than equal-sized gains are pleasurable. We model expectations-based reference points as adjusting to an agent's own effort choice and that of her rival: in the terminology of Kőszegi and Rabin (2007), they are choice-acclimating. The endogeneity of an agent's reference point is crucial: with loss aversion around a fixed reference point, even if given by a prior expectation, a Second Mover will continue to disregard First Mover effort.

Our empirical results thus address two important open questions in the literature on reference-dependent preferences: (i) what constitutes agents' reference points?; and (ii) how quickly do these reference points adjust to new circumstances? Our analysis provides evidence that when agents compete they have reference points given by their expected monetary payoff and that an agent's reference point adjusts essentially instantaneously to her own effort choice and to that of her competitor.

Abeler et al. (2011) also provide reduced-form evidence consistent with choice-acclimating reference-dependent preferences in the context of effort provision. Abeler et al. (2011) run a laboratory experiment in which subjects have a 50 percent chance of being paid piece-rate and a 50 percent chance of receiving a fixed payment, and show that effort increases in the fixed payment. To the best of our knowledge, however, we are the first to estimate the strength of loss aversion around choice-acclimating reference points when agents exert effort. Furthermore, we are able to leverage our structural analysis to provide evidence that the expectation that acts as the reference point adjusts to the agent's own choice of effort. In contrast, Abeler et al. (2011) do not distinguish between their choice-acclimating model and a more parsimonious model in which the reference point adjusts to the fixed payment but not the agent's actual effort choice. Finally, we provide evidence that choice-acclimating reference points are important in a different context to Abeler et al. (2011), namely one in which agents work to influence their probability of success. Such situations are common: in labor markets, workers often exert effort to increase their chances of winning promotions and bonuses; while agents also work to make success more likely in sports contests, examinations, patent races, and elections.

Complementary to our laboratory findings, Doran (2009), Crawford and Meng (2011), and Pope and Schweitzer (2011) find evidence of expectations–based reference-dependent preferences when cab drivers and professional golfers exert effort in the field.[2] In particular, Crawford and Meng (2011) estimate the average strength of loss aversion for cab drivers around rational expectations–based daily income and hours targets. In contrast to our model, in these papers the reference point is taken to be fixed when the agents choose how hard to work, and so is not choice-acclimating. Evidence of expectations-based reference points in the absence of effort provision includes Loomes and Sugden (1987), and Choi et al. (2007), who study choices over lotteries; Post et al. (2008), who find evidence that reference points adjust during the course of the game show "Deal or No Deal"; Card and Dahl (2011), who show that the probability of domestic violence when an NFL football team loses depends on the extent to which the loss was expected; and Ericson and Fuster (2010), who find that the valuation placed on an endowed good depends on the probability that a trading opportunity will arise. The psychology literature also supports the thesis that agents' emotional responses to the outcomes of gambles include disappointment and elation, that agents anticipate these emotions when choosing between gambles and that exerting effort, by increasing the likelihood of a good outcome, intensifies disappointment (Mellers, Schwartz, and Ritov 1999; van Dijk, van der Pligt, and Zeelenberg 1999).

Finally, our finding of a Second Mover discouragement effect adds to the existing literature on laboratory tournaments by showing that nonstandard preferences can move behavior away from that predicted by standard theory and by providing evidence of the impact of feedback during tournaments. Charness and Kuhn (2010) summarize the experimental literature. Bull, Schotter, and Weigelt (1987) study tournaments with an induced cost of effort; van Dijk, Sonnemans, and van Winden (2001) introduce

[2] An earlier literature in which reference points do not adjust to expectations explicitly also finds evidence of reference-dependent preferences in the field; for example, Camerer et al. (1997) study cab drivers and Fehr and Goette (2007) analyze bike messengers.

real effort; while Berger and Pope (2009), and Eriksson, Poulsen, and Villeval (2009), consider feedback.

The rest of the paper is structured as follows. Section I describes the slider task and the design of the experiment. Section II develops our model of disappointment aversion when agents compete. Section III presents the empirical analysis. Section IV discusses alternative behavioral explanations of the discouragement effect. Section V concludes. Appendix A derives proofs not included in the main text. Appendix B provides further details about the structural estimation method and the model's goodness of fit. Finally, Appendices C and D (in the online Appendix) lay out the instructions provided to the experimental subjects.

I. Experimental Design

We ran six experimental sessions at the Nuffield Centre for Experimental Social Sciences (CESS) in Oxford, all conducted on weekdays at the same time of day in late February and early March 2009 and lasting approximately 90 minutes.[3] Twenty student subjects (who did not report Psychology or Economics as their main subject of study) participated in each session, with 120 participants in total. The subjects were drawn from the CESS subject pool, which is managed using the Online Recruitment System for Economic Experiments (ORSEE). The experimental instructions (Appendix C in the online Appendix) were provided to each subject in written form and were read aloud to the subjects. Seating positions were randomized. To ensure subject-experimenter anonymity, actions and payments were linked to randomly allocated Participant ID numbers. Each subject was paid a show-up fee of £4 and earned an average of a further £10 during the experiment (all payments were in Pounds sterling). Subjects were paid privately in cash by the laboratory administrator. The experiment was programmed in z-Tree (Fischbacher 2007).

A. *The Slider Task*

Before setting out the experimental procedure, we first describe the novel computerized real effort task, which we call the "slider task," that we designed for the purpose of this experiment.

The slider task consists of a single screen displaying a number of sliders. The number and position of the sliders on the screen does not vary across experimental subjects or across repetitions of the task. A schematic representation of a single slider is shown in Figure 1. When the screen containing the effort task is first displayed to the subject all of the sliders are positioned at 0, as shown for a single slider in Figure 1A. By using the mouse, the subject can position each slider at any integer location between 0 and 100 inclusive. Each slider can be adjusted and readjusted an unlimited number of times and the current position of each slider is displayed to the right of the slider. The subject's "points score" in the task is the number of sliders positioned at 50 at the end of the allotted time. Figure 1B shows

[3] We also ran one pilot session without any monetary incentives whose results are not reported here.

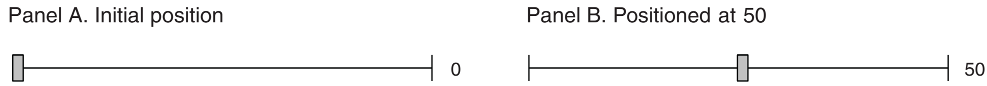

Figure 1. Schematic Representation of a Slider

a correctly positioned slider. As the task proceeds, the screen displays the subject's current points score and the amount of time remaining.

The number of sliders and task length can be chosen by the experimenter. In this experiment, we used 48 sliders and an allotted time of 120 seconds. The sliders were displayed on 22-inch widescreen monitors with a 1,680 by 1,050 pixel resolution. To move the sliders, the subjects used 800 dpi USB mice with the scroll wheels disabled.[4] Figure 2 shows a screen of sliders as shown to the subject in the laboratory. In this example, the subject has positioned three of the sliders at 50 and a points score of 3 is shown at the top of the screen. A fourth slider is currently positioned at 33 and this slider does not contribute to the subject's points score as it is not positioned correctly. To ensure that all the sliders are equally difficult to position correctly, the 48 sliders are arranged on the screen such that no two sliders are aligned exactly one under the other. This prevents the subject from being able to position the higher slider at 50 and then easily positioning the lower slider by copying the position of the higher slider.

The slider task gives a finely gradated measure of performance and involves little randomness; thus, we interpret a subject's point score as effort exerted in the task. In Section III we see that with 48 sliders and an allotted time of 120 seconds, measured effort varies from 0 to 41, so the task gives rise to substantial variation in behavior, and hence we can observe accurately how Second Movers respond to different prizes and First Mover efforts. As the task takes only two minutes to complete, we can collect repeated observations of the same Second Movers facing different prizes and First Mover efforts, allowing us to control for persistent unobserved heterogeneity. The resulting panel data enable us to use structural estimation to quantify precisely the distribution of the cost of effort and the strength of disappointment aversion across agents in the population.

The slider task also has a number of other desirable attributes: it is simple to communicate and to understand; it does not require or test preexisting knowledge; it is identical across repetitions; there is no scope for guessing; and as the task is computerized, it is easy to implement and allows flexible real-time subject interactions.

B. *Experimental Procedure*

In every session, 10 subjects were told that they would be a "First Mover" and the other 10 that they would be a "Second Mover" for the duration of the session. Each session consisted of two practice rounds followed by 10 paying rounds.

In every paying round, each First Mover was paired anonymously with a Second Mover. Each pair's prize was chosen randomly from $\{£0.10, £0.20, \dots, £3.90\}$ and revealed to the pair members. The First and Second Movers then completed our

[4] The keyboards were also disabled to prevent the subjects from using the arrow keys to position the sliders.

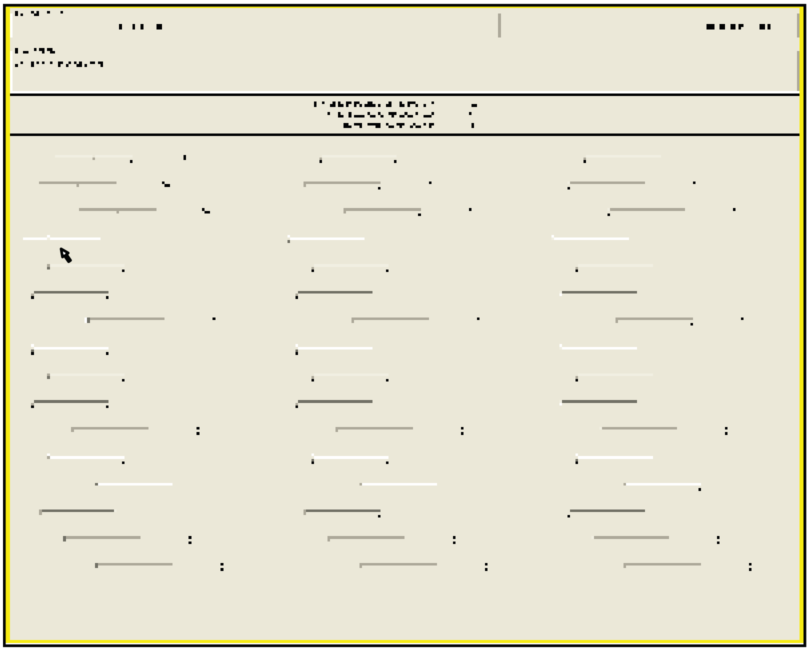

Figure 2. Screen Showing 48 Sliders

Note: The screen presented here is slightly squarer than the one seen by our subjects.

slider task sequentially, with the Second Mover discovering the points score of the First Mover she was paired with before starting the task. As explained in Section IA, we used a slider task with 48 sliders and an allotted time of 120 seconds. During the task, a number of pieces of information appeared at the top of the subject's screen: the round number; the time remaining; whether the subject was a First or Second Mover; the prize for the round; and the subject's points score in the task so far. If the subject was a Second Mover, she also saw the points score of the First Mover. Figure 2 provides an example of the screen visible to the Second Movers.

The probability of winning the prize for each pair member was 50 plus her own points score minus the other pair member's points score, all divided by 100. Thus, we imposed winning probabilities linear in the difference of the points scores, with equal points scores giving equal winning probabilities, while an increase of 1 in the difference raised the chance of winning by 1 percentage point for the pair member with the higher points score. The probability of winning function was explained verbally and using Table 6. (See Appendix C in the online Appendix.) At the end of the round, the subjects saw a summary screen showing their own points score, the other pair member's points score, their probability of winning the prize given the respective points scores, the prize for the round, and whether they were the winner or loser of the prize in that round.

After each paying round the subjects were re-paired according to the "no contagion" matching algorithm of Cooper et al. (1996). This rotation-based algorithm ensures that not only do the same subjects never meet each other more than once, but that each round is truly one-shot in the sense that a given subject's actions in one round cannot influence, either directly or indirectly, the actions of other subjects who the subject is paired with later on. The explanation to the subjects in the experimental instructions provides further detail.

Before starting the paying rounds, the subjects played two practice rounds to gain familiarity with the task and procedure and to give opportunities for questions. To prevent contamination, the subjects were made aware that during the practice rounds they were playing against automata who behaved randomly. At the end of each practice round, the subjects were informed of what their probability of winning would have been given the respective points scores, but were not told that they had won or lost in that round, and no prizes were awarded. We do not include the practice rounds in the econometric analysis.

II. Theoretical Predictions

In this section, we provide a theoretical model of the behavior of a generic pair of First and Second Movers competing for a prize v in a particular round. After describing the model, we show that in the absence of disappointment aversion the Second Mover's effort does not depend on the First Mover's effort, while a disappointment averse Second Mover will respond negatively to the effort choice of the First Mover.

A. *One-Shot Theory Model*

Two agents compete to win a fixed prize of monetary value $v > 0$ in a rank-order tournament, choosing their effort levels sequentially. The First Mover chooses her effort level e_1 from an action space $\mathcal{A} \subseteq [0, \overline{e}]$ which can be discrete or continuous. The Second Mover observes e_1 before choosing her effort level e_2 from $\mathcal{A}$. As noted in Section IA, we interpret a subject's points score in the slider task as effort exerted.[5] Agent i's probability of winning the prize $P_i(e_i, e_j)$ increases linearly in the difference between her own effort, e_i, and the other agent's effort, e_j. Assuming symmetry of the probability of winning functions,

$$P_i(e_i, e_j) = \frac{e_i - e_j + \gamma}{2\gamma}, \tag{1}$$

where we impose $\gamma \geq \overline{e}$ to ensure that $P_i \in [0, 1]$.[6] Throughout, we focus on the behavior of the Second Mover conditional on the First Mover's effort, e_1. Thus, we are able to abstract from any game-theoretic considerations, as the Second Mover faces a pure optimization problem given the First Mover effort that she observes.

[5] Our use of the term "effort" therefore corresponds to the behavior of the agent (the number of correctly positioned sliders) rather than the associated cost of positioning the sliders.

[6] Che and Gale (2000) call this a piece-wise linear difference-form success function. Note that for any First Mover effort $e_1 \in \mathcal{A}$, the Second Mover's probability of winning function is given by $P_2 = (e_2 - e_1 + \gamma)/(2\gamma)$ for the whole range of $e_2 \in \mathcal{A}$ as $e_1 \geq 0$ and $e_2 \leq \overline{e} \leq \gamma$ so $e_2 \leq \gamma + e_1$ and hence $(e_2 - e_1 + \gamma)/(2\gamma) \leq 1$, while $e_1 \leq \overline{e} \leq \gamma$ and $e_2 \geq 0$ so $e_2 \geq e_1 - \gamma$ and hence $(e_2 - e_1 + \gamma)/(2\gamma) \geq 0$. In our experiment, we set $\overline{e} = 48$ and $\gamma = 50$.

B. *No Disappointment Aversion*

Applying the canonical model in the tournament literature, the Second Mover's utility U_2 is separable into utility $u_2(y_2)$ from her tournament payoff $y_2 \in \{0, v\}$, which we call her *material utility*, and her cost of effort $C_2(e_2)$, so

$$U_2(y_2, e_2) = u_2(y_2) - C_2(e_2). \tag{2}$$

This separability assumption in the canonical model is equivalent to saying that: (i) the cost of effort does not depend on whether the agent wins the prize; and (ii) when the monetary prize is awarded, the valuation placed on the prize is independent of how hard the agent worked. The agent exerts effort in a two-minute interval before the outcome of the tournament is determined, justifying the first assumption. In relation to (ii), disappointment aversion is a microfounded explanation of why agents might indeed care about efforts exerted when evaluating the prize (our model of disappointment predicts that the Second Mover's value to winning relative to losing is increasing in her own effort).[7]

Separability implies that the Second Mover's expected utility is given by

$$\begin{aligned} EU_2(e_2, e_1) &= P_2(e_2, e_1)u_2(v) + \big(1 - P_2(e_2, e_1)\big)u_2(0) - C_2(e_2) \\ &= \big(u_2(v) - u_2(0)\big)\left(\frac{e_2 - e_1 + \gamma}{2\gamma}\right) + u_2(0) - C_2(e_2). \end{aligned} \tag{3}$$

As the winning probabilities are linear in the difference in efforts, the First Mover's effort e_1 has no effect on the marginal impact of the Second Mover's effort e_2 on her probability of winning. Thus, the Second Mover's marginal utility with respect to her own effort does not depend on e_1, giving the following result.

PROPOSITION 1: *In the canonical model without disappointment aversion the Second Mover's optimal effort* e_2^* (*or set of optimal efforts*) *does not depend on the First Mover's effort* e_1.

Note that we have not imposed any concavity, continuity, or differentiability assumptions on $u_2(y_2)$ (and nor have we assumed anything about the shape of $C_2(e_2)$). Thus, the result continues to hold if the Second Mover exhibits any degree of risk aversion over her monetary payoff, if she places a value on winning per se in addition to the value placed on the monetary payoff from winning (as $u_2(v)$ can incorporate this joy of winning), if she is inequity averse over monetary payoffs (Fehr and Schmidt 1999), or if she is loss averse around a *fixed* reference point (the last two follow as the utility to winning or losing can be redefined to incorporate a comparison to a fixed reference

[7] In their related work, Abeler et al. (2011) and Crawford and Meng (2011) make an equivalent separability assumption. In contrast to the standard labor literature where agents vary their hours of work, in our setting the time spent on the task is fixed and therefore only the intensity of effort can affect the valuation of the prize: thus standard complementarities or substitutabilities between leisure and consumption do not apply directly.

point or to the payoff of the First Mover). The result also holds if $u_2(y_2)$ incorporates an impact of winning or losing on the utility function in any later tournaments, e.g., via changes in wealth or the reference point.

C. *Disappointment Aversion*

Models of disappointment aversion (e.g., Bell 1985; Loomes and Sugden 1986; Delquié and Cillo 2006; Kőszegi and Rabin 2006, 2007) build on the idea that agents are sensitive to deviations from their expectations, suffering a psychological loss when they receive less than expected and experiencing elation when they receive more. Furthermore, agents anticipate these losses and gains when deciding how to behave.

We follow the literature in embedding disappointment aversion in a loss aversion-type framework. Suppose that the Second Mover compares her material utility $u_2(y_2)$ to a reference level of utility R_2, suffering losses when $u_2(y_2)$ is less than this reference point and enjoying gains when $u_2(y_2)$ exceeds the reference point. Specifically, total utility U_2 is given by[8,9]

$$U_2(y_2, R_2, e_2) = u_2(y_2) + 1_{u_2(y_2) \geq R_2} G_2(u_2(y_2) - R_2) + 1_{u_2(y_2) \leq R_2} L_2(u_2(y_2) - R_2) - C_2(e_2), \tag{4}$$

where the loss function $L_2(x) < 0$ for $x < 0$, the gain function $G_2(x) \geq 0$ for $x > 0$, and $G_2(0) = L_2(0) = 0$. The utility arising from the comparison of $u_2(y_2)$ to the reference point is termed *gain-loss utility*. The Second Mover is said to be *loss averse* if losses due to downward departures from the reference point are more painful than equal-sized upward departures are pleasurable, i.e., $G_2(x) < |L_2(-x)|$ for all $x > 0$. The Second Mover is *first-order loss averse* if she is loss averse in the limit as the deviations from the reference point go to zero, i.e., $\lim_{x \uparrow 0} L_2'(x) > \lim_{x \downarrow 0} G_2'(x)$, assuming differentiability of gain-loss utility except at the kink where $x = 0$.

Starting with Kahneman and Tversky (1979), most models of loss aversion take the reference point to be fixed exogenously, for example assuming it to be equal to the status quo. We noted above that the utility formulation (2) is flexible enough to incorporate loss aversion around a fixed reference point. Thus, a fixed reference point does not introduce any interdependence between the efforts of the First and Second Movers. (To see that Proposition 1 continues to hold, note that if $u_2(y_2)$ in (2) is redefined to include gain-loss utility, the analysis proceeds as before.)

[8] The Prospect Theory of Kahneman and Tversky (1979) incorporates a loss averse value function defined only over losses and gains relative to the reference point, while we follow the disappointment aversion literature in defining total utility over both material utility and gain-loss utility arising from the comparison of material utility to the reference point.

[9] By modeling each tournament as a one-shot interaction, we are assuming that our subjects frame each tournament narrowly, i.e., they compare the outcome of each tournament to their reference point in isolation. In our setting, each interaction is one-shot and uncertainty is resolved immediately: at the end of each tournament, the subjects find out whether they won or lost and then get rematched with a new rival. Models and tests of loss aversion generally incorporate narrow framing, either implicitly or explicitly (DellaVigna 2009) and the literature on narrow framing suggests that attitudes towards small gambles can only be explained by loss aversion together with the narrow framing of individual gambles (Barberis, Huang, and Thaler 2006).

Instead of holding a fixed reference point, we assume that a disappointment averse Second Mover is loss averse around an *endogenous* reference point equal to her expected material utility given the effort levels that are actually chosen, so

$$R_2 = E[u_2(y_2) \mid e_2, e_1]. \tag{5}$$

Thus, a Second Mover's reference point will be sensitive to both the effort chosen by the First Mover and her own effort, and when optimizing the Second Mover understands that her effort choice affects her reference point. Notice that the endogeneity of the expectation is crucial. If the Second Mover starts with a reference point equal to a prior expectation that is invariant to the effort levels that are actually chosen, the reference point is fixed so that, as explained above, Proposition 1 still holds. Instead, our reference point adjusts to the agents' choices: in the terminology of Kőszegi and Rabin (2007), the reference point is *choice-acclimating*.[10]

To operationalize our model, we linearize material utility and gain-loss utility.[11] We assume that $u_2(y_2) = y_2$, so material utility is linear in money and the Second Mover's reference point becomes her expected monetary payoff, i.e.,

$$R_2 = vP_2(e_2, e_1). \tag{6}$$

Furthermore, we assume that the gain-loss utility arising from the comparison of $u_2(y_2)$ to the reference point is piece-wise linear, with a constant slope of g_2 in the gain domain and l_2 in the loss domain. With piece-wise linearity, loss aversion implies that $l_2 > g_2$, so losses are more painful than same-sized gains are pleasurable.[12] Thus we define disappointment aversion as follows.[13]

DEFINITION 1: *A disappointment averse Second Mover is loss averse around her expected monetary payoff, so $\lambda_2 \equiv l_2 - g_2$, which measures the strength of disappointment aversion, is strictly positive.*

[10] Technically our game is psychological as the Second Mover's utility depends on her beliefs about the chosen efforts via the reference point. In particular, our game falls under the framework of a dynamic psychological game (Battigalli and Dufwenberg 2009) as utility depends on terminal node (ex post) beliefs, which are pinned down by the chosen efforts, so beliefs can update during the course of the game.

[11] Given that the experimental stakes are small, we believe this comes at a low cost.

[12] With piece-wise linearity, loss aversion and first-order loss aversion are equivalent. If $l_2 = g_2$, gains and losses relative to the reference point cancel out in expectation, so the agent acts as if she had standard preferences.

[13] Our model of disappointment aversion directly extends the single-agent setup of Bell (1985) to our competitive environment. Our formulation is also equivalent to that of Delquié and Cillo (2006) and the choice-acclimating model of Kőszegi and Rabin (2007, Section IV); it is straightforward to show that in a linear environment with only two outcomes, the reference lottery approach in those papers is equivalent to using a single reference point given by the endogenous expected payoff. We use the same parameterization as Bell (1985) and Delquié and Cillo (2006); Section IIIC6 explains the equivalence between our parameterization and that used by Kőszegi and Rabin (2007). Although in the same spirit, our model of disappointment is slightly different to that of Loomes and Sugden (1986) who do not allow a kink in utility, but instead use a disappointment/elation function that is nonlinear around the endogenous expected payoff. Finally, the Kőszegi and Rabin (2006) model does not predict a discouragement effect as in their "personal equilibrium" the reference point is taken to be fixed at the point of optimization.

We can then express a disappointment averse Second Mover's expected utility as

$$(7)\qquad EU_2(e_2,e_1) = P_2(v + g_2(v - vP_2)) + (1 - P_2)(0 + l_2(0 - vP_2)) - C_2(e_2) = vP_2 - \lambda_2 vP_2(1 - P_2) - C_2(e_2),$$

and we let

$$(8)\qquad \Lambda_2(e_2,e_1) \equiv -\lambda_2 v P_2(1 - P_2)$$

represent the extra term introduced into expected utility by disappointment aversion. We call Λ_2 the Second Mover's *disappointment deficit* as it is always negative for $\lambda_2 > 0$ (strictly negative for $P_i \notin \{0,1\}$). For a given prize v the disappointment deficit is proportional to $v^2P_2(1 - P_2)$, the variance of the Second Mover's two-point distribution of monetary payoffs. A disappointment averse Second Mover dislikes variance in her monetary payoff as losses relative to her expected payoff loom larger than gains. (With risk aversion, agents care only about their probability of winning as there are only two possible outcomes.)

The variance is strictly concave in P_2 and maximized at $P_2 = ½$. When efforts are such that the Second Mover has zero probability of winning, the Second Mover has a reference point of zero and her realized payoff equals her reference point; she is never disappointed and never receives more than expected. Hence, her disappointment deficit is zero. Starting at zero, a small increase in her probability of winning leads to a large increase in the variance of her monetary payoff. Further increases in the probability of winning towards ½ lead to further yet smaller increases in the variance. At $P_2 = ½$ the variance is at its highest so the disappointment deficit is at its most negative—irrespective of whether she wins or loses, the Second Mover's realized payoff is very different from her expected payoff. Starting at $P_2 = ½$, increases in the probability of winning reduce the variance, initially by small amounts, and then by larger amounts as the probability of winning approaches 1.

For any value of the Second Mover's effort, an increase in the First Mover's effort reduces the Second Mover's probability of winning. The variance therefore increases faster in P_2 (when $P_2 < ½$) or falls less fast in P_2 (when $P_2 > ½$), so the Second Mover has a lower marginal incentive to exert effort (given her effort always has the same marginal effect on her probability of winning). We thus have a *discouragement effect*, which is crucial to our identification strategy: a disappointment averse Second Mover responds negatively to the First Mover's effort, so the harder the First Mover works, the more the Second Mover shies away from exerting effort. Thus, First and Second Mover efforts are strategic substitutes.[14]

[14] It is straightforward to extend the proof of Proposition 2 to show that if λ_2 were negative, the Second Mover would respond positively to the First Mover's effort.

PROPOSITION 2: *When the Second Mover is disappointment averse, higher First Mover effort discourages the Second Mover: the Second Mover's optimal effort* e_2^* *is always (weakly) decreasing in the First Mover's effort* e_1.

PROOF:
See Appendix AA.

Up to now we have imposed no assumptions on the shape of the cost of effort function. In order to derive an analytical expression for how the Second Mover responds to the First Mover's effort, and to see how the slope of the reaction function changes in the value of the prize and the strength of disappointment aversion, we now assume a quadratic cost of effort function:

$$C_2(e_2) = be_2 + \frac{ce_2^2}{2}. \tag{9}$$

With this cost function, the Second Mover's objective function will be everywhere convex or everywhere concave. With strict convexity, the Second Mover will always set effort at a corner. Instead we focus here on the case of strict concavity, which allows interior optima, showing that the discouragement effect becomes stronger as the Second Mover becomes more disappointment averse or the value of the prize goes up.

PROPOSITION 3: *Suppose a disappointment averse Second Mover has a quadratic cost function (given by (9)) and a strictly concave objective function, i.e.,* $2\gamma^2 c - \lambda_2 v > 0$. *When the action space is continuous, the slope of the Second Mover's reaction function in the interior is given by*

$$\frac{de_2^*}{de_1} = \frac{-\lambda_2 v}{2\gamma^2 c - \lambda_2 v} < 0, \tag{10}$$

which becomes strictly more negative in the strength of disappointment aversion λ_2 *and the value of the prize* v. *When the action space is discrete, the discrete analog of the reaction function behaves similarly.*

PROOF:
See Appendix AB.

These effects are intuitive. Referring back to (8) we see that the disappointment deficit term becomes more negative in the strength of disappointment aversion λ_2 and the value of the prize v, so the Second Mover becomes more sensitive to First Mover effort as v and λ_2 go up.

III. Empirical Analysis

A. *Overview and Sample Description*

We use the dataset collected from the laboratory experiment described in Section I to test our theory of disappointment aversion. In Section IIIB we show in a reduced form setting that, as predicted by our theory of disappointment aversion, Second Movers respond negatively to the effort choice of the First Mover they are paired with and that the strength of this effect is increasing in the value of the prize. In Section IIIC we use structural modeling to estimate the strength of disappointment aversion on average and the heterogeneity in disappointment aversion across the population. As outlined in the introduction and Section IA, our estimation strategies exploit identifying variation obtained from the properties of our slider task together with the experimental design.[15]

We analyze the behavior of Second Movers conditional on the effort choices of the First Movers. As noted in Section IIA, this allows us to abstract from any game-theoretic considerations as the Second Movers face a pure optimization problem. This conditional analysis is sufficient for the purpose of identifying the presence and strength of disappointment aversion. Moreover, solving for the optimal behavior of the First Movers would require further assumptions concerning the First Movers' beliefs about the unobserved characteristics and behavior of the Second Movers. We avoid these issues, together with the associated computational complexities and potential sources of misspecification, when performing a conditional analysis of the Second Mover effort choices.[16]

As explained in Section IA, we interpret the number of sliders correctly positioned by a subject within the allotted time, i.e., the points score, as the effort exerted by the subject in the task. While the slider task provides a finely gradated measure of effort, effort is still discrete. We emphasize that this discreteness is entirely unproblematic. Indeed, the above theoretical framework encompasses both discrete and continuous effort choices, and the testable implications of our theory of disappointment aversion apply irrespective of whether effort is discrete or continuous. In addition, as detailed below, discrete effort choices are easily accommodated in our structural model.

From the laboratory sessions we collected data on 60 First Movers and 60 Second Movers, each observed for 10 paying rounds, with re-pairing between rounds as detailed in Section IB. One Second Mover appears to have been unable to position any sliders at exactly 50.[17] Throughout our analysis this subject is dropped, except for the purpose of showing that our results are robust to our sample selection. Table 1 summarizes the behavior of the 59 Second Movers and the corresponding First Movers in each round. Efforts range between 0 and 41 sliders for First Movers and 0 and 40 sliders for Second Movers. Within each round, on average First and Second Movers exert roughly the same effort, with average effort increasing from around 22 sliders to just under 27 sliders over the 10 rounds.

[15] Evidence from the field, in which agents operate in a natural environment, would complement our laboratory findings. See Levitt and List (2007) for a discussion of the relationship between laboratory and field evidence more generally.

[16] Similar problems would arise if we attempted to identify disappointment aversion from responses to the prize in a simultaneous-move tournament. Furthermore, in a simultaneous-move context, even if all subjects were identical, disappointment aversion would be difficult to identify as symmetric and asymmetric pure-strategy equilibria coexist for certain values of the prize and subjects might play mixed-strategy equilibria.

[17] The data show that this subject was moving sliders around throughout the session but failed to position any sliders at exactly 50 in either the practice rounds or in the paying rounds. This subject also experienced problems when entering his/her Participant ID number.

Table 1—Summary of First and Second Mover Efforts

Paying					Minimum		Maximum	
round	Mean e_1	SD e_1	Mean e_2	SD e_2	e_1	e_2	e_1	e_2
1	22.034	5.991	21.763	6.101	1	0	33	34
2	22.627	6.708	23.458	4.836	0	11	33	33
3	24.763	6.075	24.831	4.875	0	12	37	38
4	24.627	5.956	25.203	4.502	0	16	35	36
5	24.966	6.800	25.119	5.660	0	0	36	35
6	24.729	7.508	24.898	7.039	1	0	37	39
7	25.881	5.855	25.763	6.109	9	0	37	37
8	26.831	5.858	26.169	5.133	9	14	41	35
9	25.593	8.550	26.254	6.702	0	0	38	40
10	26.322	6.781	26.729	5.988	1	0	40	39

Note: SD denotes standard deviation and e_1 and e_2 denote, respectively, First and Second Mover effort.

B. *Reduced Form Analysis*

We use a panel data regression to examine whether Second Movers respond to the effort choice of the First Mover they are paired with. Exploiting Proposition 1, we hypothesize that if Second Movers are not disappointment averse then the observed efforts of the Second Movers will not depend on the corresponding First Mover efforts once controls for the prize and round effects are included.[18] Alternatively, if subjects are disappointment averse then Proposition 2 implies a negative dependence of observed Second Mover efforts on the corresponding First Mover efforts, again conditional on controls for the prize and round effects.

To explore how Second Movers respond to First Mover effort, we estimate the following linear random effects panel data model:

$$(11) \qquad e_{2,n,r} = \beta_1 + \beta_2 v_{n,r} + \beta_3 e_{1,n,r} + \beta_4 e_{1,n,r} \times v_{n,r} + d_r + \omega_n + \epsilon_{n,r} \quad \text{for} \quad n = 1,\ldots,N; \quad r = 1,\ldots,10,$$

where n and r index, respectively, Second Movers and paying rounds, and N denotes the total number of Second Movers. $e_{1,n,r}$ is the effort of the First Mover paired with the nth Second Mover in the rth round, and $v_{n,r}$ is the prize draw for the nth Second Mover in the rth round. The prize, the First Mover's effort, and the First Mover's effort interacted with the prize are included as explanatory variables. The inclusion of the interaction of the prize and the First Mover's effort is motivated by Proposition 3, which shows that in the case of a quadratic cost of effort function, the negative effect of the First Mover's effort on the Second Mover's optimal effort is larger at higher prizes. Additionally, the equation includes a set of round dummies denoted by d_r for $r = 1,\ldots,10$, with the first paying round providing the omitted category, to capture systematic differences between rounds that are common across Second Movers, and round invariant Second Mover specific effects denoted ω_n for $n = 1,\ldots,N$ to capture systematic differences between Second Movers. Lastly, $\epsilon_{n,r}$ is an unobservable that

[18] We note, however, that First and Second Mover efforts will not be independent unconditionally in the presence of prize and round effects that impact on both pair members.

Table 2—Random Effects Regressions for Second Mover Effort

	Preferred sample 59 Second Movers		Full sample 60 Second Movers	
	Coefficient	Standard error	Coefficient	Standard error
First Mover effort	0.044	0.049	0.047	0.049
Prize	1.639***	0.602	1.655***	0.592
Prize × First Mover effort	−0.049**	0.023	−0.050**	0.023
Intercept	19.777***	1.400	19.392***	1.447
σ_ω	4.288		5.342	
σ_ϵ	3.852		3.826	
$N \times R$	590		600	
Hausman test for random versus fixed effects	2.60 $df = 12, p = 0.998$		2.43 $df = 12, p = 0.998$	

Notes: df denotes degrees of freedom. Both specifications further include dummy variables for each of rounds 2–10 inclusive. The coefficients on these variables are between 1.7 and 5.2, significantly greater than zero, and tend to increase over the rounds.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

varies over rounds and over Second Movers and captures differences between rounds in a Second Mover's effort choice that cannot be attributed to the other terms in the model. ω_n is assumed to be identically and independently distributed over Second Movers with a variance σ_ω^2, while $\epsilon_{n,r}$ is assumed to be identically and independently distributed over rounds and Second Movers with a variance σ_ϵ^2.

Table 2 reports estimates of the parameters appearing in (11). The results for the preferred sample show a negative effect of First Mover effort on Second Mover effort. In more detail, at low prizes First Mover effort does not significantly affect Second Mover effort, while at high prizes there is a large and significant discouragement effect as predicted by our theory of disappointment aversion. Application of the Delta method reveals that the effect of First Mover effort on Second Mover effort is not significant at the 5 percent level for prizes less than £2, is significant at the 5 percent level for prizes between £2 and £2.60, and is significant at the 1 percent level for prizes of £2.70 and above. For the highest prize of £3.90, a 40 slider increase in First Mover effort decreases Second Mover effort by approximately 6 sliders, a 24 percent decrease relative to the average effort of 25 sliders.[19] Furthermore, there is a large and significant positive prize effect, and we find that the persistent unobserved individual characteristics explain more of the variation in behavior than the transitory unobservables.[20]

[19] We use a 40 slider increase as an illustrative example as First Mover efforts ranged from 0 to about 40 (see Table 1).

[20] To test whether behavior changes during the course of the experiment due to learning about the payoff function, we also estimated the model with separate coefficients on First Mover effort and First Mover effort interacted with the prize for the first five rounds of the experiment and for the final five rounds. The parameters corresponding to the first half of the experiment are not statistically significantly different from those corresponding to the second half (p-value of 0.710). We further estimated the model excluding all prize controls, and found a negative but statistically insignificant effect of First Mover effort on Second Mover effort: both First and Second Mover efforts are positively correlated with the prize, so without prize controls the coefficient on First Mover effort is biased upward as it picks up part of the effect of the prize on Second Mover effort.

We note that, although the parameters reported in Table 2 were estimated from a linear random effects model, an alternative specification in which round invariant Second Mover specific effects are treated as fixed effects yields almost indistinguishable results. This is because Second Mover specific effects are uncorrelated with the prize and the First Mover efforts due to the experimental design. Finally, Table 2 shows that including the 60th Second Mover does not change conclusions concerning significance, and neither does this have substantial effects on the coefficient estimates.

C. *Structural Modeling*

Structural modeling seeks to fit the theoretical model with disappointment aversion, presented above in Section IIC, to the experimental sample. In contrast to the reduced form analysis above, structural modeling recovers estimates of the strength of disappointment aversion on average and the population-level heterogeneity in disappointment aversion. Below, we describe our preferred empirical specification, our estimation strategy, including a discussion of identification, and our results. We then explore robustness to the specification of the reference point and of the cost of effort function. Finally, we relate our estimate of the disappointment aversion parameter to existing measures of loss aversion around fixed reference points.

1. *Preferred Empirical Specification.*—We use $\lambda_{2,n}$ to denote the disappointment aversion parameter of the nth Second Mover. In this specification the strength of disappointment aversion may vary between subjects; however, for a given subject the strength of disappointment is constant over rounds. We adopt the following specification for $\lambda_{2,n}$:

$$(12) \qquad \lambda_{2,n} \sim N(\tilde{\lambda}_2, \sigma_\lambda^2) \quad \text{for} \quad n = 1, \ldots, N,$$

and further assume that $\lambda_{2,n}$ is independent over Second Movers. The parameter $\tilde{\lambda}_2$ represents the strength of disappointment aversion on average, and σ_λ^2 denotes the variance of the strength of disappointment aversion in the population.

The cost of effort function is assumed to be quadratic, as in (9). The parameter b is assumed to be constant over rounds and common to Second Movers, while unobserved cost differences between Second Movers and learning effects enter the cost of effort function through the convexity parameter c; $c_{n,r}$ denotes the convexity parameter of the nth Second Mover in the rth round and takes the following form:

$$(13) \quad c_{n,r} = \kappa + \delta_r + \mu_n + \pi_{n,r} \quad \text{for} \quad n = 1, \ldots, N; \quad r = 1, \ldots, 10.$$

In the above, κ denotes the component of $c_{n,r}$, that is common across Second Movers and rounds. δ_r for $r = 1, \ldots, 10$ are round effects, with the first paying round providing the omitted category. These round dummies allow the marginal costs of the first and later units of effort to vary over rounds at the population level. A cost of effort that is declining over rounds due to learning is represented by values of δ_r which are negative and decreasing over rounds. μ_n denotes unobserved differences in the cost of effort functions across Second Movers that are constant over rounds.

For the purpose of estimation, μ_n is assumed to be independent over Second Movers and to have a Weibull distribution with scale parameter ϕ_μ and shape parameter φ_μ. The final term in the cost function is $\pi_{n,r}$, which represents unobserved differences in Second Movers' cost of effort functions that vary over rounds as well as over Second Movers; $\pi_{n,r}$ is assumed to be independent over Second Movers and rounds and to have a Weibull distribution with scale parameter ϕ_π and shape parameter φ_π. The Weibull distribution is a flexible two-parameter distribution that has positive support, thus allowing us to impose convex cost of effort functions on all Second Movers when estimating the model.[21]

Given this parameterization of the theoretical model, the structural model has 17 unknown parameters, corresponding to the parameters describing disappointment aversion, $\tilde{\lambda}_2$ and σ_λ, the common cost parameters b and κ, the nine round effects δ_r for $r = 2, \ldots, 10$, and the four parameters appearing in the distribution of the unobservables in the cost of effort function, namely, ϕ_μ, φ_μ, ϕ_π, and φ_π. These 17 structural parameters are collectively denoted by the vector $\boldsymbol{\theta}$.

2. *Estimation Strategy and Identification*.—We estimate the 17 unknown parameters using the Method of Simulated Moments (MSM) (McFadden 1989; Pakes and Pollard 1989). The analytic complexity of choice probabilities, due to the multiple sources of unobserved heterogeneity, precludes the use of Maximum Likelihood and Method of Moments estimation techniques. MSM, in contrast, uses easily computed features of the sample as the basis for estimating the unknown parameters. Formally, the sample observations are used to compute a $k \times 1$ dimensional vector of moments, with $k \geq 17$, denoted $\mathbf{M}$. Critically, every moment included in $\mathbf{M}$ should depend at least in part on one or more endogenous variables. The researcher has considerable discretion over the moments included in $\mathbf{M}$; however, $\mathbf{M}$ typically includes period-specific averages of endogenous variables, here the effort choices of the Second Movers in each round, together with correlations between the endogenous variables and the explanatory variables.

MSM proceeds by generating S simulated samples. Each simulated sample contains N Second Movers, each observed for 10 rounds. In each simulated sample the Second Movers face the same prizes and First Mover efforts as observed in the actual sample. The behavior of the Second Movers in the simulated samples is determined from the structural model using a trial value, $\boldsymbol{\theta}_t$, of the values of the unknown parameters, $\boldsymbol{\theta}$. In particular, unobservables are assigned to Second Movers in accordance with the above described distributions. For each Second Mover and each round, the expected utility is calculated for each feasible Second Mover effort choice, and the simulated effort choice is the action with the highest expected utility. Further details concerning the construction of the simulated samples are provided in Appendix BA.

The behavior of the Second Movers in the simulated samples is then compared to the behavior of the actual experimental subjects. Specifically, for each of the S simulated samples the vector of moments $\mathbf{M}_s(\boldsymbol{\theta}_t)$ is computed. These are the

[21] In the preferred specification we use round-specific cost of effort functions; however, we also estimated the model with a flexible cost of effort function defined over a Second Mover's total number of completed sliders during the paying rounds, x. Specifically, the cost of effort function was given by $bx + (cx^\psi)/\psi$ where, as in the preferred specification, c included Weibull distributed subject-specific random effects. The Newey test for the validity of overidentifying restrictions firmly rejected this specification.

same k moments as computed for the observed sample. The simulated moments $\mathbf{M}_s$ are a function of the parameters $\boldsymbol{\theta}_t$ used to simulate the behavior of the Second Movers, as different values of the parameters imply different optimal Second Mover effort choices. The average of $\mathbf{M}_s$ over the S simulated samples, $\left(\sum_{s=1}^{S} \mathbf{M}_s(\boldsymbol{\theta}_t)\right)/S$, provides a summary of the behavior of Second Movers in the simulated samples. The process of averaging over the S simulated samples reduces the effect of simulation noise on the simulated moments. The following metric is then formed:

$$(14) \qquad J(\boldsymbol{\theta}_t) = \left(\mathbf{M} - \frac{1}{S}\sum_{s=1}^{S} \mathbf{M}_s(\boldsymbol{\theta}_t)\right)' \mathbf{W}_N \left(\mathbf{M} - \frac{1}{S}\sum_{s=1}^{S} \mathbf{M}_s(\boldsymbol{\theta}_t)\right),$$

where $\mathbf{W}_N$ is a fixed $k \times k$ dimensional positive semidefinite weighting matrix. The quantity $J(\boldsymbol{\theta}_t)$ provides a scalar measure of the distance between the observed behavior of the actual experimental subjects and the behavior of the Second Movers in the simulated samples at the trial parameter vector $\boldsymbol{\theta}_t$. The MSM estimator of $\boldsymbol{\theta}$, denoted $\hat{\boldsymbol{\theta}}$, is the value of $\boldsymbol{\theta}_t$ that minimizes $J(\boldsymbol{\theta}_t)$: $\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}_t} J(\boldsymbol{\theta}_t)$. Thus, MSM estimates the structural parameters to be such that the behavior of Second Movers simulated on the basis of the structural model is as similar as possible to the behavior of the actual Second Movers as observed in sample.

Under the conditions of Pakes and Pollard (1989), the MSM estimator is consistent and asymptotically normal for any consistent weight matrix $\mathbf{W}_N$. We use a weight matrix with diagonal elements equal to the inverse of N times the variances of the sample moments and zeros elsewhere and use bootstrap sampling of Second Movers with replacement to estimate $\mathbf{W}_N$.[22] Further details pertaining to the properties of the MSM estimator and estimation routine are presented in Appendix BB.

We use 38 moments to estimate the 17 structural parameters. The moments are described in Table 4 in Appendix BB. Correlations between Second Mover effort and First Mover effort, and between Second Mover effort and First Mover effort interacted with the prize, provide identifying information about $\tilde{\lambda}_2$, the parameter describing the strength of disappointment aversion on average. Percentiles of Second Mover specific correlations provide information about the standard deviation of disappointment aversion in the population, σ_λ. The correlation between Second Mover effort and the prize helps to identify κ, which measures the component of the convexity of the cost of effort function common to Second Movers and rounds, while the associated percentiles help to identify the shape of the distributions of the unobserved cost differences between Second Movers. Moments pertaining to the marginal distribution of Second Mover effort, such as round specific means and the standard deviation, provide further identifying information.

3. *Results*.—The upper left panel of Table 3 reports the parameter estimates for the preferred specification. Before discussing the results, we briefly consider the goodness of fit of the preferred specification, presented in Table 5, located in Appendix BB. Table 5 shows that all fitted moments correspond closely to the values observed in the sample: in particular, the z-test statistics show that the observed and fitted

[22] Using instead the optimally weighed minimum distance estimator improves efficiency but can introduce considerable finite sample bias (see Altonji and Segal 1996).

TABLE 3—MSM PARAMETER ESTIMATES

	Preferred specification		Nonquadratic cost of effort		Normally distributed cost unobservables	
	Estimate	SE	Estimate	SE	Estimate	SE
$\tilde{\lambda}_2$	1.729***	0.532	1.758***	0.640	1.260***	0.470
σ_λ	1.823***	0.556	1.868***	0.634	1.393***	0.481
b	−0.538***	0.036	−0.407***	0.018	−0.493***	0.012
κ	1.946***	0.103	2.063***	0.135	2.427***	0.059
σ_μ	0.516***	0.062	0.902***	0.151	0.266***	0.024
σ_π	0.346***	0.127	0.716***	0.204	0.204***	0.030
α	—	—	—	—	—	—
ψ	—	—	2.534***	0.128	—	—
$de_2/de_1(v{=}£0.10, \text{low } \lambda_{2,n})$	−0.000	0.001	−0.000	0.001	−0.000	0.002
$de_2/de_1(v{=}£2, \text{average } \lambda_{2,n})$	−0.030***	0.011	−0.028**	0.013	−0.025*	0.013
$de_2/de_1(v{=}£3.90, \text{high } \lambda_{2,n})$	−0.127***	0.026	−0.107***	0.034	−0.100***	0.019
OI test	25.555 [0.224]		13.435 [0.858]		61.480 [0.000]	
	Own-choice-acclimating reference point ($g_2 = 0$)		Own-choice-acclimating reference point ($g_2 = 1$)		Full sample: 60 Second Movers	
	Estimate	SE	Estimate	SE	Estimate	SE
$\tilde{\lambda}_2$	2.070***	0.426	1.909***	0.664	1.200***	0.426
σ_λ	1.476**	0.643	1.201**	0.534	1.206*	0.654
b	−0.615***	0.017	-0.591***	0.015	−0.486***	0.024
κ	2.187***	0.103	2.102***	0.060	1.769***	0.071
σ_μ	0.526***	0.050	0.578***	0.077	0.600***	0.110
σ_π	0.410***	0.086	0.345***	0.062	0.317***	0.122
α	0.944***	0.236	0.986***	0.156	—	—
ψ	—	—	—	—	—	—
$de_2/de_1(v{=}£0.10, \text{low } \lambda_{2,n})$	−0.001	0.001	−0.001	0.001	−0.000	0.001
$de_2/de_1(v{=}£2, \text{average } \lambda_{2,n})$	−0.034***	0.012	−0.032***	0.012	−0.024**	0.011
$de_2/de_1(v{=}£3.90, \text{high } \lambda_{2,n})$	−0.106***	0.027	−0.099***	0.026	−0.096***	0.028
OI test	11.583 [0.930]		20.980 [0.398]		24.005 [0.293]	

Notes: Where applicable, standard deviations of the transitory and persistent unobservables in the cost of effort function, σ_π and σ_μ, are computed from the estimates of the parameters of the Weibull distribution. Estimates of κ, σ_π and σ_μ have been multiplied by 100. All specifications further include dummy variables for each of rounds 2–10 inclusive. In the preferred specification, the coefficients on these variables, scaled as per κ, are between -0.1 and -0.5, significantly less than zero, and tend to decrease over the rounds. Reaction functions and their gradients were obtained using simulation methods. Using the estimated parameters of the cost of effort function for round 5, we simulated a large number of hypothetical Second Mover optimal efforts conditional on specific values of First Mover effort and the prize, and computed the mean best response. The reaction functions are linear, except in the case of nonquadratic effort costs where we evaluate the gradients at $e_1 = 20$. Low, average, and high $\lambda_{2,n}$ refer to the 20th, 50th, and 80th percentiles of the distribution of $\lambda_{2,n}$. The construction of the test statistic for the validity of overidentifying restrictions (OI test) is detailed in Newey (1985). p-values are shown in brackets. Unless stated otherwise, all results were obtained using our preferred sample of 59 Second Movers.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

moments never differ by more than 1.2 bootstrapped standard deviations. Consistent with this, the Newey test for the validity of overidentifying restrictions (OI test), reported in Table 3, does not reject the validity of the preferred specification.

Turning to the parameter estimates for the preferred specification, our estimate of the strength of disappointment aversion on average, $\tilde{\lambda}_2$, is 1.729 and this is significantly different from zero at all conventional significance levels.[23] In Section IIIC.6 we place

[23] We follow most of the literature in using the term "disappointment aversion" to describe this kink in utility around a choice-acclimating expectations-based reference point; we note, however, that our empirical results provide no direct evidence about the psychological processes that might underlie the kink.

this estimate in the context of the related literature but we note here that a figure of 1.729 is in line with previous studies that estimate the strength of loss aversion around a fixed reference point. We find that σ_λ is significantly greater than zero, thus providing evidence for heterogeneity in disappointment aversion across individuals. Our parameter estimates imply that $\lambda_{2,n}$ is greater than 3.3 for 20 percent of individuals, and is less than 0.2 for 20 percent. For 17 percent of individuals, $\lambda_{2,n}$ is less than zero.

The results further show that the cost of effort function exhibits significant convexity. In addition, there is significant transitory and permanent variation over Second Movers in the cost of effort, with persistent unobserved differences being more important than transitory differences.[24] Our estimate of b, the linear component of the cost of effort function, is negative, indicating that the cost of effort is declining at low effort levels. This negative coefficient is required to fit accurately observed average Second Mover effort. The linear component of the cost of effort function, however, does not affect how Second Movers respond to the First Movers' efforts. Moreover, it is not surprising that the cost of effort is at first declining, as the experimental subjects have self-selected into participating in the experiment and the outside option during the task is to do nothing for 120 seconds. Other experiments have also found that subjects derive some utility from carrying out real effort tasks, e.g., Brüggen and Strobel (2007).

Figure 3 shows the extent to which heterogeneity in disappointment aversion translates into differences in mean Second Mover responses to First Mover effort, evaluated at the average prize of £2 and at the highest prize of £3.90. Second Movers with low values of $\lambda_{2,n}$, defined to be the 20th percentile of the distribution of $\lambda_{2,n}$, do not respond appreciably to changes in First Mover effort. In contrast, we observe a significant discouragement effect (at the 1 percent level) for Second Movers with average values of $\lambda_{2,n}$, or with high values, defined to be the 80th percentile of the distribution of $\lambda_{2,n}$. At the highest prize of £3.90, a 40 slider increase in First Mover effort decreases optimal Second Mover effort by 2.5 sliders for an individual with the average $\lambda_{2,n}$, and by 5.1 sliders for an individual with a high $\lambda_{2,n}$. In the context of an average Second Mover effort of 25, these effects represent reductions of 10 percent and 20 percent, respectively, in optimal Second Mover effort.[25]

It is important to emphasize that the impact of disappointment aversion on Second Movers' average monetary payoff is small. Given the estimated Second Mover cost of effort function and the observed distribution of First Mover efforts and prizes, the Second Movers would each have earned just under £0.01 more on average over the course of the experiment had they not been disappointment averse. Nonetheless, we do find that disappointment aversion induces a significant discouragement effect and the impact on behavior is big enough to allow us to estimate the strength of disappointment aversion on average and the heterogeneity in disappointment aversion

[24] The magnitude and significance of the parameters controlling the persistent unobservables are indicative of the importance of our assumptions about persistent unobserved heterogeneity; to confirm their importance, we also estimated a simpler model (parameter estimates not reported) in which all unobserved heterogeneity comes through a mean zero normally distributed error term, assumed to be independent over both rounds and subjects, and found that the Newey test rejects this specification (p-value of 0.000).

[25] The magnitudes of the estimated slopes are somewhat lower than the corresponding estimates implied by the reduced form analysis in Section IIIB. This is because MSM seeks to fit simultaneously a variety of different moments. If we arbitrarily put a higher weight on the moments identifying these slopes, the estimated magnitudes would be larger.

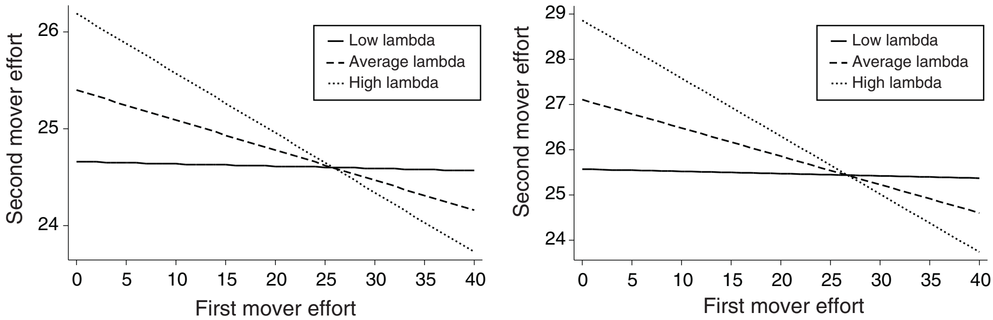

FIGURE 3. REACTION FUNCTIONS IMPLIED BY THE PREFERRED SPECIFICATION OF THE STRUCTURAL MODEL

Notes: We illustrate the reaction functions over the range 0 to 40 sliders as Table 1 shows First Mover efforts varied over this range. See notes to Table 3 for an explanation of how these reaction functions are constructed. Error bars are omitted; standard errors for the average $\lambda_{2,n}$ case in panel A and for the high $\lambda_{2,n}$ case in panel B are reported in Table 3. Low, average and high $\lambda_{2,n}$ refer to the 20th, 50th, and 80th percentiles of the distribution of $\lambda_{2,n}$.

across the population. Furthermore, given the estimated parameters of our model and the observed First Mover efforts and prizes, Second Movers would be willing to pay about £0.79 each on average per round to rid themselves of disappointment aversion. This has important implications for the design of labor market incentive schemes, as employers will be keen to design schemes that lower workers' expected disappointment in order to soften their participation constraint. For example, Gill and Stone (2010, Section 5.2) consider how worker loss aversion around expectations-based reference points impacts on an employer's choice of relative performance incentive contract.

4. *Robustness: Own-Choice-Acclimating Reference Point*.—The expectations-based reference point in our model adjusts to both the First Mover's and the Second Mover's effort choices. Thus, our finding of significant disappointment aversion provides evidence of loss aversion around choice-acclimating reference points when agents compete.

With a fixed reference point, including one given by a prior expectation, Proposition 1 shows that we should observe no discouragement effect. If the expectations-based reference point adjusted only to the First Mover's effort, however, there would still exist a discouragement effect.[26] In order to test whether the expectations-based reference point adjusts only to the First Mover's effort, we generalize the reference point (6) as follows for $\alpha \in [0,1]$:

$$R_2 = \alpha v P_2(e_2, e_1) + (1 - \alpha) v P_2(\tilde{e}_2, e_1), \tag{15}$$

where $\tilde{e}_2$ is fixed, and so does not adjust to the Second Mover's choice of effort ($\tilde{e}_2$ could for instance arise from a prior expectation).

[26] Similarly, in Abeler et al. (2011) the main empirical findings are consistent with a reference point that adjusts to the fixed payment but not the subjects' effort choices.

We reestimate our model, simultaneously estimating α as well as the 17 other parameters from the preferred specification.[27] The bottom lower and bottom middle panels of Table 3 show that we estimate α to be close to 1, so the Second Movers place little weight on the part of the reference point that does not adjust to their own effort choice. Moreover, the estimates of α are significantly different from zero at the 1 percent level. We argue that this provides strong evidence that the Second Movers' reference points are indeed own-choice-acclimatizing, as assumed in the preferred specification.

5. *Further Robustness*.—The remaining panels in Table 3 provide further robustness checks for various features of our analysis. The upper middle panel of Table 3 reports the results for a specification in which the cost of effort function is not constrained to be quadratic, but instead takes the form $C(e_2) = be_2 + (ce_2^{\psi})/\psi$, where ψ is an additional parameter to be estimated. We estimate ψ to be approximately 2.5.[28] In the upper right panel, we report estimates of a specification in which the unobservables appearing in the cost of effort function are normally distributed, rather than being drawn from Weibull distributions. The Newey test for the validity of overidentifying restrictions (OI test) rejects this specification, which illustrates the flexibility of the Weibull distribution. Finally, the bottom right panel shows results obtained for the preferred specification but estimated with the full sample of 60 Second Movers (see Footnote 17 for details concerning the omitted Second Mover).

We see that irrespective of the choice of sample and the specification of the cost of effort function, our estimate of the strength of disappointment aversion on average $\tilde{\lambda}_2$ is significantly different from zero. Also, all specifications show significant variation across individuals in the strength of disappointment aversion. Finally, the estimated response of a Second Mover to a change in First Mover effort varies little across specifications.

6. *Relationship to Existing Estimates of Loss Aversion*.—The endogeneity of the reference point means that behavior in our model is driven by the size of the kink in gain-loss utility $\lambda_2 = l_2 - g_2$. Other models of choice-acclimating reference points share the same feature. To see this, we introduce the parameterization of Kőszegi and Rabin (2006, 2007), which involves a weighting on gain-loss utility relative to material utility, $\eta \geq 0$, and a coefficient of loss aversion for gain-loss utility, λ, which measures the ratio of the slopes of gain-loss utility alone in the loss and gain domains. We estimate the size of the kink in gain-loss utility, scaled relative to material utility, and $\lambda_2 = \eta\lambda - \eta = \eta(\lambda - 1)$. In their model of single-agent effort provision, the first-order conditions of Abeler et al. (2011) also depend on $\eta(\lambda - 1)$, as

[27] With this more general reference point, the fixed $\tilde{e}_2$ and the slope of gain-loss utility in the gain domain, g_2, become relevant to the determination of the level of effort (but not to how the Second Movers respond to First Mover effort). As g_2 and $\tilde{e}_2$ are not identified under the null that $\alpha = 1$, we do not attempt to estimate these parameters, instead estimating the model for various values of g_2 and $\tilde{e}_2$. Results for $g_2 = 0$ and $g_2 = 1$ with $\tilde{e}_2$ equal to the average level of effort of 25 are reported in Table 3. We further estimated the model with $g_2 = ½$, and also with $\tilde{e}_2 = 0$ together with different values of g_2. The results were not substantially different.

[28] At the estimated parameters of this specification, the qualitative predictions of Proposition 3 continue to hold numerically. For every possible First Mover effort, the simulated Second Mover reaction function evaluated at the average $\lambda_{2,n}$ becomes steeper as the prize moves from £0.10 to £2 and from £2 to £3.90. Similarly, at the average prize of £2, the reaction function becomes steeper as we move from low $\lambda_{2,n}$ to the average $\lambda_{2,n}$ and from the average $\lambda_{2,n}$ to high $\lambda_{2,n}$. The notes to Table 3 detail the construction of these reaction functions.

do preferences over lotteries in the choice-acclimating version of the model of Kőszegi and Rabin (2007, p. 1,059 and Proposition 12(i)). The original disappointment aversion model of Bell (1985) also builds on the size of the kink. We cannot estimate λ directly, as this coefficient interacts with the weight put on gain-loss utility to determine the size of the kink in gain-loss utility; nonetheless, because we estimate that $\tilde{\lambda}_2 > 0$, it follows that $\eta > 0$ and $\lambda > 1$.

Our measure of disappointment aversion is therefore not directly comparable to previous measures of loss aversion around fixed reference points. Evidence from previous studies suggests a coefficient of loss aversion of about 2 for the value function of Kahneman and Tversky (1979) (Kahneman 2003); i.e., the value function is about twice as steep in the loss domain as it is in the gain domain. For example, from choices over lotteries, Tversky and Kahneman (1992) estimate a coefficient of 2.25 for their median subject. The value function of Kahneman and Tversky (1979) is defined only over gains and losses: if we consider this value function to include implicitly any consumption value of losses and gains as well as psychological elation and pain from deviating from the reference point, then the comparable figure in our setting to the usual loss aversion coefficient is the ratio of the slopes of total utility in the loss and gain domains, given by

$$\frac{1 + l_2}{1 + g_2} = 1 + \frac{\lambda_2}{1 + g_2}. \tag{16}$$

Given an assumption about g_2, our estimate of $\tilde{\lambda}_2$ therefore implies an estimate of the average value of (16) in the population. For example, if we assume that $g_2 \in (0, 1)$, so the elation associated with receiving more than expected is positive but less important than the associated material utility, then our estimate $\tilde{\lambda}_2 = 1.729$ implies that $(16) \in (1.865, 2.729)$, which matches previous estimates of the coefficient of loss aversion.

IV. Alternative Behavioral Explanations

Our model of disappointment aversion fits our data well: as explained in Section IIIC.3, all fitted moments correspond closely to the values observed in our experimental sample and the Newey test shows that our preferred specification is not rejected by the data. This is good statistical evidence that there are not important additional factors that would help to explain our data better. Furthermore, the burgeoning empirical literature showing the importance of expectations-based reference points in many different contexts (summarized in the introduction) lends weight to our thesis that expectations might be salient when agents compete. Nonetheless, we argue below that a number of alternative behavioral explanations of the observed discouragement effect are unconvincing.

Confusion: To test whether Second Movers understood the compensation scheme, we conducted an incentivized comprehension quiz in early December 2010 using a new sample of 60 students selected randomly from the pool of subjects at the University of Arizona's Economic Science Laboratory. The instructions and questions are reported in Appendix D (in the online Appendix). The quiz was designed to test whether Second

Movers understood that the marginal effect of an extra unit of effort on their probability of winning did not depend on the effort of the First Mover they were paired with. In particular, question 5 asked the subjects by how much a specific one unit increase in a Second Mover's points score increases her probability of winning given a First Mover points score of 20, and question 10 asked the same question given a First Mover score of 30. The data provide good evidence that most subjects understood the experimental instructions. Ninety percent of the subjects answered all 10 of the questions correctly. Of the 6 subjects who made at least one mistake, half nonetheless answered questions 5 and 10 correctly, and a further 2 subjects' answers to questions 5 and 10 were incorrect but did not vary between the two questions. A single subject gave a lower answer for question 10 than for question 5, but overall the average difference between the subjects' answers to questions 5 and 10 was not statistically significantly different from zero (p-value of 0.317). Footnote 20 also provides evidence that behavior did not change during the course of our experiment, so learning about the payoff function does not appear to have played an important role.

Peer Effects: Second Movers who imitate the behavior of their peers (Falk and Ichino 2006), or who compete by matching or beating their rival's effort level, would respond positively rather than negatively to the effort of the First Mover they are paired with. Moreover, we find no evidence of matching or beating. Specifically, if Second Movers have a tendency to match their rival's effort, we should see few Second Movers completing one slider fewer or one slider more than their rival. Similarly, if Second Movers tend to want to beat their rival's effort, we should see few Second Movers completing the same number of sliders or two sliders more than their rival. Formal tests of these hypotheses reveal no significant evidence of either matching or beating behavior.[29] Furthermore, by looking at the behavior of First Movers, we provide evidence that our estimates of disappointment aversion are not biased by a more general desire to mimic peer behavior. In particular, we look to see whether First Movers' efforts in a given round respond to the effort of their Second Mover rivals in the previous round. A general desire to mimic one's peers should lead to a positive relationship, untainted by any role for disappointment, as the uncertainty arising from the previous round's competition has been resolved. We find no evidence that First Movers respond to the effort of their peers in the previous round.[30]

Probability Weighting: Kahneman and Tversky (1979) propose that agents apply decision weights to probabilities, overweighting small probabilities and

[29] In more detail, let $\text{Prop}(x) \equiv \left(\sum_{n=1}^{59} \sum_{r=1}^{10} \mathbf{1}_{e_{2,n,r} - e_{1,n,r} = x}\right)/590$ be the proportion of Second Mover efforts for which the difference between the Second Mover's effort and that of the First Mover she is paired with is exactly x. To test for matching behavior we conduct two joint nonparametric bootstrapped tests. Our first test cannot reject the null that $\text{Prop}(-1) = \text{Prop}(-2)$ and $\text{Prop}(-1) = \text{Prop}(0)$, and the second cannot reject the null that $\text{Prop}(1) = \text{Prop}(0)$ and $\text{Prop}(1) = \text{Prop}(2)$, where in each case the one-sided alternative is that the middle proportion is significantly lower than either of the proportions above or below (or lower than both). The p-values are, respectively, 0.481 and 0.937. Similarly, when testing for beating behavior, our first test cannot reject the null that $\text{Prop}(0) = \text{Prop}(-1)$ and $\text{Prop}(0) = \text{Prop}(1)$, and the second cannot reject the null that $\text{Prop}(2) = \text{Prop}(1)$ and $\text{Prop}(2) = \text{Prop}(3)$. The p-values are, respectively, 0.832 and 0.138.

[30] Specifically, we estimate a reduced form model similar to (11), where First Mover effort depends on the prize in the current round, the prize in the previous round, the previous round effort of the Second Mover that the First Mover was paired with in the previous round and this Second Mover's effort interacted with the prize in the previous round. The coefficients on these last two are not statistically significantly different from zero (p-values of 0.433 and 0.680).

underweighting large ones. The evidence further suggests that the probability weighting function $w(P)$ is concave for probabilities smaller than about 35 to 40 percent and convex thereafter (Wu and Gonzalez 1996; Prelec 1998). A Second Mover who evaluated the prospect arising from her choice of effort using probability weighting would show a discouragement effect in the convex region to the right of the inflection point, but an opposite *encouragement effect* in the concave region to the left of the inflection point.[31] To test for such a pattern, we estimate the reduced-form model from Section IIIB with separate coefficients on First Mover effort and First Mover effort interacted with the prize for the highest 50 percent of First Mover efforts (which, on average, give rise to low Second Mover winning probabilities) and for the lowest 50 percent of First Mover efforts (which tend to give rise to higher Second Mover winning probabilities). The parameter estimates are of similar magnitude and have the same signs as those estimated in Section IIIB; furthermore, the parameters corresponding to the lower half of First Mover efforts are not statistically significantly different from those corresponding to the upper half (p-value of 0.830).[32] This provides evidence that, in contrast to the pattern predicted by probability weighting, the discouragement effect that we observe in the data operates throughout the range of Second Mover winning probabilities.

Regret: It is not straightforward to introduce a notion of regret (Bell 1982; Loomes and Sugden 1982) into our setup as the agents have more than two possible choices and uncertainty about the outcome of alternative choices is only partially resolved ex post (Bell 1983 discusses partial resolution of uncertainty in a two action model). Furthermore, we believe that, because our subjects have only limited potential to learn what would have happened in the counterfactual in which they chose a different level of effort, regret will not be particularly salient in the context of our experiment; as Larrick (1993, p. 446) puts it, "expecting vivid, concrete feedback about what definitely would have occurred produces a greater potential for regret than pallid, abstract knowledge of what statistically was likely to occur." Nonetheless, we estimate a structural model that includes both regret and disappointment to see whether omitting regret considerations from our preferred specification impacts on our estimate of disappointment aversion. We assume that regret is felt when the realized payoff $u_2(y_2) - C_2(e_2)$ is less than the ex post expected payoff from the ex post best alternative action.[33] We find that the parameter measuring the intensity of regret

[31] The prospect arising from an effort choice e_2 is "simple" (one nonzero outcome) and, with $u_2(0) = 0$, is valued at $w(P_2)u_2(v)$ (see Prelec 1998, p. 500). Given $\partial P_2 / \partial e_2 = -\partial P_2 / \partial e_1 = 1/(2\gamma) > 0$, $(\partial w / \partial P_2)(\partial P_2 / \partial e_2)u_2(v)$ is increasing (decreasing) in e_1 when $w(P_2)$ is concave (convex) in P_2, so the Second Mover's marginal incentive to exert effort rises (falls) locally in e_1 for a given e_2.

[32] As noted above, the literature suggests that the probability weighting function's inflection point lies somewhat to the left of $P = 0.5$. Thus we also estimated the reduced form model with separate coefficients for the highest 20 percent of First Mover efforts (which give rise to particularly low Second Mover winning probabilities) and for the lowest 80 percent. The results were similar.

[33] In a theoretical analysis of regret with partial resolution of uncertainty, Krähmer and Stone (2008) make the same assumption. Formally, we estimate the same structural model as in Section IIIC1, adding in an expected regret term as follows

$$P_2(e_2, e_1)\rho \min\Big\{v - C_2(e_2) - \max_{e_2^a \in \mathcal{A}\setminus e_2}\big(vP_2^W(e_2^a, e_2, e_1) - C_2(e_2^a)\big), 0\Big\}$$

$$+ \; (1 - P_2(e_2, e_1))\rho \min\Big\{-C_2(e_2) - \max_{e_2^a \in \mathcal{A}\setminus e_2}\big(vP_2^L(e_2^a, e_2, e_1) - C_2(e_2^a)\big), 0\Big\},$$

is not statistically significantly different from zero (p-value of 0.852). Furthermore, the estimate of the strength of disappointment aversion $\tilde{\lambda}_2 = 1.710$, which is similar to that from the preferred specification, and we continue to find significant heterogeneity in disappointment aversion.

Pressure: Choking under pressure (Baumeister 1984), where an agent's performance deteriorates when incentives or stakes are higher, is also an implausible explanation of the discouragement effect. First, the Second Movers' marginal incentives do not depend on First Mover effort, so an incentives-based story for choking has no bite. Furthermore, if the level of the stakes matter, Second Movers should choke more when their probability of winning is higher, i.e., when the First Mover has worked *less hard*.

Collusion and Reciprocity: Even though effort is socially wasteful, the subjects are not able to collude in our experiment. As explained in Section IB, the rotation-based matching algorithm that we use ensures that each pair of subjects plays a truly one-shot game: the pair meets only once and a subject's action in one round cannot have a direct or indirect effect on the actions of other subjects that the subject is paired with later on. A taste for reciprocity (Rabin 1993) can sometimes allow agents to cooperate even in one-shot prisoners' dilemma-type games: in our setup, low First Mover effort can be considered a kind action to be reciprocated with low effort, so positive reciprocity could conceivably allow the agents to coordinate on low effort. In Section IIIC.3, however, we find substantial variation in subjects' effort costs. This variation clouds inferences that subjects can make about a rival's intentions: subjects only meet the rival once and so will have difficulty distinguishing the rival's kind or unkind intention from a low or high cost of effort. Furthermore, our data show that Second Movers respond to a kinder action (lower First Mover effort) with a meaner action (higher effort) instead of the kinder action (lower effort) predicted by reciprocity. To the extent that the First Movers are able to signal kindness by putting in low effort, the response by the Second Movers to increase effort will destroy any incentive on the part of the First Movers to be kind.

In order to provide some more direct evidence that reciprocity does not play an important role in our experiment, we look at how Second Movers respond to unexpectedly high or low effort on the part of the First Mover they are paired with. Given that we find substantial variation in our subjects' effort costs, it is reasonable to assume that Second Movers learn about the distribution of First Movers' effort costs from the First Mover efforts they see during the course of the experiment. If reciprocity is playing a role, then the greater the average effort seen in earlier rounds by a particular Second Mover (after allowing for the effects of prizes in the earlier rounds), the lower the inferred average cost of effort, and hence the kinder or less unkind any particular First Mover effort is perceived to be in the current round. Thus, if Second Movers reciprocate perceived kindness or unkindness, they will want to

where ρ measures the intensity of regret, e_2^a represents an alternative Second Mover effort to the chosen effort e_2, $P_2^W(e_2^a, e_2, e_1) = \min\{(e_2^a - e_1 + \gamma)/(e_2 - e_1 + \gamma), 1\}$ is the updated ex post probability of winning from having chosen e_2^a for a Second Mover who in fact wins with e_2 and $P_2^L(e_2^a, e_2, e_1) = \max\{(e_2^a - e_2)/(-e_2 + e_1 + \gamma), 0\}$ is the ex post probability of winning function after losing (some updating occurs because winning or losing is somewhat informative about whether the random draws that determine the winner were favorable or unfavorable to the Second Mover).

reward any given First Mover effort more or punish it less when they have observed higher First Mover efforts in earlier rounds, and so we should see a negative relationship between the average effort a Second Mover has seen in earlier rounds and her effort in the current round. We find no evidence of such a relationship.[34]

Equity: As we explain in Section IIB, if the Second Movers disliked inequity in monetary payoffs, their marginal incentives to exert effort would continue to be independent of First Mover effort, so such inequity aversion would not lead to a discouragement effect, as Proposition 1 would continue to apply. A notion of equity in which agents like to align efforts as inputs and monetary payoffs as outputs might provide part of the psychological basis for why our subjects are disappointment averse (see Gill and Stone 2010 for details) and so exhibit a discouragement effect. (As noted in footnote 23, our empirical results provide no direct evidence about the psychological processes that might underlie disappointment aversion in our experiment.)

V. Conclusion

People compete all the time, e.g., for: promotions; bonuses; professional partnerships; elected positions; social status; and sporting trophies. In these situations, the competitors exert effort to improve their prospects of success, and clear winners and losers emerge. Our results indicate that winners are elated while losers are disappointed, and that disappointment is the stronger emotion. In particular, we show that when our experimental subjects compete in a sequential-move real-effort competition, they are loss averse around an endogenous expectations-based reference point that is conditioned on their own work effort and that of their rival. Disappointment aversion creates a discouragement effect, whereby a competitor slacks off when her rival works hard. Our results speak to the debate about the speed at which reference points adjust. Kőszegi and Rabin (2007) note that it is unclear how much time is needed between agents making their choices and the outcome occurring for the reference point to become choice-acclimating. Given the tiny temporal gap between the agents' effort choices and the outcome of the tournament, our results indicate that, at least in our competitive framework, the adjustment process is essentially instantaneous.

We hope that our theoretical model and empirical findings will provide a useful building block when predicting how people will behave in competitive situations. Furthermore, the findings may be helpful to principals when designing competitive environments. For example, employers will want to know how much they need to compensate employees for the expected disappointment implicit in different types of compensation schemes. They will also be interested in the degree to which a given compensation structure might impact on employees' work efforts, for example by creating asymmetries with some employees exerting a lot of effort and others becoming discouraged.

[34] Specifically, for every round $r > 1$, and for every Second Mover n, we calculate the average First Mover effort that the Second Mover has seen up to (but not including) that round, where previous efforts have been partialed to remove prize and round effects. As required by our strategy to identify reciprocity, in our sample we find substantial variation in these averages. We then estimate the reduced form model (11) with this average as an extra term. The estimated coefficient on this new term is not statistically significantly different from zero (p-value of 0.782).

Appendix A: Proofs

A. *Proof of Proposition 2*

Using (1) and (7),

$$EU_2(e_2, e_1) = v\left(\frac{e_2 - e_1 + \gamma}{2\gamma}\right) - \lambda_2 v\left(\frac{\gamma^2 - (e_2 - e_1)^2}{4\gamma^2}\right) - C_2(e_2). \tag{17}$$

We use a proof by contradiction. Suppose that when e_1 increases from e_{11} to $e_{12} > e_{11}$, the Second Mover's optimal effort e_2^* increases from e_{21}^* to $e_{22}^* > e_{21}^*$. By the optimality of the Second Mover's effort choices,

$$[EU_2(e_{21}^*, e_{11}) - EU_2(e_{22}^*, e_{11})] + [EU_2(e_{22}^*, e_{12}) - EU_2(e_{21}^*, e_{12})] \geq 0. \tag{18}$$

Using (17), we get the following:

$$EU_2(e_{21}^*, e_{11}) - EU_2(e_{21}^*, e_{12}) = v\left(\frac{-e_{11} + e_{12}}{2\gamma}\right) + \lambda_2 v\left(\frac{(e_{21}^* - e_{11})^2 - (e_{21}^* - e_{12})^2}{4\gamma^2}\right); \tag{19}$$

$$EU_2(e_{22}^*, e_{12}) - EU_2(e_{22}^*, e_{11}) = v\left(\frac{-e_{12} + e_{11}}{2\gamma}\right) + \lambda_2 v\left(\frac{(e_{22}^* - e_{12})^2 - (e_{22}^* - e_{11})^2}{4\gamma^2}\right). \tag{20}$$

Thus,

$$(18) = \frac{\lambda_2 v}{2\gamma^2}(- e_{21}^* e_{11} + e_{21}^* e_{12} - e_{22}^* e_{12} + e_{22}^* e_{11}) = \frac{\lambda_2 v}{2\gamma^2}(e_{21}^* - e_{22}^*)(e_{12} - e_{11}) < 0 \tag{21}$$

given $\lambda_2 > 0$ for a disappointment averse Second Mover, which contradicts $(18) \geq 0$ from above.

Note that if there are multiple optima, the proof extends naturally to show that the highest optimal effort in response to e_{12} must lie weakly below the lowest in response to e_{11}.

B. *Proof of Proposition 3*

Using (9) and (17),

$$(22)\qquad \frac{\partial EU_2(e_2,e_1)}{\partial e_2} = \frac{v}{2\gamma} + \frac{\lambda_2 v(e_2 - e_1)}{2\gamma^2} - b - ce_2;$$

$$(23)\qquad \frac{\partial^2 EU_2(e_2,e_1)}{\partial e_2^2} = \frac{\lambda_2 v}{2\gamma^2} - c.$$

We assume that $2\gamma^2 c - \lambda_2 v > 0$, so the objective function is strictly concave.

Suppose first that the action space $\mathcal{A}$ is continuous. The first-order condition gives the following reaction function:

$$(24)\qquad e_2^*(e_1) = \begin{cases} \bar{e} & \text{if } e_1 < \dfrac{\gamma v + \lambda_2 v\bar{e} - 2\gamma^2(b + c\bar{e})}{\lambda_2 v} \\[2ex] \dfrac{\gamma v - \lambda_2 v e_1 - 2\gamma^2 b}{2\gamma^2 c - \lambda_2 v} \in [0,\bar{e}] & \text{if } e_1 \in \left[\dfrac{\gamma v + \lambda_2 v\bar{e} - 2\gamma^2(b + c\bar{e})}{\lambda_2 v}, \dfrac{\gamma v - 2\gamma^2 b}{\lambda_2 v}\right] \\[2ex] 0 & \text{if } e_1 > \dfrac{\gamma v - 2\gamma^2 b}{\lambda_2 v}. \end{cases}$$

Given $\lambda_2 > 0$ and $2\gamma^2 c - \lambda_2 v > 0$, in the interior de_2^*/de_1 is clearly strictly negative and strictly decreasing in λ_2 and v.

Suppose second that the action space $\mathcal{A}$ is discrete. Take any $e_2 \in \mathcal{A}$ for which there exists a higher effort that is a best response to some $e_1 \in [0,\bar{e}]$ and a lower effort with the same property. Let e_2^+ be the next highest effort in $\mathcal{A}$ and let e_2^- be the next lowest effort in $\mathcal{A}$. Using (9) and (17), $EU_2(e_2^+,e_1) - EU_2(e_2,e_1)$

$$(25)\qquad = \frac{v(e_2^+ - e_2)}{2\gamma} + \lambda_2 v\left(\frac{(e_2^+ - e_1)^2 - (e_2 - e_1)^2}{4\gamma^2}\right) - b(e_2^+ - e_2) - \frac{c((e_2^+)^2 - e_2^2)}{2}$$

$$(26)\qquad = \frac{(2\gamma v - 4\gamma^2 b)(e_2^+ - e_2)}{4\gamma^2} + \left(\frac{\lambda_2 v - 2\gamma^2 c}{4\gamma^2}\right)\left((e_2^+)^2 - e_2^2\right) - \left(\frac{\lambda_2 v}{4\gamma^2}\right)2e_1(e_2^+ - e_2).$$

The cutoff e_1 at which $EU_2(e_2^+,e_1) = EU_2(e_2,e_1)$ is given by

$$(27)\qquad \breve{e}_1(e_2^+,e_2) = \frac{2\gamma v - 4\gamma^2 b}{2\lambda_2 v} - \left(\frac{2\gamma^2 c - \lambda_2 v}{\lambda_2 v}\right)\left(\frac{e_2^+ + e_2}{2}\right).$$

Given $\lambda_2 > 0$ and $2\gamma^2 c - \lambda_2 v > 0$ by assumption, the cutoffs are strictly decreasing in the Second Mover's effort. From Proposition 2, best responses are (weakly) falling in e_1. Thus, if e_1 was continuous but e_2 was discrete, the cutoffs would represent the points at which the Second Mover's reaction function jumped down. As both are discrete, the cutoffs define the Second Mover's reaction function in the interior: e_2 is a best response for the Second Mover for and only for any $e_1 \in [\breve{e}_1(e_2^+, e_2), \breve{e}_1(e_2, e_2^-)] \cap \mathcal{A}$. The range $[\breve{e}_1(e_2^+, e_2), \breve{e}_1(e_2, e_2^-)]$ is of size

$$\breve{e}_1(e_2, e_2^-) - \breve{e}_1(e_2^+, e_2) = \left(\frac{2\gamma^2 c - \lambda_2 v}{\lambda_2 v}\right)\left(\frac{e_2^+ - e_2^-}{2}\right), \tag{28}$$

which is strictly decreasing in λ_2 and v.

That the cutoffs are strictly decreasing in e_2 is the discrete case analog of the reaction function being strictly downward-sloping in the continuous case. That the size of the ranges between the cutoffs is strictly decreasing in λ_2 and v is the discrete case analog of the reaction function becoming strictly steeper in λ_2 and v in the continuous case. Note also the functional form similarity: supposing that the permitted e_2's increase in unit steps, $(e_2^+ + e_2)/2 = e_2 + ½$, so the rate of change of $\breve{e}_1(e_2^+, e_2)$ with respect to e_2 is the inverse of the slope of the reaction function in the continuous case.

Appendix B: MSM: Further Details

A. *Construction of Simulated Samples*

The construction of each simulated sample is conditional on the First Mover efforts and prizes observed in the actual sample. Additionally, we make random draws that will later be used to construct the unobservables appearing in the structural model. Specifically, for each simulated sample $s = 1, \ldots, S$ we construct matrices of dimensions $N \times 1$, $N \times 1$ and $N \times 10$, denoted $\mathbf{Q1}_s$, $\mathbf{Q2}_s$, and $\mathbf{Q3}_s$, respectively. Each element of $\mathbf{Q1}_s$, $\mathbf{Q2}_s$, and $\mathbf{Q3}_s$ contains a random draw from a standard uniform distribution. These matrices are held fixed throughout the estimation.[35] Given a trial parameter vector $\boldsymbol{\theta}_t$, the effort choice of the nth Second Mover in the rth round of the sth sample is determined as follows:

1. The Second Mover is assigned values of the unobservables $\lambda_{2,n}$, μ_n, and $\pi_{n,r}$ in accordance with the distributional assumptions made in Section IIIC1. Draws from the normal distribution are found by transforming $\mathbf{Q1}_s$ as follows:

$$\lambda_{2,n} = \tilde{\lambda}_2 + \sigma_\lambda \Phi^{-1}(\mathbf{Q1}_{s,n}), \tag{29}$$

[35] Thus, as the trial parameter vector $\boldsymbol{\theta}_t$ is adjusted, the simulated samples vary only due to the change in $\boldsymbol{\theta}_t$ and not due to variation in the underlying random draws. This is necessary to ensure convergence of the estimation routine (Stern 1997).

where Φ^{-1} denotes the inverse of the standard normal distribution function. Draws from the Weibull distribution are obtained by transforming $\mathbf{Q2}_s$ and $\mathbf{Q3}_s$ as follows:

$$\mu_n = \phi_\mu(-\ln(Q2_{s,n}))^{1/\varphi_\mu}; \tag{30}$$

$$\pi_{n,r} = \phi_\pi(-\ln(Q3_{s,n,r}))^{1/\varphi_\pi}. \tag{31}$$

The values of the parameters $\tilde{\lambda}_2$, σ_λ, φ_μ, φ_π, ϕ_μ and ϕ_π are obtained by extracting the relevant elements of $\boldsymbol{\theta}_t$.

2. Given the assigned values of $\lambda_{2,n}$, μ_n, and $\pi_{n,r}$ and the remaining parameters of the cost of effort function, b, κ and δ_r for $r = 2, \ldots, 10$ as given by $\boldsymbol{\theta}_t$, the expected utility associated with each feasible Second Mover effort is computed using (7), (9), and (13).

3. The Second Mover is assigned the effort choice corresponding to the highest expected utility.

Steps 1–3 are repeated for each of the 10 rounds, the N Second Movers and the S simulated samples. Note that by comparing the expected utilities associated with each of the 49 feasible effort choices, we fully account for the discreteness of effort. Additionally, the method of simulation does not rely on the objective function being well behaved.

B. *Asymptotic Properties, Numerical Methods, Moments, and Goodness of Fit*

Under the conditions in Pakes and Pollard (1989), $\hat{\boldsymbol{\theta}}$ is consistent and asymptotically normal. Specifically, with S fixed,

$$\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} N\left(\mathbf{0}, \frac{S+1}{S}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\mathbf{D}'\mathbf{W}\boldsymbol{\Omega}\mathbf{W}\mathbf{D}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\right) \text{ as } N \to \infty, \tag{32}$$

where $\boldsymbol{\Omega} = N\text{cov}(\mathbf{M})$ is the covariance matrix of the sample moments normalized by the sample size, $\mathbf{W} = plim(\mathbf{W}_N)$ and

$$\mathbf{D} = \frac{1}{S}\sum_{s=1}^{S} \frac{d\mathbf{M}_s(\boldsymbol{\theta}_t)}{d\boldsymbol{\theta}_t'}\bigg|_{\boldsymbol{\theta}_t=\boldsymbol{\theta}}. \tag{33}$$

When implementing MSM, we use $S = 30$ simulated samples and therefore simulate 17,700 pairings when using $N = 59$, and we estimate the weight matrix $\mathbf{W}_N$ using 2,000 bootstrapped samples each containing N Second Movers sampled with replacement from the original sample.

The term $\left(\sum_{s=1}^{S} \mathbf{M}_s(\boldsymbol{\theta}_t)\right)/S$ appearing in $J(\boldsymbol{\theta}_t)$ in (14) is not a continuous function of the parameter vector $\boldsymbol{\theta}_t$ as small changes in $\boldsymbol{\theta}_t$ may cause discrete changes in

Table 4—Description of Moments

Moment	Description	Primarily identifying
$\text{SD}(e_{2,n,r})$	Standard deviation of Second Mover effort	ϕ_π, ϕ_μ, φ_π, φ_μ, σ_λ
$\text{Prop}(e_{2,n,r} < 15)$; $\text{Prop}(e_{2,n,r} > 35)$	Proportions of Second Mover efforts below 15 and above 35	ϕ_π, ϕ_μ, φ_π, φ_μ, σ_λ
$\text{SD}(e_{2,n,r} - e_{2,n,r-1})$	Standard deviation of the round on round change in Second Mover effort	ϕ_π, φ_π
$\text{Corr}\,(e_{2,n,r}, e_{2,n,r-1})$; $\text{Corr}\,(e_{2,n,r}, e_{2,n,r-2})$	1st and 2nd order autocorrelations in Second Mover effort	ϕ_μ, φ_μ, σ_λ
$\text{Mean}\,(e_{2,n,r})$ for $r = 1, \ldots, 10$	Round specific means of Second Mover effort	b, δ_r for $r = 2, \ldots, 10$
$\text{Corr}\,(e_{2,n,r}, v_{n,r} \mid e_{1,n,r}, e_{1,n,r}v_{n,r}, RD, FE)$	Correlation between Second Mover effort and the prize after partialing out linear additive effects of First Mover effort, First Mover effort interacted with the prize, round dummies, and Second Mover specific fixed effects	κ
$\text{Corr}\,(e_{2,n,r}, e_{1,n,r} \mid v_{n,r}, e_{1,n,r}v_{n,r}, RD, FE)$	Correlation between Second Mover effort and First Mover effort after partialing out linear additive effects of the prize, First Mover effort interacted with the prize, round dummies, and Second Mover specific fixed effects	$\tilde{\lambda}_2$
$\text{Corr}\,(e_{2,n,r}, v_{n,r}e_{1,n,r} \mid e_{1,n,r}, v_{n,r}, RD, FE)$	Correlation between Second Mover effort and First Mover effort interacted with the prize after partialing out linear additive effects of First Mover effort, the prize, round dummies, and Second Mover specific fixed effects	$\tilde{\lambda}_2$
$\text{Pc}_j\text{Corr}_n(e_{2,n,r}, v_{n,r} \mid e_{1,n,r}, e_{1,n,r}v_{n,r}, RT)$ for $j = 17, 33, 50, 66, 83$	jth percentile of the Second Mover specific correlation between Second Mover effort and the prize after partialing out linear additive effects of First Mover effort, First Mover effort interacted with the prize, and a linear round trend	ϕ_π, ϕ_μ, φ_π, φ_μ
$\text{Pc}_j\text{Corr}_n(e_{2,n,r}, e_{1,n,r} \mid v_{n,r}, e_{1,n,r}v_{n,r}, RT)$ for $j = 17, 33, 50, 66, 83$	jth percentile of the Second Mover specific correlation between Second Mover effort and First Mover effort after partialing out linear additive effects of the prize, First Mover effort interacted with the prize, and a linear round trend	σ_λ
$\text{Pc}_j\text{Corr}_n(e_{2,n,r}, v_{n,r}e_{1,n,r} \mid e_{1,n,r}, v_{n,r}, RT)$ for $j = 17, 33, 50, 66, 83$	jth percentile of the Second Mover specific correlation between Second Mover effort and First Mover effort interacted with the prize after partialing out linear additive effects of First Mover effort, the prize, and a linear round trend	σ_λ
$\text{Mean}(e_{2,n,r} \mid e_{1,n,r} < 23 \cap v_{n,r} < 1.33)$	Mean of Second Mover effort conditional on low First Mover effort and low prize	κ, $\tilde{\lambda}_2$
$\text{Mean}(e_{2,n,r} \mid e_{1,n,r} < 23 \cap v_{n,r} > 2.55)$	Mean of Second Mover effort conditional on low First Mover effort and high prize	κ, $\tilde{\lambda}_2$
$\text{Mean}(e_{2,n,r} \mid e_{1,n,r} > 28 \cap v_{n,r} < 1.33)$	Mean of Second Mover effort conditional on high First Mover effort and low prize	κ, $\tilde{\lambda}_2$
$\text{Mean}(e_{2,n,r} \mid e_{1,n,r} > 28 \cap v_{n,r} > 2.55)$	Mean of Second Mover effort conditional on high First Mover effort and high prize	κ, $\tilde{\lambda}_2$

Notes: *RD*, *FE*, and *RT* denote round dummies, Second Mover specific fixed effects and a linear round trend. Partialing out is accomplished by working with the residuals from regressions of the dependent variables on the control variables.

some Second Movers' optimal effort choices. Consequently, gradient and Hessian-based optimization methods are unsuitable for minimizing $J(\boldsymbol{\theta}_t)$. Instead, we use Simulated Annealing in the form suggested by Goffe, Ferrier, and Rogers (1994) to solve for the MSM estimates.

Table 5—Goodness of Fit of the Preferred Specification

	Observed moment	Bootstrapped SD	Fitted moment	z-test for difference
$\mathrm{SD}(e_{2,n,r})$	5.875	0.497	5.496	−0.760
$\mathrm{Corr}\ (e_{2,n,r}, e_{2,n,r-1})$	0.652	0.062	0.633	−0.307
$\mathrm{Corr}\ (e_{2,n,r}, e_{2,n,r-2})$	0.596	0.085	0.603	0.087
$\mathrm{SD}(e_{2,n,r} - e_{2,n,r-1})$	4.828	0.719	4.645	−0.255
$\mathrm{Mean}\ (e_{2,n,1})$	21.763	0.784	21.602	−0.205
$\mathrm{Mean}\ (e_{2,n,2})$	23.458	0.633	23.264	−0.305
$\mathrm{Mean}\ (e_{2,n,3})$	24.831	0.650	24.933	0.158
$\mathrm{Mean}\ (e_{2,n,4})$	25.203	0.585	25.360	0.268
$\mathrm{Mean}\ (e_{2,n,5})$	25.119	0.737	24.927	−0.260
$\mathrm{Mean}\ (e_{2,n,6})$	24.898	0.897	25.233	0.373
$\mathrm{Mean}\ (e_{2,n,7})$	25.763	0.798	25.968	0.258
$\mathrm{Mean}\ (e_{2,n,8})$	26.169	0.673	26.310	0.208
$\mathrm{Mean}\ (e_{2,n,9})$	26.254	0.860	26.401	0.171
$\mathrm{Mean}\ (e_{2,n,10})$	26.729	0.774	26.592	−0.177
$\mathrm{Corr}\ (e_{2,n,r}, v_{n,r} \mid e_{1,n,r}, e_{1,n,r} v_{n,r}, RD, FE)$	0.124	0.044	0.084	−0.905
$\mathrm{Corr}\ (e_{2,n,r}, e_{1,n,r} \mid v_{n,r}, e_{1,n,r} v_{n,r}, RD, FE)$	0.041	0.042	0.003	−0.916
$\mathrm{Corr}\ (e_{2,n,r}, v_{n,r} e_{1,n,r} \mid e_{1,n,r}, v_{n,r}, RD, FE)$	−0.095	0.047	−0.038	1.200
$\mathrm{Pc}_{17}\mathrm{Corr}_n(e_{2,n,r}, v_{n,r} \mid e_{1,n,r}, e_{1,n,r} v_{n,r}, RT)$	−0.275	0.098	−0.179	0.975
$\mathrm{Pc}_{33}\mathrm{Corr}_n(e_{2,n,r}, v_{n,r} \mid e_{1,n,r}, e_{1,n,r} v_{n,r}, RT)$	0.033	0.079	0.043	0.124
$\mathrm{Pc}_{50}\mathrm{Corr}_n(e_{2,n,r}, v_{n,r} \mid e_{1,n,r}, e_{1,n,r} v_{n,r}, RT)$	0.222	0.071	0.212	−0.145
$\mathrm{Pc}_{66}\mathrm{Corr}_n(e_{2,n,r}, v_{n,r} \mid e_{1,n,r}, e_{1,n,r} v_{n,r}, RT)$	0.388	0.041	0.360	−0.675
$\mathrm{Pc}_{83}\mathrm{Corr}_n(e_{2,n,r}, v_{n,r} \mid e_{1,n,r}, e_{1,n,r} v_{n,r}, RT)$	0.469	0.051	0.523	1.051
$\mathrm{Pc}_{17}\mathrm{Corr}_n(e_{2,n,r}, e_{1,n,r} \mid v_{n,r}, e_{1,n,r} v_{n,r}, RT)$	−0.328	0.052	−0.386	−1.118
$\mathrm{Pc}_{33}\mathrm{Corr}_n(e_{2,n,r}, e_{1,n,r} \mid v_{n,r}, e_{1,n,r} v_{n,r}, RT)$	−0.218	0.061	−0.204	0.224
$\mathrm{Pc}_{50}\mathrm{Corr}_n(e_{2,n,r}, e_{1,n,r} \mid v_{n,r}, e_{1,n,r} v_{n,r}, RT)$	0.019	0.089	−0.027	−0.514
$\mathrm{Pc}_{66}\mathrm{Corr}_n(e_{2,n,r}, e_{1,n,r} \mid v_{n,r}, e_{1,n,r} v_{n,r}, RT)$	0.179	0.064	0.141	−0.585
$\mathrm{Pc}_{83}\mathrm{Corr}_n(e_{2,n,r}, e_{1,n,r} \mid v_{n,r}, e_{1,n,r} v_{n,r}, RT)$	0.361	0.064	0.350	−0.169
$\mathrm{Pc}_{17}\mathrm{Corr}_n(e_{2,n,r}, v_{n,r} e_{1,n,r} \mid e_{1,n,r}, v_{n,r}, RT)$	−0.194	0.080	−0.208	−0.175
$\mathrm{Pc}_{33}\mathrm{Corr}_n(e_{2,n,r}, v_{n,r} e_{1,n,r} \mid e_{1,n,r}, v_{n,r}, RT)$	0.019	0.067	0.001	−0.268
$\mathrm{Pc}_{50}\mathrm{Corr}_n(e_{2,n,r}, v_{n,r} e_{1,n,r} \mid e_{1,n,r}, v_{n,r}, RT)$	0.146	0.057	0.169	0.395
$\mathrm{Pc}_{66}\mathrm{Corr}_n(e_{2,n,r}, v_{n,r} e_{1,n,r} \mid e_{1,n,r}, v_{n,r}, RT)$	0.298	0.068	0.305	0.101
$\mathrm{Pc}_{83}\mathrm{Corr}_n(e_{2,n,r}, v_{n,r} e_{1,n,r} \mid e_{1,n,r}, v_{n,r}, RT)$	0.474	0.050	0.483	0.170
$\mathrm{Mean}(e_{2,n,r} \mid e_{1,n,r} < 23 \cap v_{n,r} < 1.33)$	23.821	0.867	24.121	0.346
$\mathrm{Mean}(e_{2,n,r} \mid e_{1,n,r} < 23 \cap v_{n,r} > 2.55)$	25.485	0.814	25.595	0.135
$\mathrm{Mean}(e_{2,n,r} \mid e_{1,n,r} > 28 \cap v_{n,r} < 1.33)$	25.836	0.975	25.265	−0.586
$\mathrm{Mean}(e_{2,n,r} \mid e_{1,n,r} > 28 \cap v_{n,r} > 2.55)$	25.050	1.208	25.665	0.509
$\mathrm{Prop}(e_{2,n,r} < 15)$	0.029	0.010	0.033	0.401
$\mathrm{Prop}(e_{2,n,r} > 35)$	0.015	0.008	0.017	0.206

Notes: See Table 4 for a description of the moments. Observed moments are computed from the sample and fitted moments are computed using parameter estimates from the preferred specification.

REFERENCES

Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman. 2011. "Reference Points and Effort Provision." *American Economic Review,* 101(2): 470–92.

Altonji, Joseph G., and Lewis M. Segal. 1996. "Small-Sample Bias in GMM Estimation of Covariance Structures." *Journal of Business and Economic Statistics,* 14(3): 353–66.

Barberis, Nicholas, Ming Huang, and Richard H. Thaler. 2006. "Individual Preferences, Monetary Gambles, and Stock Market Participation: A Case of Narrow Framing." *American Economic Review,* 96(4): 1069–90.

Battigalli, Pierpaolo, and Martin Dufwenberg. 2009. "Dynamic Psychological Games." *Journal of Economic Theory*, 144(1): 1–35.

Baumeister, Roy F. 1984. "Choking under Pressure: Self-Consciousness and Paradoxical Effects of Incentives on Skillful Performance." *Journal of Personality and Social Psychology*, 46(3): 610–20.

Bell, David E. 1982. "Regret in Decision Making under Uncertainty." *Operations Research*, 30(5): 961–81.

Bell, David E. 1983. "Risk Premiums for Decision Regret." *Management Science*, 29(10): 1156–66.

Bell, David. E. 1985. "Disappointment in Decision Making under Uncertainty." *Operations Research*, 33(1): 1–27.

Berger, Jonah, and Devin G. Pope. 2009. "Can Losing Lead to Winning?" http://faculty.chicagobooth.edu/devin.pope/research/pdf/berger_pope_manuscript.pdf.

Brüggen, Alexander, and Martin Strobel. 2007. "Real Effort versus Chosen Effort in Experiments." *Economics Letters*, 96(2): 232–36.

Bull, Clive, Andrew Schotter, and Keith Weigelt. 1987. "Tournaments and Piece Rates: An Experimental Study." *Journal of Political Economy*, 95(1): 1–33.

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard H. Thaler. 1997. "Labor Supply of New York City Cabdrivers: One Day at a Time." *Quarterly Journal of Economics*, 112(2): 407–41.

Card, David, and Gordon B. Dahl. 2011. "Family Violence and Football: The Effect of Unexpected Emotional Cues on Violent Behavior." *Quarterly Journal of Economics*, 126(1): 103–43.

Charness, Gary, and Peter Kuhn. 2010. "Lab Labor: What Can Labor Economists Learn from the Lab?" In *Handbook of Labor Economics,* Vol. 4, ed. Orley Ashenfelter and David Card, 229–330. Amsterdam: Elsevier, North-Holland.

Che, Yeon-Koo, and Ian Gale. 2000. "Difference-Form Contests and the Robustness of All-Pay Auctions." *Games and Economic Behavior*, 30(1): 22–43.

Choi, Syngjoo, Raymond Fisman, Douglas Gale, and Shachar Kariv. 2007. "Consistency and Heterogeneity of Individual Behavior under Uncertainty." *American Economic Review*, 97(5): 1921–38.

Cooper, Russell, Douglas V. DeJong, Robert Forsythe, and Thomas W. Ross. 1996. "Cooperation without Reputation: Experimental Evidence from Prisoner's Dilemma Games." *Games and Economic Behavior*, 12(2): 187–218.

Crawford, Vincent P., and Juanjuan Meng. 2011. "New York City Cab Drivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income." *American Economic Review*, 101(5): 1912–32.

DellaVigna, Stefano. 2009. "Psychology and Economics: Evidence from the Field." *Journal of Economic Literature*, 47(2): 315–72.

Delquié, Philippe, and Alessandra Cillo. 2006. "Expectations, Disappointment, and Rank-Dependent Probability Weighting." *Theory and Decision*, 60(2–3): 193–206.

Doran, Kirk B. 2009. "Reference Points, Expectations, and Heterogeneous Daily Labor Supply." http://www.nd.edu/~kdoran/Doran_Labor_Supply.pdf. (accessed December 3, 2011).

Ericson, Keith M. M., and Andreas Fuster. 2010. "Expectations as Endowments: Evidence on Reference-Dependent Preferences from Exchange and Valuation Experiments." http://www.people.fas.harvard.edu/~kericson/papers/EricsonFuster-ExpectationsAsEndowments.pdf.

Eriksson, Tor, Anders Poulsen, and Marie Claire Villeval. 2009. "Feedback and Incentives: Experimental Evidence." *Labour Economics*, 16(6): 679–88.

Falk, Armin, and Andrea Ichino. 2006. "Clean Evidence on Peer Effects." *Journal of Labor Economics*, 24(1): 39–57.

Fehr, Ernst, and Lorenz Goette. 2007. "Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment." *American Economic Review*, 97(1): 298–317.

Fehr, Ernst, and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics*, 114(3): 817–68.

Fischbacher, Urs. 2007. "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics*, 10(2): 171–78.

Gill, David, and Rebecca Stone. 2010. "Fairness and Desert in Tournaments." *Games and Economic Behavior*, 69(2): 346–64.

Gill, David, and Victoria Prowse. 2012. "A Structural Analysis of Disappointment Aversion in a Real Effort Competition: Dataset." *American Economic Review.* http://dx.doi.org/10.1257/aer.102.1.469.

Goffe, William L., Gary D. Ferrier, and John Rogers. 1994. "Global Optimization of Statistical Functions with Simulated Annealing." *Journal of Econometrics*, 60(1–2): 65–99.

Kahneman, Daniel. 2003. "A Psychological Perspective on Economics." *American Economic Review*, 93(2): 162–68.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, 47(2): 263–91.

Kőszegi, Botond, and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences." *Quarterly Journal of Economics*, 121(4): 1133–65.

Kőszegi, Botond, and Matthew Rabin. 2007. "Reference-Dependent Risk Attitudes." *American Economic Review*, 97(4): 1047–73.

Krähmer, Daniel, and Rebecca Stone. 2008. "Regret in Dynamic Decision Problems." http://www.wiwi.uni-bonn.de/kraehmer/dyn-reg-thy-171008.pdf.

Larrick, Richard P. 1993. "Motivational Factors in Decision Theories: The Role of Self-Protection." *Psychological Bulletin*, 113(3): 440–50.

Levitt, Steven D., and John A. List. 2007. "What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World?" *Journal of Economic Perspectives*, 21(2): 153–74.

Loomes, Graham, and Robert Sugden. 1982. "Regret Theory: An Alternative Theory of Rational Choice under Uncertainty." *Economic Journal*, 92(368): 805–24.

Loomes, Graham, and Robert Sugden. 1986. "Disappointment and Dynamic Consistency in Choice under Uncertainty." *Review of Economic Studies*, 53(2): 271–82.

Loomes, Graham, and Robert Sugden. 1987. "Testing for Regret and Disappointment in Choice under Uncertainty." *Economic Journal*, 97: 118–29.

McFadden, Daniel. 1989. "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration." *Econometrica*, 57(5): 995–1026.

Mellers, Barbara, Alan Schwartz, and Ilana Ritov. 1999. "Emotion-Based Choice." *Journal of Experimental Psychology: General*, 128(3): 332–45.

Newey, Whitney K. 1985. "Generalized Method of Moments Specification Testing." *Journal of Econometrics*, 29(3): 229–56.

Pakes, Ariel, and David Pollard. 1989. "Simulation and the Asymptotics of Optimization Estimators." *Econometrica*, 57(5): 1027–57.

Pope, Devin G., and Maurice E. Schweitzer. 2011. "Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes." *American Economic Review*, 101(1): 129–57.

Post, Thierry, Martijn J. van den Assem, Guido Baltussen, and Richard H. Thaler. 2008. "Deal or No Deal? Decision Making under Risk in a Large-Payoff Game Show." *American Economic Review*, 98(1): 38–71.

Prelec, Drazen. 1998. "The Probability Weighting Function." *Econometrica*, 66(3): 497–527.

Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics." *American Economic Review*, 83(5): 1281–1302.

Stern, Steven. 1997. "Simulation-Based Estimation." *Journal of Economic Literature*, 35(4): 2006–39.

Tversky, Amos, and Daniel Kahneman. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty." *Journal of Risk and Uncertainty*, 5(4): 297–323.

van Dijk, Frans, Joep Sonnemans, and Frans van Winden. 2001. "Incentive Systems in a Real Effort Experiment." *European Economic Review*, 45(2): 187–214.

van Dijk, Wilco W., Joop van der Pligt, and Marcel Zeelenberg. 1999. "Effort Invested in Vain: The Impact of Effort on the Intensity of Disappointment and Regret." *Motivation and Emotion*, 23(3): 203–20.

Wu, George, and Richard Gonzalez. 1996. "Curvature of the Probability Weighting Function." *Management Science*, 42(12): 1676–90.