Stackelberg in the Lab: The Effect of Group Decision Making and "Cooling-off" Periods*

Eric Cardella†
University of Arizona

Ray Chiu‡
University of Arizona

April 19, 2012

Abstract

The Stackelberg duopoly is a fundamental model of sequential output competition amongst firms. The equilibrium outcome of the model results in a first mover advantage where the first-moving firm produces more output, and earns larger profits, relative to the second-moving firm. Huck, Müller, and Normann (2001) and Huck and Wallace (2002) test the Stackelberg duopoly in a lab setting and find that behavior is largely inconsistent with the equilibrium predictions of the model. We hypothesize that this inconsistency is a result of differences between the decision making environment implemented in the lab and firm environments in the field. In this paper, we experimentally investigate whether group decision making and a "cooling-off" period before the decision lead to more profit maximizing Stackelberg behavior in the lab. To do so, we re-test the Stackelberg duopoly in the lab while implementing (i) two-person decision making groups, and (ii) a 10-minute cooling-off period for second movers. Our results suggest group decision making leads to more profit maximizing behavior for first movers, while the 10-minute cooling-off period has very little effect on behavior of second movers.

Keywords: Stackelberg, Group Decision Making, Cooling-off Periods

JEL Codes: C72, C91, C92, D20, D43, L10

*We would like to thank Anna Breman, Martin Dufwenberg, Price Fishback, Tamar Kugler, Stan Reynolds, Mark Stegeman, John Wooders and seminar participants at the University of Arizona for helpful comments. We are grateful to the Economic Science Laboratory at the University of Arizona for providing financial support.

†Department of Economics, University of Arizona, McClelland Hall 401, PO Box 210108, Tucson, AZ 85721-0108, Telephone: (858) 395-6699; Email: ecardell@email.arizona.edu.

‡Department of Economics, University of Arizona, McClelland Hall 401, PO Box 210108, Tucson, AZ 85721-0108, Telephone: (520) 343-0224; Email: raychiu@email.arizona.edu.

1 Introduction

The Stackelberg model is a fundamental and frequently applied model of sequential oligopolistic output competition amongst firms. The Subgame Perfect equilibrium (SPE) outcome of the model, assuming a symmetric duopoly, is asymmetric with the first-moving firm producing a larger output level, and earning larger profits, relative to the second-moving firm; A phenomenon referred to as the first mover advantage. However, the results from previous laboratory experimental investigations of the Stackelberg duopoly are, in general, inconsistent with the SPE predictions (Huck, Müller, and Normann, 2001 (HMN henceforth); Huck and Wallace, 2002 (HW henceforth)).[1] Specifically, HMN find that second-movers fail to best respond by choosing non profit maximizing output levels. Similarly, first-movers fail to choose their profit maximizing output level, relative to both the SPE prediction and the empirical response function of second-movers.

HMN and HW cite social preferences and/or emotional motivations as the primary explanations for the observed deviations in the experimental data from the SPE outcome. In particular, the authors indicate that inequality aversion, e.g. Fehr and Schmidt (1999), would lead to lower than predicted output levels for first-movers and flatter than predicted best-response functions by second-movers.[2] HW also point toward reciprocity motivations as an alternative plausible explanation for the theoretically inconsistent best responses of second-movers.[3] How should such insights be interpreted in the context of firm decision making in the field? Would we expect firms to fail to exploit their first-mover advantage because of preferences for equality? Would we expect firms to fail to choose the profit maximizing best-response because of preferences for equality or reciprocity?

We hypothesize that one possible reason for the inconsistency between the SPE predictions of the Stackelberg model and the experimental results of HMN and HW is the differences between the lab environment and firm decision making environments in the field. In particular, we concentrate on two dimensions along which the decision

[1]Endogenous timing variation of the Stackelberg model (Hamilton and Slutsky 1990) have been also tested experimentally and the results are also, in general, inconsistent with the theoretical predictions (see Huck, Müller, and Nornamm, 2002; Fonseca, Huck and Normann, 2005; and Fonseca, Müller, and Normann, 2006).

[2]Lau and Leung (2010) re-examine the data from HMN and show that the data is consistent with a simplified version of the Fehr and Schmidt model of ineqaulity aversion. Specifically, the authors find that more than 1/3 of the subjects exhibit disadvantageous ineqaulity aversion.

[3]The authors note that second-movers “quite calmly plan to punish leaders in case they try to exploit their strategic adavantage and, at the same time, they are willing to not to exploit cooperative moves by the leader” (p. 1). Although the authors do not explicitly refer to the second-movers’ preferences for reciprocity, this explanation for the behavior of second-movers is consistent with the notion of reciprocity modeled by Dufwenberg and Kirchsteiger (2004). Hence, we refer to this type of second-mover behavior as reciprocity for the remainder of the paper.

making environment in the lab likely differs from that of firms in the field: (i) the size of the decision making unit, and (ii) the length of the deliberation period before decisions are made. Both HMN and HW use a lab environment where individual decision makers act as firms, and subjects have very little time to deliberate about their decision.[4] However, we contend that important firm decisions in the field, e.g. market output decisions, are likely to be discussed (formally or informally) and jointly decided upon by a committee of executive (see Messick, Moore, and Bazerman, 1997; Cox, 2002; Kocher and Sutter, 2005; Cox and Hayne, 2006, and Kugler, Kausel, and Kocher, 2012). Similarly, important firm decisions are likely to be carefully consider and made after some period of deliberation to ensure that the decisions are well-thought. The conventional wisdom being that over this deliberation period, "cooler heads will prevail" which will result in more "rational" decisions.[5]

The motivation of this paper is to investigate the impact of implementing group decision making and deliberation periods on Stackelberg behavior in the lab. Specifically, we test whether group decision making and a deliberation period lead to more profit maximizing Stackelberg behavior in the lab. In doing so, we experimentally re-test a Stackelberg duopoly market using a lab environment that has been augmented along one of the following two dimensions: (1) *firm* decisions are made by a 2-person unitary group, or (2) the second-moving *firm* makes its decision after a 10-minute "cooling-off" period.[6] By implementing a lab protocol that uses group decision making and cooling-off periods, we aim at creating a lab environment that comes *closer* to firm decision making environments in the field. In turn, we hypothesize that this will lead to decisions in the lab that are more in line with profit maximization and, consequently, more in line with the SPE predictions of the Stackelberg model.

Clearly, the size of the decision making unit and the length of the deliberation period between decisions do not fully exhaust the set of differences between decision making environments in the lab and firm environments in the field.[7] However, we concentrate on the impact of these two dimensions for two important reasons. First, a growing body of literature (discussed thoroughly in the subsequent section) suggests that both group decision making and cooling-off periods can mitigate the influence of social preferences and emotional motivations, which results in more *selfish* decision making. In the context of a Stackelberg duopoly, selfish decision making, which corresponds to profit maximizing decision making, would lead to behavior more in line with the SPE predictions. Second, both group decision making and cooling-off periods are protocols that can be practically implemented in a lab environment. Hence,

[4]The environment used by HMN and HW is prototypical of many lab experiments and is consistent with what Harrison and List (2004) would consider a "conventional lab experiment".

[5]Thefreedictionary.com defines the idiom "cooler head prevail" as: the ideas or influence of less emotional people prevail.

[6]We note that social psychologists often refer to 2-people as a dyad rather a group.

[7]See Harrison and List (2004), and Levitt and List (2007) for a discussion of other environmental dimesions along which the lab and the field differ.

the possibility to experimentally investigate the impact of group decision making and cooling-off periods on Stackelberg behavior in the lab.

After we began this study, we became aware of related work in progress by Müller and Tan (2011) (MT henceforth), which was conducted independently, that similarly explores topics related to group decision making in an experimental Stackelberg duopoly market. While there exists some overlap in the motivation and experimental design between their study and ours, there are important differences. Namely, MT investigate the effect of group decision making in both a one-shot and repeated Stackelberg game, while we investigate the effect of a cooling-off period in a Stackelberg game. Additionally, their design features 3-person groups, while we our study features 2-person groups. We will provide more discussion about the comparisons of the two studies, when relevant, in the corresponding sections of the paper. We ultimately view this study as complementary to that of MT, and we refer interested readers to their paper for additional insightful discussion on the impact of group decision making in Stackelberg markets.

Laboratory experiments provide a controlled environment and, as a result, are a useful research tool for gaining valuable insights regarding behavior in naturally occurring economic environments (Falk and Heckman, 2009). Such experimental insights may be of particular value when investigating industrial organization models of firm behavior (see Normann and Ruffle, 2011 for a discussion). However, one of the major concerns with lab experiments is the limited extent with which the results from the lab can be extrapolated to behavior in the field (Levitt and List, 2007). Levitt and List note that "perhaps the most fundamental question in experimental economics is whether findings from the lab are likely to provide reliable inferences outside of the laboratory" (pp. 170); A concept Levitt and List refer to as the *generalizability* of a lab experiment.[8] We assert that the generalizability of lab results relating to firm behavior is particularly tenuous due to the substantial differences in the decision making environment between the lab and the field, which include the size of the decision making unit and the length of the decision deliberation period.

Gneezy and List (2006) argue that, "before we can begin to make sound arguments that behavior observed in the lab is a good indicator of behavior in the field, we must explore whether certain dimensions of the laboratory environment might cause differences in behavior across these domains" (p. 1381). We take a first step, in relation to a Stackelberg model, by investigating the effect of group decision making and cooling-off periods in the lab. Both of which, we contend, are representative characteristics of firm decision making environments in the field. Implementing a lab environment that is *closer* to firms in the field can, in turn, help increase the generalizability of lab results (Friedman and Sunder, 1994). Although we study an

[8] Other terms have been used in reference to the extrapolation of lab results to the field, including *external validity* (Campbell and Stanley, 1963) and *parallelism* (Wilde, 1981; Smith, 1982).

experimental Stackelberg duopoly, the results from this study can provide insights regarding the effect of group decision making and cooling-off periods in other experiments that investigate models of firm specific decision making. Thus, as a broader methodological contribution, we hope this study can be informative to the design of future laboratory experiments relating to other IO models of strategic firm behavior, e.g. entry, pricing, mergers, R&D, and advertising.

This paper proceeds by discussing relevant literature in Section 2. We present the experimental design and develop testable research hypotheses in Section 3. The results are presented in Section 4, and Section 5 concludes with discussion.

2 Related Literature

2.1 Group Decision-making

The literature relating to group decision making, and the comparison with individual decision making, is extensive and spans many disciplines including economics and social psychology. Several studies have documented significant differences between group behavior and individual behavior for a wide range of decision making environments. In general, the results from these studies suggest that decisions made in groups are more self-interested compared to decisions made by individuals. In what follows, we provide a brief outline of some of the relevant literature. This is not intended as a comprehensive survey of all prior group decision making literature. Instead, we refer interested readers to Insko and Schopler (1987) or, more recently, Bornstein (2008) and Kugler et al. (2012) for more thorough reviews of the literature relating to group decision making.

Social psychologist refer to the difference between group decision making and individual decision making as the "discontinuity effect" (Brown 1954). This discontinuity effect has been extensively documented is many experimental studies. In a series of related studies, McCallum et al. (1985), Insko et al. (1987, 1988, 1990, 1994), and Schopler et al. (1991, 1993) find that groups, in general, exhibit significantly more competitive behavior than individuals in various prisoners' dilemma games. Insko et al. (1987) posit two plausible hypotheses to explain the more competitive behavior exhibited by groups. The first, "social support of self-interested competitiveness", suggests that groups provide shared support to other members in the group for acting in a self-interested manner. The second, "schema-based distrust", suggests that groups form beliefs that other groups will behave in a more self-interested manner. Consequently, because groups perceive other groups to be more self-interested , they themselves will behave in a more self-interested manner.[9]

[9]These two hypotheses are in contrast to the Social Comparison Theory referred to by Cason and Mui (1997). This theory suggests that people are motivated to present themselves to the group in

Several economic studies have similarly documented differences between group decisions and individual decisions, across a broad range of games. The results from these studies also generally find that group decisions are more self-interested than individual decisions; hence, group decisions are usually found to be closer to the standard game theoretic predictions. For example, Robert and Carnevale (1997) and Bornstein and Yaniv (1998) find that groups make significantly higher demands in an ultimatum game, relative to individuals. Bornstein et al. (2004) considered a centipede game and find that groups chose to end the game at an earlier stage, relative to individuals, in both the increasing and constant sum versions of the game. Kocher and Sutter (2005) find that groups converge to lower guesses and earn higher profits in a repeated guessing game, relative to individuals.[10] Cooper and Kagel (2005) considered several versions of a limit pricing signaling game and find that groups exhibit more "strategic" behavior than individuals. Kugler et al. (2007) and Cox (2002) compare groups decisions with individual decisions in a trust game; the former study finds that group senders send less than individuals, while the latter study finds that group receivers return less than individuals.

To summarize, most of the prior experimental studies have found that group decision making is more self-interested than individual decision making.[11] Recall, HMN and HW find Stackelberg behavior in the lab that is inconsistent with theoretical predictions, which the authors argue is in large part a result of other-regarding motivations. In light of the psychological theories posited by Insko et al. (1987) and the empirical results from the studies described above, we hypothesize that implementing group decision making in a Stackelberg game will lead to more self-interested behavior; this, in turn, will then lead to outcomes that are more consistent with the SPE. Our paper contributes to this body of experimental group decision making literature by investigating the impact of group decision making in a Stackelberg game.

2.2 Cooling-off Periods

Classical economic models often assume that agents are calm, self-interested, flawless decision makers. However, a growing body of behavioral economics research suggests

a socially desirable way. Depending on the stucture of the game and what is considered the social norm, this motivation may push the behavior of the group toward more other-regarding. We address this theory in more detail in the conclusion, and how it relates to our results.

[10] In a follow up study, Sutter (2005) finds that groups of four make significantly lower guesses, relative to 2-person groups and individuals in the guessing game. However, the author does not find a significant difference between the 2-person groups and individuals.

[11] There exists one notable exception. Cason and Mui (1997) found that group behavior was less self-interested in a dictator game. That is, groups gave significantly more as the dictator compared to individuals. However, Luhan et al. (2009) implemented a very similar design to that of Cason and Mui and found contradictory results, i.e. group dictators gave significantly less than individuals.

that psychological and emotional factors can influence decision makers and, consequently, economic outcomes (Loewenstein, 2000; Sanfey et al., 2003). Emotional influences in decision making have been explored formally using dual-system models (see Kahneman, 2003 for a review). The general idea is that human decision making is governed by an interaction between a "hot" system that responds to emotions and a "cold" system that responds to reason.

One possible way to mitigate the influence of emotions in decision making is to delay the decision, i.e. take a break and deliberate (Ury et al., 1988; Goleman, 1995; Adler et al., 1998). The idea is that delaying the decision allows time for emotional motivations to "cool-off" so that, ultimately, "cooler heads prevail". Or alternately, the deliberation period allows time for the cold system to analyze the problem and make a well thought decision. The idea that emotional motivations cool-off when agents are allowed time to deliberate has been documented experimentally. In a recent study, Oechssler et al. (2008) investigate how a cooling-off period affects rejection rates in an ultimatum game. The authors find that after a 24-hour cooling-off period, a significant number of subjects who had initially rejected an unfair offer switch and accept the offer. Similarly, Grimm and Mengel (2011) find that a 10-minute cooling-off period reduces rejection rates by about 1/2 in an ultimatum game; rejection rates fall from 80% in standard treatments to around 40-60% when the 10-minute cooling-off period for responders was implemented.

Recall, HW cite negative reciprocity as one explanation for why second-movers in their study do not best respond. The general idea behind reciprocity (Dufwenberg and Kirchsteiger 2004) is that agents are motivated to respond to the kind actions of others with kind actions (positive reciprocity), and respond to unkind actions with unkind actions (negative reciprocity), even at the expense of their own material payoff. We hypothesize that delaying the decision of second-movers, via a 10-minute cooling-off period, will mitigate reciprocal motivations, which in turn will lead to more profit maximizing behavior by second-movers. The results from this study can be viewed as complementary to the work of Oechssler et al. (2008) and Grimm and Mengel (2011), and as further contributing to our understanding of how cooling-off periods impact strategic economic decision making.

3 Experimental Design

In this section, we first describe the experimental Stackelberg duopoly around which the design is based. We then describe the experimental treatments and outline the experimental procedure. Lastly, we develop a set of testable hypotheses aimed at answering the primary the research questions of this paper. Namely, does implementing a group decision making and cooling-off period protocol lead to more profit-maximizing Stackelberg behavior in the lab?

3.1 The Model

We consider a symmetric, exogenous timing, Stackelberg duopoly with the same parameterization used by HMN, HW, and MT. In the model, there are two quantity setting firms, call them Firm A and Firm B. Let q_a and q_b denote each firm's output choices, respectively, and let $Q = q_a + q_b$ be the total market output. The market price is given by the following inverse demand function:

$$P(Q) = \max\{30 - Q, 0\} \text{ where } Q = q_a + q_b$$

and each firm faces a linear cost of production given by:

$$C_i(q_i) = 6q_i, i = a, b$$

The firms choose their quantities sequentially. Firm A (the first-mover) begins by choosing its output level, q_a. Then, after observing q_a, Firm B (the second-mover) responds by choosing its output level, q_b. The SPE is given by ($q_a = 12$, $q_b(q_a) = 12 - q_a/2$). Thus, the Stackelberg equilibrium outcome, abbreviated SE, is $q_a = 12$ and $q_b = 6$ yielding equilibrium profits of $\pi_a = 72$ and $\pi_b = 36$. The Cournot equilibrium outcome, abbreviated CE, is $q_a = q_b = 8$ yielding profits of $\pi_a = \pi_b = 64$, and the *symmetric* joint profit maximizing outcome, abbreviated JPM, is $q_a = q_b = 6$ yielding profits of $\pi_a = \pi_b = 72$.

Table 1: Discrete Stackelberg Payoff Matrix

		Firm B								
		5	6	7	8	9	10	11	12	13
	5	70,70	65,78	60,84	55,88	50,89	45,90	40,88	35,84	29,78
	6	78,65	72,72	66,77	60,80	54,81	48,80	41,77	36,72	30,65
	7	84,60	77,66	70,70	63,72	55,71	49,70	42,66	35,60	28,52
	8	88,55	80,60	72,63	64,64	56,63	48,60	40,55	32,48	24,39
Firm A	9	89,50	81,54	71,55	63,56	54,54	45,50	36,44	27,36	18,26
	10	90,45	80,48	70,49	60,48	50,45	40,40	30,33	20,24	10,13
	11	88,40	77,41	66,42	55,40	44,36	33,30	22,22	11,12	0,0
	12	84,35	72,36	60,35	48,32	36,27	24,20	12,11	0,0	-12,-13
	13	78,29	65,30	52,28	39,24	26,18	13,10	0,0	-13,-12	-26,-26

Similar to HMN, HW, and MT we use a discretized action set of the above Stackelberg game with nine possible quantity choices $q_i \in \{5, 6, 7, 8, 9, 10, 11, 12, 13\}$ for $i = a, b$. This action space allows for the possibility of the SE, CE, and JPM outcomes. Table 1 below displays the corresponding payoff matrix. Note, discretizing the action space induces multiple equilibria in this Stackelberg game. To ensure uniqueness of the Stackelberg and Cournot equilibrium, we employ the same method as HMN and HW and manipulate the payoff table slightly by subtracting one from 10 of the 162 entries.

3.2 Experimental Treatments

The experiment consisted of three treatments, which we refer to as: (i) Baseline, (ii) Group, and (iii) Cooling-off. We implement a between groups design and each subject participated in only one treatment. The three treatments are as follows:

Baseline (Treatment B) The baseline treatment involves subjects playing the discretized Stackelberg game once in either the role of Firm A or Firm B. The baseline treatment is intended to establish a baseline measure of the departures from equilibrium under an environment similar to the one used by HMN and HW, i.e., a standard lab environment protocol.

Group (Treatment G) Here we used the same setup and procedure as Treatment B, but, the decision-making units consisted of 2-person unitary groups. Each 2-person group was responsible to make one quantity decision for the group. No explicit instructions or rules were provided to govern how the group made its decision.

Cooling-off (Treatment C) Here we used the same setup and procedure as Treatment B except subjects playing the role of Firm B made their response decisions after a 10-minute "cooling-off" period.

A simple questionnaire of approximately 10 minutes in length was administered in all treatments to all subjects (a copy of the questionnaire in included in the Appendix). In Treatment B and Treatment G, the questionnaire was administered to subjects after both Firm A and Firm B had made their output decisions. However, in Treatment C, the 10 minute questionnaire was administered to subjects after Firm A's output decision had been selected and revealed to Firm B, but *before* Firm B chose its response quantity. Then, after the questionnaire was complete, subjects playing the role of Firm B were allowed to choose their response quantity. Hence, the 10-minutes spent answering the questionnaire served as the cooling-off period for subjects playing the role of Firm B in Treatment C. In all treatments, the subjects were informed in the instructions that they would be required to complete a questionnaire.

We chose to implement a questionnaire, and feel it is a suitable choice to serve as the cooling-off period, for several reasons. First, we wanted to control for the amount of time that subjects spent in the lab. Having all subjects fill out the 10-minute questionnaire ensured that all subjects from all three treatments spent approximately the same amount of total time in the lab. While altering when the questionnaire was administered still allowed us to implement a cooling-off period. Second, we strived to implement a cooling-off period that was rather innocuous and not out of the context of the experiment, in order to minimize any experimenter demand effects. A questionnaire is an appropriate choice since many experiments include a post decision questionnaire, and filling out a questionnaire would be very natural for subjects. Third, using a 10-minute questionnaire to serve as the cooling-off period is in line with the protocol implemented by Grimm and Mengel (2011).

3.3 Experimental Procedure

All experimental sessions were conducted in the Economic Science Laboratory (ESL) at the University of Arizona. The subjects were undergraduates recruited via an online database. All sessions were programmed using z-Tree (Fischbacher 2007).[12] In total, 192 subjects participated, 50 in Treatment B (25 markets), 92 in Treatment G (23 markets), and 50 in Treatment C (25 markets). Each subject participated in only one treatment, and each experimental session consisted of only one treatment. Subject payments were converted at a rate of 10:1 from the payoffs displayed in Table 1 into dollars. The average experimental earnings per subject, including a $3 show-up payment, was $9.49 and each session lasted less than 25 minutes.

Upon entering the lab, subjects in Treatment B and Treatment C were seated at individual computer carrels. In Treatment G, two subjects were randomly assigned to a carrel and these two subjects formed a unitary decision making group. Subjects in Treatment G were allowed to communicate, face-to-face, freely with their decision making partner during the decision task. Subjects were randomly assigned either the role of Firm A or Firm B. The instructions were read aloud to all subjects (A copy of the experimental instructions can be found in the Appendix). Subjects were also given the payoff matrix and made aware of the 10:1 conversion rate into dollars. In Treatment G, each of the two members of the group earned the amount that their corresponding Firm earned. That is, the payoffs in Treatment G were essentially doubled to ensure that the per subject incentives remained constant across all treatments.

After the experimenter read through the instructions, subjects in all treatments were given 2-minutes to familiarize themselves with the payoff matrix. It is possible that given a 9 X 9 payoff matrix this may not have been ample time. However, to ensure that subject had reached an adequate understanding of the structure of the

[12] We are grateful to Urs Fischbacher for providing the software for these experiments.

game and the corresponding payoff matrix, subjects were required to correctly answer two questions about the payoff matrix prior to beginning the experiment. These two understanding check questions asked subjects what the payoff would be to each firm for two different hypothetical combinations of quantity choices. These two quantity combinations were (9,7) and (7,11), and were chosen because they did not correspond to either the SE, CE, or JPM outcomes, and did not seem focal or suggestive in any way.

The framing of the decision task was in line with HMN. Namely, participants were told that they were to act as a firm that, together with another firm, produces a homogeneous product to serve the market demand. Their output decision, along with the output decision of their rival firm, would determine the profits to each firm. Because our primary research motivation is to test the Stackelberg model in an environment more consistent with that of real firms, we feel it is appropriate to include firm specific framing as part of the experimental design.

3.4 Testable Hypotheses

The motivation of this study is to investigate whether group decision making and cooling-off periods lead to decision making that is more in line with the profit maximizing Stackelberg behavior. To do so, we analyze the behavior of first-movers and second-movers separately. For second-movers, profit maximizing equilibrium behavior is rather straightforwardly characterized by best-responding to the first-mover. Thus, if group decision making leads to more profit maximizing Stackelberg behavior, then groups acting as second-movers would be more likely to best-respond. Similarly, if a cooling-off period leads to more profit maximizing Stackelberg behavior, then second-movers who have a 10-minute cooling-off period would be more likely to best-respond. This leads to the following two testable hypotheses:

H1: The response quantities of Firm Bs in Treatment G are *closer* to the best-response output quantity, compared to Treatment B.

H2: The response quantities of Firm Bs in Treatment C are *closer* to the best-response output quantity, compared to Treatment B.

For first-movers, profit maximizing Stackelberg behavior is less straightforward. Recall, that the SPE prescribed that the first-mover choose a quantity of $q_a = 12$. Clearly, choosing $q_a = 12$ would characterize profit maximizing behavior *if* second-movers choose the profit maximizing best-response. However, if second-movers do not best respond, then choosing $q_a = 12$ will likely not be the profit maximizing output level. To account for the possibility that second-movers in our experiment may not be best-responding, we characterize profit maximizing behavior for first-movers as follows: First, we estimate the aggregate empirical response function of second-movers from the observed second-mover decisions for each treatment. Second, we calculate the profit maximizing first-mover output level for each treatment, given that second-movers will respond according to the empirical response function; we refer to this as the conditional profit maximizing first-mover output level, which we henceforth denote as $\widehat{q}_a$. If group decision making leads to more profit maximizing Stackelberg behavior, then groups acting as first-movers would choose quantities that are *closer* to $\widehat{q}_a$. This leads to the following hypothesis:

H3: Firm A output choices in Treatment G are *closer* to $\widehat{q}_a$, compared to Firm A choices in Treatment B.

Because we are only imposing the 10-minute cooling-off period for second-movers (Treatment C), we do not develop any testable prediction regarding first-mover behavior for Treatment C. We hypothesize that the cooling-off period will mitigate the influence of emotional motivations, e.g., reciprocity and/or spite. Hence, our motivation to only consider the 10-minute cooling-off period for the second-movers, who are susceptible to these motivations. In general, there could exist an indirect effect of the second-mover cooling-off period that feeds back to first-mover behavior in Treatment C. Although, as we will see, the data presented in the next section reveals little support for such an indirect effect.

4 Results

We begin by looking at the breakdown of market outcomes across the three treatment. We classify the outcomes into the following 4 categories: Stackelberg Equilibrium (SE) where $q_a = 12$ and $q_b = 6$, Cournot Equilibrium (CE) where $q_a = 8$ and $q_b = 8$, the symmetric Joint Profit Maximizing (JPM) where $q_a = 6$ and $q_b = 6$, and any other outcome as (OTHER). Table 2 presents the breakdown of market outcomes by treatment, as well as the average total Stackelberg duopoly payoff per treatment.

From Table 2, we can see that of the 73 total markets, there were 2 SE outcomes, 7 CE outcomes, 2 JPM outcomes, and 62 OTHER outcomes. Comparing across treatments, there are no significant differences in any of the four types of outcomes considered in Table 2.[13] However, from Table 2 we can see that the average total payoff in Treatment G is 117.22 compared to 104.64 in Treatment B, which is significant using a Mann-Whitney test (p = 0.069). Whereas, the average total payoff of 101.80 in Treatment C is not significantly different from Treatment B (p = 0.732).

[13]Specifically, a Fisher's Exact test on each of the $4 \times 3 = 12$ pairwise comparisons across the three treatments yields p-values that are greater than 0.10.

Table 2: Market Outcomes: All Treatments

Outcome	Treatment B	Treatment G	Treatment C	Total
SE	1	0	1	2
CE	2	3	2	7
JPM	0	2	0	2
OTHER	22	18	22	62
Total	25	23	25	73
Avg Total Payoff	104.64	117.22	101.80	

It is clear from Table 2 that the data is largely inconsistent with the theoretically predicted Stackelberg Equilibrium outcome, for all three treatments. Recall, however, that deviations from the Stackelberg Equilibrium outcome can result if either the first-mover fails to choose $q_a = 12$, or the second-mover fails to best respond. We proceed by separately investigating the behavior of first-movers and second-movers across the three treatments, which allows us to test H1-H3. We will begin by first looking at the second-mover behavior.

4.1 Second Mover Response Data

The theoretical prediction of the Stackelberg duopoly is that second-movers will choose the profit maximizing response quantity. Our first two hypotheses, H1 and H2, stated that second-movers in Treatment G and Treatment C will choose quantities that are closer to the best-response quantity than second-movers in Treatment B, respectively. To test each of these hypotheses, we use the following two metrics: (i) the percentage of second-movers who chose the best-response quantity, which we denote as BR Rate, and (ii) the absolute deviation from the best-response quantity, denoted as Abs Dev from BR, which measures how far the second-movers were from the profit maximizing best-response quantity. Table 3 presents the relevant second-mover date for the each of the three treatments.

Table 3: Second Movers (SM) Response Data: All Treatments

Panel A: All Second Movers	Treatment B	Treatment G	Treatment C
Avg Quantity (q_b)	7.96	8.30	9.00
BR Rate	12/25 (48%)	10/23 (43%)	10/25 (40%)
Abs Dev from BR	1.16	1.30	1.80
Avg Payoffs	47.24	59.57	50.16
Panel B: Conditional on $q_a \leq 8$	Treatment B	Treatment G	Treatment C
BR Rate	4/8 (50%)	7/15 (47%)	5/9 (56%)
Abs Dev from BR	0.86	1.10	1.00
Panel C: Conditional on $q_a > 8$	Treatment B	Treatment G	Treatment C
BR Rate	7/16 (43%)	3/8 (38%)	5/16 (31%)
Abs Dev from BR	1.36	1.75	2.25

Notes: BR Rate was tested using a Fisher's Exact test and Abs Dev from BR and % of BR Profits were tested using a Mann-Whitney U-test. All tests were in relation to Treatment B.

Table 3 – *Panel A* presents the aggregate response data for all second-movers by treatment. Comparing the second-mover response data from Treatment B and Treatment G, we see that 10/23 (43%) best responded from Treatment G and 12/25 (48%) best responded from Treatment B, which is not significant using a 1-sided Fisher's Exact test (p = 0.755). Similarly, the average absolute deviation from the best-response quantity is 1.30 in Treatment G compared to 1.16 in Treatment B, which is also not significant using a 1-sided Mann-Whitney test (p = 0.653). Taken together, these results suggest that second-movers from Treatment G do not choose quantities

that are closer to the profit maximizing best-response, compared to second-movers from Treatment B. Hence, the data fails to support H1.

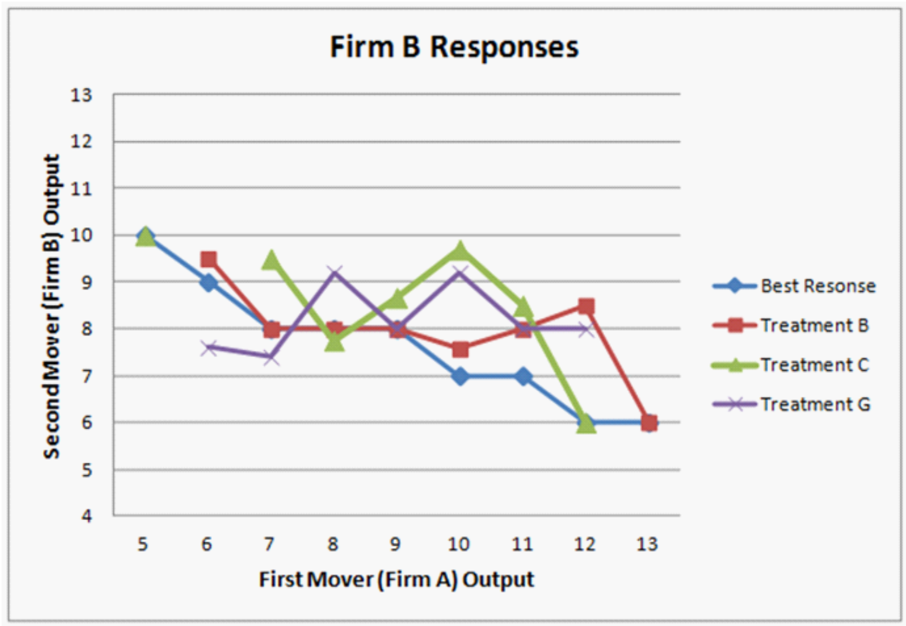
Comparing the second-mover data from Treatment C and Treatment B, we see from Table 3 – *Panel A* that 10/25 (40%) best responded in Treatment C and 12/25 (48%) in Treatment B, which is not significant using a 1-sided Fisher's Exact test (p = 0.612). Similarly, the average absolute deviation from the best-response quantity is 1.80 in Treatment G compared to 1.16 in Treatment B, which is not significant using a 1-sided Mann-Whitney test (p = 0.891). Again, these results suggest that second-movers from Treatment C do not choose quantities that are closer the best-response that Treatment B. Hence, the data fails to support H2.

Recall that the Cournot outcome is $q_a = q_b = 8$, which results in profits to each firm of 64. At all $q_a > 8$ the best-response by the second-mover would yield unequal payoffs in favor of the first-mover, and $q_a > 8$ could be viewed as an unkind action that may invoke possible negative reciprocity by the second-mover. Conversely, at all $q_a < 8$ the best-response would yield unequal payoffs in favor of the second-movers, and $q_a < 8$ could be viewed as a kind actions that may invoke positive reciprocity by the second-mover. To allow for the possibility that second-movers have different sensitivities to positive/negative inequality and positive/negative reciprocity, we look at second-mover response data in each of these two cases separately. Table 3 *Panel B* displays only the data for second-movers who's corresponding first-mover chose $q_a \leq 8$, and *Panel C* displays only the data for second-movers who's corresponding first-mover chose $q_a > 8$.

Looking at the conditional second-mover data, we see that across all treatments second-movers tend to best respond less, choose quantities further away from the best-response, and earn a lower percentage of BR profits when $q_a > 8$ compared to when $q_a \leq 8$. This would be consistent with the idea that agents have stronger sensitivities to disadvantageous inequality (Fehr and Schmidt 1999) and negative reciprocity (Charness and Rabin 2002; Dufwenberg, Smith, and Van Essen *forthcoming*). However, when we compare second-mover response data between the treatments, there are no significant differences between Treatment B and Treatments G or Treatment C, regardless of whether $q_a \leq 8$ or $q_a > 8$.

To get a better picture of the response data by Firm Bs, we plot the observed average response functions for each of the three treatments and the profit maximizing best-response function in Figure 1. From Figure 1, we can see there does not appear to be any clear differences in the empirical average response functions across the three treatments. However, one pattern that does emerge is that the response functions for each of the three treatments are *flatter* than the profit maximizing best-response function, which is consistent with the response data observed in HMN, HW, and MT. In particular, Firm Bs in all three treatments produce more than the profit maximizing best-response for $q_a > 9$.

Figure 1: Second-Mover Responses



To help quantify this, we estimate the empirical response function, which we denote $\widehat{q}_b(q_a)$, via a simple linear regression of the observed second-mover's response quantity on the first-movers quantity and a constant, i.e., $\widehat{q}_b(q_a) = \beta_0 + \beta_1 q_a$. Table 4 presents the results. Given the parameterization of the Stackelberg duopoly we consider, the theoretical best-response function is given by: $q_b^{BR}(q_a) = 12 - .5q_a$. However, from Table 4, we can see that $\widehat{q}_b(q_a) \neq q_b^{BR}(q_a)$ for each of the three treatments. Specifically, using a Wald test, we can reject the null hypotheses that $\beta_0 = 12$ and $\beta_1 = -.5$ at the 1% level for each of the three treatments, which further confirms the pattern observed in Figure 1 that the empirical response functions are flatter than profit maximizing best-response function.

To summarize, the data reveals that subjects playing the role of Firm B exhibited response decisions that was largely inconsistent with profit maximizing behavior, regardless of the treatment. In particular, Firm Bs best responded less than 50% of the time, and their estimated response functions did not coincide with the profit maximizing best-response function. In addition. the data revealed very little difference in the response data from Firm Bs in Treatment B compared to Firm Bs in Treatment G or Treatment C. Hence, the data fails to support H1 and H2 that group decision making units and a 10-minute cooling-off period will lead to Stackelberg follower decisions that are more in line with profit maximization. Our comparison of 2-person group and individual second-mover behavior is largely consistent with the comparison of 3-person group and individual second-mover behavior from MT; namely, in their 1-shot setting that most closely matches our setting, MT similarly find no significant difference between group and individual response behavior of second-movers.

Table 4: Empirical Response Function of Second Movers

	Treatment B	Treatment G	Treatment C
Response Function	$\widehat{q}_b = 9.67 - .17q_a$	$\widehat{q}_b = 6.22 + .25q_a$	$\widehat{q}_b = 10.01 - .11q_a$
p-value[a]	0.008	0.002	0.002

a: the reported p-value is from a Wald test of the joint hypothesis that $\beta_0 = 12$ and $\beta_1 = -.5$ from the regression of $\widehat{q}_b(q_a) = \beta_0 + \beta_1 q_a$

4.2 First Mover Data

We proceed by presenting the aggregate first-mover data, and then testing H3. Figure 2 displays a histogram of the Firm A output choices and the first row of Table 5 shows the average output level by treatment. From Figure 2 we can see that in all three treatments, the Firm A output choices we spread over the action space, i.e., output choices were not concentrated around the Stackelberg leader equilibrium quantity of $q_a = 12$. However, Figure 2 reveals some differences across the treatments. In particular, there were higher frequencies of $q_a \leq 10$ in Treatments G and C compared to Treatment B, while there was a higher frequency of $q_a = 12$ choices in Treatment B compared to Treatments G and C. This is confirmed in Table 5 where we see that the average Firm A output was 9.88 in Treatment B compared to 8.13 in Treatment G and 9.04 in Treatment C; the difference between Treatment B and Treatment G is significant using a Mann Whitney test (p = 0.005). However, as we previously noted, relatively higher output choices and more $q_a = 12$ choices are not necessarily consistent with more profit maximizing behavior of Firm As, as this ultimately depends on the response of Firm B.

Recall, H3 stated that Firm A output choices in Treatment G will be closer to $\widehat{q}_a$ than in Treatment B, where $\widehat{q}_a$ was the profit maximizing output quantity *given* the empirical response function of Firm Bs. We have already shown in the previous section that the response functions of Firm Bs in both Treatments B and G were not consistent with the profit maximizing best-response function (see Table 4). Therefore, $q_a = 12$ may no longer be the profit maximizing Firm A output quantity. We calculate

$\widehat{q}_a$ for each treatment, given the empirical response functions of Firm B in Table 4, which are shown in the second row of Table 5. From Table 5, we can see that choosing $\widehat{q}_a = 7.83$, $\widehat{q}_a = 7.90$, and $\widehat{q}_a = 7.41$ is optimal for first-movers, given $\widehat{q}_b(q_a)$, in Treatment B, Treatment G, and Treatment C, respectively.

Figure 2: First-Mover Output Choices

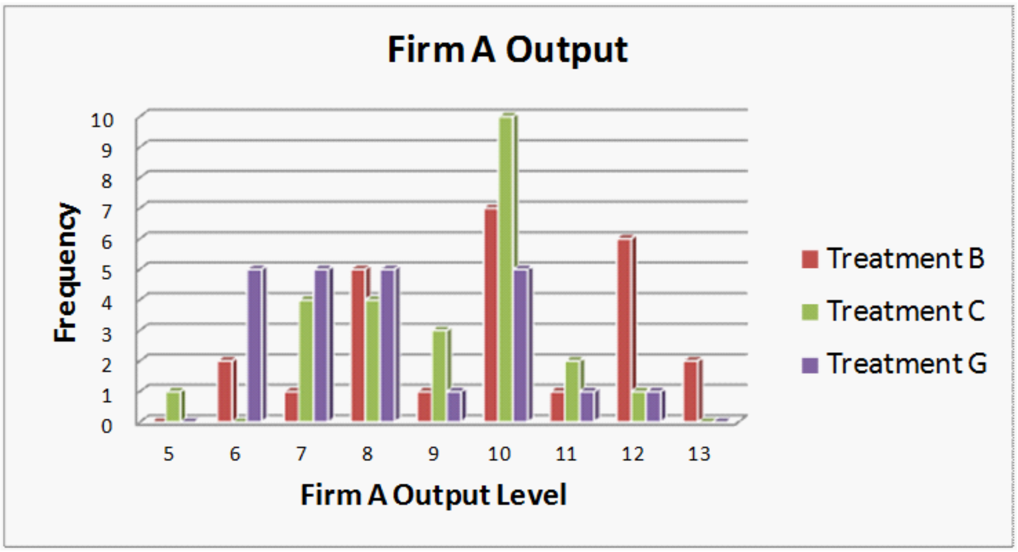


Table 5: First Movers (FM) Output Data: All Treatments

	Treatment B	Treatment G	Treatment C
Avg Output (q_a)	9.88	8.13***	9.04
$\widehat{q}_a \mid \widehat{q}_b(q_a)$	7.83	7.90	7.41
Abs Dev from $\widehat{q}_a$	2.41	1.45***	1.95

Notes: Avg Output and Abs Dev from $\widehat{q}_a$ were tested using a Mann-Whitney U-test, in relation to Treatment B
***denotes significance at the 1% level

Similar to testing H1 and H2, we test H3 using the absolute deviation from $\widehat{q}_a$, denoted Abs Dev from $\widehat{q}_a$, which measures how far away the Firm A output choices were away from the conditional profit maximizing output level. The third row of Table 5 presents the average Abs Dev from $\widehat{q}_a$ for each of the three treatments. From Table 5, we can see that the average Abs Dev from $\widehat{q}_a$ is 1.45 in Treatment G compared to 2.41 in Treatment B. This difference is significant at the 1% level using a Mann-Whitney Test (p = 0.004). Thus, the data provides support for H3, namely, first-movers in Treatment G choose output levels that are closer to the conditional profit maximizing output level, relative to first-movers in Treatment B.

To summarize, the data reveals that groups acting as Firm As choose significantly lower output levels than individuals acting as Firm As. However, after conditioning on the actual aggregate empirical response function of Firm Bs, the group Firm As choose output levels that are significantly closer to the conditional profit maximizing output level than individual Firm As. Hence, the data generally supports out hypothesis (H3) that group first-movers exhibit Stackelberg behavior that is more consistent with profit maximization. Our comparison of 2-person group and individual first-mover output choices differs somewhat from the comparison of 3-person groups and individuals in MT; namely, in their 1-shot setting that most closely matches our setting, MT find that average group first-mover choices of 9.33 and different from the average individual first-mover choices of 9.11.

5 Conclusion

The previous experimental tests of the Stackelberg duopoly model have found little support for the Subgame Perfect equilibrium predictions of the model (HMN and HW). We hypothesize that the theoretically inconsistent results from HMN and HW are a result of systematic differences between the lab environments used and naturally occurring firm environments. Additionally, these differences in the decision making environments may limit the generalizability of the lab results from HMN and HW to firm behavior in the field. The motivation of the study is twofold. First, this study experimentally investigates whether unitary decision making groups exhibit more profit maximizing Stackelberg behavior than individuals. Second, this study experimentally investigates whether a cooling-off period leads second-movers to make decisions that are more in line with the predicted profit maximizing best-response. We argue that group decision making units and cooling-off periods are both representative of decision-making units in the field, hence, our motivation to investigate their impact on decision making in an experimental Stackelberg duopoly in the lab.

In relation to the influence of group decision making, the data reveals very little difference in the decision making of second-movers between individuals and 2-person unitary groups. Specifically, both group and individual second-movers exhibit behavior that is inconsistent with profit maximizing best-response. However, group first-movers choose output levels that are significantly lower than individual first-movers. Initially, one might think that this pattern in the data reveals that group first-movers are exhibiting behavior that in more inconsistent with profit maximization. However, after conditioning on the empirical non best-response of second-movers, the data reveals that the chosen output levels of group first-movers are significantly *closer* to the predicted profit maximizing quantity than individual first-movers. Additionally, total duopoly profits are significantly larger in the markets where 2-person groups acts as firms compared to individuals.

Note, group first-movers choosing quantity levels that are lower than the predicted Stackelberg leader quantity ($q_a = 12$) is consistent with the idea that group first-movers exhibited more "clever" decision making by collectively correctly anticipated the non best-response of second-movers and, consequently, maximized profits by choosing lower quantity levels, relative to individual first-movers. However, an alternative explanation is that group first-movers where simply exhibiting more cooperative behavior, which would also correspond to choosing lower quantities. Groups exhibiting more cooperative behavior would be consistent with the findings of Cason and Mui (1997), and what they refer to as the "Social Comparison Theory" (SCT). The idea is that group members want to be perceived in a socially desirable way by other group members and, as a result, group decision making is more cooperative. However, this is contrast with the "social support of self-interested competitiveness" and the "schema-based distrust" hypotheses that predict more competitive behavior by groups. Given the body of experimental literature (discussed in Section 2) that generally finds more competitive and self-interested behavior by groups, we are inclined to think that the former is the more plausible explanation, i.e., group first-movers appear to exhibit behavior that is more consistent with rational profit maximization.

In relation to the affect of a cooling-off periods, the data reveals that the 10-minute cooling-off period did not lead to significantly different second-mover response decisions, relative to when there was no cooling-off period. In particular, second-movers with and without the cooling-off period exhibited behavior that was inconsistent with the predicted profit maximizing best-response. In fact, second-movers that had the 10-minute cooling-off period actually choose response quantities that were *further* from the best-response output level, although these differences were not significant at standard levels. One possibility is that the "cooling-off" period may have actually acted as more of a "heating-up" period for harboring emotions, rather than cooling-off emotions as it was intended to do.

We acknowledge that the observed lack of effectiveness of the cooling-off period, in the sense that it did not lead to more rational profit maximizing best responses, may be a by-product of the chosen length, 10 minutes. However, Grimm and Mengel (2011) found evidence that a 10-minute cooling-off period was long enough to have an effect on response behavior in an ultimatum game. Additionally, to the best of our knowledge, the literature related to the effect of cooling-off periods does not postulate any formal model linking the length of the cooling-off period to its effectiveness. Thus, there is no reason to suspect, a priori, that a longer cooling-off period would produce different results. Additionally, using a longer cooling-off period, e.g. 24-hours, would be cumbersome and impractical to implement in a lab, thus undermining one of our initial motivations to use a 10-minute cooling-off period. Certainly, future research that investigates the effectiveness of cooling-off periods, with respect to duration and the context of the decision making environment, is warranted.

By investigating the impact of group decision making and cooling-off periods in a Stackelberg game, this study contributes to the mature body of literature on group decision making and growing body of literature on cooling-off periods. Overall, we find that unitary groups earn higher profits, and that group first-movers in the Stackelberg game exhibit decision making that is more consistent with rational profit maximization, which suggests that group decision making might foster collective critical thought and more clever decision making in games. Whereas, we find that group second-movers exhibit decision making similar to individuals, which suggests that groups decision making may be less effective at mitigating the influence of non-selfish motivations. Additionally, a 10-minute cooling-off period did not result in second-mover decisions that were more consistent with profit maximization. As a broader methodological contribution, we hope the results of this study can provide insights regarding the design of future lab experiments that seek to investigate behavior in models of firm decision making. In particular, when testing models that require relatively higher levels of critical thought and strategic sophistication, implementing group decision making in the lab might result in behavior that is more consistent with profit maximization. However, when testing models where social preferences and emotional motivations are likely to be salient, implementing a protocol that features group decision making or a *short* cooling-off period might have little effect on behavior in the lab.

6 Appendix

6.1 Player Instructions - Baseline Treatment[14]

PLAYER INSTRUCTIONS

Welcome to our experiments! Please read these instructions carefully! Do not talk to your neighbors and please remain quiet during the entire experiment. Raise your hand if you have a question. We will answer them privately. In our experiment, you can earn different amounts of money, depending on your decisions and the decision of the other participants who are matched with you.

You play the role of a firm which produces the same product as another firm in the market. Both firms always have to make a single decision, namely the amount of output they want to produce in this market. The profit to each firm will depend on the level of output chosen by each of the firms. In the table on the other sheet that is given to you, you can see the profits of each firm for all possible output combinations of the two firms. The table reads as follows: the header of the row represents one firm's output decision (Firm-A) and the header of the column represents the output decision of the other firm (Firm-B). Inside the little box where row and column intersect, Firm-A's profit corresponding to this combination of outputs is the number to the left. Firm-B's profit corresponding to this combination of outputs is the number to the right. The profit is denoted in a fictitious unit of money which we call Taler. Before the experiment begins, you will have a few minutes to look over the payoff table. You will then be asked two control questions about the matrix to ensure your understanding of it.

You have been randomly assigned either the role of Firm-A or Firm-B, and randomly matched to another participant of the opposite role. After the two control questions, your Firm role will be revealed to you. The experiment will proceed in three stages.

Stage 1: Firm-A will begin by choosing an output level to produce. Firm-A's output level will then be revealed to Firm-B.

Stage 2: Firm-B will then respond by choosing an output level to produce.

Stage 3: The output decisions of both Firms and the corresponding profits of each Firm will be displayed to both Firms. You will then be asked to complete a simple questionnaire that will take approximately 10 minutes to complete.

[14]The instructions for the Group Treatment were essentially identical to the Baseline Treatment. Except, subjects were informed that they, along with another subject, comprise a two person group that will be playing the role of a firm.

After all three stages are complete; you will be privately paid your experimental earnings. Your profit in Talers from the decision task will be converted to $ at a rate of 10-1. That is, every 10 Talers correspond to $1 USD. In addition to your profit, you will receive a $3 USD show-up payment for participating in the experiment. All decisions and answers to the questionnaire will be kept anonymous among the participants and the experimenters.

6.2 Player Instructions - Cooling-off Treatment

PLAYER INSTRUCTIONS

Welcome to our experiments! Please read these instructions carefully! Do not talk to your neighbors and please remain quiet during the entire experiment. Raise your hand if you have a question. We will answer them privately. In our experiment, you can earn different amounts of money, depending on your decisions and the decision of the other participants who are matched with you.

You play the role of a firm which produces the same product as another firm in the market. Both firms always have to make a single decision, namely the amount of output they want to produce in this market. The profit to each firm will depend on the level of output chosen by each of the firms. In the table on the other sheet that is given to you, you can see the profits of each firm for all possible output combinations of the two firms. The table reads as follows: the header of the row represents one firm's output decision (Firm-A) and the header of the column represents the output decision of the other firm (Firm-B). Inside the little box where row and column intersect, Firm-A's profit corresponding to this combination of outputs is the number to the left. Firm-B's profit corresponding to this combination of outputs is the number to the right. The profit is denoted in a fictitious unit of money which we call Taler. Before the experiment begins, you will have a few minutes to look over the payoff table. You will then be asked two control questions about the matrix to ensure your understanding of it.

You have been randomly assigned either the role of Firm-A or Firm-B, and randomly matched to another participant of the opposite role. After the two control questions, your Firm role will be revealed to you. The experiment will proceed in three stages.

Stage 1: Firm-A will begin by choosing an output level to produce. Firm-A's output level will then be revealed to Firm-B.

Stage 2: Both Firms will then be asked to complete a simple questionnaire that will take approximately 10 minutes to complete.

Stage 3: After the questionnaire, Firm-B will then respond by choosing an output level to produce. Then output decisions of both Firms and the corresponding profits of each Firm will be displayed to both Firms.

After all three stages are complete; you will be privately paid your experimental earnings. Your profit in Talers from the decision task will be converted to $ at a rate of 10-1. That is, every 10 Talers correspond to $1 USD. In addition to your profit, you will receive a $3 USD show-up payment for participating in the experiment. All decisions and answers to the questionnaire will be kept anonymous among the participants and the experimenters.

6.3 Questionnaire

1. What is your gender?
2. How old are you?
3. What is your class level?
4. What is your major?
5. What is your approximate GPA?
6. Have you ever taken an economics course?
7. Are you currently employed?
8. Is your current job in a business related industry?
9. How many total years of work experience do you have?
10. Have you ever participated in an experiment?
11. How did you hear about the Economic Science Lab?
12. Have you ever referred a friend to the Economic Science Lab?
13. Is English your first language?
14. Are you an Arizona resident?
15. Are you currently carrying more than $10 in cash?
16. Suppose a bat and a ball cost a total of $1.10, and the bat cost $1.00 more than the ball. How much does the ball cost? (Frederick 2005; CRT #1)
17. Suppose it takes 5 machines 5 minutes to make 5 gadgets. How many minutes does it take for 100 machines to make 100 gadgets? (Frederick 2005; CRT #2)
18. In a lake, there is a patch of lillypads. Everyday the patch doubles in size. It takes 48 days for the patch to cover the entire lake. How many days does it take to cover 1/2 of the lake? (Frederick 2005; CRT #3)

References

[1] Adler, R., Rosen. B., & Silverstein, E. (1998). "Emotions in Negotiation: How To Manage Fear and Anger." Negotiation Journal 14, 161-179.

[2] Bornstein, G., Kugler, T. & Ziegelmeyer, A. (2004). "Individual and group decisions in the centipede game: are groups more 'rational' players?" Journal of Experimental Social Psychology 40, 599-605.

[3] Bornstein, G. (2008). A Classification of Games by Player Type. In A. Biel, D. Eek, T. Gärling, & M. Gustafsson (Eds.), New Issues and Paradigms in Research on Social Dilemmas, pp. 27-42. New York: Springer.

[4] Bornstein, G. & Yaniv, I. (1998). "Individual and group behavior in the ultimatum game: are groups more 'rational' players?" Experimental Economics 1, 101-108.

[5] Brown, R. (1954). Mass phenomena. In G. Lindzey (Ed.), Handbook of Social Psychology, pp. 833-876. Cambridge, MA: Addison- Wesley.

[6] Campbell, D. T., & Stanley J.C. (1963). Experimental and Quasi-Experimental Designs for Research. Boston: Houghton Mifflin.

[7] Cason, T. & Mui, V. (1997). "A laboratory study of group polarization in the team dictator game." Economic Journal 107, 1465-1483.

[8] Charness, G., & Rabin, M. (2002). "Understanding Social Preferences with Simple Tests." The Quarterly Journal of Economics 117, 817-869.

[9] Cooper, D. & Kagel, J. (2005). "Are two heads better than one? Team versus individual play in signaling games." American Economic Review 95, 477-509.

[10] Cox, J. (2002). "Trust, reciprocity, and other-regarding preferences: groups vs. individuals and males vs. females." In R. Zwick & A. Rapoport (Eds.), Advances in experimental business research, pp. 331-350. Dordrecht: Kluwer Academic Publishers.

[11] Cox, J., and Hayne, S. (2006). "Barking up the Right Tree: Are Small Groups Rational Agents?" Experimental Economics 3, 209-222.

[12] Dufwenberg, M. & Kirchsteiger, G. (2004). "A Theory of Sequential Reciprocity." Games and Economic Behavior 47, 268-298.

[13] Dufwenberg, M., Smith, A., & Van Essen, M. (forthcoming). "Hold-up: With a Vengeance." Economic Inquiry.

[14] Falk, A., Heckman, J. (2009). "Lab Experiments are a Major Source of Knowledge in the Social Sciences." Science 326, 535-538.

[15] Fehr, E. & Schmidt, K. (1999). "A Theory of Fairness, Competition, and Cooperation." The Quarterly Journal of Economics 114, 817-868.

[16] Fischbacher, U. (2007). "z-Tree, Toolbox for Readymade Economic Experiments." Experimental Economics 10, 171-178.

[17] Fonseca, M., Huck, S. & Normann, H. (2005). "Playing Cournot Although Shouldn't: Endogenous Timing in Experimental Duopolies with Asymmetric Costs." Economic Theory 25, 669-677.

[18] Fonseca, M., Muller, W. & Normann, H. (2006). "Endogenous Timing in Duopoly: Experimental Evidence." International Journal of Game Theory 34, 443-456.

[19] Frederick, S. (2005). "Cognitive Reflection and Decision Making." Journal of Economic Perspectives 19, 25-42.

[20] Friedman, D., & Sunder, S. (1994). "Experimental Methods: A Primer for Economists." Cambridge University Press.

[21] Gneezy, U., & List, J. (2006). "Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments." Econometrica 74, 1365-1384.

[22] Goleman, D. (1995). Emotional intelligence. New York: Bantam Books.

[23] Grimm, V., & Mengel, F. (2011). "Let Me Sleep on it: Delay Reduces Rejection Rates in Ultimatum Games." Economic Letters 111, 113-115.

[24] Hamilton, J. & Slutsky, S. (1990). "Endogenous Timing in Duopoly Games: Stackelberg or Cournot Equilibria." Games and Economic Behavior 2, 29-46.

[25] Harrison, G. & List, J. (2004). "Field Experiments." Journal of Economic Literature 42, 1009-1055.

[26] Huck, S., Muller, W. & Normann, H. (2001). "Stackelberg Beats Cournot: On Collusion and Efficiency in Experimental Markets." The Economic Journal 111, 749-765.

[27] Huck, S., Muller, W. & Normann, H. (2002). "To Commit or Not to Commit: Endogenous Timing in Experimental Duopoly Markets." Games and Economic Behavior 38, 240-264.

[28] Huck, S. & Wallace, B. (2002). "Reciprocal Strategies and Aspiration Levels in a Cournot-Stackelberg Experiment." Economics Bulletin 3, 1-7.

[29] Insko, C., & Schopler, J. (1987). Categorization, competition and collectivity. In C. Hendrick (Ed.), Group processes. Vol. 8, pp. 213-251. New York: Sage.

[30] Insko, C., Hoyle, R., Pinkley, R., Hong, G., Slim, R., Dalton, G., Lin, Y., Ruffin, P., Dardis, G., Berthal, P., & Schopler, J. (1988). "Individual-group discontinuity: The role of a consensus rule." Journal of Experimental Social Psychology 24, 505-19.

[31] Insko, C. A., Schopler, J., Hoyle, R., Dardis, G., & Graetz, K. (1990). "Individual-group discontinuity as a function of fear and greed." Journal of Personality and Social Psychology 58, 68-79.

[32] Insko, C., Schopler, J., Graetz, K., Drigotas, S., Currey, D., Smith, S., Brazil, D., & Bornstein, G. (1994). "Interindividual-Intergroup Discontinuity in Prisoner's Dilemma Game." The Journal of Conflict Resolution 38, 87-116.

[33] Kahneman, D. (2003). "Perspective on Judgment and Choice: Mapping Bounded Rationality." American Psychologist 58, 697-720.

[34] Kocher, M. G., & Sutter, M. (2005). "The decision maker matters: individual versus group behavior in experimental beauty-contest games." Economic Journal 115, 200–223.

[35] Kugler, T., Bornstein, G., Kocher, M. G., & Sutter, M. (2007). "Trust between individuals and groups: groups are less trusting than individuals but just as trustworthy." Journal of Economic Psychology 28, 646-657.

[36] Kugler, T., Kausel, E., & Kocher M. (2012). "Are Groups more Rational than Individuals? A Review of Interactive Decision Making in Groups." CESifo Working Paper.

[37] Lau, S. & Leung, F. (2010). "Estimating a Parsimonious Model of Inequality Aversion in Stackelberg Duopoly Experiments." Oxford Bulletin of Economics and Statistics 72, 669-686.

[38] Levitt, S., & List, J. (2007). "What do Laboratory Experiments Measuring Social Preferences Reveal About the Real World." Journal of Economic Perspectives 21, 153-174.

[39] Loewenstein, G. (2000). "Emotions in Economic Theory and Economic Behavior." The American Economic Review 90, 426-432.

[40] Luhan, W., Kocher, M., & Sutter, M. (2009). "Group Polarization in the Team Dictator Game Reconsidered." Experimental Economics 12, 26-41.

[41] McCallum, D. M., Harring, K., Gilmore, R., Drenan, S., Chase, J., Insko, C. A., & Thibaut, J. (1985). "Competition Between Groups and Between Individuals." Journal of Experimental Social Psychology 21, 301-320.

[42] Messick, D., Moore, D., & Bazerman, M. (1997). "Ultimatum Bargaining with a Group: Underestimating the Importance of the Decision Rule." Organizational Behavior and Human Decision Processes 69, 87-101.

[43] Normann, H-T., Ruffle, B. (2011). "Introduction to the Special Issue on Experiments in Industrial Organization." International Journal of Industrial Organization 29, 1-3.

[44] Oechssler, J., Roider, A., & Schmitz, P. (2008). "Cooling-Off in Negotiations – Does it Work?" Working paper.

[45] Robert, C., & Carnevale, P. (1997). "Group Choice in Ultimatum Bargaining." Organizational Behavior and Human Decision Processes 72, 256-279.

[46] Sanfey, A., Rilling, J., Aronson, J., Nystrom, L., & Cohen, J. (2003). "The Neural Basis of Economic Decision-Making in the Ultimatum Game." Science 300, 1755-1758.

[47] Schopler, J., Insko, C., Graetz, K., Drigotas, S., & Smith, V. (1991). "The Generality of the Individual-group Discontinuity Effect: Variations in Positivity-negativity of Outcomes, Players' Relative Power, and Magnitude of Outcomes." Personality and Social Psychology Bulletin 17, 612-24.

[48] Schopler, J., Insko, C., Graetz, K., Drigotas, S., Smith, V., & Dahl, K. (1993). "Individual-group Discontinuity: Further Evidence for Mediation by Fear and Greed." Personality and Social Psychology Bulletin 19, 419-31.

[49] Smith, V.L. (1982). "Microeconomic Systems as an Experimental Science." The American Economic Review 72, pp. 923-955

[50] Sutter, M. (2005). "Are four heads better than two? An experimental beauty-contest game with teams of different size." Economic Letters 88, 41-46.

[51] Ury, W., Brett, J., & Goldberg, S. (1988). "Designing and Effective Dispute Resolution System." Negotiation Journal 4, 413-431.

[52] Wilde, L. (1981). "On the use of Laboratory Experiments in Economics." In P. Joseph (Ed.), The Philosophy of Economics, pp. 137-143. Dordrecht: Reidel.