Gaming Against Managers in Incentive Systems: Experimental Results with Chinese Students and Chinese Managers

By DAVID J. COOPER, JOHN H. KAGEL, WEI LO, AND QING LIANG GU*

We examine strategic interactions between firms and planners in China, comparing behavior between: (i) students and managers with field experience with this situation, (ii) standard versus increased monetary incentives, and (iii) sessions conducted "in context," making explicit reference to interactions between planners and managers, and those without any such references. The dynamics of play are similar across treatments with play only gradually, and incompletely, converging on a pooling equilibrium. A fivefold increase in incentives significantly increases initial levels of strategic play. Games played in context generated greater levels of strategic play for managers, with minimal impact on students. (JEL D23, D8, C92)

Students of centrally planned economies have long posited that these systems are prone to a "ratchet effect" (Joseph S. Berliner, 1976; Martin L. Weitzman, 1980). Planners are never as well informed about the production capacity of any *individual* firm as are the managers of the firm. The central planner knows that firms have different production capacities, with some firms having higher productivity (lower marginal costs) than oth-

ers, but cannot identify, a priori, which firms are of which type. As a result, the central planner sets initial targets for input usage and output levels that are of broad applicability and uses the information gathered from the firm's performance to identify what is feasible in future planning periods. If firms respond naively to this initial incentive plan, high-productivity firms will outperform low-productivity firms. The central planner will then "reward" these high-productivity firms by reducing their input allocations and/or increasing their required output levels.[1] However, firms should anticipate this ratcheting up of targets, cutting their production in an attempt to mimic low-productivity types. This reduction of output with high-productivity firms imitating low-productivity types is known as the "ratchet effect."

Potentially, the ratchet effect can seriously reduce production in centrally planned economies. Moreover, the ratchet effect is not just limited to planned economies. Similar strategic interactions arise in any principal–agent interaction where the principal is uninformed about the agent's type and cannot (or will not) commit to a long-term contract. Examples include environmental regulation (Dennis A. Yao, 1988), job evaluation and incentive schemes (Barry W. Ickes and Larry Samuelson, 1987; James Dearden et al., 1990), piece-work wage

[1] The Chinese have a rather colorful term for this— "whipping the fast ox."

rate schedules (Robert Gibbons, 1987), regulation of natural monopolies (Michael A. Crew, 1994), and procurement contracting (Jean-Jacques Laffont and Jean Tirole, 1993).

In this paper, we report an experiment which addresses several open questions about the development of the ratchet effect in centrally planned economies and, more generally, about the evolution of strategic play in games. We study a game that captures the essential features of the strategic interactions between planners and firms in a centrally planned economy. In all equilibria of this game, a ratchet effect emerges with high-productivity firms imitating low-productivity types, thereby resulting in a pooling equilibrium.

We compare the dynamics of adjustment to equilibrium between: (i) "standard" versus substantially increased monetary incentives, (ii) the "standard" experimental subject population (students) and an "expert" subject population with extensive field experience with the ratchet effect (managers and white-collar workers in state enterprises in the People's Republic of China), and (iii) experimental sessions conducted "in context" versus a "generic" setting. The latter is no doubt the most novel of our experimental treatment conditions, and therefore warrants further explanation.

For sessions played in context, the language is deliberately designed to relate to subjects' field experiences with similar games. For example, we refer to subjects as "planners" and "managers" with "planners deciding what kind of production target to assign to managers— easy or tough." To the extent that the laboratory game and its equilibria are similar to experiences outside the laboratory, the use of context *may* make subjects more sensitive to the strategic implications of the game, facilitating convergence to the pooling equilibrium. In contrast, in generic sessions, the language is kept deliberately abstract with firm managers referred to as "A" players and planners referred to as "B" players, with B players deciding between "X" and "Y" (tough or easy contracts) in response to A players' "choices." The generic language is intended to minimize any relation to subjects' field experiences and *may* be expected to reduce sensitivity to the strategic implications of the game, retarding convergence to a pooling equilibrium.

Several conclusions can be drawn from our experimental data. At a general level, in all versions of the game, and with all subject populations, there is a strong initial tendency for high-productivity firms to ignore the strategic implications of the ratchet effect. Only gradually (and incompletely) does play converge to a pooling equilibrium. This replicates results from other signaling game experiments (Jordi Brandts and Charles A. Holt, 1992; Zeinab Partow and Andrew Schotter, 1993; Cooper et al., 1997a, b), indicating that this general pattern of play is unaffected by culture, subject population, and incentives.

This general pattern of play also suggests an answer to an important question about the ratchet effect: Why should the ratchet effect ever be observed in field settings? If high-productivity firms anticipate that planners will ratchet up their production targets and therefore imitate low-productivity types, central planners should also anticipate firms' mimicry and select incentive schemes that do not elicit this strategy (for example, precommiting to a long-term contract). Our results show that incentive schemes that eventually produce the ratchet effect may actually work quite well *initially* for the central planners, with inefficiencies arising only gradually over time as high-productivity firms learn to imitate low-productivity types. It has been suggested that, "One possibility [why a planner might not precommit to a long-term contract], of course, is that the central planner is not aware of the benefits of commitment" (Xavier Freixas et al., 1985). Our results suggest a different explanation: the planner may not precommit *because initially there are minimal benefits to commitment.*

Looking at the detailed pattern of play, several interesting observations emerge. First, increasing payoffs for People's Republic of China (PRC) students induced substantially higher pooling rates by high-productivity firms in early plays of the game (students in standard-pay sessions only caught up after 20 or more repetitions of the game). That is, more money induced heightened initial levels of strategic play, suggesting that money can serve as a partial substitute for experience. This provides some of the first unambiguous evidence that increasing payoffs leads to increased strategic behavior, consistent with the intuition of most theorists,

and a result of fundamental importance for the study of learning in games.

Second, context facilitated the development of strategic play among PRC managers, but had minimal effect on PRC students. PRC managers were affected both in their roles as planners and as firms, with the impact on their role as planners more pronounced than as firms. This result provides strong (albeit, indirect) evidence that PRC firms have been subject to a ratcheting up of production targets in response to superior performance.[2] The fact that context had a much stronger effect on PRC managers than on students suggests that context must be eliciting something from managers' experiences *as managers*. It is difficult to see what this experience could be other than experience with the ratchet effect.[3]

Finally, contrary to our expectations, PRC students in their role as firms exhibited significantly higher initial levels of strategic play than older, more experienced PRC managers. Subsequent manipulations indicate that this surprising result was due to the age and/or lower educational levels of older managers. This serves as a warning to experimenters as they expand their studies to an ever larger array of subject populations: although experimenters are used to assuming that subjects can make elementary calculations, follow simple chains of logic, etc., facility with such "test-taking" skills may not extend to older and/or less-educated subject populations.

The plan of the paper is as follows. Section I briefly outlines the theory underlying the ratchet effect. Drawing heavily on the cognitive psychology literature, we present the theoretical

underpinnings of our treatments in Section II. This section also outlines the procedures used in running the experiments. Section III presents our experimental results. Section IV summarizes our results and offers concluding remarks.

I. The Ratchet Effect Game

Payoff tables were generated to capture the essential strategic interactions of the ratchet effect game in the simplest possible setting. Generating the ratchet effect requires two separate insights from players. First, central planners (CPs) must realize that they can use the information revealed by firms' (Fs') choices in resetting targets. Second, Fs must realize that CPs will use any information they reveal *and* that they should use a pooling strategy to avoid revealing harmful information. Our game therefore concentrates on Fs' output decisions and the CPs' responses. The Appendix to our working paper (Cooper et al., 1998) gives full details on the underlying model of the ratchet effect, its relationship to Chinese institutions, its relationship to the laboratory game, and how the payoff tables were generated from the model.

The payoff tables are based on a simplified version of the model in Freixas et al. There are two types of Fs, high productivity (FH) and low productivity (FL). The game begins with the firm learning its type. The firm then chooses its first-period output, while facing a preexisting target.[4] After observing first-period output (but not the firm's type), the planner sets a target for the second period. The game ends with the firm choosing its second-period output. The firm attempts to maximize the discounted sum of its profits. The planner's objective function is the total surplus generated by the firm minus a penalty proportional to the firm's profits.[5]

In our laboratory game, the firm's second-period output decision is taken as given—this is a straight maximization problem with no interesting strategic elements. The remaining game is a standard signaling game with the firm acting as the sender (the first-period output is its

---

[2] Not withstanding theorists' concern with the ratchet effect, empirically verifying its existence in centrally planned economies has been subject to considerable controversy (see the exchange between David Granick [1980, 1983] and Michael Keren [1982]).

[3] Psychologists' research on learning generalizability indicates that prior experience is typically *not* called on in solving a new, related problem unless subjects have been explicitly told to consider the similarities between the two problems, have been trained to identify similarities between problems, and/or the problems are quite similar (Mary L. Gick and Keith J. Holyoak, 1980; David N. Perkins and Gabriel Salomon, 1988; Salomon and Perkins, 1989). Given that we get substantial transfer with context, this suggests that the problem PRC managers face in the ratchet effect game must be quite similar to their experiences in the field.

[4] In Freixas et al. the game begins with the period 1 target being set by the CP. We eliminate this step to simplify the game while concentrating on the ratchet effect.

[5] Objective functions of this form are standard in the ratchet effect literature.

TABLE 1—FIRM PAYOFF AND PLAYER PAYOFF

	Firm payoff				
	Low-productivity firm Production target		High-productivity firm Production target		
Output	EASY	TOUGH	EASY	TOUGH	Output
1	710	542	1,108	815	1
2	730	561	1,145	852	2
3	697	528	1,230	937	3
4	527	357	1,308	1,015	4
5	273	103	1,328	1,035	5
6	220	48	1,298	1,005	6
7	190	15	1,250	966	7

	Planner payoff	
Production target	Facing low-productivity firm	Facing high-productivity firm
EASY	645	764
TOUGH	528	820

message) and the planner acting as the receiver (the second-period target is its action).[6] The timing of the laboratory game is as follows:

1. Prior to the start of the game Fs learn their productivity level, FH or FL (it was common knowledge that the prior probability of each type was 50 percent).
2. Fs choose among several different output levels. An F's payoff is a function of F's type, the output level chosen, and the CP's actions (see Table 1). For any given output level, payoffs are higher if the CP responds with an EASY rather than a TOUGH production target. If an F did not anticipate that his choices affected the CP's response, he would choose 2 as an FL and 5 as an FH; we refer to these as Fs' full information output levels.
3. CPs see the output level of the F with whom they have been paired, but not F's type. CPs choose between a TOUGH or EASY production target. The CP's payoffs are a function of F's type and the target chosen. Payoffs are

greater for choosing EASY than TOUGH when facing an FL type and greater for choosing TOUGH than EASY when facing an FH type.
4. Fs learn CPs' choices and CPs learn Fs' type. Payoffs are then determined.[7]

The game has three possible pure strategy sequential equilibria which are pooling at output levels 1, 2, or 3. All of these equilibria have a ratchet effect, with FHs imitating FLs in efforts to receive an easy target. To see why these are equilibria, consider the pooling equilibrium at output level 2. Since both types of F choose the same output level, the CP must believe each type is equally likely following play of 2. Given these beliefs, the CP maximizes her expected payoff by choosing EASY (expected value of 704.5 versus 674 for TOUGH given the 50-50 prior distribution of F's type). For all other output levels, there are no constraints on the CP's possible beliefs. The out-of-equilibrium beliefs that support pooling at 2 are that any other output level must come from an FH type, so that the CP assigns a TOUGH contract. Given that the CP assigns EASY following 2 and TOUGH following all other output levels,

[6] Ananish Chaudhuri (1996) investigates a full-blown two-period ratchet effect game with a unique pure strategy separating equilibrium using U.S. students as subjects. Sessions had 10–12 plays of the game which, given our experience, is far too few to observe reliable steady-state results even in a one-period signalling game.

[7] As noted above, we eliminate F's end-period response to the target. Instead we impose the optimal response, adjusting the payoffs accordingly.

neither type of F can gain by deviating from 2. The pure strategy pooling equilibria at output levels 1 and 3 are supported by similar strategies and beliefs.[8]

The set of equilibria can be narrowed down by using the equilibrium refinements which have been proposed for signalling games. Unfortunately, most of the standard refinements do not yield a unique outcome. For example, the intuitive criterion (In-Koo Cho and David M. Kreps, 1987) leaves both pooling at 1 and 2. Only perfect sequential equilibrium (Sanford J. Grossman and Motty Perry, 1986), an extremely strong refinement, is able to make a unique prediction, pooling at 2.

There is considerable experimental research indicating that none of the standard equilibrium refinements adequately organizes data from signalling games (see Brandts and Holt, 1992; Partow and Schotter, 1993; Cooper et al., 1997a). As an alternative, we ran simulations using the simple adaptive learning model developed in Cooper et al. (1997b). This is essentially a multiplayer version of fictitious play with initial beliefs fitted from the data. Cooper et al. (1997a, b) demonstrate that this model does a good job of characterizing play in signaling game experiments similar to the ratchet effect game. For our game, simulations of the adaptive learning model overwhelmingly converge to the pooling equilibrium at 2. (This result holds whether initial beliefs are fitted from the data or flat priors are imposed.) Convergence to this pooling equilibrium follows a typical pattern of play: simulated players' early beliefs fail to adequately account for the relationship between Fs' choices and CPs' responses, so that Fs' initial choices are clustered around their full information output levels (2 for FLs and 5 for FHs). These initial choices induce CPs to give the easy contract for output 2 and the tough contract for higher output levels. As FHs learn that the likelihood of an easy contract is much higher at 2 than at higher output levels, play converges on a pooling equilibrium at 2.

II. Experimental Treatments, Research Hypotheses, and Procedures

The motivation for our experimental treatments, as well as the predicted impact of these treatments, can be best summarized within the framework of a belief-based learning model such as the modified fictitious play model of Cooper et al. (1997a, b). Play in this model consists of two distinct stages. Prior to any actual play, agents must form initial beliefs about the likely actions of their opponents. As play takes place, players modify their initial beliefs in response to the observed play of others and choose actions to maximize their payoffs subject to their updated beliefs. We propose that incentives, expertise, and context can affect agents' initial beliefs and/or the evolution of these beliefs over time in the following ways.

A. Incentives

Subjects who put themselves into the other players' shoes in forming initial beliefs have a substantial advantage over players who simply assume a distribution (such as a uniform distribution over all strategies). However, such deductive reasoning involves some effort on players' part which, ceteris paribus, they prefer to avoid.[9] Thus, the amount of effort players expend on reasoning about the game before forming initial beliefs should be sensitive to the payoffs resulting from their actions. Since increased deductive reasoning should lead to more strategic play, increasing the payoffs should, if anything, increase the initial levels of strategic play.

While there are strong theoretical reasons to expect increased incentives to yield more strategic play, past experiments reveal that incentives do not typically play as strong a role in behavior as anticipated. There are numerous reports of experimental phenomena that are essentially unaffected by incentives (see the many examples cited in Colin Camerer, 1995). Further, when incentives have impacted on behavior, they typically reduce the variance of

---

[8] This game has no pure strategy separating equilibria. There do exist mixed strategy equilibria; these equilibria require all FHs to act strategically, choosing lower output than their full information output.

[9] Our logic in this section is analogous to the standard analysis of the allocation of time between work and leisure (Gary S. Becker, 1965).

outcomes around the mean, with little or no impact on mean performance (see Vernon L. Smith and James M. Walker [1993b] for a review of the literature). In cases where mean performance is altered, the observed changes can usually be attributed to risk aversion (for example, Yoram Kroll et al., 1988; Steven J. Kachelmeier and Mohamed Shehata, 1992; Smith and Walker, 1993a).

Thus, we did not originally expect to find an incentive effect which could not be explained by reduced variance in outcomes or increased risk aversion. However, as will be seen in Section III, subsection B, such an effect did emerge. We discuss why the ratchet effect game is particularly likely to be responsive to incentives following documentation of these effects.

B. Expertise and Context

There are strong reasons to presume an interaction effect between expertise (as reflected in managers past experience with the ratchet effect) and the context treatment. Psychologists have done extensive work analyzing differences in performance and cognitive processes between experts and novices (see Robert Glaser [1990] and Colleen Zeitz and Glaser [1994] for summaries of the many differences between experts and novices identified to date). Experts typically employ more sophisticated approaches in solving problems and see larger patterns in data than novices. For example, in classic experiments involving chess pieces arranged as they might be in the course of an actual game, experts recalled the position of many more pieces than novices (William G. Chase and Herbert A. Simon, 1973; Adrianus D. DeGroot, 1978). This greater recall was not based on greater memory capacity but rather on the ability to organize groups of pieces into subpatterns which could then be recalled as a chunk (e.g., a castled king, a chain of pawns). As a result, when pieces were arranged on the board in random positions, experts were no longer able to recall the positions with any greater accuracy than the novices. Experts also tend to approach problems in terms of basic principles, rather than focusing on surface details as novices tend to do. One classic example involves the approach to simple physics problems (D. P. Simon and Simon, 1978; Michelene T. Chi et al.,

1981). Novices tended to classify problems on the basis of superficial details (problems with balls or problems with springs) while experts concentrated on the physical principles involved (Newton's Second Law or conservation of momentum).

However, expertise is narrowly defined and only likely to be transferred when situations are perceived to be quite similar (Gick and Holyoak, 1980; Perkins and Solomon, 1988; Solomon and Perkins, 1989). For example, several studies have examined the ability of students trained in computer programming to solve other logical problems (Roy D. Pea and D. M. Kurland, 1984; Kurland et al., 1986; Perkins et al., 1988). Even though these problems drew upon skills used in computer programming, such as planning ahead or breaking up problems into manageable pieces, little transfer was found, except when generalizability was stressed in the training (David Klahr and Sharon M. Carver, 1988).

Context plays an important role in fostering transfer because of its impact on perceptions of similarity. Information and knowledge are learned within a specific context. Even when context is not directly relevant to what is being learned, it can serve as a trigger for recall. In one dramatic example, Duncan R. Godden and Alan D. Baddeley (1975) had divers learn lists of words either on dry land or under 20 feet of ocean water. One day later, subjects were randomly assigned to one of these two environments and asked to recall the lists of words. Subjects in the same environment where they had learned the list were able to recall 59 percent of the words, while subjects in a different environment only recalled 46 percent of the list!

Turning to our experiment, we expect that managers, based on their past experiences, would be expert in recognizing and dealing with the ratchet effect. However, this expertise can only have an effect if managers recognize the similarity between our laboratory game and their field experiences. Context can be expected to play an important role in triggering this connection, resulting in heightened levels of strategic play to begin with and/or more rapid adjustments towards equilibrium compared to managers playing the generic version of the same game.

However, the existence of a context effect

requires two necessary conditions: (1) the existence of a ratchet effect in field settings, and (2) the ability of subjects to transfer knowledge from field settings to a laboratory experiment. Neither of these points can be taken for granted. As noted previously, there exists no general consensus that CPs actually ratcheted up quotas (see footnote 1), and psychologists have found that transfer tends to be narrow in scope. As such, it cannot be assumed *ex ante* that a context effect will exist for our manager subjects. Consequently, if such an effect exists, it suggests that the strategic situation in the field must be quite close to the model implemented in the lab.

Unlike managers, students do not have the relevant field experience to become expert in the ratchet effect game. However, to the extent that central planning and the ratchet effect permeate Chinese culture, there may exist a context effect within the student subject population.[10] But we expect that any observed context effect will be larger for managers than for students because of their greater experience in dealing with the problem.

Along similar lines, our prior beliefs were that managers would exhibit at least as much (if not more) strategic behavior than students, particularly if we confined our attention to context sessions. What we did not anticipate was that recent experience as a student provides a significant advantage in our experiment for reasons not directly related to understanding the strategic problem. The experiment involved figuring out a relatively complex payoff table and a fair amount of record keeping. Students are used to dealing with arcane computational tasks designed by professors and did not seem bothered by these details. In contrast, many of the older managers were having so much trouble with the paperwork that it is hard to imagine they were even thinking about the actual game. On a more mundane level, bifocals are rare in China. This meant that many of our older managers had to use glasses to read the blackboard but removed their glasses to complete their record sheets. They were so busy taking their glasses on and off that they scarcely had time to think about the

game. Although we made some effort to control for these problems, the education effect clearly makes it difficult to compare the absolute levels of strategic play between students and managers.

C. *Summary of Research Hypotheses and Implementation*

To summarize, *ex ante* we had three general hypotheses to test: (1) Increased incentives lead to increased strategic play. Based on past experimental outcomes, we anticipated at best a weak effect. (2) Context promotes heightened levels of strategic play among managers compared to generic sessions. Necessary conditions for this effect are the existence of the ratchet effect in field settings and similarity, actual and perceived, between managers' field experiences and our laboratory game. By the same token, context should have a much less pronounced (or nonexistent) effect for PRC students. (3) Field experience with the ratchet effect results in heightened levels of strategic play for managers compared to students, particularly in context sessions. *Ex post,* we discovered an education effect among older managers which weakens our last prediction. Our experimental design tests these three predictions and attempts to control for the education effect.

Differing payoff levels were used to determine the effect of incentives on strategic play. All sessions were conducted with payoffs in an experimental currency called "francs." In standard-pay PRC student sessions, francs were converted into yuan so that, for the pooling equilibrium at 2, the expected value of subject earnings was 30 yuan (approximately \$3.75 at the official exchange rate). These earnings were thought to be the minimum necessary to insure a good flow of recruits in Shanghai (which is a high-wage area in China) and to be reasonably consistent, in terms of purchasing power and outside income, to earnings in typical U.S. student experiments. For high-pay PRC student sessions, payoffs were scaled up by a factor of five, giving an average payoff of 150 yuan for the pooling equilibrium at 2. For a college student in the PRC this was a sizable amount of money. As points of comparison, the monthly wage was about 550 yuan for an instructor and 1,200 yuan for an associate professor at China

---

[10] Context may also improve performance as it can aid comprehension by illustrating the meaning of otherwise abstract concepts.

Textile University during the time period our experiments were run, and the vast majority of our manager sample earned less than 2,000 yuan per month.[11]

To test for the effects of expertise, we used three subject populations: college students, older managers, and younger managers. The college students were from China Textile University, which is considered a high-quality university in China. (Less than 0.7 percent of China's population attend any university, and China Textile University is considered a top 20 university.) Few of these students would have had any direct experience with the ratchet effect.

Whenever possible all workers in China, including foremen, are subject to production plans. Therefore, all of the managers in our experiment should have had extensive field experience with the strategic situation our ratchet effect game was designed to capture. Older manager sessions employed relatively high-ranking managers (some of whom were participating in a two-year, part-time M.B.A-style program at China Textile University), mid-level managers, and senior foremen from textile factories operating in Shanghai. About 25 percent of the older managers had university degrees; another 48 percent had some post-high-school education, typically a two- or three-year degree similar to community college or trade school in the United States. Among those who had post-high-school training, most had not been students any time recently as 40 percent were over the age of 50 and another 40 percent were between 40–50 years old. Moreover, age is inversely correlated with quality of formal education in China (even for the college educated). These factors probably all contributed to a substantial education effect.[12]

Under our original procedures, older PRC managers started out with substantially less stra-

tegic play than the other subject groups and never caught up. To compensate for this education effect, we simplified the procedures in two ways. First, for each output level which was actually selected by an F, the experimenter announced the Fs' payoffs conditional on the CPs' responses and the payoffs of both CPs and Fs for these choices, so that subjects no longer had to compute their own payoffs. Second, forms for recording past play were eliminated, with the entire history posted in several places (in extremely large print), making it easier for subjects to see and think about the outcomes.[13]

In addition, we recruited younger managers who were all university educated and under age 30.[14] This group included graduate students at China Textile University who had spent at least five years in factories before returning to the university to obtain a graduate (M.B.A-type) degree and China Textile University alumni working in the area, with at least two years of experience. Thus, all of these subjects should have had some experience as managers confronting the ratchet effect. Because all of these managers were relatively recent college graduates, we would not expect them to be subject to the education effect.

To capture the interaction between expertise and context, we ran approximately half of the sessions as generic sessions and half as context sessions. Generic sessions employed the same payoffs as the context sessions. The only difference was in the instructions. The instructions for generic sessions used abstract language designed to minimize the similarity between the experiment and field settings. In context

---

[11] The conversion rate was 0.0026 yuan = 1 franc after subtracting 18,000 francs from total earnings. This total included subjects' show-up fee. All sessions were conducted in May–June, 1995.

[12] All of the PRC manager sessions employed the same conversion rate as the high-pay student sessions. Given the higher incomes of the managers (as compared to students), these sessions might best be viewed as standard-pay sessions.

[13] Under our original procedures, Fs' payoff table included two additional output choices attached to either the beginning or end of Table 1. These choices were designed to shift the precise location of the equilibrium between sessions (to offset possible information leakage across replications of student sessions). These two choices were dominated by one or more of the outputs in Table 1, so they do not alter the equilibrium predictions. These additional choices were eliminated in the modified treatments.

[14] University degree holders in China can become foremen after six months of training. After one or two years, they will normally be promoted to a junior manager position or be in charge of a department in the firm. Average earnings for the young managers were substantially higher than for the older managers, as many worked in joint venture enterprises involving some private capital and some state enterprise or government capital.

sessions, the instructions deliberately used language emphasizing similarities between the experiment and field settings. For example, the terms "easy contract" and "tough contract" in context sessions were replaced with X and Y in generic sessions, "output" was replaced with "choice," "high-productivity firms" were replaced with "A1" players, etc.[15]

Table 2 summarizes our experimental treatments.

D. *Procedures*

Each experimental session employed between 12 and 16 subjects. Subjects were randomly divided into two groups, Fs and CPs. A total of 36 games were played in each session, with subjects switching roles after every 6 games. Within each 6 plays of the game a given CP was matched with a given F once, and only once, with different random matchings within each 6 plays of the game. Subject identification numbers were always suppressed so that subjects never knew the identify of the person with whom they were paired.

Each subject was given a copy of the instructions including the payoff tables for both CPs and Fs. The instructions were also read out loud to all subjects. Any substantive questions were repeated and answered to the entire group to ensure common knowledge. Before playing, subjects were required to fill out a short questionnaire to insure they understood the payoffs associated with different actions (instructions are available from the authors upon request). The experiments were conducted manually with barriers separating CPs and Fs so that they could not see each other.

Prior to each play of the game one of the experimenters, standing on the Fs' side of the room, flipped a coin to decide Fs' productivity level for that game. The outcome of the coin flip was shown to all Fs who then chose an output level. Fs recorded their outputs on their subject record sheets and on a "communication" sheet,

TABLE 2—EXPERIMENTAL TREATMENTS

Game structure	Subject population	
	PRC students[a]	PRC managers[a]
Context	Standard pay (3) High pay (2)	Older managers Original procedures (3) Modified procedures (2) Young managers (1)
Generic	Standard pay (3) High pay (2)	Older managers Original procedures (2) Modified procedures (2) Young managers (1)

[a] Numbers in parentheses indicate number of experimental sessions.

which was given to the CP with whom they were paired for that period.

CPs recorded Fs' outputs and chose between "tough" and "easy" production plans. These choices were entered on the communication sheet, which was then passed back to the other side of the room. Fs' type was publicly announced and subjects recorded their payoffs for that play of the game.[16]

Outcomes following each play of the game were posted, in a matrix format, for all pairings on a blackboard along with Fs' type. Thus, subjects knew the overall choice frequencies for both Fs (by type) and CPs. Subjects were encouraged to keep their own records of past play on forms provided (which virtually all subjects did).

We conducted several "standard-pay" sessions in the United States, both to shake down our procedures and as a potential control condition (for cultural effects) in case we observed a strong context effect among PRC students. We did not observe the latter and found no significant differences between standard-pay PRC sessions and the U.S. sessions. An earlier version of this paper reporting the results of these U.S. sessions is available upon request.

---

[15] The order of payoffs for choices were inverted in generic sessions relative to the context games, so that for FHs (A1 players) to mimic FLs (A2 players) they would have to choose higher rather than lower numbers. This was done to further disguise the relationship between the laboratory game and field experience.

[16] Our assistants monitored the recording of payoffs (to prevent cheating) and the communication sheets (to insure that no additional information was sent).

III. Experimental Results

A. General Patterns of Play over Time

Figures 1 and 2 depict the development of play over time. These figures show pooled results for standard-pay PRC students and for older managers, modified procedures. Separate graphs are shown for generic and context treatments in both cases. These sessions are representative of the overall data set, since the temporal pattern of play is similar across treatments. Detailed data for all treatments can be found in Tables 3–5. Data are pooled for each 12 plays of the game, so that each subject has switched roles, playing 6 times as an F and 6 times as a CP. The histograms show Fs' choices; tough target rates are shown just below each output level.

CONCLUSION 1: *The learning/adjustment process was similar across all treatments. Initially (games 1–12), Fs' choices were clustered around their type's full information output levels. CPs responded by giving far more easy contracts for output levels 1–3 than for higher output levels. Experience generated increased levels of strategic play by FHs (choice of output levels 1–3), with play converging on a pooling equilibrium at 2. However, in all treatments there was a sizable frequency of nonstrategic play by FHs (output levels 4–7) even in the last 12 games.*

In the first 12 plays of the game (rounds), full information output levels were the modal choice for both F types across all treatments. In early rounds strategic behavior by FHs, defined as a choice of output levels 1–3, often involved choices of 3 more frequently than choices of 2.[17] CPs' choices suggest they did a far better job initially of anticipating Fs' action than vice versa. Even in early rounds, the vast majority of CPs reliably chose EASY following play of output levels 1–3 and TOUGH following higher outputs, best responses for these choices coming from FLs and FHs respectively. With one

notable exception (older managers, original procedures, generic instructions), the difference in the percentage of tough contracts selected by CPs at 2 versus 5 was large enough to support pooling at 2 by FHs compared to the nonstrategic choice of 5. More often than not, 3 was also more profitable than 5 in these initial plays of the game.[18] In spite of this, FHs overwhelmingly chose 5 initially. In what follows (Section III, subsection B) we will discuss why CPs may have anticipated Fs' actions in early rounds, while the opposite did not hold.

Experience generated increased levels of strategic play in games 13–24, with it not being uncommon for pooling rates to have nearly doubled for FHs. The frequency of strategic play continued to grow in games 25–36, albeit at a slower pace. Strategic play by FHs, particularly pooling at 2, was supported by CPs' choices throughout. FLs' choices typically became more concentrated at 2 over time, consistent with the emergence of a pooling equilibrium at 2.

In all treatments, there was a sizable frequency of nonstrategic play by FHs, reflected in the fact that the overall pooling rate averaged 53.1 percent even in the last 12 plays of the game. In contrast, a substantially higher percentage of FHs tried pooling at least once (85.9 percent of all subjects). This failure to continue to pool, once having tried it, no doubt reflects an important stochastic component to choice, as well as responses to short-run variations in the relative profitability of pooling.[19]

B. Incentive Effects: High versus Standard-Pay PRC Students

The top section of Table 3 reports for each 12 plays of the game frequency counts of

---

[17] We define strategic play as FHs choosing outputs 1–3 because: (i) all three outputs are possible pooling equilibria, and (ii) both the adaptive learning model described in Section II and observed behavior indicate that strategic play by FHs involves all 3 outputs.

[18] The break-even point to make pooling at 2 profitable is a 62.5-percent higher percentage of tough targets at 5 than 2. The break-even point for pooling at 3 is a 33.5-percent higher proportion of tough targets at 5 compared to 3.

[19] Cooper (1997) has begun to investigate the relative ability of the belief-based model of Cooper et al. (1997b) and the reinforcement based model of Alvin E. Roth and Ido Erev (1995) to characterize the data. He finds that neither model does extraordinarily well at fitting the behavior of FHs; pooling emerges too slowly in the Roth-Erev model and too fast in the Cooper et al. model compared to the data. However, both models capture the central trend towards pooling at 2.
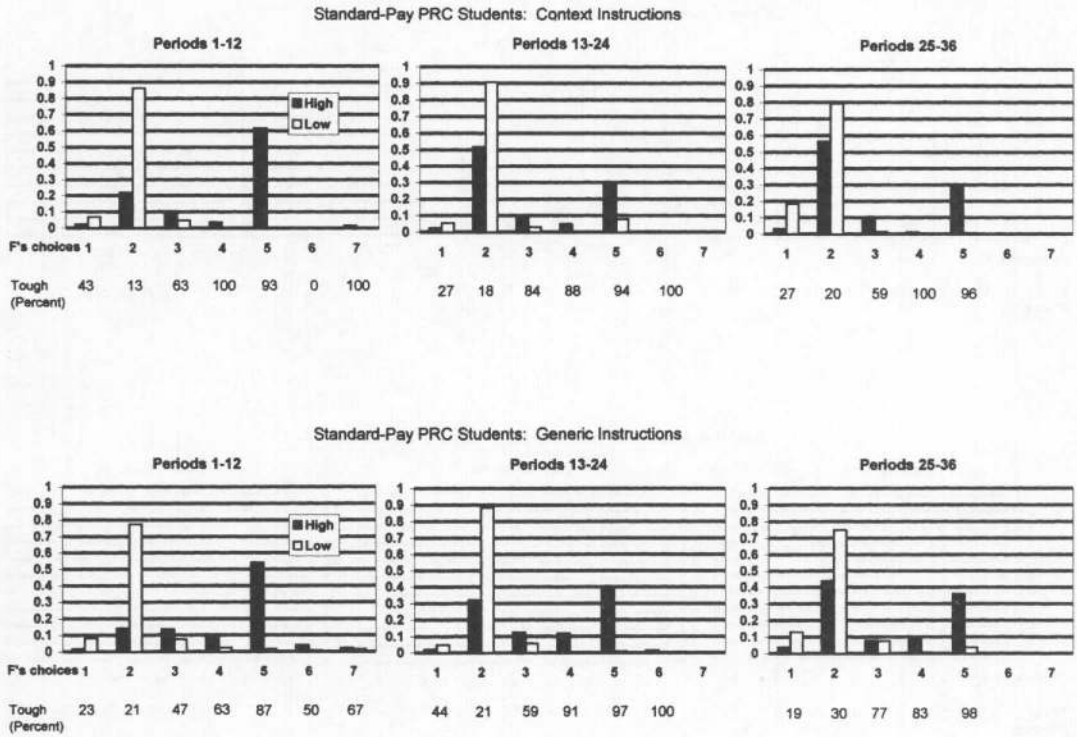
FIGURE 1. FIRM'S (F'S) OUTPUT CHOICES AND CP'S RESPONSES (PERCENTAGE OF TOUGH CONTRACTS) FOR STANDARD-PAY PRC STUDENTS

Note: Output frequencies of high- and low-productivity firms taken separately sum to 1.0 within each 12 plays of the game.

strategic play by FHs for the PRC student sessions. Also shown are differences in the percentage of tough targets in response to pooling (outputs 1–3) versus not pooling (outputs 4–7). The latter will be referred to as the target rate differential.

CONCLUSION 2: *Increased pay promoted significantly more strategic play by Fs initially. However, the gap between standard- and high-payoff sessions narrowed over time so that there were no significant differences in the level of strategic play by the last 12 games. Higher pay had no significant effect on CPs' choices.*

In the first 12 games, FHs played strategically much more frequently in high-pay sessions than in standard-pay sessions (49.5 percent versus 31.1 percent). In contrast, the target rate differential was similar for high- and standard-pay treatments during the first 12 games (69.8 per-

cent for high pay versus 62.7 percent for standard pay). FHs increased their level of strategic play after the first 12 games in all 10 sessions, with more rapid increases in the standard-pay sessions. As a result, by the final 12 rounds of play, the probability of pooling was virtually the same across treatments (63.2 percent and 60.5 percent in standard- and high-pay sessions, respectively). In comparison, target rate differentials changed little over time, rising in some cases and falling in others, with no notable differences emerging between standard- and high-pay sessions at any time.

We use probits to further explore the data, controlling for each subject's history of opponents' play and for individual effects. These are reported in the bottom section of Table 3.

Our firm probits employ FH choices as the dependent variable, coded as 1 for strategic play and 0 for nonstrategic play. Three alternative specifications are shown. To capture changes in
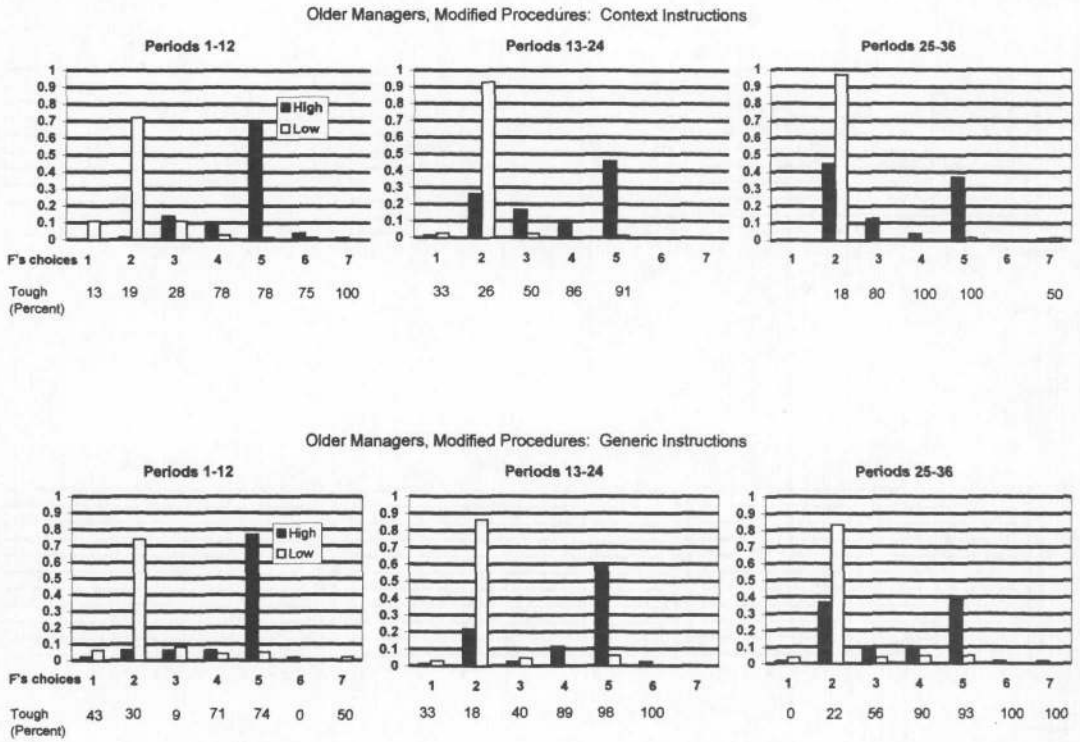
Older Managers, Modified Procedures: Context Instructions

Older Managers, Modified Procedures: Generic Instructions

FIGURE 2. FIRM'S (F'S) OUTPUT CHOICES AND CP'S RESPONSES (PERCENTAGE OF TOUGH CONTRACTS)
FOR OLDER MANAGERS, MODIFIED PROCEDURES

Note: Output frequencies of high- and low-productivity firms taken separately sum to 1.0 within each 12 plays of the game.

behavior over time we employ cycle dummies, with the cycle 2 dummy coded as 1 for games 13–24 and 0 otherwise, and the cycle 3 dummy coded as 1 for games 25–36 and 0 otherwise. Treatment effects are interacted with the cycle dummies with the HIi dummy coded as 1 for the high-pay treatment in cycle i and 0 otherwise. To capture the effect of CPs' responses on Fs' behavior we include the variable $\Delta 23$. For each session, this variable gives the observed target rate differential using all previous games. Because this variable is unreliable in early rounds, it is interacted with a dummy variable whose value is 0 for games 1–12 and 1 for games 13–36.[20]

Planners' probits employ "mistakes" as the dependent variable, with mistake taking on a value of 1 when a CP assigns a tough target following pooling (outputs 1–3) or an easy target following not pooling (outputs 4–7). Dummy variables corresponding to those employed in the F probits are used to capture the effect of time and payment levels on behavior. The impact of Fs' choices on CPs' actions is captured with the variable PPO23, the observed percentage of past pooling outcomes accounted for by FH types. All past periods are averaged in the PPO23 calculations and it is interacted with a dummy variable whose value is 0 for games 1–12 and 1 for games 13–36.

[20] Intuitively, $\Delta 23$ is serving as a proxy for Fs' expectations about CPs' responses. For early plays of the game, the observed target rate differential is probably a poor proxy for Fs' expectations, both because initial expectations and the observed target rate differential are likely to differ and

because the observed target rate differential is quite volatile before Fs have had several chances to play as both FHs and FLs. $\Delta 23$ employs session-level frequencies, information that was provided to all subjects.

TABLE 3—INCENTIVE EFFECTS: STANDARD VERSUS HIGH-PAY PRC STUDENTS

Raw data: Frequency counts

Games	High-productivity firms frequency of strategic play						Planners differential frequency of tough targets: No pooling-pooling					
	Standard pay			High pay			Standard pay			High pay		
	All	Context	Generic	All	Context	Generic	All	Context	Generic	All	Context	Generic
1–12	0.311	0.328	0.298	0.495	0.509	0.479	0.627	0.725	0.543	0.698	0.759	0.633
13–24	0.547	0.638	0.467	0.701	0.727	0.677	0.691	0.692	0.680	0.637	0.665	0.623
25–36	0.632	0.697	0.555	0.605	0.583	0.625	0.673	0.725	0.618	0.747	0.713	0.783

Statistical tests

Variable	Firm probits			Variable	Planner probits		
	Model 1	Model 2	Model 3		Model 1	Model 2	Model 3
Constant	−0.605 (0.112)**	−0.365 (0.084)**	−0.628 (0.115)**	Constant	−0.925 (0.063)**	−0.958 (0.051)**	−0.916 (0.062)**
Cycle 2	0.722 (0.089)**	−0.841 (0.323)**	−0.752 (0.334)*	Cycle 2	0.026 (0.078)	−0.496 (0.175)**	−0.510 (0.182)**
Cycle 3	1.041 (0.089)**	−0.786 (0.325)*	−0.478 (0.340)	Cycle 3	0.153 (0.068)*	−0.582 (0.178)**	−0.523 (0.202)**
HI1	0.589 (0.170)**	—	0.621 (0.175)**	HI1	−0.101 (0.106)	—	−0.103 (0.103)
HI2	0.485 (0.166)**	—	0.382 (0.165)†	HI2	0.217 (0.112)†	—	−0.018 (0.119)
HI3	−0.122 (0.165)	—	−0.144 (0.170)	HI3	−0.176 (0.115)	—	−0.197 (0.113)†
Δ23	—	2.305 (0.481)**	2.350 (0.501)**	PPO23	—	1.812 (0.412)**	1.754 (0.457)**
Log-likelihood	−925.51	−930.61	−919.15	Log-likelihood	−1,434.25	−1,428.89	−1,427.18
Number subjects	160	160	160	Number subjects	160	160	160
Number observations	1,512	1,512	1,512	Number observations	2,880	2,880	2,880

† Significantly different from 0 at the 10-percent level.
* Significantly different from 0 at the 5-percent level.
** Significantly different from 0 at the 1-percent level.

To correct for individual effects, all of the probits employ a "random effects" error specification. In other words, rather than assuming all error terms are independently and identically distributed, we assume there is correlation between the error terms for an individual, including an additional term for this in the likelihood function. For all of our regressions, this random effects term is significant at the 1-percent level. Since it is of little direct interest, this parameter is suppressed in all of the tables.

Looking at the F probits, model 1 employs cycle and the high-pay dummies without any direct accounting for variations in CPs' responses via Δ23. In this specification the cycle dummies are both positive, large, and statistically significant at the 1-percent level, reflecting

the large growth in strategic play in the last two cycles of play. The high-pay treatment dummies HI1 and HI2 are positive, large, and statistically significant at the 1-percent level, reflecting the substantially higher levels of strategic play observed early on in the high-pay treatment. The HI3 dummy is negative, but small in absolute value and not significant at any standard level, indicating no systematic differences in the level of strategic play in the last 12 games. Model 2 drops the high-pay dummies and introduces the Δ23 variable capturing the impact of variation in CPs' responses to Fs' behavior. The coefficient for Δ23 is positive, large, and statistically significant at the 1-percent level. The cycle dummies take on negative signs in model 2, resulting from the fact that Δ23 is 0 in cycle 1

and only takes on positive values for cycles 2 and 3. (The net effect of Δ23 and the cycle dummies still implies substantial growth in strategic play over time.) Model 3 reintroduces the high-pay dummies along with the Δ23 variable. There is little change from the coefficient values for the high-pay dummies reported in model 1. The chi-square value for the high-pay dummies is significant at the 1-percent level ($X^2 = 22.9$, d.f. $= 3, p < 0.01$) with the high-pay dummies positive and significant in cycles 1 and 2, and negative, but not significant, in cycle 3.[21]

The differences in pooling rates across treatments and over time predicted by model 3 are substantial. Setting Δ23 equal to its average value over all observations to control for any differences in CPs' actions, model 3 predicts 50-percent strategic play by FHs in cycle 1 compared to 27 percent for FLs and 70-percent strategic play by FHs in cycle 2 versus 50 percent for FLs, values that are quite close to the frequency data reported in the top part of Table 3.[22]

The CP probits show no effect of incentives on behavior. Model 1 employs cycle dummies and high-pay dummies. The signs of the cycle dummies are positive, suggesting increasing mistakes over time. The high-pay dummies have mixed signs, and are jointly significant at the 5-percent level ($X^2 = 9.64$, d.f. $= 3, p < 0.025$). Model 2 drops the high-pay treatment dummies and introduces the variable PPO23. The sign of the PPO parameter is positive, large, and statistically significant at the 1-percent level, indicating more mistakes by CPs in re-

sponse to pooling by FHs. These mistakes primarily took the form of more tough targets in response to outputs 1–3.[23] With the addition of PPO23 to the model, the cycle dummies become negative and significant at the 1-percent level, indicating that the increase in mistakes over time found in model 1 was in response to FHs' increased strategic play. Reintroducing the high-pay dummies in model 3 results in an insignificant reduction in the log-likelihood function relative to model 2 ($X^2 = 3.42$, d.f. $= 3, p > 0.25$), so that we fail to reject a null hypothesis that higher incentives had no effect on CPs' choices. That is, the high-pay effects reported in model 1 reflect CPs' responses to the increased strategic play of FHs rather than any direct effect of increased incentives on CPs' behavior.

Our results show a puzzling dichotomy: incentives strongly impact the behavior of Fs but not CPs. Recall that we hypothesized that increased incentives might generate more strategic play by inducing more deductive reasoning. Our intuition immediately suggests that for Fs to play strategically right from the start requires a deeper level of insight into the properties of the game than it does for CPs. This intuition can be formalized by thinking about the game in terms of levels of deductive reasoning in the spirit of Stahl (1993). A zero-level reasoner is completely naive, picking strategies at random. In practice, a zero-level reasoner generates a uniform distribution over strategies. A first-level reasoner is slightly more sophisticated, picking the strategy which maximizes his/her payoff versus a zero-level reasoner. Intuitively, a first-level reasoner is sophisticated enough to think of maximizing, but not sophisticated enough to think about what his/her opponent might be doing. For the ratchet effect game, a first-level reasoner F chooses his/her full information output, and a first-level reasoner CP chooses EASY regardless of F's output. A second-level reasoner is more sophisticated yet, picking the best response to a first-level rea-

---

[21] An alternative model specification with Δ23 interacted with a high-pay treatment dummy was tried but this variable failed to achieve statistical significance at conventional levels ($p > 0.10$). Throughout this paper, all X^2 statistics are from log-likelihood ratio tests.

[22] The initial difference between treatments is sufficiently robust that even relatively weak tests are able to capture it. For example, using session values as the unit of observation, the average frequency of strategic play in the first 12 games under standard pay was 0.307 (with a standard error of the mean of 0.043) compared to 0.503 (with a standard error of the mean of 0.044) for the high-pay treatment ($t = 3.08$, d.f. $= 8, p < 0.01$, 1-tailed t-test). Further, there was no overlap in the average frequency of strategic play between standard- and high-pay sessions in these first 12 games, so that a Mann-Whitney test for differences between means is significant at the 0.01 level as well (1-tailed test).

[23] Given that virtually all FLs selected output levels 1–3, assignments of TOUGH following play of these outputs could not have reflected expected payoff maximization. Moreover, these choices do not reflect risk aversion in any conventional sense since EASY always had greater expected value *and* less variance in payoffs than TOUGH.

soner. A second-level reasoner both maximizes and takes into account that his/her opponent is maximizing (albeit in a naive fashion). Like a first-level reasoner, a second-level reasoner F chooses his/her full information output. However, second-level reasoning changes behavior for CPs; a second-level reasoner CP chooses EASY for output 2 and TOUGH for output 5. Going up one more level, a third-level reasoner maximizes versus a second-level reasoner, using a more sophisticated model for his/her opponent than a second-level reasoner. With third-level reasoning, both types of F choose 2 and CPs choose EASY following 2 and tough following 5. Higher levels of reasoning can be formulated, but are unnecessary as we are already in equilibrium.[24]

This logic not only formalizes our intuition that strategic play requires a deeper level of insight for Fs than it does for CPs, but also meshes with the observed effects of incentives. Suppose that most PRC students were able to get to the second level of reasoning with little prompting, but needed some additional factor to push them to the third level. One such factor might be more experience with the game; as subjects have an opportunity to observe the play of others, more sophisticated models of other players' actions naturally suggest themselves. (It is not actually necessary that players understand that others are using relatively sophisticated models, only that they act "as if" they realize that others are employing relatively sophisticated strategies. For example, players in fictitious play models eventually end up acting like our third-level reasoners, even though they never do anything more sophisticated than maximizing against observed frequencies of past actions.) As suggested previously, another factor in moving subjects to third-level reasoning might be higher pay. In either case, the switch from second-level to third-level reasoning will be observed as an increase in Fs' strategic play *but not as a decrease in CPs' mistakes.* Thus, higher pay acts as a substitute for experience in a very natural fashion; the greater reasoning required of Fs (as compared to CPs) to play strategically requires some catalyst, and either higher pay or more experience can play this role.

Other researchers have reported incentive effects on performance in experiments, but the mechanism(s) underlying these effects are typically different from the one identified here. One general tendency is for incentives to reduce the variance of outcomes around the mean, with no effect on mean performance (see Smith and Walker [1993b] for a review of the literature).[25] In contrast, we observe a clear effect on mean behavior of Fs. In other experiments where mean outcomes have been affected by increased incentives, the effect can usually be rationalized in terms of risk aversion (see for example, Kroll et al., 1988; Kachelmeier and Shehata, 1992; Smith and Walker, 1993a). In our game, increased traces of risk aversion brought about by increases in payoffs would cause an expected utility-maximizing CP to initially choose more easy targets under the high-pay treatment (EASY had higher expected value and less spread than TOUGH). This outcome was *not* observed. Rather, our results suggest that increased payoffs can affect subjects' behavior in a more fundamental manner—inducing heightened levels of deductive reasoning. Unlike many games, increased reasoning in the ratchet effect game leads to clear, easily measurable changes in behavior. This may explain why we observe the incentive effects that we do.

C. Learning and Context Effects

1. *Student Subjects.*—We hypothesize that context affects managers' behavior by catalyzing transfer of strategic insights gained from field experience. However, context might affect subject behavior regardless of any field experience. We therefore first examine the effect of context on student behavior before turning to managers.

CONCLUSION 3: *Context had no effect on the behavior of PRC students as Fs, regardless of the level of payoffs. Context reduced the initial level of mistakes by PRC students as CPs, with this effect confined to standard-pay sessions.*

---

[24] Our conclusions still hold even if we treat players as maximizing versus an arbitrary mix of lower-level reasoners—pooling only emerges with third-level (or higher) reasoners.

[25] There are a number of studies where incentives have no effect on performance (see the many examples cited in Camerer, 1995).

To test for context effects with students in their role as Fs, we modified model 3 for Fs in Table 3 by replacing the three high-pay dummies with three context dummies (one for each cycle). Comparing this regression with manager model 2 in Table 3 yields $X^2 = 1.38$ (d.f. $= 3$, $p > 0.5$), with none of the context dummies individually significant at even the 10-percent level. Further, applying the same model specification to standard-pay and high-pay sessions separately, none of the individual context dummies are significant at the 10-percent level. Thus, there is no evidence for a context effect with students in their role as Fs.

A similar test for context effects for students in their role as CPs yields $X^2 = 7.96$, slightly above the cutoff of 7.82 for significance at the 5-percent level. Most of the action here is captured in the dummy for context in cycle 1, which is negative and significant at the 1-percent level.[26] Looking at standard-pay and high-pay treatments separately, the context dummies are significant at the 10-percent level for the standard-pay treatment ($X^2 = 6.34$), but not for the high-pay treatment ($X^2 = 3.80$). Further, the only context dummy which is significant at the 5-percent level is for cycle 1 in the standard-pay sessions. Thus, there is some evidence for an initial context effect among students in their role as CPs, with this effect confined to standard-pay sessions.

2. *Manger Subjects.*—Turning from students to managers, Tables 4 and 5 summarize results from sessions with PRC managers and white-collar workers. Table 4 gives the frequency of strategic play by Fs and Table 5 gives the target rate differential for CPs. Data are reported for the managers as a whole as well as by sub-

---

[26] The estimated probit (with estimated standard errors reported in parentheses) is

Mist = $\quad$ −0.831 $\quad\quad$ −0.655 Cycle2 $\quad$ −0.646 Cycle3
$\quad\quad\quad$ (0.069)** $\quad\quad$ (0.192)** $\quad\quad\quad$ (0.193)**

$\quad\quad\quad$ −0.266 Con1 $\quad$ +0.078 Con2 $\quad$ −0.113 Con3
$\quad\quad\quad$ (0.100)** $\quad\quad$ (0.109) $\quad\quad\quad$ (0.107)

$\quad\quad\quad$ +1.795 PPO23
$\quad\quad\quad$ (0.418)**

with the variables defined as in the text and notation for statistical significance is the same as in Table 3. The log-likelihood value of the probit is −1,424.90 with 160 subjects and 2,880 total observations.

groups. Several results emerge even with casual examination of the raw data. First, the frequency of strategic play by Fs was markedly lower to begin with under both old manager treatments compared to the standard-pay student sessions. In contrast, the younger managers exhibited far more strategic play initially than standard-pay students, comparing favorably to high-pay students. Second, strategic play by Fs increased over time for almost all groups of managers, with generic sessions for older managers, original procedures a notable exception. Third, strategic play increased more rapidly in context sessions than in generic sessions. Fourth, and most notably, the target rate differential was substantially and persistently higher in context compared to the generic sessions.

CONCLUSION 4: *Controlling for the tough target differentials, context sped up the learning/adjustment process, promoting increased levels of strategic play for PRC managers in their role as Fs in later cycles of play.*

Looking at the bottom half of Table 4, model 1 of the all-manager probits indicates growing levels of strategic play over time for both context and generic treatments (positive cycle dummies and positive context dummies for cycles 2 and 3). Model 1 also suggests an initial (cycle 1) negative effect of context on strategic play by Fs which reverses by cycle 3. This negative effect of context on initial strategic play is, however, not supported by model 3 which accounts for CPs' actions. In this specification, the parameter estimate for Δ23 is positive and statistically significant at the 1-percent level, with the Con1 dummy insignificant at conventional levels ($p = 0.20$). Comparing the all-manager model 2 specification with model 3 shows the context dummies jointly significant at the 10-percent level ($X^2 = 6.62$, d.f. $= 3$). The context dummies are both positive in cycles 2 and 3 indicating faster growth in the overall level of strategic play in context relative to generic sessions, with the Con3 dummy close to achieving significance at conventional levels ($p = 0.12$).

Older managers under our original procedures show a slightly different pattern. Comparing model 3 to model 2 for this subgroup, the three context dummies are jointly significant at the 1-percent level ($X^2 = 15.74$). Neither the

TABLE 4—PRC MANAGERS IN THE ROLE OF FIRMS

						High-productivity firms frequency of strategic play						
	All managers			Old managers Original procedures			Old managers Modified procedures			Young managers		
Games	All	Context	Generic	All	Context	Generic	All	Context	Generic	All	Context	Generic
1–12	0.186	0.172	0.206	0.108	0.080	0.144	0.150	0.153	0.146	0.468	0.476	0.457
13–24	0.304	0.346	0.246	0.188	0.202	0.159	0.347	0.444	0.256	0.629	0.829	0.429
25–36	0.388	0.479	0.261	0.258	0.376	0.096	0.536	0.567	0.483	0.738	0.786	0.643

				Statistical tests: Firm probits						
	All managers			Old managers Original procedures			Old managers, modified procedures, and young managers		
Variable	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Constant	−1.144	−1.360	−1.201	−1.897	−1.928	−1.859	−1.034	−1.135	−1.060
	(0.171)**	(0.120)**	(0.181)**	(0.317)**	(0.173)**	(0.319)**	(0.264)**	(0.182)**	(0.266)**
Cycle 2	0.165	−1.025	−0.856	0.060	−0.737	−0.177	0.502	0.145	0.234
	(0.158)	(0.306)**	(0.326)**	(0.214)	(0.414)†	(0.480)	(0.258)†	(0.502)	(0.549)
Cycle 3	0.373	−0.792	−0.731	−0.269	−0.554	−0.577	1.098	0.536	0.818
	(0.146)*	(0.326)*	(0.341)*	(0.248)	(0.470)	(0.524)	(0.236)**	(0.515)	(0.543)
Con1	−0.602	—	−0.306	−0.412	—	−0.417	−0.013	—	0.000
	(0.227)**		(0.238)	(0.369)		(0.369)	(0.287)		(0.286)
Con2	0.179	—	0.135	0.391	—	0.158	0.805	—	0.770
	(0.212)		(0.234)	(0.316)		(0.420)	(0.241)**		(0.243)**
Con3	0.433	—	0.371	1.450	—	1.220	0.409	—	0.378
	(0.223)†		(0.240)	(0.340)**		(0.445)**	(0.287)		(0.293)
Δ23	—	2.958	2.201	—	2.266	0.719	—	1.425	0.526
		(0.530)**	(0.597)**		(0.767)**	(1.026)		(0.817)†	(0.912)
Log-likelihood	−640.69	−639.96	−636.65	262.35	−269.83	−261.96	−341.27	−345.6	341.14
Number subjects	150	150	150	74	74	74	76	76	76
Number observations	1,340	1,340	1,340	743	743	743	597	597	597

† Significantly different from 0 at the 10-percent level.
* Significantly different from 0 at the 5-percent level.
** Significantly different from 0 at the 1-percent level.

Con1 or Con2 dummies are significant in their own right in model 3, but the Con3 dummy is positive, relatively large, and statistically significant ($p < 0.01$). This indicates that major differences as a function of context emerged in the last cycle of play. Further, in contrast to the all-manager probits, the Δ23 variable in model 3 fails to achieve statistical significance at any standard level.

Probits for the remaining manager subgroups pool the data from older managers, modified procedures and young managers.[27] Comparing model 3 with model 2 for this subgroup, the three context dummies are jointly significant at the 5-percent level ($X^2 = 8.92$), with positive coefficients in cycles 2 and 3 (and coefficient value of zero in cycle 1). In this case context impacts primarily in cycle 2 as the Con2 dummy in model 3 is the only one of the context dummies significant in its own right. Here, too, the Δ23 variable fails to achieve statistical significance at conventional levels in model 3.[28]

[27] There are two reasons for pooling data from these subgroups: (i) treating each group separately, the analysis suffers from a paucity of individual subject observations, and (ii) both groups have in common substantially less confusion about the underlying game than older managers, original procedures. These probits all include a single dummy variable whose value

is 1 for young managers and 0 otherwise. This variable is positive and significant in all the F specifications and negative and significant in all the CP specifications. Given that we are not directly interested in this dummy, it has been suppressed in all relevant regressions.

[28] This is the one F probit with a significant interaction effect between Δ23 and context, with the coefficient for the interaction term strongly positive and significant at the

TABLE 5—PRC MANAGERS IN THE ROLE OF CENTRAL PLANNERS

Differential frequency of tough targets: No pooling-pooling

Games	All managers			Old managers, Original procedures			Old managers, Modified procedures			Young managers		
	All	Context	Generic	All	Context	Generic	All	Context	Generic	All	Context	Generic
1–12	0.485	0.569	0.384	0.410	0.529	0.232	0.506	0.582	0.429	0.665	0.706	0.624
13–24	0.666	0.724	0.597	0.712	0.794	0.578	0.693	0.611	0.773	0.581	0.818	0.372
25–36	0.664	0.768	0.544	0.700	0.822	0.505	0.708	0.732	0.685	0.454	0.615	0.322

Statistical tests: Planner probits

Variable	All managers				Old managers, Original procedures				Old managers, modified procedures, and young managers			
	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
Constant	-0.570	-0.757	-0.569	-0.577	-0.340	-0.662	-0.306	-0.316	-0.564	-0.727	-0.636	-0.644
	(0.089)**	(0.068)**	(0.090)**	(0.091)**	(0.148)*	(0.098)**	(0.155)*	(0.158)*	(0.128)**	(0.108)**	(0.129)**	(0.127)**
Cycle 2	-0.334	-0.837	-0.921	-1.150	-0.606	-0.858	-0.925	-1.151	-0.295	-0.834	-1.020	-1.520
	(0.071)**	(0.098)**	(0.099)**	(0.144)**	(0.133)**	(0.157)**	(0.173)**	(0.209)**	(0.107)**	(0.150)**	(0.170)**	(0.240)**
Cycle 3	-0.205	-0.843	-0.765	-0.981	-0.493	-0.811	-0.744	-0.918	-0.152	-0.885	-0.948	-1.479
	(0.074)**	(0.096)**	(0.100)**	(0.132)**	(0.127)**	(0.149)**	(0.173)**	(0.214)**	(0.106)	(0.160)**	(0.170)**	(0.249)**
Con1	-0.329	—	-0.320	-0.311	-0.520	—	-0.565	-0.558	-0.217	—	-0.207	-0.197
	(0.130)*	—	(0.130)*	(0.132)*	(0.208)*	—	(0.213)**	(0.216)**	(0.205)	—	(0.197)	(0.194)
Con2	-0.280	—	-0.330	0.055	-0.578	—	-0.556	-0.143	-0.035	—	-0.346	0.608
	(0.123)*	—	(0.129)*	(0.193)	(0.227)*	—	(0.223)*	(0.303)	(0.166)	—	(0.177)†	(0.441)
Con3	-0.443	—	-0.676	-0.270	-0.704	—	-0.833	-0.413	-0.228	—	-0.686	0.388
	(0.131)**	—	(0.136)**	(0.199)	(0.224)**	—	(0.232)**	(0.336)	(0.182)	—	(0.211)**	(0.498)
PPO23	—	2.647	3.100	4.278	—	1.343	1.728	2.908	—	2.942	4.111	6.653
	—	(0.326)**	(0.316)**	(0.603)**	—	(0.596)*	(0.528)**	(0.776)**	—	(0.543)**	(0.760)	(1.241)**
PPO23*CON	—	—	—	-1.874	—	—	—	-2.438	—	—	—	-4.175
	—	—	—	(0.722)*	—	—	—	(1.372)†	—	—	—	(1.644)*
Log-likelihood	-1,279.26	-1,265.48	-1,254.56	-1,252.43	-576.95	-580.39	-573.53	-571.82	-682.43	-674.10	-668.35	-664.91
Number subjects	150	150	150	150	74	74	74	74	76	76	76	76
Number observations	2,700	2,700	2,700	2,700	1,332	1,332	1,332	1,332	1,368	1,368	1,368	1,368

† Significantly different from 0 at the 10-percent level.
* Significantly different from 0 at the 5-percent level.
** Significantly different from 0 at the 1-percent level.

Overall, context consistently affected PRC managers' level of strategic play in their role as Fs, regardless of which manager population we look at. The effects are, in fact, stronger when looking at the different manager groups separately than when pooling all sessions together. This no doubt reflects heterogeneity between the different subject populations: for older managers, original procedures context had its largest impact in cycle 3 while the impact occurs earlier (in cycle 2) for older managers, modified procedures, and young managers. Pooled together these timing differences partially cancel each other

out. Nonetheless, there is clearly a context effect with PRC managers in their role as Fs, while there is none with PRC students.

CONCLUSION 5: *Context had a strong effect on PRC managers in their role as CPs, promoting uniformly higher, at times substantially higher, target rate differentials than in generic sessions.*

The role of context in inducing higher tough target differentials is evident from the raw data in Table 5; the target rate differential is higher everywhere in the context sessions with the exception of the second cycle of play (games 13–24) for older managers, modified procedures.

The CP probits reported at the bottom of Table 5 confirm the statistical significance of these effects as well as indicating their source. Consider the all manager probits. In model 1 the

---

5-percent level and the coefficient for $\Delta 23$ turning negative in sign but far from statistically significant ($p > 0.30$). Thus FHs in context sessions were more sensitive to CPs' actions, with this heightened sensitivity the source of the higher pooling rates observed.

cycle and context dummies are all negative and significant at better than the 5-percent level, indicating significantly fewer mistakes over time and consistently fewer mistakes in context than in generic sessions. In model 2, the variable PPO23 (average percentage FH types pooling) is positive and significant at better than the 1-percent level, indicating that FHs' increased pooling resulted in CPs making more mistakes (these mistakes primarily involved responding with tough targets for pooling). The change in the log-likelihood function from model 2 to model 3, along with the negative signs of the context dummies in model 3, confirms model 1's suggestion of consistently fewer mistakes in context than in generic sessions ($X^2 = 21.8$, d.f. $= 3$, $p < 0.01$). Model 4 demonstrates a consistent interaction effect between the PPO variable and context. The coefficient for this interaction term is relatively large, negative, and significant at better than the 5-percent level, indicating significantly fewer mistakes in response to FHs' pooling in context than in generic sessions.[29] The introduction of this interaction term wipes out the statistical significance of the context dummies for cycles 2 and 3, indicating that most of the differences in CP behavior between context and generic sessions can be traced to differing sensitivity to FHs' pooling.[30]

Similar results are found for the different manager subgroups looked at in isolation. The magnitude of the predicted differences in the proportion of mistakes implied by model 4, evaluated at the mean value of PPO23 over all observations, is quite dramatic in a number of cycles of play, and is well represented by the tough target rate differentials reported in the top half of Table 5.[31]

Context had a significant and relatively large impact on PRC managers in their role as CPs in all cycles of play. In contrast, context impacted on manager behavior as Fs only after the first cycle of play, emerging in full-blown force for older managers, original procedures in the third cycle of play and for older managers, modified procedures, and young managers in the second cycle of play (with the differences narrowing, and no longer significant in the last cycle of play). This leads us to the following unanticipated conclusion.

CONCLUSION 6: *For PRC managers, context had a more immediate and stronger impact on the level of strategic play in their role as CPs compared to their role as Fs.*

In accounting for this surprising result, recall that the predicted occurrence of a context effect depends in part on the similarity between managers' field experiences and our laboratory game. As it happens, there is substantial outside evidence that the strategic option open to CPs in our experiment is much closer to the option exercised in field settings than it is for Fs:

> Managers who greatly overfulfil this tax quota harm their interests. . . . First, they provoke a subsequent upward revision of the tax quota. Secondly, profits that are much higher than expected may convince superiors that this is an efficient enterprise, this may also convince them that they gave the enterprise excessively favorable treatment, and that they do not deserve such treatment in the future. As managers commonly put it, their superiors are in the habit of "lashing the fast oxen": the profitable and efficient are the first to be denied favourable treatment (Andrew G. Walder, 1989 p. 259).

In sum, CPs in the field actually do ratchet up targets of firms that reveal themselves as high-productivity types, or, amounting to the same thing, give them less favorable treatment in future resource allocations.

In contrast, managers of high-productivity

---

[29] There were no significant interaction effects of this sort for the students.

[30] Jointly, the three context dummies are still significant at the 5-percent level ($X^2 = 10.13$, d.f. $= 3$, $p < 0.025$).

[31] These context effects are sufficiently robust that they show up in cycle 1 using session as the unit of observation: A Mann-Whitney test applied to old managers, original procedures shows significantly higher tough target rate differences in context versus generic sessions ($p = 0.05$, 1-tailed test). Pooling all old managers sessions yields a significant difference at the 10-percent level using a 1-tailed Mann-Whitney test. *t*-tests using session frequencies as the unit of observation are significant at better than the 10-

percent level (1-tailed test) with or without the inclusion of the two young manager sessions.

firms have considerably more behavioral options at their disposal in field settings than simply producing at the same output level as low-productivity firms:

> "How do you prevent yourself from over-fulfilling by too much? You put your money into the development of production. You need to prepare for next year's production." In other words, the rational manager instead of showing high profits on the books, will spend these funds, working them into the costs of production in various ways. . . . We know from Chinese analysts, however, that the practice does not stop here. Chinese managers are past masters at regulating their profits on the book, while working surpluses back into the cost of production. "The vast majority of enterprises" are reputed to keep two different sets of books: one for the upper levels, one for internal use only. "Slush funds" (*xiaojin ku*) are commonplace, often built up by speculating in scarce materials and selling products produced within state plans for inflated outside-plan prices. These funds are used for expenditures that are limited by state regulation if made on the books: for housing construction, distributions in kind to employees, and other benefits that enhance the prosperity of the unit (Walder, 1989 p. 259).

In other words, managers of high-productivity firms in the field have alternative strategies available to them which were not available to FHs in our game. Moreover, managers of high-productivity firms typically use these other strategies to hide their true type.

Thus, there is a fundamental dissimilarity between our experimental game and the field setting for Fs. This is less true for CPs; the strategies used in the field closely match those available in the experiment. Given the close ties psychologists have found between similarity and transfer, we conjecture that the greater impact of context on managers in their role as CPs than Fs is a function of the greater similarity of CPs' strategic options between the experiment and field settings.

D. *PRC Managers versus Students*

As anticipated, context had a much greater impact on play of PRC managers than PRC students. Evidence for this is clear-cut: (1) In the role of Fs,

there is a small but consistent context effect for PRC managers as opposed to no discernable context effect for PRC students, and (2) In the role of CPs, the context effect for PRC managers is strong in all cycles and in all subject populations. For students this effect is confined to the first cycle of play in standard-pay sessions.[32]

What we had not anticipated was that older managers would exhibit substantially lower levels of strategic play than any other group in their role as Fs.

CONCLUSION 7: *Older managers started out with the lowest frequency of strategic play, much lower than for PRC students. This can be partially, but incompletely, accounted for by lower incentives to pool. The difference largely results from the older age and/or inferior education of older managers.*

Older managers started out with the lowest levels of strategic play (10.8 percent and 15.0 percent under original and modified procedures, respectively). Although initial incentives for strategic play were less than with the students, this difference was largely confined to the generic sessions. Yet lower levels of strategic play were observed in both context and generic sessions.[33] In contrast, young managers had close to the highest initial level of strategic play (46.8 percent). Probit analysis indicates that initial levels of strategic play by young managers, controlling for $\Delta 23$, are indistinguishable from levels in high-pay student sessions. And by rounds 25–36, young managers had higher levels of strategic play than in either type of student session.[34]

All managers had extensive field experience, but young managers had more years of schooling on average than older managers (and no doubt received better quality schooling). This suggests that education, not field experience,

---

[32] Probits comparing all managers with all students confirm these results, with the context by manager interaction effects particularly strong for CPs.

[33] Probits comparing old managers (context only) with standard-pay students confirm that the managers started with significantly lower levels of strategic play under both our original and modified procedures, whether or not we control for differing incentives to pool via $\Delta 23$.

[34] The estimated probit (with standard errors in parentheses) is

played the critical role in determining low initial levels of strategic play for older managers. Consistent with this conjecture, in the old manager, modified procedure sessions, which substantially reduced the paperwork involved in playing the game, the frequency of strategic play was substantially higher in the last 12 games (53.6 percent) than in the old manager, original procedure sessions (25.8 percent), and was only marginally below (and not significantly different from) the level of strategic play observed for PRC students in the last 12 games.

Turning to play as CPs, the differences between students and managers are less dramatic. Concentrating only on sessions with context, managers' behavior in the CP role was not notably different from students'. While the older managers generated a lower target rate differential than the students in games 1–12 (54.8 percent versus 65.4 percent for the students), probits indicate that this difference was not significant at the 10-percent level.[35] Although by the last cycle of play older managers actually achieved a higher target rate differential than students ever did, differences between students and older managers were not significant in any cycle of play.

IV. Summary and Conclusions

We have examined a simplified version of the ratchet effect game in an experimental setting.

Strat = -0.127	-0.955 Cycle2	-0.62 Cycle3
(0.201)	$(0.374)*$	(0.380)
-0.505 SP1	-0.369 SP2	-0.429 SP3
$(0.232)*$	(0.233)	(0.271)
$+0.120$ HP1	$+0.009$ HP2	-0.576 HP3
(0.242)	(0.254)	$(0.281)*$
$+2.461$ $\Delta23$		
$(0.479)**$		

The variables are defined as in the text, notation for statistical significance is the same as in Table 3, and the new variables SP1–SP3 and HP1–HP3 are dummies for standard-pay and high-pay students in cycles 1–3 respectively. $X^2 = 26.3$, d.f. $= 6$, $p < 0.01$ for the student cycle dummies.

[35] The probit specification here compares older managers (context only) with standard-pay students and includes cycle dummies, the PPO23 variable, and student cycle dummies. The X^2 statistic for including the student cycle dummies is 3.74 (3 d.f., $p > 0.25$).

Our treatment variables were subject population (managers and white-collar workers in textile factories versus students), the level of incentives, and context versus generic instructions. We reach several general conclusions.

First, the pattern of play was similar over time across all treatments. In early games, a majority of high-productivity firms played nonstrategically, revealing themselves as high-productivity types, and received tough targets as a consequence. The response to these tough targets was increased mimicry of low-productivity firms' choices. This pattern of play is consistent with a number of simple learning models (for example, Cooper et al., 1997b).

Second, there were relatively large and statistically significant incentive effects on firms' play, with significantly more strategic play in early rounds of high-pay student sessions compared to standard-pay student sessions. Over time, the standard-pay students were able to erase this difference, suggesting that experience and incentives act as substitutes for each other. There were no significant incentive effects on planners' play. Unlike other experiments which have exhibited incentive effects, our results cannot be explained by decreased variance in outcomes or by risk aversion. We argue that our results are best explained by increased levels of reasoning with heightened incentives. Although this is the impact most theorists would predict, it has rarely been reported. More work is clearly needed to examine incentive effects in environments subject to strong and easily measured learning effects to see if our results generalize.

Third, context had a much larger and more consistent effect on PRC managers than on students. In addition, this impact materialized faster and was more robust for managers in their role as central planners than in their role as firms. These results are consistent with our hypothesis that context can catalyze strategic play by clarifying the similarity of the laboratory game to subjects' past experience; managers should have the most relevant field experience, and the strategic options available to planners in our laboratory game were much closer to the options exercised in field settings than those available to firms.

Our experiment has several substantive implications for understanding the ratchet effect in field settings. First, this study provides evidence

for central planners ratcheting up quotas in response to increased performance in China. As already noted, clear evidence of this effect can sometimes be difficult to obtain using field data. Second, it takes time for managers to respond strategically to this ratcheting up of incentives and there may well be some nonlearners or very slow learners. Thus, despite its negative long-run consequences, the ratchet principle provides strong short-run benefits to the planners, which may well explain its popularity. Finally, a major concern with the ratchet effect is that it creates economic inefficiencies through underproduction. The empirical literature suggests that managers use a number of alternate strategies (e.g., keeping two sets of books) rather than cutting output. These strategic alternatives also generate inefficiencies, but they are likely to be smaller than matching the output levels of low-productivity types. The relatively small context effect for managers playing as firms provides indirect evidence for the use of alternative strategies in the field.

Our results on the relationship between context and expertise, along with the cognitive psychology literature on this topic, generate several insights into the methodology of experiments comparing experts with students. First and foremost, it appears that the experiment must be embedded in a context which is familiar to the experts for expertise to have a chance to significantly affect behavior. Otherwise there is little basis for experts to call on their superior knowledge base in determining what actions to pursue in the experiment. In this respect it is worth noting that most economic experiments comparing experts with students have employed generic settings or something very close to it. Second, we would anticipate no significant differences between expert and student performance if the strategic options available in the experiment and/or contextual framework differ significantly from the field setting. In fact, subpar performance may result if the differences are too large; context may elicit previously learned knowledge from experts which, although helpful in the field setting, is actually counter productive in the experiment (see P. Burns [1985] for a notable example of this).

Our context results also have implications for understanding learning processes in games and convergence to equilibrium. In using learning models to explain convergence to equilibrium, one of the concerns theorists have had is that unless learning processes generalize across similar games, we are likely to be observing out-of-equilibrium behavior for quite some time, as learning processes can, at times, be exceedingly slow (Drew Fudenberg and Kreps, 1988). The psychological research on learning generalizability indicates that: (i) learning generalizability is typically incomplete, and (ii) the more dissimilar environments are, the less likely learning is to generalize. Our results with context and a subject population quite familiar with the underlying structure of our game suggest that these findings from the psychology literature are likely to generalize to economic environments with financial incentives.

REFERENCES

Becker, Gary S. "A Theory of the Allocation of Time." *Economic Journal*, September 1965, *75*(299), pp. 493–517.

Berliner, Joseph S. *The innovation decision in Soviet industry*. Cambridge: MIT Press, 1976.

Brandts, Jordi and Holt, Charles A. "An Experimental Test of Equilibrium Dominance in Signaling Games." *American Economic Review*, December 1992, *82*(5), pp. 1350–65.

Burns, P. "Experience and Decision Making: A Comparison of Students and Businessmen in a Simulated Progressive Auction," in Vernon L. Smith, ed., *Research in experimental economics*, Vol. 3. Greenwich, CT: JAI Press, 1985, pp. 139–53.

Camerer, Colin. "Individual Decision Making," in John H. Kagel and Alvin E. Roth, eds., *Handbook of experimental economics*. Princeton, NJ: Princeton University Press, 1995, pp. 587–703.

Chase, William G. and Simon, Herbert A. "Perception in Chess." *Cognitive Psychology*, January 1973, *4*(1), pp. 55–81.

Chaudhuri, Ananish. "The Ratchet Principle in a Principle Agent Game with Unknown Costs: An Experimental Analysis." Mimeo, Rutgers University, 1996.

Chi, Michelene T.; Feltovich, Paul J. and Glaser, Robert. "Categorization and Representation of Physics Problems by Experts and Novices." *Cognitive Science*, April–June 1981, *5*(2), pp. 121–52.

Cho, In-Koo and Kreps, David M. "Signaling Games and Stable Equilibria." *Quarterly Journal of Economics*, May 1987, *102*(2), pp. 179–221.

Cooper, David J. "A Comparison of Learning Models in Ratchet Effect Games." Mimeo, University of Pittsburgh, 1997.

Cooper, David J.; Garvin, Susan and Kagel, John H. "Adaptive Learning vs. Equilibrium Refinements in an Entry Limit Pricing Game." *Economic Journal*, May 1997a, *107*(442), pp. 553–75.

_____. "Signaling and Adaptive Learning in an Entry Limit Pricing Game." *Rand Journal of Economics*, Winter 1997b, *28*(4), pp. 662–83.

Cooper, David J.; Kagel, John H.; Lo, Wei and Gu, Qing Liang. "An Experimental Study of the Ratchet Effect: The Impact of Incentives, Context and Subject Sophistication on Behavior." Mimeo, University of Pittsburgh, 1998.

Crew, Michael A. *Incentive regulation for public utilities*. Boston: Kluwer Academic Publishers, 1994.

Dearden, James; Ickes, Barry W. and Samuelson, Larry. "To Innovate or Not Innovate: Incentives and Innovation in Hierarchies." *American Economic Review*, December 1990, *80*(5), pp. 1105–24.

DeGroot, Adrianus D. *Thought and choice in chess*, 2nd Ed. Mouton: The Hague, 1978.

Freixas, Xavier; Guesnerie, Roger and Tirole, Jean. "Planning under Incomplete Information and the Ratchet Effect." *Review of Economic Studies*, April 1985, *52*(2), pp. 173–91.

Fudenberg, Drew and Kreps, David M. "A Theory of Learning, Experimentation, and Equilibrium in Games." Mimeo, MIT, 1988.

Gibbons, Robert. "Piece-Rate Incentive Schemes." *Journal of Labor Economics*, October 1987, *5*(4), pp. 413–29.

Gick, Mary L. and Holyoak, Keith J. "Analogical Problem Solving." *Cognitive Psychology*, July 1980, *12*(3), pp. 306–55.

Glaser, Robert. "Expertise," in Michael W. Eysenck, Andrew Ellis, and Earl B. Hunt, eds., *The Blackwell dictionary of cognitive psychology*. Oxford: Blackwell, 1990, pp. 139–42.

Godden, Duncan R. and Baddeley, Alan D. "Context-Dependent Memory in Two Natural Environments: On Land and Under Water." *British Journal of Psychology*, August 1975, *66*(3), pp. 325–31.

Granick, David. "The Ministry as the Maximizing Unit in Soviet Industry." *Journal of Comparative Economics*, September 1980, *4*(3), pp. 255–73.

_____. "The Ministry and the Ratchet: Response to Keren." *Journal of Comparative Economics*, December 1983, *7*(4), pp. 432–48.

Grossman, Sanford J. and Perry, Motty. "Perfect Sequential Equilibrium." *Journal of Economic Theory*, June 1986, *39*(1), pp. 97–119.

Ickes, Barry W. and Samuelson, Larry. "Job Transfers and Incentives in Complex Organizations: Thwarting the Ratchet Effect." *Rand Journal of Economics*, Summer 1987, *18*(2), pp. 275–86.

Kachelmeier, Steven J. and Shehata, Mohamed. "Examining Risk Preferences Under High Monetary Incentives: Experimental Evidence from the People's Republic of China." *American Economic Review*, December 1992, *82*(5), pp. 1120–144.

Keren, Michael. "The Ministry, Plan Changes and the Ratchet in Planning." *Journal of Comparative Economics*, December 1982, *6*(4), pp. 327–42.

Klahr, David and Carver, Sharon M. "Cognitive Objectives in a LOGO Debugging Curriculum: Instruction, Learning, and Transfer." *Cognitive Psychology*, July 1988, *20*(3), pp. 362–404.

Kroll, Yoram; Levy, Haim and Rapoport, Amnon. "Experimental Tests of the Separation Theorem and the Capital Asset Pricing Model." *American Economic Review*, June 1988, *78*(3), pp. 500–19.

Kurland, D. M.; Pea, Roy D.; Clement, Catherine A. and Mawby, Ronald. *A Study of the development of programming ability and thinking skills in high school students*. New York: Bank Street College of Education, Center for Children and Technology, 1986.

Laffont, Jean-Jacques and Tirole, Jean. *A theory of incentives in procurement and regulation*. Cambridge: MIT Press, 1993.

Partow, Zeinab and Schotter, Andrew. "Does Game Theory Predict Well for the Wrong Reasons: An Experimental Investigation." C. V. Starr Center for Applied Economics Research Report No. 93-46, New York University, 1993.

Pea, Roy D. and Kurland, D. M. "On the Cognitive Effects of Learning Computer Programs." *New Ideas in Psychology*, 1984, *2*(2), pp. 137–68.

Perkins, David N. and Salomon, Gabriel. "Teaching for Transfer." *Educational Leadership*, September 1988, *46*(1), pp. 22–32.

Perkins, David N.; Schwartz, S. and Simmons, R. "Instructional Strategies for the Problems of Novice Programmers," in Richard E. Mayer, ed., *Teaching and learning computer programming: Multiple research perspectives.* Hillsdale, NJ: Erlbaum, 1988, pp. 153–78.

Roth, Alvin E. and Erev, Ido. "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term." *Games and Economic Behavior*, January 1995, *8*(1), pp. 164–212.

Salomon, Gabriel and Perkins, David N. "Rocky Roads to Transfer: Rethinking Mechanisms of a Neglected Phenomenon." *Education Psychologist*, Spring 1989, *24*(2), pp. 113–42.

Simon, D. P. and Simon, Herbert A. "Individual Differences in Solving Physics Problems," in R. Siegler, ed., *Children's thinking: What develops?* Hillsdale, NJ: Erlbaum, 1978, pp. 325–48.

Smith, Vernon L. and Walker, James M. "Rewards, Experience and Decision Costs in First Price Auctions." *Economic Inquiry*, April 1993a, *31*(2), pp. 237–45.

_____. "Monetary Rewards and Decision Cost in Experimental Economics." *Economic Inquiry*, April 1993b, *31*(2), pp. 245–61.

Stahl, Dale. "Evolution of Smartn Players." *Games and Economic Behavior*, October 1993, *5*(4), pp. 604–17.

Walder, Andrew G. "Factory and Manager in an Era of Reform." *China Quarterly*, June 1989, *18*(1), pp. 242–64.

Weitzman, Martin L. "The 'Ratchet Principle' and Performance Incentives." *Bell Journal of Economics*, Spring 1990, *11*(1), pp. 302–08.

Yao, Dennis A. "Strategic Responses to Automobile Emissions Control: A Game-Theoretic Analysis." *Journal of Environmental Economics and Management*, December 1988, *15*(4), pp. 419–38.

Zeitz, Colleen and Glaser, Robert. "Expert Level of Understanding," in Torsten Husen and T. N. Postelthwaite, eds., *The international encyclopedia of education*, 2nd Ed. Oxford: Pergamon Press, 1994, pp. 2194–99.