How Hard to Fight? Cross-Player Effects and Strategic Sophistication in an Asymmetric Contest Experiment

February 16, 2017

Abstract

Many political phenomena—from wars to elections and lobbying—involve winner-take-all contests in which the value of the prize differs across the actors involved and from one issue to the next. To better understand competitive behavior in such environments, we conduct a controlled laboratory experiment in which participants face a series of asymmetric prize values in a lottery contest game. We find support for some, but not all, of the game's comparative static predictions. Most subjects respond to changes in their own values, but few subjects conditionally respond to cross-player changes. Our data suggest a new type of heterogeneity in the degree of strategic sophistication, one that differs substantially from the commonly used level-k model. We also administer two information based treatments, feedback and a calculator, finding that feedback on past play has a stronger effect on increasing subjects' strategic sophistication than a payoff calculator. (Abstract: XX words)

JEL Codes: C72, D72, D74

In domestic and international politics, political actors compete with one another to achieve mutually exclusive goals. These situations are often thought of as contests in which the chance of winning the prize is a function of the effort or resources that each side commits, but where increased competition is socially wasteful. War, for example, is a contest in which devoting more manpower, materiel, and industrial capacity increases the chance of winning but costs human lives and wasted economic productivity (e.g., Slantchev, 2010; Besley and Persson, 2011). Less violently, NGOs influence human rights policy by mobilizing and counter-mobilizing for and against reforms (Bob, 2012; Sell and Prakash, 2004). Regulations governing foreign and domestic interactions, such as tariff or environmental policy, can be thought of as the outcome of competition between special interest groups with opposing preferences who make campaign contributions (e.g., Grossman and Helpman, 1994; Goldstein and Martin, 2000). More generally, rent-seeking and lobbying are thought of as classic examples of contests (Tullock, 1967; Krueger, 1974; Becker, 1983). Similarly, the struggle for control of government between incumbents and challengers has been modelled as a contest in both autocracies (Myerson, 2008; Svolik, 2009) and democracies (Iaryczower and Mattozzi, 2013; Meirowitz, 2008; Serra, 2010).

Our study is motivated by the observation that contests occurring in the real world are often *asymmetric*. That is, parties to a conflict can differ in the value they place on winning the contest or in the effectiveness with which they transform their resources into competitive advantages. Winning access to a new ocean port is much more valuable to a landlocked country than a coastal one. Likewise, winning influence over tariff policy is much more valuable to an import-competing firm facing bankruptcy from foreign competition than it is to a consumer who dislikes marginally higher prices. Hiring a lobbying firm with extensive networks may be more effective than spending the same amount with a less connected firm. In electoral contests, incumbency can affect a politician's ability to turn campaign resources into electoral success while campaign spending has greater diminishing marginal returns for non-incumbents than for incumbents.

Just as contests vary across participants, features of the competitive environment also change between contests and across time. The discovery of natural resources in disputed regions has significant implications for the value of controlling a piece of territory (Ross, 2004*a*,*b*), as with the presence of oil in the conflict between North and South Sudan or the discovery of alluvial diamonds in Sierra Leone. Actions by international institutions can change costs and valuations in a domestic political contest over whether to comply with the rules of an international organization (Chaudoin 2015). Economic shocks, like the Great Recession, affect a firm or special interest group's urgency of obtaining a protective tariff or favorable regulatory ruling (Henn and McDonald, 2014; Davis and Pelc, 2012). Climate change can affect competition for scarce resources by raising the value of those resources (see Salehyan, 2008; Nordås and Gleditsch, 2007). In the electoral arena, an unforeseen scandal or sudden crisis can shift the advantage from one candidate to another (Abramowitz, 1991; Welch and Hibbing, 1997; Levitt, 1994), while a judicial ruling on campaign finance reform can unexpectedly increase the costs of influencing voters' perceptions (Meirowitz, 2008).

We use a laboratory experiment to study how these features of real-world contests—asymmetry in and changes to players' valuations of the prize—affect behavior in a Tullock-style lottery contest. The laboratory setting is appropriate because we can precisely control these valuations and observe participants' effort levels, quantities which would be difficult to measure in observational settings.[1] While there is a large body of experimental work studying contests in the laboratory, we are the first (to our knowledge) to test the effects of asymmetric and changing valuations on behavior.[2] Specifically, we test three comparative static predictions derived from Nash equilibrium

[1]As with any method, there are trade-offs. Lab experiments are similar to models in that the design focuses attention on a small set of key variables—the setting is not meant to perfectly replicate reality or generalize to observational phenomena. Yet, the experiment is valuable because it allows us to explicitly test behavioral predictions in a controlled setting in which we manipulate the salient features of many real-world situations.

[2]For an extensive, recent survey, see: Dechenaux, Kovenock and Sheremeta (2014). Much of this work focuses on explaining the total effort levels and heterogeneity between individuals in contests with stable, symmetric valuations. Work on variation across individuals has focused on a wide array of individual-level explanations, such as demographic

analysis. First, an increase in one player's own prize valuation directly increases their own effort. However, that increase also has second-order cross-player effects, indirectly causing an increase or decrease in her opponent's effort levels, depending on the players' relative valuations. If it increases her opponent's effort, because her opponent seeks to discourage or deter her from further competition, we call this "doing the deterring." If it decreases her opponent's effort, because the contest becomes more lopsided and the marginal return to effort decreases for her opponent, we call this "getting deterred," a concept similar to the discouragement effect documented in existing experimental work (Gill and Prowse, 2012*b*; Deck and Sheremeta, 2012).

In addition to investigating how these asymmetries specifically affect effort choices, we are also generally concerned with understanding the social costs of contest behavior. "Effort" in a war consists of purchasing armaments which can increase the human cost of war and worsen the "guns versus butter" tradeoff. In contests over policy, effort in the form of lobbying and campaign contributions is socially inefficient since that money is a rent captured by politicians. How actors respond to asymmetry and changes in valuations of the prizes in these contests, e.g. whether they respond as predicted by comparative statics or some other fashion, thus directly affects welfare through any subsequent changes in effort. ~~For example,~~ suppose that Country A is challenging Country B over the status quo division of territory. If an election is called for in Country A, this ~~change~~ might change A's leaders' valuation of the territorial prize. Knowing how that change affects welfare goes beyond simply assessing the likelihood of war and which side wins. We should also care about whether A *and B* change their effort levels, since additional effort in war usually means more deaths and less resources for domestic spending. Similar claims could be made about changes to electoral and lobbying contests that affect welfare-wasting expenditures. To the extent that effort and resources devoted to winning a political contest are not only wasted, but often destructive, anything that lowers contest effort represents societal gain.

Our study therefore assesses the degree to which information enhances strategic behavior and

characteristics, preferences toward risk, or other behavioral phenomena such as "the hot hand" fallacy.

improves social efficiency. Specifically, we are interested in how individuals respond to two different kinds of information about their strategic environments: experience and detailed calculations from hypothetical scenarios. Studying how these sources of information affect behavior has important implications for understanding political contests writ large. For example, if experience reduces contest expenditures, then we should expect leaders and politicians with longer tenure in office or who face long-term adversaries to be less likely to waste resources than novices or those who constantly face new strategic threats. We can think of precise calculations from hypothetical scenarios as computational aids to rational decision making. This is akin to a leader who grasps the strategic nature of contest actions in broad terms relying on detailed briefings and by intelligence and military officials who have worked out the consequences of different plans of action.

Our design directly manipulates these sources of information. In the baseline condition, players know only the rules of the game and the value of the prize to each player. In the *feedback* treatment, players also observe the effort levels of their opponent and the outcome of the contest after each round. In the *calculator* treatment, players are given a payoff calculator which allows them to search the action space and observe their own and their opponent's expected utilities for pairs of effort levels. Whereas the feedback treatment gives them *empirical* data about how their opponents have played and their resulting payoffs, the calculator treatment allows them access to *hypothetical* data about payoffs, a tool which is potentially very powerful, but which puts the onus on the participant to take advantage of it.

We find strong support for two comparative static predictions, and mixed support for a third. First, increasing a player's valuation increases their own effort under all treatment conditions. Second, we find support for the "getting deterred" effect, whereby increasing player i's valuation *decreases* player j's effort in all but two treatment conditions. In contrast, there is more modest support for the "doing the deterring" effect, as players do not always increase their effort levels in response to increases by their opponents.

In terms of information effects, we find that feedback and the payoff calculator both lower effort levels, which brings observed effort levels closer to the Nash predictions. However, the magnitude and statistical significance of the feedback effect are much higher than that of the calculator's effect. Even accounting for players' learning over time, the feedback treatment significantly decreases the distance between observed behavior and the Nash prediction. Additionally, we find that the feedback and calculator treatments increase the number of respondents whose behavior comports with the Nash comparative statics.

These findings shed light on how individuals gain knowledge about their strategic environments, which in turn, affects their effort levels. Tangible, experiential information, as embodied by the feedback treatment, more effectively induces strategic behavior than the abstract information embodied in the calculator treatment. This suggests that adversaries who engage each other frequently across multiple contests in the real world will behave more strategically, and with less wasteful effort, than competitors in one-off interactions.

Finally, in our analysis, we also assess the degree of variation in strategic sophistication displayed by the subjects. A large body of literature, such as work based on level-k models of iterated reasoning (Nagel, 1995; Stahl and Wilson, 1995) or models of cognitive hierarchies (Camerer, Ho and Chong, 2004), classifies individuals based on their degree of strategic thinking. We document a form of variation in strategic thinking that diverges from that of existing work. Specifically, we assess whether there is variation in the degree to which individuals behave according to the comparative statics, and we find that there is indeed significant heterogeneity across individuals in their strategic responses to differences in valuations. A small number of subjects display behavior that is consistent with all of the comparative static predictions while approximately forty percent display behavior that is consistent with only one, but not both, of the cross-player hypotheses.

Our characterization of a distinct type of heterogeneity in individual behavior is important since a growing body of literature is interested in an individual's level of strategic sophistication as an explanatory and as an outcome variable. Iterated reasoning and the ability to anticipate your opponent's moves are only two aspects of "strategic thinking." Our research highlights how individuals vary in their ability to understand how changes to the game affect themselves and their opponents incentives. This is a type of dynamic strategic thinking, where individuals vary in how their degree of understanding when it comes to changes that influence strategic interactions. This aspect of strategic thinking is also empirically distinct from the features captured by level k models. Subjects' levels are poorly correlated with the degree to which their behavior matches comparative statics in our game, and individual level predictors of levels do not similarly predict an individual's level of comparative static behavior. This dynamic aspect of strategic thinking may be more appropriate in real world settings where the parameters of the situation are fluid and strategic reactions are paramount.

Contest Model

We consider a simple contest model in which two players can each exert costly effort in order to increase their chances of winning a prize. Each of the two players, i and j, has a strictly positive value to winning the prize, V_i and V_j, where the prize values are distinct, $V_i \neq V_j$. Each player chooses an effort level, denoted e_i and e_j, and they have constant marginal costs of effort, $c_i = c_j = 1$. The contest is a function which maps their effort levels into the probability of winning the prize. The probability that player i wins the contest is $\phi_i(e_i, e_j) = \frac{e_i}{e_i+e_j}$, and we assume that no one wins the prize if neither player exerts any effort, $\phi_i(0,0) = 0$. This is the familiar ratio or Tullock (1967) contest success function. Player i's objective function is

$$\Pi(e_i, e_j) = \phi_i(e_i, e_j)V_i - e_i,$$

and the Nash equilibrium effort level obtained from the players' accompanying first order conditions is

$$e_i^* = \frac{V_i^2 V_j}{(V_i + V_j)^2}.$$

How do optimal effort levels change as players' valuations change? The simplest effect of changing valuations is that Player i's optimal effort level is monotonically increasing in her own valuation to winning the contest.[3] As the contest prize becomes more valuable to Player i, she is willing to exert more effort to win the prize, regardless of Player j's valuation. We call this the "own value" (OV) effect.

The effect of V_j on Player i's optimal effort level, however, depends on the two players' *relative* valuations.[4] When Player j values the prize more than Player i, increasing V_j decreases the marginal return to effort for Player i, which *decreases* i's optimal effort. We call this the "getting deterred" (GD) effect. When Player j values the prize less than Player i, increasing V_j *increases* i's optimal effort. As V_j increases, the marginal utility to effort, which helps Player i retain the prize she values so highly, also increases. We call this the "doing the deterring" (DD) effect.

Player i's optimal effort level thus varies non-monotonically with Player j's valuation. These two effects are akin to deterrence. The player with the higher valuation responds to increases in her opponent's valuation and subsequent effort levels with reciprocal increases in her own effort. The player with the lower valuation responds to increases in her opponent's valuation and subsequent effort levels with decreases in her own effort.

The three dimensions to the comparative statics show that the strategic interaction between players is more complicated than a simple discouragement effect.[5] Consider a scenario where the players start with equal valuations (symmetry) and then change so that i's value increases while j's decreases (asymmetry), which is one of the treatments in the contest experiment by Anderson and Stafford (2003). Comparing effort levels under these versions of symmetry against asymmetry, however, conflates the three dimensions of the shock's effect on effort that we identify.

[3]Formally, $\frac{\partial e_i^*}{\partial V_i} = \frac{2V_i V_j^2}{(V_i+V_j)^3} > 0$.

[4]This is because the sign of $\frac{\partial e_i^*}{\partial V_j} = \frac{V_i^2(V_i-V_j)}{(V_i+V_j)^3}$ depends on the $V_i - V_j$.

[5]Until now, we have discussed only the effects of changing valuations but have not discussed changing marginal costs to effort. However, increasing one player's valuation is isomorphic to decreasing their marginal costs of effort. The three effects identified above also obtain for changes to marginal costs of effort. The effect of decreasing c_i on the optimal effort of both players is the same as the effect of increasing V_i (Corchon, 2007).

Our experimental protocol is designed to detect and decompose all three of these effects.

Experimental Design and Procedures

We described the task to subjects as a "Lottery Contest Game" in which two players compete for a prize. Subjects played the game multiple times, and we referred to each play of the game as a "round." We informed subjects that the prize would be worth different amounts to each player in each round and that they would know their exact values and their opponent's values when making their decision. Their decision was described in terms of "purchasing contest tickets," with the probability of winning the prize equal to one's share of total tickets purchased in the round. Each ticket cost 1 point, and subjects received a fresh endowment of 1000 points in every play of the game and kept whatever portion of their endowment they didn't spend.

Each experimental session was divided into two parts. As described below, we divided the rounds into parts so that in some of the sessions we could vary the information available to subjects between the two parts. At the beginning of Part 1, subjects received written instructions explaining the contest game (see the Appendix). After the experimenter read the instructions out loud, subjects took a brief comprehension quiz. In Part 1, subjects played 17 rounds of the two-player asymmetric contest game, 16 rounds with positive valuations and a 17th round where both players' value to winning the prize was zero. In Part 2, any changes in the instructions were distributed and read and then subjects played another 16 rounds of positive valuations and a final zero value round. In every round, subjects were randomly matched with another player and were informed of each player's valuation of the prize.

We selected eight distinct pairs of prize valuations (as shown in Table 1) to test the comparative static predictions. We refer to the set of valuations $S = \{200, 900\} \times \{300, 800\}$ as *single valuations*, and the set $D = \{400, 1800\} \times \{600, 1600\}$ as *double valuations* since each pair in the latter set is twice the value of a pair from the former. We also refer to prize values of $200, 300, 400,$

Table 1: Valuations and Nash equilibrium predictions

	Valuations		Nash predictions		Expected payoffs	
	v_1	v_2	e_1^*	e_2^*	EU_1	EU_2
Single valuations	200	300	48	72	32	108
	200	800	32	128	8	512
	900	300	169	56	507	19
	900	800	224	199	253	177
Double valuations	400	600	96	144	64	216
	400	1600	64	256	16	1024
	1800	600	338	112	1014	38
	1800	1600	448	398	506	354

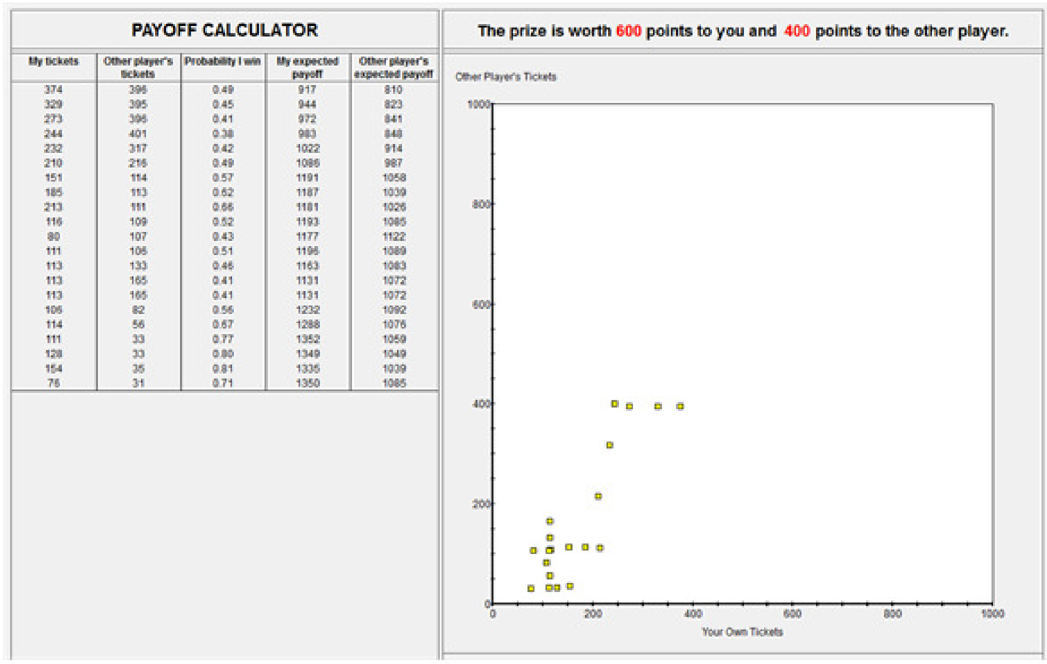
and 600 as *low values* and to prize values of $800, 900, 1600$, and 1800 as *high values*.

If v_i is low, then increasing v_j from low to high generates the "getting deterred" comparative static prediction (whereby e_i^* decreases), while if v_i is high, then increasing v_j from low to high generates the "doing the deterring" effect (whereby e_i^* increases). Each subject played each of these pairs twice within each part, once as Player 1 and once as Player 2. We randomly generated a sequence of valuation pairs prior to the first session, with the order independent across parts 1 and 2, and held the sequence fixed across sessions. The purpose of the "zero value" round (in which $v_i = v_j = 0$) measures whether any player has an unobserved preference for winning the contest (i.e., "joy of winning").

In addition to manipulating valuations, our design also manipulates the information available to the subjects in two ways: availability of feedback and the presence of a payoff calculator to test whether varying information or providing a computational aid might enhance subjects' strategic thinking and encourage choices closer to the equilibrium predictions. In the *Feedback* treatment, each round provided respondents with a screen that included the effort level of each player, the probability that each player would win the prize given the chosen effort levels, which player won the contest, and each player's payoff (denominated in points) for all of the previous rounds. We refer to the absence of feedback as the *No feedback* treatment.

We also varied whether subjects had access to a computational aid in each round. In the *Calcu-*

Figure 1: Screenshot of graphical interface for payoff calculator



lator treatment, we provided subjects with a graphical interface in every round (shown in Figure 1), while no such calculator was available in the *Baseline* treatment. The graphical interface for the calculator allows subjects to search the strategy space quickly and easily. To do so, subjects clicked on a point in the white square on the right side of the screen (which represents the strategy space). Each time a subject clicked on the calculator, the pair of effort levels was displayed in a list on the left side of the screen along with the probability of winning and each player's expected payoff. Each subject saw a list of all of their previous searches in that round, and at the beginning of every round the calculator was reset.

Crossing the feedback and calculator manipulations yields four conditions: *Baseline-No feedback* (BN), *Baseline-Feedback* (BF), *Calculator-No feedback* (CN), and *Calculator-Feedback* (CF). As summarized in Table 2, we structured the sessions and treatments to allow for both within- and between-subject comparisons. In four of the sessions, we varied feedback within-session, with *No feedback* in Part 1 and *Feedback* in Part 2, holding constant *Baseline* or *Calculator*. In the other six sessions, we held the condition constant for both parts of the entire session (*BN*, *BF*, or *CF*

Table 2: Experimental design

Condition			
Part 1	Part 2	# Sessions	# Subjects
BN	BF	2	32
BN	BN	2	26
BF	BF	2	34
CN	CF	2	28
CF	CF	2	30

). The sessions and treatments were structured so that information was not ever taken away from a subject, e.g. there are no sessions where a subject starts with feedback and then feedback is removed.[6]

We programmed the experiments in z-tree (Fischbacher, 2007) and conducted them in the [experimental laboratory at authors' institution]. A total of 150 subjects participated in the experiment. At the end of each session, we randomly selected one round for payment and converted points to cash at a rate of $1 per 75 points. Each session lasted less than an hour and a half, and subjects earned an average of $21.50 (including a $5 show-up fee).

The laboratory experiments most closely related to ours investigate asymmetry across players and the so-called discouragement effect, where a stronger player (one with higher valuations, lower costs to effort, or better effort technology) induces the weaker player(s) to decrease effort. Anderson and Stafford (2003) vary participants' costs to effort and find that costs are negatively associated with effort levels. Fonseca (2009), Anderson and Freeborn (2010), and Kimbrough, Sheremeta and Shields (2014) investigate games where players can have different effort technology. In these games, a unit of effort by a "strong" player has a greater marginal effect on her winning probability than a unit of effort from a weak player. They find that weak bidders gener-

[6]We thank the reviewers for emphasizing the importance of the between-subjects comparisons for disentangling the presence of feedback from experience, since our original data featured only within-subject comparisons. Analyzing only the within-subjects comparisons raises the possibility that subjects learn over time so that any effect of feedback may be confounded with learning in sessions where we introduced feedback in part 2. While we provide details in the Appendix that the estimated treatment effects are not attributable to learning, the between-subject data from rounds in Part 1 provide us with cleaner comparisons.

ally exert less effort. Deck and Sheremeta (2012) analyze an experiment where a defender must defend against a sequence of attacks, and they vary the defender's value to successfully defending all attacks. They find mixed support for a discouragement effect.

Experimental literature on contests also consistently identifies the phenomenon of "overbidding," where players' effort levels are much higher than the Nash equilibrium prediction. Dechenaux, Kovenock and Sheremeta (2014) observe that the degree of overbidding is sometimes high enough to give the players negative payoffs, meaning that they would have been better off not participating in the contest at all. There are likely many contributing factors to overbidding, such as if a player derives non-monetary utility from winning the contest or if the player simply makes mistakes.[7]

Results

Comparative Statics

We first analyze the results with respect to the comparative static predictions of the asymmetric contest game. To test the predictions, we estimate the following regression model of effort choice

$$\text{Effort}_i = \sum_{t \in T} \beta_t \text{High}_{it} + \sum_{t \in T} \gamma_t \text{GD}_{it} + \sum_{t \in T} \delta_t \text{DD}_{it} + \sum_{t \in T} \alpha_t + \varepsilon_i$$

where i indexes observations and t indexes the set of treatments crossed with the set of valuations, $T = \{BN, BF, CN, CF\} \times \{S, D\}$. This specification allows us to estimate the comparative static predictions for each treatment and valuation separately. The treatment-specific dummy variables α_t allow for the baseline effort levels when own valuations are low to vary across treatments (i.e., varying intercepts). High_{it} is a dummy variable indicating that the player's value is high (where high and low values are defined as in the previous section) and that observation was under treatment t. The set of coefficients β_t measure the effect of increasing i's own valuation separately for each

[7] Dechenaux, Kovenock and Sheremeta (2014) contains a section reviewing overbidding.

of the treatments t. In theory, effort is increasing in one's own valuation, so we expect all $\beta_t > 0$.

The dummy variable GD_{it} indicates that j's value is high and i's value is low (and that the observation is under treatment t), so the coefficients γ_t measure the getting deterred effect (how increasing j's value affects i's effort when i's value is low). Similarly, the dummy variable DD_{it} indicates that j's value is high and i's value is high. We expect from our theoretical analysis that all $\gamma_t < 0$. The coefficients δ_t measure the doing the deterring effect (how increasing j's value affects i's effort when i's value is high). We expect all $\delta_t > 0$.

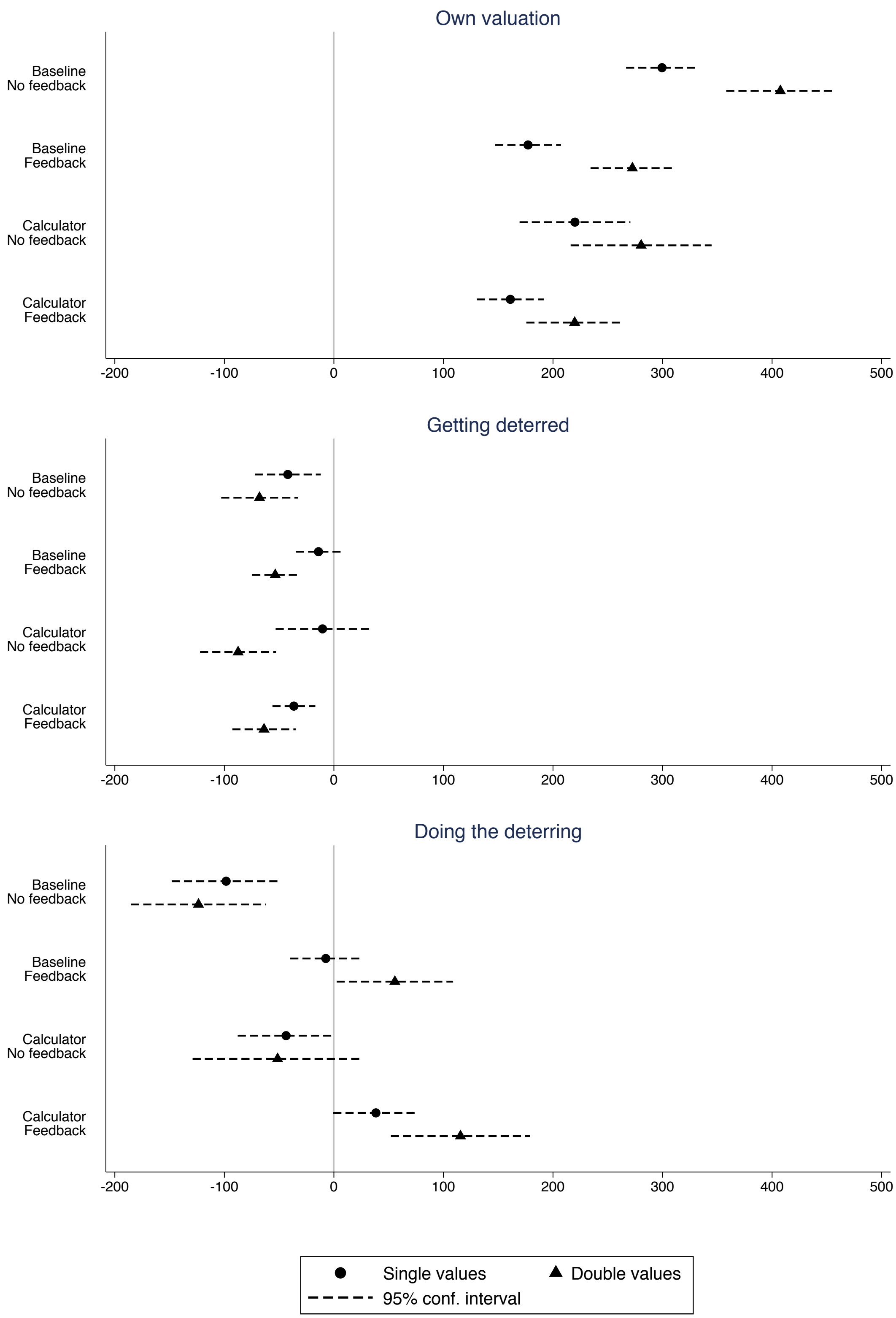
Figure 2 presents ordinary least squares regression estimates of the coefficients in this model for rounds with single valuations (circles) and with double valuations (triangles).[8] To account for within-subject dependence, we use robust standard errors clustered at the subject-level. Each pane corresponds to a particular comparative static coefficient: β_t for own value effect (top), γ_t for getting deterred (middle), and δ_t for doing the deterring (bottom).

The first thing to note about our results is that subjects respond naturally to increases in their own valuations. Looking at the top pane of Figure 2, all of the coefficients are positive and significantly different from zero. In substantive terms, for single valuation rounds, subjects purchase about 96 tickets when their valuations are low, compared to 309 tickets when their valuations are high. The exact magnitudes vary somewhat across treatment conditions, but the overall effect of increasing one's own valuation is consistently positive. The same holds for double valuation rounds, with effort levels that are nearly doubled in the direction that equilibrium theory predicts: subjects purchase an average of 173 tickets when their valuations are low and an additional 507 tickets when their valuations are high. While not entirely surprising, these results provide assurance that subjects respond rationally to changes in the size of their own prize, consistent with the predicted own valuation effect.

More interestingly, the results provide evidence for the cross-player comparative statics—specifically, there is strong evidence for the getting deterred effect. In the middle pane of Figure 2,

[8] This analysis excludes the zero valuation rounds.

Figure 2: Comparative static coefficient estimates



all of the estimated coefficients are negative across treatment conditions. The estimated coefficients are significant at the 0.01 level for all treatment conditions in the double valuation rounds, and they are negative and significant in half of the treatment conditions for the single valuations (in the *BN* and *CF* conditions).[9] Overall, consistent with the comparative static prediction, we find that when subjects' own valuations are low, they respond to increases in their opponent's valuation by reducing their effort.

The evidence for the doing the deterring effect is mixed, as subjects respond differently across treatments. When subjects do not receive feedback (the BN and CN conditions), they respond to increases in their opponent's valuation by *decreasing* their effort—opposite of what the comparative static analysis predicts—and these effects are largest in the BN condition for both single and double valuations. This effect is significant in the BN condition, but not in the CN condition. In contrast, behavior is generally consistent with the theoretical predictions with feedback, as subjects respond to their opponent's valuation with increases in effort. These effects are statistically significant in the CF condition with both single and double valuations, and for the BF condition with double valuations. The data suggest that the change in subjects' effort levels depend on their information and are consistent with the theoretical predictions only when feedback information about the history of play is available. We return to this in later sections, showing the effect of treatment on the strategic sophisticaion of different subjects.

In general, it seems that players engaged in incomplete strategic reasoning. They correctly recognized that valuations affected their own, and their opponents' effort levels. In the case of the getting deterred prediction, they correctly inferred the second order effect of valuation changes when they were in positions of relative weakness, namely that they might decrease their own effort in response to increases in their opponents'. However, they did not make an analogous second-order calculation when they were in positions of relative strength, recognizing that they should

[9]In regressions that pool the observations across treatment conditions and do not include interaction terms, the own value and getting deterred coefficients are negative and significant for both the single valuations and the double valuations. These results are in the appendix.

increase their own effort to deter players with lower valuations.

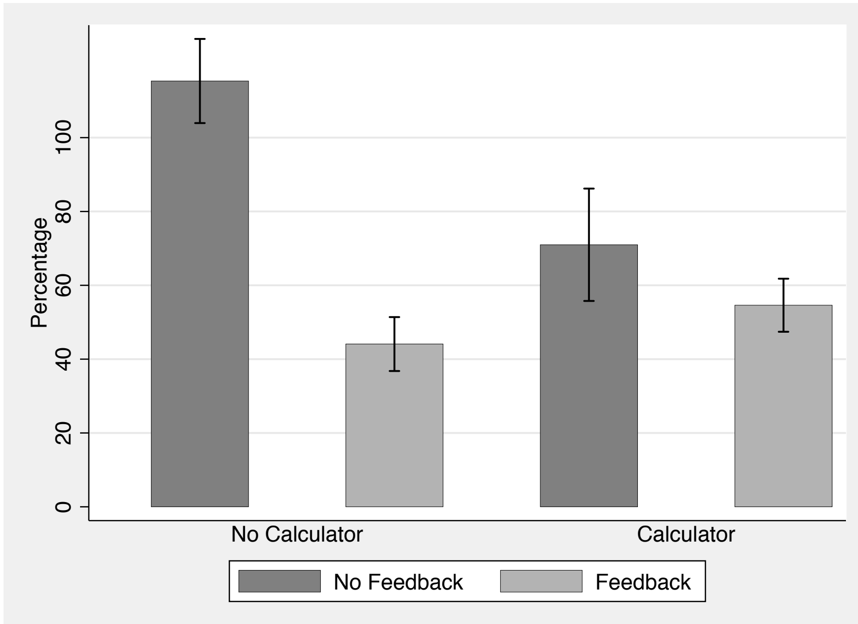
Information Effects

How do the information treatments affect the players' effort choices? Does feedback or access to the payoff calculator reduce socially wasteful contest effort, bringing expenditures closer to Nash equilibrium predictions and increasing social efficiency? As a simple assessment of the treatment effects, we first analyzed whether providing feedback or the payoff calculator promotes effort levels closer to the Nash equilibrium predictions. Here, the dependent variable of interest is the percentage difference between observed effort and the equilibrium prediction: Pct. Difference $= \frac{\text{Effort}-\text{Prediction}}{\text{Prediction}}$. Smaller magnitudes indicate behavior that is closer to the predictions. We pool the data across valuation pairs in each condition and exclude the zero valuation round from analysis.

The results show that both feedback and the calculator have negative treatment effects, though in general, the feedback effects are stronger. Figure 3 presents the average difference by experimental condition. There is a high level of effort, relative to the Nash prediction, in the baseline BN condition (115%). Introducing feedback has a substantial effect in the *Baseline* treatment, reducing effort by more than half (to 44%). Introducing the payoff calculator (holding the absence of feedback constant) also decreases the amount of effort (to 71%), although not as much as introducing feedback alone does. The effect of introducing feedback in the calculator treatments also decreased effort, though to a smaller degree (from 71% to 55%). However, introducing the calculator in the feedback treatments seems to increase the degree of effort somewhat (from 44% to 55%).

To account for differences between rounds that might affect our estimates of the treatment effects, we estimated a series of OLS models regressing the percentage difference from Nash predictions on indicators for the availability of feedback, the availability of the calculator, and their interaction. We use robust standard errors clustered by subject to account for dependence across observations and report models with and without a set of additional controls. The results are presented in Table 3.

Figure 3: Effort as percentage of Nash prediction



The first two columns provide estimates using all of the data, from both parts 1 and 2 of each session. The first column presents the model specification without controls, while the second column includes controls for the equilibrium effort level for the player in that round (*Nash effort*, which effectively controls for prize values), an indicator for *double valuations*, a counter for the number of previous rounds the player has played (*experience*, excluding zero valuation rounds), and the average effort chosen by the subject in the two zero valuation rounds (*zero value effort*), which controls for competitiveness or a preference for winning.[10]

The regression analysis in Table 3 reinforces our interpretation of Figure 3. We find that feedback has a sizable, statistically significant effect on reducing effort. The calculator also reduces effort, but the magnitude of this effect is smaller and only marginally significant at the 0.10 level. The positive coefficient on the interaction term implies that the effect of the different kinds of in-

[10]We also estimated a specification that included an additional set of individual-level traits. These included gender, an aggression scale (citation), and a "Machiavellian" scale (citation). Including the additional controls does not change the point estimates by much, so we include only the smaller set of controls for ease of presentation. The Appendix presents the additional specifications.

Table 3: Regression analysis of information treatments

	All data		Part 1 only		Part 2 after BN	
	b/se	b/se	b/se	b/se	b/se	b/se
Feedback	-71.25**	-66.44**	-56.15**	-51.76*	-77.93**	-132.99**
	(17.75)	(17.79)	(21.04)	(20.90)	(27.85)	(45.68)
Calculator	-44.36+	-47.10+	-42.62+	-41.20+		
	(24.71)	(23.91)	(24.76)	(23.78)		
Feed. × Calc.	54.88*	55.98*	54.84+	51.39+		
	(24.10)	(23.65)	(30.27)	(29.70)		
Nash effort		-0.13**		-0.20**		-0.05
		(0.04)		(0.04)		(0.05)
Double valuation		-10.94*		-14.21*		-12.32
		(4.91)		(6.22)		(10.23)
Experience		-1.00*		-3.61**		-3.63*
		(0.40)		(0.97)		(1.63)
Zero value effort		0.25**		0.34**		0.12**
		(0.05)		(0.06)		(0.03)
Constant	115.34**	151.28**	113.60**	178.22**	119.21**	215.61**
	(17.64)	(22.65)	(17.70)	(23.51)	(23.43)	(54.44)
N	4,800	4,800	2,400	2,400	928	928
R^2	0.03	0.06	0.02	0.06	0.05	0.07

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

formation, experiential and hypothetical, are not additive. That is, the combined effect of feedback and the calculator is not effectively different than feedback by itself. If anything, the estimates suggest that having both tools available is somewhat worse than having either tool alone. These results hold up in column two when we include the control variables. Notably, while over-effort decreases over time with experience, the estimated effect of feedback diminishes only slightly, suggesting that the effect of feedback is due to informational content rather than merely increased familiarity playing the game.

The results in the remaining columns of Table 3 provide additional assurance that the effect of feedback is not merely due to increased familiarity with the strategic environment gained through learning and experience. In columns three and four, we estimate models using data only from part 1 of each session, thus omitting observations from part 2 that could have been influenced by experience with different environments (e.g., *BN* versus CN). When we place each treatment on an equal footing (purely between-subjects) with respect to prior experience, the results are similar. The coefficient for feedback remains negative and statistically significant at the .05 level while the coefficients for the calculator and interaction remain significant at the .10 level. The magnitude of the feedback coefficient is smaller when we use only part 1 data, which suggests that the effect of feedback in columns 1 and 2 cannot be entirely accounted for by learning.

In the last two columns, we use only part 2 data from sessions in which subjects play the contest game in the Baseline-No feedback (BN) condition in part 1. That is, we compare part 2 Baseline-No feedback (BN) with part 2 Baseline-Feedback (BF) holding constant experience with BN in part 1. This version of the experiment allows us to further distinguish learning from feedback in that subjects' part 1 experience allows them to gain familiarity with the game *before* they are exposed to feedback. Interestingly, we find that prior experience with the BN condition *increases* the effect of feedback, which further bolsters our conclusion that feedback information is distinct from experience.

Overall, the effects of both the feedback and calculator treatments are consistent with the idea that having more information about the underlying structure of the game and how other players have behaved decreases wasteful effort and increases efficiency. The stronger feedback effects are consistent with the idea that tangible, experiential information provides better strategic information—information about how others' play—that individuals find more useful than the more abstract calculator tool. Furthermore, we speculate that the interactive effects could simply be the result of cognitive overload: Having both sets of information decreases the effectiveness of each because subjects spend mental energy trying to figure out which tool to use rather than how to use the tools most effectively.

Search Quality

The calculator is potentially a much powerful analytical tool than feedback because a player can quickly learn a lot about the underlying payoff surfaces for herself and her opponent. If she had a hypothesis about her opponent's effort levels, she could use the calculator to find her best response. If she were willing to search the space extensively, iteratively finding best responses in a way that traces out Cournot-style best response dynamics, she could identify the Nash equilibrium of the game. Used effectively, a player would also be able to recognize the cross-player comparative statics or at least compute and choose the effort levels consistent with them. So why don't we observe behavior closer to the equilibrium predictions in the calculator treatments?

One possible explanation is that the quality of searches was generally low. Rather than iteratively searching for the players' best responses, strategically naive subjects might instead use the calculator in relatively simple ways: for example, to check the probability of winning or the expected payoffs associated with a given number of contest tickets, perhaps searching for an effort level until it exceeds an unobserved threshold in a manner that akin to satisficing. Since we programmed our interface to store all of the subjects' clicks in the calculator tool, we can investigate these possibilities by constructing and analyzing several measures of search quality.

For an initial measure of search quality, we code whether each click or guess yields net positive expected utility relative to purchasing 0 tickets and ensuring a payoff of 1000 points for each player. This is a "minimal" measure of search quality in that it only requires that subjects search an area of the strategy space that is minimally rational for one or both players. Another measure of search quality relates to the direction of search. To compute the direction, we calculate the angle of the vector defined by two successive clicks. Horizontal searches reflect a subject's attention to her own payoffs, while holding constant her opponent's effort, in a manner consistent with searching for one's own best response. Likewise, vertical searches reflect attention to her opponent's payoffs, consistent with searching for her opponent's best response.[11]

We find that the quality of subjects' searches according to these measures tends to be fairly poor. Across all of the calculator sessions (*CN* and *CF* conditions), subjects clicked a total of $12,010$ times in the calculator tool, with an average of 6.5 clicks per round.[12] According to our positive expected payoff measures, at most half of subjects' searches in the calculator treatment can be classified as minimally rational: 53% of guesses involve positive net expected utility for the subject's own payoffs and 50% of guesses for their opponent's payoffs. However, fewer than one-third of guesses (32%) involve positive expected payoffs for both the subject and their opponent. The prevalence of negative expected payoff guesses suggests to us that most searches are of low quality.

We also find that horizontal and vertical searches comprise between 50% and 70% of all the guesses entered into the calculator, suggesting that searchers did tend to focus on varying one dimension of their search at a time. Searches along one dimension also tend to be more horizontal ($30-40\%$) than vertical ($23\%-32\%$). This pattern of search behavior suggests that subjects tend to focus on their own payoffs rather than their opponents', which is consistent with our finding that the own value effects are generally much stronger and consistent with the theoretical predictions

[11]In our analysis, we allow for two levels of error tolerance in how we classify horizontal and vertical searches, with a relatively narrow tolerance of $\pm 10°$ and a wider tolerance of $\pm 22°$. Both levels yield similar results.

[12]See the Appendix for full details about our measures of search quality.

than the cross-player comparative statics.

In general, better searching yielded effort levels that were closer to Nash prediction.[13] In regressions of the percentage of over-effort on various measures of search quality, searching in the Both Positive region yielded lower effort levels. Searches the were only vertical or horizontal resulted in higher effort levels, thought this effect was not significant. More extensive searching, measured by the number of clicks and the distance covered, was not associated with effort levels closer to Nash predictions.

Strategic Sophistication

While we observed behavior consistent with several of the predicted comparative statics (the own value and cross-player getting deterred effects), effort levels remain well above the Nash equilibrium predictions even with information and experience. To gain a better understanding of this behavior, we draw from a large and growing body of research in behavioral game theory that explains and organizes experimental subjects' departures from Nash equilibrium in terms of heterogeneous levels of strategic sophistication (Agranov et al., 2012; Arad and Rubinstein, 2012; Crawford, 2003; Gill and Prowse, 2012*a*). In this section, we describe and then apply this approach.

The basic idea is that some players are more strategic than others: more sophisticated subjects engage in higher orders of reasoning or form more accurate forecasts of others' behavior. The two most commonly used models, level-k and cognitive hierarchy (CH) models, were developed in the context of the "beauty contest" game, where players choose a number between 1-100 and are rewarded if they choose the value that is closest to $2/3$ times the average value the players chose (Nagel, 1995). The game is solvable by iterated reasoning, so players who engage in more steps of reasoning are described as having higher levels of strategic behavior. The level-k model describes the most naive, level-0, players as choosing randomly over the full support of possible actions. A

[13] Regression results are in the appendix.

level-1 player chooses the best response to a population of level-0 players. A level-2 player best responds to level-1 players, and so on. Players are classified by the level that is most consistent with their observed behavior, and players with higher levels are thought to have greater ability to anticipate their opponents' actions and incentives.

Experimental research has used a player's level as both an explanatory variable and an explanator. While there is substantial research in economics on strategic behavior in experimental games (Crawford, Costa-Gomes and Iriberri, 2013), we will focus on the small but growing literature in political science that uses these methods. For example, Hafner-Burton et al. (2014) use a beauty contest game to measure participants' levels of strategic reasoning. They find that players with higher observed levels choose more strategically in an international negotiation and cooperation setting. Hafner-Burton, Hughes and Victor (2013) hypothesize that variation in the level of strategic thinking explains the difference between experimental results using college student convenience samples and those using more experienced policy elites. Bausch and Zeitzoff (2014) find variation in individuals' level of strategic thinking in a terrorist/counter-terrorist laboratory game. Loewen, Hinton and Sheffer (2015) find similar variation in a game about strategic voting in an election. Bassi (2015) analyzes how variation in electoral rules and candidate profiles can affect subjects' degree of strategic sophistication as measured by their level in a level-k model. Bassi and Williams (2014) analyze how variation in financial incentives affects the level of subjects' strategic thinking. Minozzi and Woon (2013) find that a level-k model seems to fit non-equilibrium behavior in a communication game with competing experts.

We take an alternative, broader view of strategic sophistication, which we believe encompasses a variety of differences in strategic thinking beyond the number of steps of iterated reasoning.[14] Thinking strategically means understanding a game's underlying incentives as well as understand-

[14]Although political scientists tend to treat strategic thinking as if it were a stable trait (Hafner-Burton, Hughes and Victor, 2013; Loewen, Hinton and Sheffer, 2015), a growing literature suggests that observed levels are contextual, varying with changes in beliefs and across games in inconsistent ways (Agranov et al., 2012; Georganas, Healy and Weber, 2015).

ing the behavior of others. One might be able to formulate the best response to an ideal opponent but misjudge how close one's opponent comes to that ideal. Conversely, one might accurately anticipate opponent's behavior but fail to recognize the optimal response.

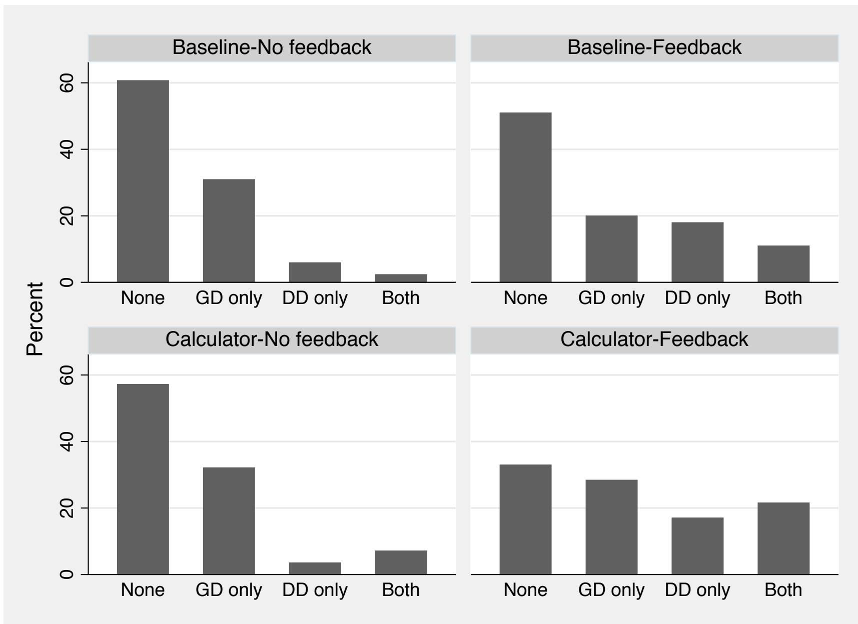
Strategic reasoning *also* encompasses an individual's ability to react to changes in the strategic setting. In the canonical beauty contrast game, the parameters of the game and the incentives facing the players are static. In contrast, most ~~off~~ the theoretical models of interest to political scientists and economists generate predictions about how actors respond to changes in their strategic environment. Unlike the canonical beauty contest experiments, varying the prize values in our experiment means that subjects face a changing competitive landscape.

Good strategic thinking also involves understanding comparative statics: how changes in the structure of the game affect incentives, behavior, and the feedback between these aspects of strategic interaction. Players with lower levels of strategic sophistication might understand how changes in their own prize values affect their own payoffs and incentives in a decision-theoretic sense (i.e., holding the behavior of others constant) but fail to think about second order effects on the behavior of others. At higher levels of sophistication, players understand how others' prize values affect their opponents' behavior, thereby affecting their own behavior even if their own value of the prize remains constant. In the models of interest to political scientists, this dimension of strategic sophistication matters a great deal more than the number of steps of iterated reasoning.

Here, we focus only on the cross-player doing the deterring (DD) and getting deterred (GD) comparative statics since there is already strong support for the own-player comparative statics across subjects. We construct subject-level measures of consistency with each of these cross-player comparative statics the following way. We first take each pair of rounds for which player i's value is the same and compute the change in effort as the opponent's value increases from low to high. We then code whether the change in i's effort is consistent with getting deterred (decreases when i's value is low) or doing the deterring (increases when i's value is high). We then compute the percentage of changes for each subject that are consistent with each comparative static and code a subject as consistent if their consistency rate exceeds 50% for each comparative static, respectively.

We found substantial heterogeneity across respondents in the degree to which their behavior is consistent with the Nash comparative static predictions. Figure 4 shows the distribution of ~~subjects~~ consistency with one or both comparative statics for each information condition.

Figure 4: Subject classification of comparative static consistency



The feedback treatment increased the subjects degree of strategic behavior. The feedback treatment increases the number of individuals whose behavior is consistent with each prediction, especially the DD effect. Looking from left to right in each row, adding feedback shifts the distributions rightward, towards an increasing mass of individuals displaying behavior consistent with one or both comparative statics. Without feedback, only 8% of individuals' behavior is consistent with the DD prediction, compared to 35% with feedback.[15] The feedback treatment also increases the number of individuals whose behavior is consistent with the GD prediction as well, 37% without feedback compared to 44% with feedback. The percent of individuals whose behavior is consistent with both predictions increases from 3% without feedback to 18% with feedback.

[15] These numbers pool the data across the Calculator and No Calculator treatments.

Figure 4 also shows how the calculator treatment increased the degree to which subjects' behavior was consistent with predictions. Looking from top to bottom in each column, adding the calculator shifts the distribution to the right, increasing the number of respondents whose behavior is consistent with one or both comparative statics. The calculator treatment increased the percentage of individuals whose behavior was consistent with the DD effect from 21% to 34%.[16] For the GD effect, it increases the percentages from 37% to 55%. The percentage of individuals whose behavior is consistent with both triples, from 7% to 21%.

Interestingly, the two treatments have mutually reinforcing positive effects on the degree of subjects' strategic behavior. In the above analysis of the distance between observed behavior and Nash predictions, having both treatments yielded a negative interactive effect. The presence of one treatment tended to lessen the impact of the other. Here, however, the two treatments seem to have positive interactive effects. The highest proportion of subjects displaying behavior consistent with both predictions occurs in the Calculator - Feedback treatments.

Conclusion

Many real-world situations are like contests where players value the prize differently. In wars, lobbying battles, campaigns, and other contests, the stakes of the contest can also change for one or more participants across issues, levels, or time. This paper has focused on how players respond to these changes and to asymmetry. Under certain conditions, equilibrium theory predicts these shocks increase or decrease one or the other player's effort levels. Effort is important because the amount of effort exerted in a contest incurs opportunity cost—such effort and resources could have contributed to societal welfare, but did not. Understanding the effect of these shocks on effort is worthy of attention.

In a laboratory setting, we found support for some but not all of the comparative static pre-

[16]These numbers pool the data across the Feedback and No Feedback treatments.

dictions regarding asymmetry and valuation shocks. Intuitively, players increase their effort levels in response to positive valuation shocks, and vice versa. In terms of the predicted cross-player effects, we find that players "get deterred": when player j values the prize less highly than player i, positive shocks to i's valuation decreases j's effort level. However, we find mixed support for "doing the deterring." That is, theory also predicts that, when j values the prize more than i, increases in i's valuation should increase j's effort. We find that this prediction finds support only under treatments in which players have feedback about past rounds.

There is also significant heterogeneity across subjects in the degree to which their behavior is consistent with the Nash predictions, with some subjects displaying some, none, or all of the predictions. Both the feedback and calculator treatments increased the strategic sophistication of the subjects' behavior. We speculate that feedback helps players better understand the strategic incentives when they respond to valuation shocks. This potential mechanism may stem from cognitive complexity because the own-value effects are easy to understand: players should exert more effort when the prize is worth more to them. The cross-player dynamics are subtler and more difficult to understand. It is possible that players have an easier time understanding the "getting deterred" effect, but only understand the "doing the deterring" effect with greater experience, having observed past play via feedback.

The differential findings for the calculator and feedback treatments have implications for real world contests, where some contests are characterized by repeated interactions with the possibility for gaining experiential feedback and others are characterized by changes in the principals and decision-makers. In international relations, for example, a large body of work focuses on leadership turnover and its effects on crisis bargaining.[17] Wolford (2007) argues that leadership turnover can affect crisis bargaining because new leaders may have differing levels of resolve or willingness to fight than their predecessors. Relatedly, Chiozza and Goemans (2004) and Gelpi

[17] Similar variation occurs in American politics, where some contests are over repeated topics, like appropriations, or regulatory battles over yearly quotas or rules, and other contests are one-off.

and Grieco (2001) argue that newly- and long-tenured leaders differ in their conflict behavior. In general, existing work on leadership changes argues that transitions are destabilizing because of the uncertainty they create. Our findings suggest an alternate mechanism: where players have gained experiential feedback about the contest and their opponents over time, players' behavior may be more consistent with Nash predictions, and they may be less inclined to over-invest in contestation.

Other topics in existing literature are qualitatively similar to the calculator treatment. For example, some scholars of crisis bargaining have analyzed the effects of intelligence and information gathering. Radtke (2016) argues that autocratic leaders receive poorer information due to cronyism amongst their advisors. Lindsey and Hobbs (2015) argue that the Presidential electoral cycle diverts leaders' attention from diplomacy, yielding worse foreign policy outcomes. The weaker treatment effect of the payoff calculator is consistent with the argument that history and experiential interactions are more relevant than the degree to which the actors invest in analyzing a particular situation, with reports, technical consultation, spying, or other information-gathering mechanisms. Although, that argument clearly comes with the caveat that laboratory subjects may process information about the payoff surface differently than elites or experts.

Our finding may also further explain divergence in behavior between undergraduate laboratory subjects and elites or politicians. The two groups may be similar in their ability to gather and process information about the payoff space, but the former group lacks the experiential learning needed for more advanced strategic thought. Elites and politicians may behave more strategically or respond to asymmetry and shocks in more strategically sophisticated ways because they have experience with a particular contest.

Finally, our findings suggest the importance of broadening our conception of how strategic thinking varies across individuals. Level-k models based on iterative reasoning capture an important facet of strategic thinking, especially in games where iterative elimination of dominated strategies yields the Nash equilibrium. But individuals also vary in their ability to understand the underlying dynamics of interactions. They vary in their understanding of comparative statics: how changes to the features of the game pertaining to themselves and their opponents affect their incentives and choices. In many games that represent real world phenomena, this novel concept of strategic heterogeneity may more accurately describe an individual's degree of strategic sophistication. Our conception of heterogeneity in strategic thinking provides new avenues for exploring the individual level characteristics that explain subjects' behavior as well as a new metric for assessing the effects of interventions and treatments on subjects' degree of strategic sophistication.

References

Abramowitz, Alan I. 1991. "Incumbency, campaign spending, and the decline of competition in US House elections." *The Journal of Politics* 53(01):34–56.

Agranov, Marina, Elizabeth Potamites, Andrew Schotter and Chloe Tergiman. 2012. "Beliefs and endogenous cognitive levels: An experimental study." *Games and Economic Behavior* 75(2):449–463.

Anderson, Lisa R and Beth A Freeborn. 2010. "Varying the intensity of competition in a multiple prize rent seeking experiment." *Public Choice* 143(1-2):237–254.

Anderson, Lisa R. and Sarah L. Stafford. 2003. "An Experimental Analysis of Rent Seeking Under Varying Competitive Conditions." *Public Choice* 115(1-2):199–216.

Arad, Ayala and Ariel Rubinstein. 2012. "The 11–20 Money Request Game: A Level- kReasoning Study." *The American Economic Review* 102(7):3561–3573.

Bassi, Anna. 2015. "Voting systems and strategic manipulation: An experimental study." *Journal of Theoretical Politics* 27(1):58–85.

Bassi, Anna and Kenneth C Williams. 2014. "Examining Monotonicity and Saliency Using Level-k Reasoning in a Voting Game." *Games* 5(1):26–52.

Bausch, Andrew W and Thomas Zeitzoff. 2014. "Citizen Information, Electoral Incentives, and Provision of Counter-Terrorism: An Experimental Approach." *Political Behavior* pp. 1–26.

Becker, Gary S. 1983. "A Theory of Competition Among Pressure Groups for Political Influence." *The Quarterly Journal of Economics* 98(3):371–400.

Besley, Timothy and Torsten Persson. 2011. "The Logic of Political Violence." *The Quarterly Journal of Economics* 126(3):1411–1445.

Bob, Clifford. 2012. *The Global Right Wing and the Clash of World Politics*. Cambridge University Press.

Camerer, Colin F., Teck-Hua Ho and Juin-Kuan Chong. 2004. "A Cognitive Hierarchy Model of Games." *Quarterly Journal of Economics* 119(3):861–898.

Chiozza, Giacomo and Hein E Goemans. 2004. "International conflict and the tenure of leaders: Is war still ex post inefficient?" *American Journal of Political Science* 48(3):604–619.

Corchon, Luis C. 2007. "The theory of contests: a survey." *Review of Economic Design* 11(2):69–100.

Crawford, Vincent P. 2003. "Lying for Strategic Advantage: Rational and Boundedly Rational Misrepresentation of Intentions." *American Economic Review* 93(1):133–149.

Crawford, Vincent P, Miguel A Costa-Gomes and Nagore Iriberri. 2013. "Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications." *Journal of Economic Literature* 51(1):5–62.

Davis, Christina and Krzysztof J Pelc. 2012. "Cooperation in Hard Times: Self-Restraint of Trade Protection." *Princeton University manuscript, June* .

Dechenaux, Emmanuel, Dan Kovenock and RomanM. Sheremeta. 2014. "A survey of experimental research on contests, all-pay auctions and tournaments." *Experimental Economics* pp. 1–61.

Deck, Cary and Roman M Sheremeta. 2012. "Fight or flight? Defending against sequential attacks in the game of Siege." *Journal of Conflict Resolution* 56(6):1069–1088.

Fischbacher, Urs. 2007. "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental economics* 10(2):171–178.

Fonseca, Miguel A. 2009. "An experimental investigation of asymmetric contests." *International Journal of Industrial Organization* 27(5):582–591.

Gelpi, Christopher and Joseph M Grieco. 2001. "Attracting Trouble Democracy, Leadership Tenure, and the Targeting of Militarized Challenges, 1918-1992." *Journal of Conflict Resolution* 45(6):794–817.

Georganas, Sotiris, Paul J. Healy and Roberto A. Weber. 2015. "On the persistence of strategic sophistication." *Journal of Economic Theory* 159:369.

Gill, David and Victoria L. Prowse. 2012*a*. "Cognitive ability and learning to play equilibrium: A level-k analysis.".

Gill, David and Victoria Prowse. 2012*b*. "A structural analysis of disappointment aversion in a real effort competition." *The American economic review* pp. 469–503.

Goldstein, Judith and Lisa L. Martin. 2000. "Legalization, Trade Liberalization, and Domestic Politics: A Cautionary Note." *International Organization* 54:603–632.

Grossman, Gene M. and Elhanan Helpman. 1994. "Protection for Sale." *The American Economic Review* 84(4):833–850.

Hafner-Burton, Emilie M, Brad L LeVeck, David G Victor and James H Fowler. 2014. "Decision maker preferences for international legal cooperation." *International Organization* 68(04):845–876.

Hafner-Burton, Emilie M, D Alex Hughes and David G Victor. 2013. "The cognitive revolution and the political psychology of elite decision making." *Perspectives on Politics* 11(02):368–386.

Henn, Christian and Brad McDonald. 2014. "Crisis Protectionism: The Observed Trade Impact." *IMF Economic Review* 62(1):77–118.

Iaryczower, Matias and Andrea Mattozzi. 2013. "On the nature of competition in alternative electoral systems." *The Journal of Politics* 75(03):743–756.

Kimbrough, Erik O, Roman M Sheremeta and Timothy W Shields. 2014. "When parity promotes peace: Resolving conflict between asymmetric agents." *Journal of Economic Behavior & Organization* 99:96–108.

Krueger, Anne O. 1974. "The Political Economy of the Rent-Seeking Society." *The American Economic Review* 64(3):pp. 291–303.

Levitt, Steven D. 1994. "Using repeat challengers to estimate the effect of campaign spending on election outcomes in the US House." *Journal of Political Economy* pp. 777–798.

Lindsey, David and William Hobbs. 2015. "Presidential Effort and International Outcomes: Evidence for an Executive Bottleneck." *The Journal of Politics* 77(4):1089–1102.

Loewen, Peter John, Kelly Hinton and Lior Sheffer. 2015. "Beauty contests and strategic voting." *Electoral Studies* 38:38–45.

Meirowitz, Adam. 2008. "Electoral contests, incumbency advantages, and campaign finance." *The Journal of Politics* 70(03):681–699.

Minozzi, William and Jonathan Woon. 2013. "Lying aversion, lobbying, and context in a strategic communication experiment." *Journal of Theoretical Politics* 25(3):309–337.

Myerson, Roger B. 2008. "The autocrat's credibility problem and foundations of the constitutional state." *American Political Science Review* 102(01):125–139.

Nagel, Rosemarie. 1995. "Unraveling in Guessing Games: An Experimental Study." *American Economic Review* 85(5):1313–1326.

Nordås, Ragnhild and Nils Petter Gleditsch. 2007. "Climate change and conflict." *Political Geography* 26(6):627–638.

Radtke, Mitchell. 2016. "When War is Bad Advice: The Effect of Ministerial Cronyism on Crisis Bargaining." Manuscript.

Ross, Michael L. 2004*a*. "How Do Natural Resources Influence Civil War? Evidence from Thirteen Cases." *International Organization* 58(01):35–67.

Ross, Michael L. 2004*b*. "What do we know about natural resources and civil war?" *Journal of peace research* 41(3):337–356.

Salehyan, Idean. 2008. "From climate change to conflict? No consensus yet." *Journal of Peace Research* 45(3):315–326.

Sell, Susan K. and Aseem Prakash. 2004. "Using Ideas Strategically: The Contest Between Business and NGO Networks in Intellectual Property Rights." *International Studies Quarterly* 48(1):143–175.

Serra, Gilles. 2010. "Polarization of what? A model of elections with endogenous valence." *The Journal of Politics* 72(02):426–437.

Slantchev, Branislav L. 2010. "Feigning Weakness." *International Organization* 64(03):357–388.

Stahl, D.O. and P.W. Wilson. 1995. "On Players' Models of Other Players: Theory and Experimental Evidence." *Games and Economic Behavior* 10(1):218–254.

Svolik, Milan W. 2009. "Power sharing and leadership dynamics in authoritarian regimes." *American Journal of Political Science* 53(2):477–494.

Tullock, Gordon. 1967. "The Welfare Cost of Tariffs, Monopolies and Theft." *Western Economic Journal* 5(3):224–232.

Welch, Susan and John R Hibbing. 1997. “The effects of charges of corruption on voting behavior in congressional elections, 1982–1990.” *The Journal of Politics* 59(01):226–239.

Wolford, Scott. 2007. “The turnover trap: New leaders, reputation, and international conflict.” *American Journal of Political Science* 51(4):772–788.