세무와회계저널
제12권 제2호 2011년 6월 pp.229～256
한국세무학회

Strategic Thinking in Tax Reporting

Chung Kweon Kim*

<Abstract>

Behavioral game theory experimentally investigates strategic interaction by combining insights from game－theoretic models and models of human information processing. Given limited rationality, what level of strategic thinking do individuals achieve in a tax reporting game? The experiment in this study reveals that tax auditors and firms adjust their strategy based on a belief about the opponent's response to parametric change, but do not correctly recognize that the opponent does likewise, indicating that tax auditor's strategic decision is biased by normatively irrelevant information.

<Key words> Game theory, Tax compliance, Limited rationality, Tax auditing, Strategic thinking

Ⅰ. Introduction

This study investigates strategic behavior in an experimental audit setting that corresponds to a game－theoretic model of tax compliance (Graetz et al. 1986). The game－theoretic model derives the players' optimal strategies at equilibrium given a common knowledge assumption of individual rationality. Experimentation complements the mathematical modeling by systematically observing strategic behavior when this assumption does not hold (Camerer 2003). Recognizing that humans normally cannot work through the higher levels of strategic thinking in game－theoretic models, my study demonstrates a procedure for inferring the level of strategic thinking that subjects actually achieve. I assign subjects to the roles of an auditor (*verifier*) and auditee (*reporter*) with monetary incentives based on the model's payoffs and repeated trials to allow for

* Professor, College of Business Administration, Ewha Womans University, acckim@ewha.ac.kr

논문접수일 2011년 4월　　1차수정일 2011년 5월　　게재확정일 2011년 5월

learning. I manipulate a model parameter and observe the subjects' responses in terms of both their strategy decision and their belief about the opponent's strategy. The model parameter is the base rate for unethical reporters who under-report income given the incentive to do so. The observed response pattern allows me to infer the achieved level of strategic thinking.

The "level of strategic thinking" issue is important for descriptive and prescriptive reasons. Many accounting settings involve non-cooperative strategic interaction, e.g., tax auditor-taxpayer and external auditor-manager. Although formal game theory provides insight into strategic behavior in such settings, the empirical validity of the theoretical results depends on the assumption noted above. If individuals fail to act in accord with the assumption, then it is important to understand their strategic decision biases and the implications for accounting institutions and policy formation. Institutional features sometimes mitigate strategic decision biases by helping individuals to anticipate and incorporate the responses of others (Haynes and Kachelmeier 1998 ; Plott 1996). In addition, an understanding of strategic decision biases can inform policy that improves the functioning of accounting institutions (Kachelmeier and King 2002 ; Wilks and Zimbelman 2004). For example, a goal of audit policy makers is to guide practitioners in identifying and processing normatively relevant information when assessing and responding to the risk of intentional misstatement.

The experiment in this study is a horserace to determine which of four levels of strategic thinking is most consistent with observed behavior. At one extreme, *sophisticated strategic thinking* means that subjects mimic the game-theoretic model when responding to parametric change, adjusting their strategy and belief about the opponent's strategy in the direction that the model implies (level 3). At the other extreme, a *base-rate fallacy* reflects the crudest form of reasoning : subjects are insensitive to parametric change (level 0). Between the extremes, I distinguish two levels of strategic thinking. *Non-strategic thinking* means that subjects consider how contextual factors directly affect them, but fail to consider how contextual factors indirectly affect them through the opponent's action (level 1). Non-strategic thinkers implicitly assume that the opponent is a level 0 thinker or mindless part of the context. *Simple strategic thinking* means that subjects adjust their strategy based on a belief, or "other mind" inference, about the opponent's response to parametric change, but do not correctly recognize that the opponent also makes a belief-based adjustment (level 2). Correct recognition amounts to sophisticated strategic thinking. Instead, simple strategic thinkers implicitly assume that the opponent is a level 1 thinker.[1)]

1) The four levels are mutually exclusive, but not collectively exhaustive. Arguably, there are infinitely many steps between levels 2 and 3, although past studies in experimental economics suggest that level

I base my inferences about the level of strategic thinking on aggregate－and individual－level analyses. The aggregate－level analysis indicates that both verifiers and reporters achieve simple strategic thinking. As the base rate for unethical reporters increases (from 0.25 to 0.50 to 0.75), the verifiers generally respond with a higher estimate about the reporters' rate of under－reporting and with a higher rate of investigating reports of low income. The reporters generally respond with a higher estimate about the verifiers' rate of investigation and with a lower rate of under－reporting. These base－rate effects are stronger for reporters than for verifiers. The individual－level analysis indicates that the role of reporter versus verifier moderates the achieved level of strategic thinking. The majority of reporters achieve simple strategic thinking, whereas the verifiers are more evenly split between non－strategic and simple strategic thinkers. No subject in either role achieves sophisticated strategic thinking.[2)]

I also analyze the consequent biases in the subjects' decisions and beliefs. A normative implication of the game－theoretic model is that the verifiers should choose a constant rate of investigation without adjusting for changes in the base rate, whereas the reporters should adjust their rate of under－reporting inversely with changes in the base rate. In other words, the reporters absorb the full impact of the parametric change. Instead, the verifiers commit a strategic decision bias by conditioning the rate of investigation on the base rate and consequently under－auditing when the base rate is low. For their part, the reporters cheat too much at all levels of the base rate. Regarding beliefs, the reporters correctly believe that the verifiers increase the rate of investigation as the base rate increases, whereas the verifiers incorrectly believe that the reporters increase the rate of under－reporting as the base rate increases. The verifiers would be better off if they ignored changes in the base rate.

Ⅱ. Background

1. Strategic decision biases and experimentation

Past research distinguishes three perspectives on judgment and decision making in accounting (Ashton 1983). Descriptively, the goal is to observe behavior and provide explanations in terms

2 is the practical upper limit in many games (Camerer, 2003).

2) It is important to distinguish achieved versus achievable levels of strategic thinking. An expert matched with a novice might choose a level of strategic thinking that is only one step above the novice's level even though the expert could achieve higher levels.

of underlying processes. Normatively, the goal is to evaluate behavior relative to a formal model that assumes individual rationality. Prescriptively, the goal is to suggest means for improving judgments and decisions. The perspectives interrelate in the following way. Experimenters can produce observations of “bias” by comparing judgments and decisions with the results of a relevant normative model. To obtain many observations over contrasting conditions, experimenters manipulate a model parameter and assess whether subjects are properly sensitive to the parametric change. The observed bias provides a basis for both inferences about underlying processes and prescriptions that de－bias future judgments and decisions. This interrelationship is central to the heuristic－and－bias paradigm in psychology (Kahneman et al. 1982) and related research in accounting (Ashton and Ashton 1995).

Experimenters can take a similar approach to examine judgment and decision making in non－cooperative strategic settings (Camerer 2003). But, because strategic settings involve interaction between individuals with conflicting interests, additional considerations are necessary. One consideration is the relevant normative model. The heuristic－and－bias paradigm employs normative models that assume a rational individual in isolation. For strategic settings, game－theoretic models that derive the best－response strategies at equilibrium are the appropriate benchmark. Such models go beyond the assumption of a rational individual and make a common knowledge assumption about rationality (i.e., each player is rational, each player knows that the opponent is rational, and so on). Another consideration is that one player’s strategic decision bias is likely to affect the opponent’s strategic decision. Each player makes an “other mind” inference about the opponent, and suspected decision biases are part of the inference. These considerations raise a question for game theory : what are the implications for strategic interaction when common knowledge of individual rationality does not hold, specifically, when at least one player commits a strategic decision bias? Experiments shed light on this question by observing strategic decision making by humans in laboratory games.

2. Level of strategic thinking

Many studies in experimental economics examine subjects’ achieved level of strategic thinking. An early study by Nagel (1995) observes subjects’ behavior in a guessing game. Each subject in a group picks a number between 0 and 100. The experimenter computes a target number equal to the group’s mean response multiplied by a pre－announced proportion. The subject whose number is closest to the target number wins a prize. Assume that the pre－announced proportion is 0.80. If the group's mean response is 80, then the prize goes to the subject whose number is closest to 64 (i.e., 0.80 × 80). What number should subjects pick? What number do subjects pick? If subjects pick a number at random, then the mean response is 50 (*zero−order thinking*). If a subject believes that other subjects employ zero−order thinking, then the subject should pick 40, or 0.80 × 50 (*first−order thinking*). If a subject believes that other subjects employ first−order thinking, then the subject should pick 32, or 0.80 × 40 (*second−order thinking*). At the highest level of strategic thinking, subjects should pick 0, which is the theoretical best response at equilibrium. Using this procedure, Nagel (1995) infers that subjects generally achieve between first−and second−order thinking.

Following Nagel (1995), many studies infer subjects' level of strategic thinking in dominance−solvable games (Camerer 1997 ; Camerer et al. 2004 ; Duffy and Nagel 1997 ; Ho et al. 1998 ; Nagel, 1999 ; Ohtsubo and Rapoport 2006 ; Rapoport and Amaldoss 2000 ; Slonim 2005). Imagine a hierarchy of k levels or steps of strategic thinking (k = 1, 2, ⋯, n). Returning to the guessing game, assume a pre−announced proportion of 0.80 and an initial interval of [0,100]. Taking the first step, a response of 80 (i.e., 0.80 × 100) or below dominates responses above 80 (k = 1). Taking additional steps, a response of 64 (i.e., 0.80 × 80) or below dominates responses above 64 (k = 2), a response of 51 (i.e., 0.80 × 64) or below dominates responses above 51 (k = 3), and so on to the theoretical best response of 0. Relying on the principle of iterative dominance, experimenters infer that a response between 65 and 80 reflects first−order thinking, a response between 52 and 64 reflects second−order thinking, and so on. Summarizing past studies, Ohtsubo and Rapoport (2006) conclude that subjects generally achieve between first − and second−order thinking, but there often are differences across individuals and games. Subjects who achieve second−order thinking in the guessing game might achieve higher−order thinking in another game.

3. Strategic thinking and economic institutions

Economists address issues of limited rationality, not merely as psychological phenomena, but because of the implications for the design and functioning of economic institutions. Drawing from experimental economics, Plott (1996) defines rationality as a property of behavior that emerges in three stages. In the first stage, myopia dominates as subjects struggle with unfamiliar tasks, and experimenters observe inconsistencies in subjects' preferences. In the second stage, with monetary incentives and repeated trials, myopia gives way to more considered and stable choice behavior.

In the third stage, social institutions sometimes help interacting subjects to anticipate and incorporate the rationality (or biases) of others into their individual decisions. In "lemons" markets, for example, warranties offered by sellers help buyers to incorporate sellers' intentions into buying decisions. Plott (1996) stresses, however, that the third stage is problematic, and cites examples (including experimental games with mixed strategies) in which subjects fail to consider the rationality of others despite monetary incentives and repeated trials.

Experiments on the guessing game illustrate the "other mind" problem for economic institutions. The guessing game is a stylization of John Maynard Keynes' discussion about the similarity of strategic thinking in the stock market and a newspaper beauty contest in which the readers guess which of several faces ranks as most beautiful in the opinion of readers :

> It is not a case of choosing those which, to the best of one's judgment, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest. We have reached the third degree, where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practice the fourth, fifth, and higher degrees (Keynes 1936, 156).

In the guessing game, suppose that a subject knows that the theoretical best response is zero, but also believes that other subjects in the group lack this knowledge. Based on this belief, the knowledgeable subject reasonably picks a response greater than zero, i.e., intentionally commits a strategic decision bias relative to the normative benchmark. Even with monetary incentives and repeated trials, subjects rarely pick the theoretical best response (Camerer 2003). In the stock market, a smart trader who knows a stock's intrinsic value, but also believes that other noise traders commit decision biases, reasonably distorts portfolio decisions to incorporate noise trader risk (Shleifer 2000).

4. Strategic thinking and past auditing experiments

Only two past auditing experiments address strategic thinking when common knowledge of individual rationality does not hold. Bloomfield (1997) examines the sources of inaccurate risk assessments in a simple audit game. The auditee potentially misstates an asset value, and the auditor accepts or rejects the asserted asset value. The sources of inaccuracy relate to *strategic dependence*, i.e., the degree to which the auditor's belief about the auditee's strategy affects the auditee's strategy. Each subject makes a strategy decision and estimates the opponent's strategy. The subjects play four game variations that cross two factors, i.e., rate of unintentional misstatements and auditor's liability ratio. The inaccuracy and variability of subjects' risk assessments increase as the manipulated factors increase strategic dependence. Zimbelman and Waller (1999) examine strategic interaction in a more complex audit game. Given private information, the auditee asserts an asset's value. The auditor collects a sample and decides whether to accept or reject the asserted asset value. The authors hypothesize that subjects consider factors that directly affect them (e.g., auditees consider their payoff for undetected misstatement) and factors that directly affect the opponent (e.g., auditors consider the auditee's payoff for undetected misstatement), but cannot achieve higher－order thinking. The authors manipulate such factors and interpret the main and interactive effects to test their hypotheses.

My study demonstrates a new procedure for inferring subjects' achieved level of strategic thinking. Unlike Nagel (1995) and other studies in experimental economics that rely on dominance－solvable games, my procedure more generally applies to any game in which the players potentially adjust their strategy based on a belief about the opponent's strategy.[3] This advantage is important in settings where iterative dominance fails to identify the unique best response. Like Bloomfield (1997), I collect data for both the subjects' strategy decision and their belief about the opponent's strategy, but my main concern is with identifying the achieved level of strategic thinking rather than the sources of inaccurate risk assessment. Like Zimbelman and Waller (1999), I distinguish direct and indirect effects, but my procedure employs a game－theoretic model as a normative benchmark. This step is necessary to establish expectations about subjects' behavior in the absence of any strategic decision bias (Bloomfield 1999). Unlike past experiments, this study involves a tax audit game in which reporters potentially understate income rather than an external audit game in which auditees potentially overstate income or an asset value.

Ⅲ. Audit game setting and hypotheses

1. Setting

The setting involves the strategic interaction of two players, a reporter and a verifier. Reporter

3) Another limitation of the guessing game concerns the starting point for a chain of reasoning. Assume a pre－announced proportion of 0.80 and an initial interval of [0,100]. Is first－order thinking validly measured by a response of 40 (relative to zero－order thinking that implies a mean of 50) or by a response between 65 and 80 (relative to zero－order thinking that fails to eliminate dominated responses above 80)?

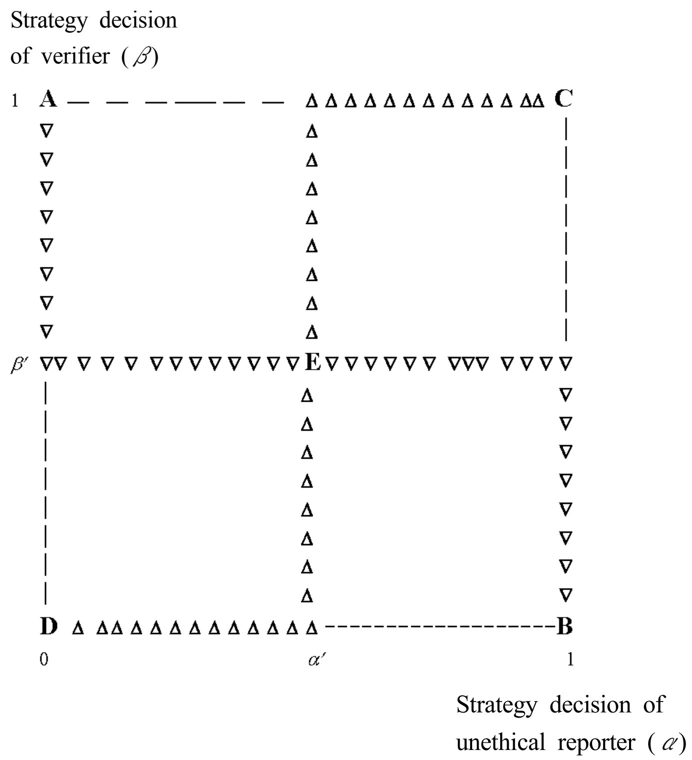
and verifier are the terms that I use in my experiment. Strategic taxpayer and tax auditor are the corresponding terms that Graetz et al. (1986) use in their model. The reporter is either of two types : (1) unethical reporters who under-report income given the incentive to do so, and (2) ethical reporters who never under-report income. The benefit of under-reporting income, if undetected by the verifier, is a lower tax. The cost of under-reporting income, if detected by the verifier, is a fine plus a higher tax. There is no incentive to over-report income.

Let ρ denote the base rate for unethical reporters, and $1-\rho$ denote the base rate for ethical reporters. The base rate is common knowledge, although the verifier does not know any reporter's type. The reporters know their own type and whether their true income is high or low. If true income is low, then all reporters report accurately. If true income is high, then the reporting strategy depends on the reporter type. The ethical reporters always report high income accurately. The unethical reporters under-report income with a probability, α. In the game-theoretic model, α is an unethical reporter's mixed strategy for under-reporting income when true income is high.

After observing reported income, the verifier decides whether to conduct a costly investigation that perfectly reveals true income. If reported income is high, then the verifier has no reason to investigate. If reported income is low, then the verifier's net gain from investigating depends on the investigation cost, the probability that true income is high, and the fine plus higher tax when true income is high. The verifier investigates a report of low income with a probability, β. In the game-theoretic model, β is the verifier's mixed strategy for investigating whether a report of low income is accurate.

The optimal strategies of the verifier and unethical reporter are interdependent. Figure 1 shows each player's best responses to the opponent's strategies. At one extreme, if the verifier always investigates when reported income is low, then the unethical reporter's best response is never to under-report (point A). At the other extreme, if the verifier never investigates when reported income is low, then the unethical reporter's best response is always to under-report (point B). Generally, if the verifier's strategy is greater (less) than β', then the unethical reporter's best response is never (always) to under-report. The downward-facing triangles that connect points A and B depict the unethical reporter's best-response function.

〈Figure 1〉 Best responses



1) The figure shows the best−response functions of the unethical reporter (▽) and verifier (△). The players' best−response functions intersect at the equilibrium strategies, α' and β'. The figure highlights four disequilibrium points (A, B, C, D) and the equilibrium point (E).

From the verifier's perspective, if the unethical reporter always under−reports when true income is high, then the verifier's best response is always to investigate a report of low income (point C). At the other extreme, if the unethical reporter never under−reports, then the verifier's best response is never to investigate (point D). Generally, if the unethical reporter's strategy is greater (less) than α', then the verifier's best response is always (never) to investigate a report of low income. The upward−facing triangles that connect points C and D depict the verifier's best−response function.

The players' best−response functions intersect at point E : α' is the unethical reporter's optimal reporting strategy, and β' is the verifier's optimal investigation strategy. The game−theoretic model derives the optimal strategies as functions of the following parameters :

$$\alpha' = C\ (1\ -\ P)\ /\ \rho\ P\ (F\ +\ T\ -\ C) \tag{1}$$

$$\beta' = T\ /\ (F\ +\ T) \tag{2}$$

where C is the verifier's investigation cost, P is the prior probability of high versus low true income, F is the fine for detected under-reporting, T is the tax difference for high versus low income, and ρ is the base rate for unethical reporters.

Table 1 shows the players' optimal strategies for the conditions that I examine in the experiments. I manipulate the base rate for unethical reporters at three levels (0.25, 0.50, 0.75), holding constant the verifier's investigation cost (10 francs), fine for detected under-reporting (50 francs), prior probability of high income (50 percent), and tax difference for high versus low income (50 francs). Francs are the experimental currency. I use these parameter values to simplify the experimental task and equate the expected payoff for the roles of verifier and reporter. Entering the parameter values into Equation 1, α' equals 0.44 when ρ is 0.25, 0.22 when ρ is 0.50, and 0.15 when ρ is 0.75. Entering the parameter values into Equation 2, β' equals 0.50. In Experiment 2, I also manipulate the verifiers' knowledge about the base rate (see below).

〈Table 1〉 Optimal strategies given experimental parameter values

Base rate for reporters (ρ)	Reporter's optimal strategy (α')	Verifier's optimal strategy (β')
0.25	0.44	0.50
0.50	0.22	0.50
0.75	0.15	0.50

The game-theoretic model of Graetz et al. (1986) has attractive features. The model is a seminal contribution to the tax compliance literature (Andreoni et al. 1998). Recognizing that some ethical taxpayers do not cheat despite an economic incentive to do so, a model parameter is the base rate for unethical reporters. I choose to manipulate this parameter as part of my procedure for inferring the level of strategic thinking, because psychology contains many studies that examine base-rate effects, but in single-person settings (Kahneman et al. 1982 ; Koehler 1996). Finally, by modifying the payoff structure, the model generalizes to external audit settings (Jones et al. 2004).

2. Hypotheses for verifiers

The hypotheses for the verifiers state four mutually exclusive response patterns (Table 2). The crudest form of reasoning is a *base−rate fallacy* : the verifiers' strategy decision and belief about the reporters' strategy are insensitive to variation in the base rate. Many past studies in psychology (Kahneman et al. 1982 ; Koehler 1996) and accounting (Butt 1988 ; Ganguly et al. 1994 ; Joyce and Biddle 1981) demonstrate that individuals are not sufficiently sensitive to variation in base rates, relative to a Bayesian model. Tversky and Kahneman (1984) propose that individuals make probability judgments (e.g., what is the probability that a case belongs to one category versus another category?) by assessing the similarity between the case and categories. Although easy to employ, this heuristic ignores variation in the category base rates. If the verifiers employ a similar form of reasoning, then I expect no base−rate effect on the verifiers' strategy decision or their belief about the reporters' strategy.4)

H1a. There is no base−rate effect on the verifiers' strategy or their belief about the reporters' strategy.

〈Table 2〉 Alternative response patterns for verifiers (Hypotheses H1a−d)

Levels of strategic thinking	Base−rate effect on :	
	Verifiers' strategy (β)	Verifiers' belief about reporters' strategy ($E[\alpha]$)
Base−rate fallacy (H1a)	None	None
Non−strategic thinking (H1b)	Positive	None
Simple strategic thinking (H1c)	Positive	Positive
Sophisticated strategic thinking (H1d)	None	Negative

Non−strategic thinking means that players consider how contextual factors directly affect them, but fail to consider how contextual factors indirectly affect them through the opponent's actions. Non−strategic thinkers treat a strategic setting as if it were a single−person setting. When

4) H1a is not exactly the same as the base−rate fallacy in psychological studies. H1a refers to the effect of base rates on the verifiers' strategy or belief about the reporters' strategy, and not to the effect of base rates on posterior probability judgments. Further, psychological studies reveal considerable variation in base−rate usage depending on the experimental task and setting (Hertwig and Ortmann, 2001 ; Koehler, 1996).

variation in base rates is a contextual factor, non－strategic thinkers implicitly assume that the opponent mindlessly commits a base－rate fallacy. If the verifiers employ this form of reasoning, then I expect a positive base－rate effect on the verifiers' strategy, and no base－rate effect on the verifiers' belief about the reporters' strategy. The verifiers increase their rate of investigation, simply because of an increase in the proportion of potential cheaters in the population. The verifiers fail to consider how the reporters adjust the rate of under－reporting when the base rate varies.

H1b. There is a positive base－rate effect on the verifiers' strategy and no base－rate effect on the verifiers' belief about the reporters' strategy.

Simple strategic thinking means that players form a belief about how the opponent responds to a contextual factor, and adjust their strategy on the basis of this belief. This form of reasoning is strategic, because it recognizes the interdependence of the players' strategies. This form of reasoning is simple, however, because it does not correctly recognize that the opponent also makes a belief－based adjustment. Simple strategic thinkers implicitly assume that the opponent is a non－strategic thinker. If the verifiers employ this form of reasoning, then I expect positive base－rate effects on the verifiers' strategy and belief about the reporters' strategy. This response pattern involves a norm for reporters : a cheater cheats more when there are more cheaters. The verifiers believe that the rate of under－reporting by each unethical reporter increases as the base rate for such reporters increases, and increase the rate of investigation on the basis of this belief.

H1c. There are positive base－rate effects on the verifiers' strategy and belief about the reporters' strategy.

Sophisticated strategic thinking means that players mimic the game－theoretic model when responding to parametric change. Although sophisticated strategic thinkers do not necessarily analyze the game as a theorist would, they nevertheless adjust their strategy and belief about the opponent's strategy in the direction that a theorist would.[5] Graetz et al. (1986) explain that an increase in the base rate for unethical reporters has only one effect in their model : to decrease the optimal reporting strategy for an unethical reporter. As the base rate increases, a report of low income is more likely to come from an unethical reporter than from an ethical reporter with low

5) This form of reasoning leaves room for bias. A player might specify a mixed strategy that is higher or lower than the optimal mixed strategy in a game－theoretic model.

true income. Although there are more potential cheaters, each cheater is more likely to report honestly in equilibrium. These effects exactly cancel each other, so that the aggregate number of reporters who under−report does not change and the verifier's optimal rate of investigating a report of low income does not change. If the verifiers employ this form of reasoning, then I expect no base−rate effect on the verifiers' strategy and a negative base−rate effect on their belief about the reporters' strategy. This response pattern corresponds to the optimal strategies in Table 1.

H1d. There is no base−rate effect on the verifiers' strategy and a negative base−rate effect on their belief about the reporters' strategy.

3. Hypotheses for reporters

Table 3 summarizes the alternative response patterns for the reporters, which mirror those for the verifiers. The crudest form of reasoning leads to a base−rate fallacy. If the reporters employ a similar form of reasoning, then I expect no base−rate effect on the reporters' strategy or their belief about the verifiers' strategy.

H2a. There is no base−rate effect on the reporters' strategy or their belief about the verifiers' strategy.

〈Table 3〉 Alternative response patterns for reporters (Hypotheses H2a−d)

Levels of strategic thinking	Base−rate effect on :	
	Reporters' strategy (α)	Reporters' belief about verifiers' strategy ($E[\beta]$)
Base−rate fallacy (H2a)	None	None
Non−strategic thinking (H2b)	Positive	None
Simple strategic thinking (H2c)	Negative	Positive
Sophisticated strategic thinking (H2d)	Negative	None

The second level is non−strategic thinking. If the reporters employ this form of reasoning, then I expect a positive base−rate effect on the reporters' strategy and no base−rate effect on their belief about the verifiers' strategy. This pattern involves the norm for reporters discussed above. Each unethical reporter increases the rate of under−reporting as the base rate for such reporters increases. However, this form of reasoning fails to consider that the verifiers adjust their rate of investigation as the base rate varies.

H2b. There is a positive base－rate effect on the reporters’ strategy and no base－rate effect on their belief about the verifiers’ strategy.

The third level is simple strategic thinking. If the reporters employ this form of reasoning, then I expect a negative base－rate effect on the reporters’ strategy and a positive base－rate effect on their belief about the verifiers’ strategy. The reporters who are simple strategic thinkers believe that the verifiers increase the rate of investigation as the base rate for unethical reporters increase, and inversely adjust the rate of under－reporting on the basis of this belief.

H2c. There is a negative base－rate effect on the reporters’ strategy and a positive base－rate effect on their belief about the verifiers’ strategy.

The fourth level is sophisticated strategic thinking. If the reporters employ this form of reasoning, then I expect a negative base－rate effect on the reporters’ strategy and no base－rate effect on their belief about the verifiers’ strategy. This response pattern corresponds to the optimal strategies in Table 1.

H2d. There is a negative base－rate effect on the reporters’ strategy and no base－rate effect on their belief about the verifiers’ strategy.

Ⅳ. Experiment

1. Method

A. Overview

The experiment operationalizes the game setting described in Section 3. The subjects are 40 undergraduate students from a major university in Hong Kong. They are 2nd and 3rd year business major undergraduate students.[6] The data are from two sessions, each with 12 subjects in the role of reporter and 8 subjects in the role of verifier.[7] Each session consists of 20 rounds.

6) The subjects are 2nd－ and 3rd－year undergraduate students mostly in business major and consist of 22 female students and 18 male students.

I manipulate the base−rate for unethical reporters at three levels (0.25, 0.50, 0.75) in random order over the 20 rounds. The procedure also randomly and anonymously re−pairs reporters and verifiers round−by−round in order to minimize reputation formation.

The experiment uses a computer network. The computer assigns subjects to the role of reporter or verifier. The instructions tell subjects about the task, incentives, and information for each role. Each subject receives 120 Hong Kong dollars (equivalent to US$15) for completing the experiment plus the outcome of a lottery with a cash prize of 120 Hong Kong dollars (equivalent to US$15). Each subject's chance of winning the prize is a linear function of earned francs. The expected chance of winning the prize is the same for verifiers and reporters, based on the optimal strategies in Table 1. For subjects who maximize the expected chance of winning the prize, these monetary incentives are consistent with the risk−neutrality assumption in Graetz et al. (1986).

B. Procedure for round

The procedure consists of eight steps in each round. The subjects provide their responses in steps 3 and 4. The computer executes all remaining steps. In the first two rounds, the subjects have a maximum of four minutes to enter their responses. In later rounds, the subjects have a maximum of 1.5 minutes to enter their responses.

Step 1. The computer determines the base rate for reporters. The procedure varies the base rate without changing the number of active subjects by using *automatic reports*. The instructions state that there are a number of automatic reports in addition to the 12 reports from subjects in the role of reporter. The number of automatic reports varies among three equally likely values (4, 12, 36). Reported income always matches actual income in an automatic report. In contrast, reported income in a report from a reporter can be lower than actual income depending on the reporter's strategy decision. The instructions include a schedule detailing the three possibilities (Table 4).

〈Table 4〉 Schedule showing base rate for reporters (ρ)

Number of reports from reporters	Number of automatic reports	Proportion of reports from reporters (ρ)	Chance of occurrence
12	36	25%	1/3
12	12	50%	1/3
12	4	75%	1/3

7) The data are part of a larger data set. A report is available from the author.

Step 2. The computer assigns all reports to the eight verifiers, without disclosing the assignment to any subject. The computer assigns 6 reports to each verifier when the number of automatic reports is 36 (i.e., 6 = [12 + 36] / 8), 3 reports to each verifier when the number of automatic reports is 12 (i.e., 3 = [12 + 12] / 8), and 2 reports to each verifier when the number of automatic reports is 4 (i.e., 2 = [12 + 4] / 8). The instructions state that the verifiers do not know whether the reports assigned to them are automatic reports or reports from reporters without conducting an investigation for each report, and that the reporters do not know which of the verifiers have their report.

Step 3. The verifiers provide responses to two questions. One question asks each verifier to assume that a reporter's actual income is 100 francs and to estimate the probability that the reporter reports income of 0 franc. This response measures the verifier's belief about the reporter's strategy ($E[\alpha]$). The other question asks each verifier to assume that reported income is 0 franc and to state the probability of conducting an investigation for this report. The latter response measures the verifier's strategy (β).

Step 4. The reporters provide responses to two questions. One question asks each reporter to assume that a verifier receives a report of 0 franc and to estimate the probability that the verifier conducts an investigation. This response measures the reporter's belief about the verifier's strategy ($E[\beta]$). The other question asks each reporter to assume that their actual income is 100 francs and to state the probability that they report income of 0 franc. The latter response measures the reporter's strategy (α). Steps 3 and 4 take place simultaneously.

Step 5. The computer determines whether actual income is 0 or 100 francs, separately for each reporter and automatic report. There is a 50-50 chance for the two levels of income.

Step 6. The computer determines the amount of reported income, separately for each reporter. If a reporter's actual income is 0 franc, then the computer always produces a report of 0 franc. If a reporter's actual income is 100 francs, then the computer applies the reporter's strategy response from Step 4. Suppose that a reporter's strategy response is 80 percent. The computer implements a chance device with a probability of 0.80 of reporting income of 0 franc and a probability of 0.20 of reporting income of 100 francs.

Step 7. The computer determines whether the verifier conducts an investigation, separately for each report. If reported income is 100 francs, then the computer always produces the decision not to investigate. If reported income is 0 franc, then the computer applies the verifier's strategy response from Step 3. Suppose that a verifier's strategy response is 25 percent. For each report assigned to the verifier, the computer implements a chance device with a probability of 0.25 of an investigation and a probability of 0.75 of no investigation.

Step 8. The computer tallies the results for the round and provides feedback to the subjects. The instructions include a schedule detailing each player's payoff under all possible scenarios (Table 5).

〈Table 5〉 Schedule showing the players' payoffs in experimental currency(in francs)

Reporter's true income	Reported Income	Is there an investigation?	Reporter's payoff	Verifier's payoff
0	0	No	0	0
0	0	Yes	0	－10
100	100	No	50	50
100	0	No	100	0
100	0	Yes	0	90

C. Computer screen

Each subject's screen displays an information window, a history window, and a message window. The information window shows the conditions that are in effect for the round. The history window shows the results for prior rounds. The message window includes the two questions described earlier (Step 3 for verifiers or Step 4 for reporters) and spaces for providing responses.

2. Results

A. Descriptive statistics

Table 6 presents descriptive statistics for the subjects' strategy decisions and beliefs. The verifiers' mean rate of investigation increases from 0.34 to 0.45 to 0.51, as the base rate increases from 0.25 to 0.50 to 0.75, indicating a positive base－rate effect on the verifiers' strategy decisions. The reporters' mean rate of under－reporting decreases from 0.67 to 0.46 to 0.33, as the base rate increases from 0.25 to 0.50 to 0.75, indicating a negative base－rate effect on the reporters' strategy decisions.[8] The verifiers' mean belief about the reporters' strategy increases from 0.63 to 0.66 to 0.74, as the base rate increases from 0.25 to 0.50 to 0.75, indicating a

8) No subject acted like an ethical reporter who always sets a zero rate of under－reporting.

positive base−rate effect on the verifiers' belief about the reporters' strategy. The reporters' mean belief about the verifiers' strategy increases from 0.34 to 0.54 to 0.68, as the base rate increases from 0.25 to 0.50 to 0.75, indicating a positive base−rate effect on the reporters' belief about the verifiers' strategy.

〈Table 6〉 Descriptive statistics

	Verifiers (n = 16)		Reporters (n = 24)	
Base rate for reporters (ρ)	Strategy (β)	Belief ($E[\alpha]$)	Strategy (α)	Belief ($E[\beta]$)
0.25	0.34 (0.20) [−0.16]	0.63 (0.19) [−0.04]	0.67 (0.24) [0.23]	0.34 (0.17) [0.00]
0.50	0.45 (0.22) [−0.05]	0.66 (0.11) [0.20]	0.46 (0.25) [0.24]	0.54 (0.12) [0.09]
0.75	0.51 (0.21) [0.01]	0.74 (0.09) [0.41]	0.33 (0.27) [0.18]	0.68 (0.17) [0.17]

1) This table shows the mean (standard deviation) for the subjects' strategy decision and belief about the opponent's strategy. The table also shows in brackets the mean bias in the subjects' strategy decisions and beliefs. The strategic decision bias is the signed difference between a subject's strategy decision and the corresponding game−theoretic strategy. The belief bias is the signed difference between a subject's belief about the opponent's strategy and the strategy selected by the opponent.

Table 6 shows in brackets the mean bias in the subjects' strategy decisions and beliefs. The strategic decision bias is the signed difference between the subject's strategy decisions and the corresponding game−theoretic strategy. The belief bias is the signed difference between the subject's belief about the opponent's strategy and the strategy selected by the opponent. Overall, the verifiers' mean strategic decision bias is −0.07, which indicates under−auditing. This bias is most extreme when the base rate is low (−0.16). Overall, the verifiers' mean belief bias is 0.19, which indicates over−estimation of the reporters' rate of under−reporting. This bias is most extreme when the base rate is high (0.41). Overall, the reporters' mean strategic decision bias is 0.22, which indicates too much cheating. This bias is about the same regardless of the base rate. Overall, the reporters' mean belief bias is 0.09, which indicates over−estimation of the verifiers' rate of investigation. This bias is most extreme when the base rate is high (0.17).[9]

9) To check for changes over rounds, I compute the descriptive statistics for the last two observations at each level of the base rate. The results (not tabled) are similar to the results in Table 6, although variability over subjects is slightly higher.

B. Aggregate－level analysis

We conduct a series of repeated－measures analyses of variance to test the hypotheses stated in Section 3 (Panel A of Table 7). As to strategy decisions, there are significant base－rate effects on the verifiers' rate of investigation ($F = 5.88$, $p = 0.007$) and the reporters' rate of under－reporting ($F = 18.98$, $p < 0.001$). As noted earlier, the mean responses in Table 6 suggest a positive base－rate effect on the verifiers' rate of investigation and a negative base－rate effect on the reporters' rate of under－reporting. These results contradict a base－rate fallacy with respect to strategy decisions for verifiers (H1a) and reporters (H2a). These results also contradict sophisticated strategic thinking for verifiers (H1d), because the verifiers adjust their strategy depending on the base rate, and non－strategic thinking for reporters (H2b), because the reporters adjust their strategy inversely with changes in the base rate.

〈Table 7〉 Aggregate－level analysis of base－rate effects

Panel A.	Verifiers ($n = 16$)		Reporters ($n = 24$)	
Source of variance	Strategy (β)	Belief ($E[\alpha]$)	Strategy (α)	Belief ($E[\beta]$)
Base rate for reporters (ρ)	5.88^{**}	4.57^{*}	18.98^{**}	37.80^{**}
Panel B.	Verifiers ($n = 16$)		Reporters ($n = 24$)	
Source of variance	Strategic decision bias	Belief bias	Strategic decision bias	Belief bias
Base rate for reporters (ρ)	5.88^{**}	67.28^{**}	1.38	9.80^{**}

1) * Significant at $p = 0.05$, two－tailed.
2) ** Significant at $p = 0.01$, two－tailed.
3) This table shows the F ratio and its significance from repeated－measures analyses of variance using the base rate for reporters as the within－subject factor. In Panel A, the dependent variables are the subjects' strategy decisions and beliefs. In Panel B, the dependent variables are biases in the subjects' strategy decisions and beliefs. The strategic decision bias is the signed difference between the subject's strategy decision and the corresponding game－theoretic strategy. The belief bias is the signed difference between the subject's belief about the opponent's strategy and the strategy selected by the opponent.

As to beliefs, there are significant base－rate effects on the verifiers' belief about the reporters' strategy ($F = 4.57$, $p = 0.018$) and the reporters' belief about the verifiers' strategy ($F = 37.80$, $p < 0.001$). As noted earlier, the mean responses in Table 6 suggest positive base－rate effects on both the verifiers' belief about the reporters' strategy and the reporters' belief about the verifiers' strategy. These results contradict a base－rate fallacy with respect to beliefs for verifiers

(H1a) and reporters (H2a). These results also contradict non-strategic thinking for verifiers (H1b), because the verifiers adjust their belief about the reporters' strategy depending on the base rate, and sophisticated strategic thinking for reporters (H2d), because the reporters adjust their belief about the verifiers' strategy depending on the base rate.[10)]

Taken together, the aggregate-level results contradict all of the hypotheses stated in Section 3 except for simple strategic thinking for verifiers (H1c) and reporters (H2c). The verifiers estimate an increase in the rate of under-reporting as the base rate increases, and respond by increasing the rate of investigation. The reporters estimate an increase in the rate of investigation as the base rate increases and respond by decreasing the rate of under-reporting.[11)]

C. Individual-level analysis

The aggregate-level analysis implicitly assumes that the mean response or change in response is typical and that variability around the mean is noise. Another possibility is individual differences as to response tendencies. Collecting many observations for each subject allows us to perform an individual-level analysis. The individual-level analysis implicitly assumes that each subject applies the same reasoning process over rounds and that variability around the subject's mean response or change in response is noise. We do not view either the aggregate- or individual-level analysis as superior. Instead, we view the two levels of analysis as complementary and useful in avoiding oversimplified conclusions about the subjects' behavior.

To classify each subject's level of strategic thinking, we perform a pair of simple regressions using the subject's strategy and belief responses over 20 rounds : regression of strategy responses on the base rate and regression of belief responses on the base rate. The observed base-rate effects are positive, negative, or none, depending on the sign and significance of the t-value for the base rate's regression coefficient. The combination of the base-rate effects for the pair of regressions allows us to classify the subject as to the level of strategic thinking.

Table 8 presents the results for the verifiers. For example, the results for verifier 1 indicate a positive base-rate effect on the strategy responses ($t = 4.84$, $p < 0.001$) and a positive base-rate effect on the belief responses ($t = 2.99$, $p = 0.008$). This combination of base-rate effects

10) The results (not tabled) are similar for the last two observations at each level of the base rate, although slightly weaker because of more variability over subjects.

11) Panel B of Table 7 presents analyses of biases in the subjects' strategy decisions and beliefs, which serve to confirm the patterns in the descriptive statistics noted earlier. There are base-rate effects on biases in the verifiers' strategy decisions and beliefs, in the reporters' beliefs, but not in the reporters' strategy decisions.

implies simple strategic thinking (H1c). Using the same procedure to classify each verifier, the results contradict sophisticated strategic thinking for all verifiers (H1d). Only two verifiers are consistent with a base−rate fallacy (H1a). Of the remaining 14 verifiers, five are consistent with non−strategic thinking (H1b), four are consistent with simple strategic thinking (H1c), and five have other response patterns. There is a clear tendency for the verifiers to increase the rate of investigation as the base rate increases (regression 1), but no clear tendency in their belief about the reporters' strategic response to parametric change (regression 2).

〈Table 8〉 Individual−level analysis for verifiers

Verifier	Regression 1	Regression 2	H1a	H1b	H1c	H1d	Other
1	4.84**	2.99**			×		
2	6.31**	1.92		×			
3	−1.52	3.76**					×
4	3.43**	0.00		×			
5	1.17	3.90**					×
6	3.61**	3.49**			×		
7	−3.15**	1.30					×
8	3.04**	4.01**			×		
9	1.67	0.32	×				
10	4.39**	0.00		×			
11	6.19**	−0.50		×			
12	4.27**	1.62		×			
13	−1.08	−1.31	×				
14	2.65*	4.82**			×		
15	4.55**	−2.27*					×
16	5.20**	2.90**					×

1) * Significant at $p = 0.05$, two−tailed.
2) ** Significant at $p = 0.01$, two−tailed.
3) This table shows t−values and their significance from two simple regressions for each verifier : regression of the verifier's strategy decision on the base rate for reporters (regression 1), and regression of the verifier's belief about the reporter's strategy on the base rate for reporters (regression 2). The remaining columns classify each verifier according to alternative levels of strategic thinking : base−rate fallacy (H1a), non−strategic thinking (H1b), simple strategic thinking (H1c), sophisticated strategic thinking (H1d), or other.

Table 9 presents the results for the reporters. For example, the results for reporter 1 indicate a negative base−rate effect on the strategy responses (t = −4.97, $p < 0.001$) and a positive base−rate effect on the belief responses (t = 17.74, $p < 0.001$). This combination of base−rate effects implies simple strategic thinking (H2c). Using the same procedure to classify each reporter, the results contradict sophisticated strategic thinking for all reporters (H2d). Only three reporters are consistent with a base−rate fallacy (H2a). Of the remaining 21 reporters, one is consistent with non−strategic thinking (H2b), 14 are consistent with simple strategic thinking (H2c), and six have other response patterns.12) There is a clear tendency for the reporters to increase their estimates about the rate of investigation as the base rate increases (regression 2), and respond by decreasing the rate of under−reporting (regression 1). As with the aggregate−level analysis, the individual−level analysis supports the hypothesis of simple strategic thinking for the reporters (H2c).

〈Table 9〉 Individual−level analysis for reporters

Reporter	Regression 1	Regression 2	H2a	H2b	H2c	H2d	Other
1	−4.97**	17.74**			×		
2	0.30	−2.90**					×
3	0.07	0.14	×				
4	−4.34**	4.19**			×		
5	−3.98**	7.68**			×		
6	−2.82**	2.61*			×		
7	−6.04**	6.84**			×		
8	−3.60**	3.27**			×		
9	−2.38*	8.47**			×		
10	−7.00**	15.86**			×		
11	−1.48	4.17**					×
12	−0.41	−0.74	×				
13	0.97	17.74**					×
14	−9.72**	15.55**			×		

12) Four of the six reporters in the “other” category are similar to the modal pattern, showing a positive base−rate effect on belief responses and a negative (albeit insignificant) base−rate effect on strategy responses.

〈Table 9〉 Individual－level analysis for reporters (Continued)

Reporter	Regression 1	Regression 2	H2a	H2b	H2c	H2d	Other
15	－0.90	2.60*					×
16	－6.34**	7.79**			×		
17	－4.89**	2.99**			×		
18	－5.50**	6.48**			×		
19	－1.62	3.76**					×
20	－0.17	0.26	×				
21	4.80**	0.38		×			
22	－4.37**	5.02**			×		
23	－1.43	4.25**					×
24	－2.35*	9.59**			×		

1) * Significant at p = 0.05, two－tailed.
2) ** Significant at p = 0.01, two－tailed.
3) This table shows t－values and their significance from two simple regressions for each reporter : regression of the reporter's strategy decision on the base rate for reporters (regression 1), and regression of the reporter's belief about the verifier's strategy on the base rate for reporters (regression 2). The remaining columns classify each reporter according to alternative levels of strategic thinking : base－rate fallacy (H2a), non－strategic thinking (H2b), simple strategic thinking (H2c), sophisticated strategic thinking (H2d), or other.

Ⅴ. Summary and conclusion

Mathematical game theory provides insight into strategic behavior in accounting settings, but for tractability assumes common knowledge of rationality. Relaxing this assumption, experiments contribute added insight by systematically observing strategic decisions by humans in laboratory games. Using game－theoretic models as normative benchmarks, experimenters observe subjects' strategic decision biases for three purposes : (1) infer reasoning processes, (2) prescribe specific ways to eliminate observed biases, and (3) explain how economic institutions mitigate or accommodate strategic decision biases.

My study mainly focuses on the first purpose and demonstrates a new procedure for inferring subjects' achieved level of strategic thinking. The experimental recipe has the following ingredients : (1) create a laboratory setting that corresponds to a game－theoretic model, without taking for granted common knowledge of rationality, (2) assign subjects to the opposing roles in the model and ask them to a make a strategy decision after estimating the opponent's strategy, (3) manipulate a model parameter over rounds, and (4) use the observed response pattern to infer the achieved level of strategic thinking. Unlike past studies that rely on dominance-solvable games, my procedure applies to any game in which players potentially adjust their strategy based on a belief about the opponent's strategy. This advantage is important for auditing games that are not solvable by iterative dominance (Bloomfield 1995).

The experiment examines strategic behavior in a tax audit game (Graetz et al. 1986) with subjects in the roles of a reporter who under-reports income and a verifier who decides whether to investigate reported income. We vary the base rate for unethical reporters over rounds and observe the pattern of subjects' strategy decisions combined with beliefs about the opponent's strategy. Based on aggregate- and individual-level analyses, we conclude that the subjects do not mimic the normative implications of the game-theoretic model. Instead, the reporters generally achieve simple strategic thinking, meaning that they adjust their rate of under-reporting based on a belief (or "other mind" inference) about how the verifiers respond to changes in the base rate. Simple strategic thinkers try to anticipate the responses of their opponent, but do not correctly recognize that the opponent does likewise.

The verifiers include non-strategic and simple strategic thinkers. Non-strategic thinkers respond to parametric change, but treat the opponent as a mindless part of the context. Either non-strategic or simple strategic thinking leads the verifiers to commit a strategic decision bias : they wrongly condition the rate of investigation on the base rate and consequently under-audit when the base rate is low. These findings have potential policy implications. The tax authority may consider the negative effects of limited rationality shown by verifiers in providing its on jot training programs to achieve more efficient tax auditing system.

There are several potential limitations regarding the generalizability of findings, though. One possible limitation of generalizability is the use of students as subjects. The use of student subjects may limit the generalizability of the results beyond the student groups, although the use of student subjects has been common in most previous tax experiments (e.g., Beck et al. 1991). Also, as I use Hong Kong Chinese subjects, it is an open question that the results can be generalized in Korean tax setting. The experimental tasks are simple, and parameters used in the experiments do not necessarily reflect those parameters in actual tax reporting environments. Finally, while the experimental procedure randomly and anonymously re-pairs reporters and verifiers round-by-round in order to minimize reputation formation, the tax authority in practice is likely to consider each taxpayer's reporting history in its tax audit decisions. Thus, how reputation affects decisions made by tax auditors as well as by taxpayers is a potentially valuable future experimental research issue.

Finally, my study demonstrates a new procedure for inferring subjects’ achieved level of strategic thinking in game settings. These results strongly suggest that the future experimental research focus on prescribing specific ways to eliminate observed biases in strategic thinking and also explaining how economic institutions mitigate or accommodate strategic decision biases.

References

Andreoni, J., Erard, B., and Feinstein, J. 1998. Tax compliance. *Journal of Economic Literature* 36 : 818 – 860.

Ashton, A., & Ashton, R. 1995. *Judgment and decision–making research in accounting and auditing*. Cambridge : Cambridge University Press.

Ashton, R. 1983. *Research in audit decision making : Rationale, evidence, and implications*. Research Monograph No. 6. Vancouver : Canadian Certified General Accountants Research Foundation.

Beck, P., J. Davis, and W. Jung. 1991. Experimental Evidence on Taxpayer Reporting Under Uncertainty. *The Accounting Review* 66 (July) : 535 – 558.

Bloomfield, R. 1995. Strategic dependence and inherent risk assessments. *The Accounting Review* 70 : 71 – 90.

____________, 1997. Strategic dependence and the assessment of fraud risk : A laboratory study. *The Accounting Review* 72 : 517 – 538.

____________, 1999. Discussion of an experimental investigation of auditor – auditee interaction under ambiguity. *Journal of Accounting Research* 37 (Supplement) : 157 – 165.

Butt, J. 1988. Frequency judgments in an auditing – related task. *Journal of Accounting Research* 26 : 315 – 330.

Camerer, C. 1997. Progress in behavioral game theory. *Journal of Economic Perspectives* 11 : 167 – 188.

__________, 2003. *Behavioral game theory : Experiments in strategic interaction*. Princeton : Princeton University Press.

__________, Ho, T., & Chong, J. 2004. A cognitive hierarchy model of games. *Quarterly Journal of Economics* 119 : 861 – 898.

Duffy, J., & Nagel, R. 1997. On the robustness of behavior in experimental "beauty contest" games. *Econometrica* 107 : 1684 – 1700.

Ganguly, A., Kagel, J., & Moser, D. 1994. The effects of biases in probability judgments on market prices. *Accounting, Organizations and Society* 19 : 675 – 700.

Graetz, M., Reinganum, J., & Wilde, L. 1986. The tax compliance game : Toward an interactive theory of law enforcement. *Journal of Law, Economics, and Organization* 2 : 1 – 32.

Haynes, C., & Kachelmeier, S. 1998. The effects of accounting contexts on accounting decisions : A synthesis of cognitive and economic perspectives in accounting experimentation. *Journal of Accounting Literature* 17 : 97 – 136.

Hertwig, R., & Ortmann, A. 2001. Experimental practices in economics : A methodological challenge for psychologists? *Behavioral and Brain Sciences* 24 : 383 – 451.

Ho, T., Camerer, C., & Weigelt, K. 1998. Iterated dominance and iterated best responses in experimental p－beauty contests. *American Economic Review* 88 : 947－969.

Jones, K., Tan, H., & Waller, W. 2004. Improving auditors' fraud risk assessments : Business measures and strategic thinking. University of Arizona working paper.

Joyce, E., & Biddle, G. 1981. Are auditors' judgments sufficiently regressive? *Journal of Accounting Research* 19 : 323－349.

Kachelmeier, S., & King, R. 2002. Using laboratory experiments to evaluate accounting policy issues. *Accounting Horizons* 16 : 219－232.

Kahneman, D., Slovic, P., & Tversky, A. 1982. *Judgment under uncertainty : Heuristics and biases*. Cambridge : Cambridge University Press.

Keynes, J. 1936. *The General Theory of Employment, Interest, and Money*. London : Macmillan.

Koehler, J. J. 1996. The base rate fallacy reconsidered : Descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences* 19 : 1－53.

Nagel, R. 1995. Unraveling in guessing games : An experimental study. *American Economic Review* 85 : 1313－1326.

_______, 1999. A survey of experimental guessing games : A study of bounded rationality and learning. In *Games and Human Behavior : Essays in Honor of Amnon Rapoport*, edited by Budescu, D., Erev, I., & Zwick, R. Hillsdale, NJ : Erlbaum.

Ohtsubo, Y., & Rapoport, A. 2006. Depth of reasoning in strategic form games. *The Journal of Socio－Economics* 35 : 31－47.

Plott, C. 1996. Rational individual behavior in markets and social choice processes : The discovered preference hypothesis. In *The Rational Foundations of Economic Behavior*, edited by Arrow, K., Colombatto, E., Perlamn, M., & Schmidt, C. London : Macmillan.

Rapoport, A., & Amaldoss, W. 2000. Mixed strategies and iterative elimination of strongly dominated strategies : An experimental investigation of states of knowledge. *Journal of Economic Behavior and Organization* 42 : 483－521.

Shleifer, A. 2000. *Inefficient Markets : An Introduction to Behavioral Finance*. Oxford : Oxford University Press.

Slonim, R. 2005. Competing against experienced and inexperienced players. *Experimental Economics* 8 : 55－75.

Tversky, A., & Kahneman, D. 1974. Judgment under uncertainty : Heuristics and biases. *Science* 185 : 1124－1131.

Wilks, T., & Zimbelman, M. 2004. Using game theory and strategic reasoning concepts to prevent and detect fraud. *Accounting Horizons* 18 : 173－184.

Zimbelman, M., & Waller, W. 1999. An experimental investigation of auditor－auditee interaction under ambiguity. *Journal of Accounting Research* 37 (Supplement) : 135－155.

세금보고의 전략적 사고

김정권*

<요 약>

행동게임이론은 게임모형과 정보분석연구 모형을 결합하여 세운 전략적인 상호작용에 관한 가설들을 실험방법론을 통해 실증한다. 제한된 합리성을 가진 개인이 세금보고 게임에서 어느 수준의 전략적 사고능력을 보여주는지를 관찰하는 것이 이 연구의 주된 관점이다. 이 연구의 실험에 따르면, 세무 감사인과 기업이 정책 변수의 변화에 대한 상대방의 반응을 예측하여 그들의 전략을 수정하고 있지만 상대방도 또한 그런 수정을 한다는 사실을 충분히 인식하여 자신의 전략을 수정하지는 못한다는 결과를 보여 준다. 특히, 이러한 실험 결과는 세무 감사인의 전략적 결정은 게임이론상 의사결정에 관련이 없는 정보에 대해 편의를 보인다는 의미를 가지고 있다.

〈주제어〉 게임이론, 조세순응, 제한적 합리성, 세무조사, 전략적 사고

* 이화여자대학교 경영대학 교수, acckim@ewha.ac.kr, 02－3277－4039