

© 2002 American Accounting Association
Accounting Horizons
Vol. 16 No. 3
September 2002
pp. 219–232

COMMENTARY

Using Laboratory Experiments to Evaluate Accounting Policy Issues

Steven J. Kachelmeier and Ronald R. King

Steven J. Kachelmeier is a Professor at The University of Texas at Austin, and Ronald R. King is a Professor at Washington University.

INTRODUCTION

Accounting scholars have long been challenged to demonstrate the relevance and timeliness of their research to the issues faced by accounting policymakers. Among others, Schipper (1994), Swieringa (1998), and Barth et al. (2001) provide examples of the progress made by empirical-archival studies in addressing this challenge. Notwithstanding this progress, the ability of empirical-archival accounting research to offer policy-directed insights is inherently limited, for two reasons. First, real-world data can only inform us of reactions to policies that already exist, whereas the policymaker's perspective demands *ex ante* insights of likely reactions to policies that *could* exist. Second, the variation among firms in accounting practices coincides with a multitude of other factors that vary at the same time, making it difficult to isolate the incremental effect of accounting alternatives. Given these limitations, several observers, such as Maines (1994), Beresford and Johnson (1995), and Swieringa (1996), suggest a role for true experiments—designed studies in which participants make decisions under experimenter-controlled conditions—in addressing accounting policy questions. Only in a true experiment can we vary an existing or proposed policy alternative of interest and measure its incremental influence on decision makers, holding all other influences constant.

This commentary explores this idea in greater depth, especially with respect to the different strengths and weaknesses of individual judgment and decision-making studies, often termed *behavioral research*, and multiperson interactive studies with competitive incentives, often termed *experimental economics*, for producing useful policy insights.¹ Our

¹ For simplicity, we characterize behavioral judgment and decision-making research as involving individual judgments in settings without explicit performance-contingent incentives, and experimental economics as involving interactive settings with explicit incentives. However, the literature also includes psychological studies of group decisions and economic studies of individual decisions, making the differences we explain more of a continuum than a dichotomy.

We are grateful for helpful comments from an anonymous reviewer, Amy Choy, Nick Dopuch, Eric Hirst, Ella Mae Matsumura, and Kristy Towry.

Submitted: May 2001

Accepted: April 2002

Corresponding author: Steven J. Kachelmeier
Email: kach@mail.utexas.edu

premise is that, viewed as a whole, these two styles of experimentation help to mitigate the concern that experiments are of limited value to policy debates because they lack realism (Abdel-khalik 1994). Behavioral studies of individual judgment and decision making often lack the realism of explicit performance incentives and the multiperiod, multiperson nature of economic competition. For different reasons, experimental economic studies also lack realism, because they often involve student subjects performing stylized tasks that are abstracted from the real-world environment. When viewed as a whole, however, these two approaches complement each other, addressing the other's weaknesses and offering unique strengths.

We provide some historical perspective below, followed by examples from each approach that illustrate our premise. For the sake of brevity, we focus our examples on financial-accounting policy questions, although we could offer similar observations for policy questions involving auditing, management accounting, and taxation. We then comment on the strategic implications of accounting policy questions, distinguishing between experiments that manipulate policy alternatives and those that treat accounting choices as strategic decisions by opportunistic managers. Our concluding comments return to our overall premise and suggest ways to integrate the benefits of different experimental styles.

SOME HISTORY

Although marked by recent interest, the use of experiments to address accounting policy questions dates back several decades. Swieringa and Weick (1982) review several experimental studies of accounting policy questions conducted in the 1960s and early 1970s. The popularity of such experiments ebbed significantly by the mid-1970s, concurrent with the critical view of this literature that emerged by that time (Gonedes 1972; Gonedes and Dopuch 1974). Criticism reflected a two-pronged attack, noting both the lack of theory in the experimental literature of the day and the inconsistency of this literature with the prevailing belief in informational market efficiency. The following passage from Gonedes and Dopuch (1974, 106) is telling:

Since laboratory studies concentrate on individual behavior rather than competitive market phenomena, their relevance to the issues at hand seems nonexistent....Even if these studies were based upon theory,...any generalizations made about the aggregate behavior of capital market agents on the basis of results from laboratory studies are extremely tenuous. Specifically, given an efficient capital market, studies of the behavior of particular types of investors are not likely to lead to reliable generalizations about...capital market equilibrium.

Given views such as these, it is not surprising that by the late 1970s through the 1980s, experimental work in accounting largely evolved to studies of auditor judgment, as reviewed by Solomon and Shields (1995), with occasional forays into management accounting, as reviewed by Sprinkle (2002).

By the early 1990s, however, experimentalists had reasons to question both criticisms raised in the early 1970s. First, experiments evolved to reflect not just practical questions and intuitive conjectures, but also rigorous theoretical underpinnings from both psychology and economics (Maines 1995; Libby et al. 2002). Second, emerging empirical evidence suggested that capital markets were not as informationally efficient as was once believed (Berg et al. 1995; Lee 2001; Bloomfield 2002). Accordingly, recent years witnessed a resurgence in policy-directed accounting experimentation. Examples

of policy questions addressed in recent accounting experiments include:

- Classification of hybrid debt-equity securities (Hopkins 1996)
- The movement to reflect current values in financial reporting (Kachelmeier 1996)
- The controversy surrounding industry vs. operating classifications for segment reporting (Maines et al. 1997)
- Environmental liability disclosures (Kennedy et al. 1998)
- Alternative formats for reporting comprehensive income (Hirst and Hopkins 1998; Maines and McDaniel 2000)
- The purchase vs. pooling controversy (Hopkins et al. 2000)
- Alternative disclosures of risks and opportunities (Dietrich et al. 2001)

It is clear that accounting experimentalists are actively responding to the call for *ex ante* (or at least concurrent) policy-directed research, tapping the comparative advantage of experiments to contrast decisions under both existing and plausible alternative accounting policies in a *ceteris paribus* environment.

Libby et al. (2002) provide a comprehensive, up-to-date review of this literature. Our goal in this commentary is to examine the relative strengths and limitations for policy-directed research of two different experimental perspectives evidenced in these studies. The first, and oldest perspective—often termed “behavioral judgment and decision-making” research—generally elicits the hypothetical questionnaire-based judgments of experimental participants in a single-person setting using realistic case materials. Although these studies strive to capture contextual realism (Gibbins and Jamal 1993), they remain partially subject to Gonedes and Dopuch’s (1974) criticism that individual judgments do not necessarily imply aggregate behavior in competitive markets.

This observation suggests a role for the interactive, incentive-driven style of experimentation often termed “experimental economics,” which includes laboratory markets and other designs involving strategic interaction among participants. Experimental economic studies measure behavior from trading decisions and other actions with meaningful monetary consequences, and explicitly incorporate the discipline of economic competition. These features enable researchers to address market efficiency issues as empirical questions rather than assumptions. However, as we explain later, experimental economic studies suffer from a different set of limitations. We believe there is little value to debating which style of experimentation is better or worse. There are valuable synergies from the different strengths and limitations of different experimental approaches. We offer examples of each approach to illustrate the point.

EXAMPLES OF POLICY-DIRECTED INDIVIDUAL-JUDGMENT STUDIES: REPORTING COMPREHENSIVE INCOME

The FASB (1997) defines comprehensive income as the change in an entity’s net assets, exclusive of transactions with owners. Comprehensive income includes income as traditionally defined as well as items not reported on the traditional income statement, such as unrealized gains and losses on available-for-sale securities and foreign exchange translation adjustments under the current-rate method (FASB 1997). The FASB felt that firms should report comprehensive income, but determining where to do so was controversial. At one extreme, firms could report comprehensive income in a separate performance statement that begins with net income as traditionally defined, appending additional items to arrive at comprehensive income. At the other extreme, firms could report the components of comprehensive income as reconciling items within

the Statement of Stockholders' Equity (SSE). Empirical-archival studies are ill-equipped to address the implications of these alternative formats because of timeliness, and because the substantial majority of companies opted for the SSE format, leaving little basis for comparison.²

Related experiments by Hirst and Hopkins (1998) and by Maines and McDaniel (2000) addressed the valuation implications of alternative comprehensive income formats. Hirst and Hopkins (1998) assigned buy-side analysts to different comprehensive income format conditions and asked them to value a company that (1) did or (2) did not manage earnings by selling and repurchasing marketable securities that experienced unrealized gains.³ Maines and McDaniel (2000) assigned evening M.B.A. students to experimental conditions that varied in (1) comprehensive income reporting format and (2) the underlying volatility of unrealized security gains and losses, asking them to assess volatility and management effectiveness. Both studies found that users are better able to determine the firm's underlying economic activity when the firm reports comprehensive income in a separate performance statement. From a policy perspective, these studies call into question the FASB's compromise decision to allow both alternatives, and help to explain why the substantial majority of companies choose the less transparent SSE option.

Limitations of Individual Judgment Studies

The fact that these two studies derive similar conclusions from different elicited judgments and different types of participants helps to reinforce the studies' policy implications. Nevertheless, these studies illustrate some inherent limitations of this style of experimentation. First, without performance-contingent compensation, it is unknown whether participants in these and related individual-judgment studies exert as much information-processing effort in the experimental setting as in an environment with explicit incentives. Participants exerting less information-processing effort would fare better with a less informationally taxing reporting format, which is consistent with the observed findings.⁴ Second, even with meaningful incentives, the market aggregation question remains. An important tenet of the market efficiency argument is that even if most market participants are biased processors of information, a few informed traders at the margin can arbitrage any resultant security mispricing, such that even the uninformed become price-protected in equilibrium. Lee (2001) and Bloomfield (2002) explore challenges to this argument, and Libby et al. (2002) provide ample evidence that judgment biases can persist in market settings. Nevertheless, when an experiment measures only individual judgments, the potentially mitigating force of market competition remains an open question.

² Even if a meaningful number of companies chose both formats, the self-selection into these categories severely limits the ability to separate the valuation implications of the different reporting formats from the myriad firm characteristics likely associated with the choice of such formats. By contrast, experimentation enables random assignment to treatment conditions.

³ Hirst and Hopkins (1998) further subdivided the SSE format condition into two subconditions where comprehensive income was or was not explicitly labeled and reported as such within the SSE.

⁴ The literature on incentive effects in experiments is far from conclusive (Camerer and Hogarth 1999), and it is certainly unwarranted to assume that uncompensated participants are unmotivated or that the findings from such participants are invalid. Our only observation is that the absence of performance-contingent compensation can leave open the possibility of an alternative explanation for observed findings.

Could Laboratory Markets Address These Limitations?

One might conjecture that these limitations could be overcome by replicating the Hirst and Hopkins (1998) and/or Maines and McDaniel (2000) experiments in a laboratory market, using the tools of experimental economics. Unfortunately, such a goal is far easier said than done. The laboratory market researcher must first assign “true values” to the securities underlying the firms in the experimental materials and corresponding financial statements. Laboratory market participants redeem the securities they hold at the end of the session for terminal values defined by the researcher, providing the incentive to buy securities below true value or sell them above true value. By contrast, Hirst and Hopkins’ (1998) participants provided *judgments* of underlying security values. These judgments can differ in predictable ways between treatment conditions, but they cannot be compared to any absolute sense of true value.

The difficulty of assigning true values for something as complex as the materials used by Hirst and Hopkins (1998) and Maines and McDaniel (2000) is compounded by a second fundamental challenge of laboratory markets—the firms traded in a laboratory setting are not going concerns. By necessity, laboratory securities pay a liquidating dividend at the end of the experimental session, enabling participants to be compensated and dismissed. Put differently, participants in a laboratory market cannot take securities with them when they leave the experiment. The conventional approach is to liquidate experimental securities for well-defined terminal values at the end of the session, such that participants’ final positions are expressed only in cash. By contrast, real-world firms do not generally liquidate on a predictable basis, such that the value of a security is circularly defined by its estimated realizable value upon future sale.⁵ Participants in an individual-judgment setting can grasp this notion and estimate a realistic stock price, but their judgments do not correspond to those of laboratory market participants who expect a cash payoff from their asset holdings at the end of the session.

Suppose, for the sake of argument, that an experimenter can devise some well-defined process that ends in a calculable liquidation value based on gains and losses from marketable securities reported in comprehensive income under different formats. The reader might wish to conjecture how this could be done. Any solution undoubtedly involves a great deal of abstraction away from the rich experimental materials used by Hirst and Hopkins (1998) and Maines and McDaniel (2000). Laboratory market participants need to know exactly what is and is not relevant to the liquidation process that generates a terminal value from holding a target security. To be operational in a laboratory economy, any such process must be much simpler and hence more easily understood by market participants than the process that operates in real-world economies. It is reasonable to surmise that this abstraction makes it more difficult to find an information-processing effect of different comprehensive income formats. This is not just because of the forces of market efficiency; rather, it is because the judgment task as a whole becomes more transparent than that presented in the materials used by Hirst and Hopkins (1998) and Maines and McDaniel (2000). These challenges suggest why an individual-judgment experiment is a reasonable approach to questions of comprehensive income reporting format, notwithstanding the limitations. Nevertheless, we can gain some synergy by applying experimental economic approaches to different questions involving similar behavioral phenomena, as we illustrate presently.

⁵ Even the dividend-discount model of firm valuation, which is based in theory on the present value of all future dividends, is operationalized in practice by cutting the time series at some point, assuming that a projected future stock price represents the present value of all remaining dividends (Botosan and Plumlee 2001).

EXAMPLES OF POLICY-DIRECTED EXPERIMENTAL-ECONOMICS STUDIES: DISCLOSING UNCERTAINTY

As explained above, it is quite a challenge to design a laboratory market investigation of something as contextually rich as the format for reporting comprehensive income. However, the methods of experimental economics are well suited to other policy questions that can be attacked at a more basic level. A case in point is the investigation of user reactions to alternative forward-looking disclosures of a firm's risks and opportunities. From a policy perspective, the Special Committee on Financial Reporting of the American Institute of Certified Public Accountants calls for increased attention to disclosures that convey "more information with a forward-looking perspective, including management's plans, opportunities, risks, and measurement uncertainties" (AICPA 1994, 5). The experimental economist has the ability to construct assets with well-defined dimensions of downside risk and upside opportunity, enabling investigation of alternative disclosures that vary in the perceived degree of measurement uncertainty.

We illustrate two recent studies that apply laboratory market approaches to these issues. The first is Bloomfield and Wilks (2000), in which laboratory market participants traded risky securities with *ex ante* expected values held constant. Experimental manipulation of the precision of supplemental disclosures helped participants gauge the degree of risk associated with these values. Findings indicate that more precise disclosure facilitates greater liquidity in the market by narrowing bid-ask spreads. The authors also report differential effects of disclosure at different levels of market *depth*, referring to the prices at which traders are willing to buy or sell larger blocks of securities. The study's emphasis on the effects of disclosure on market liquidity and depth illustrates the ability of a laboratory market approach to capture phenomena not accessible in an individual-judgment experiment.

A second example is a study by Dietrich et al. (2001), who use a setting adapted from the oil industry to examine supplemental forward-looking disclosures in a laboratory economy. Their study is related to Bloomfield and Wilks (2000) insofar as both studies examine the relative effects of alternative disclosures of risk. However, unlike the Bloomfield and Wilks (2000) focus on disclosure *precision*, Dietrich et al. (2001) focus on the biasing effects of forward-looking disclosures that reveal partial information about future values, even if users can infer the information that is not explicitly revealed. They find that a disclosure analogous to the estimated proved reserves note required by the FASB (1982) in the oil and gas industry helps market traders achieve prices more consistent with expected reserve values. This finding holds even though the allowance for depletion in the primary financial statements enables traders to infer the amount of reserves, making the supplemental disclosure redundant.

Regarding risks and opportunities, Dietrich et al. (2001) find that one-sided disclosure of upside opportunities in oil reserves has an upward biasing effect on prices only for securities that previously declined in value. This finding is consistent with the prospect theory prediction of risk-seeking tendencies when decision makers experience losses from an initial reference point (Kahneman and Tversky 1979; Tversky and Kahneman 1992). Namely, the risk seeker is influenced by a disclosure of upside opportunity that presents a possible way out of a loss position, whereas the same disclosure does not have a similar biasing effect when the individual becomes risk-averse to protect a gain position. Beyond the oil and gas industry, the study suggests policy implications for other partial disclosures of risks and opportunities in financial accounting, including derivative financial

instruments (Linsmeier and Pearson 1997; Hodder et al. 2001) and contingent environmental liabilities (Kennedy et al. 1998).

Simplifications in the Experimental Economics Approach

The Dietrich et al. (2001) laboratory market study is noteworthy because it incorporates a literal, albeit fictional, income statement and balance sheet in the experimental materials. This feature moves a step closer to the traditions of the behavioral judgment literature in accounting, where materials are typically patterned after the natural environment (Gibbins and Jamal 1993; Haynes and Kachelmeier 1998). Nevertheless, the study reflects the considerable simplification and abstraction that is typical of an experimental economics approach. The study's balance sheet contains only two assets: cash, which is paid out as a liquidating dividend at the end of the session, and a single hypothetical oil well with reserves characterized by a well-defined liquidation process. Similarly, the income statement has one revenue account—sales from oil production—and the only expense is depletion. These simplifications enable the authors to operationalize a liquidating dividend for the securities traded by participants. That is, each firm in Dietrich et al. (2001) evolves after three trading periods into a single cash value, with distributional parameters that are described in the experimental instructions.

Bloomfield and Wilks' (2000) laboratory market disclosure experiment involves an even simpler process. Participants traded securities with a known expected value of 2,100 experimental "francs," with the actual value determined after adding or subtracting seven random numbers that each range from -25 to $+25$. In this setting, disclosure precision is manipulated by revealing either two or six of the seven random numbers, thereby resolving either a low or high degree of uncertainty. This manipulation captures the essence of differential precision in accounting disclosure, but at a considerable level of abstraction.

Could We Enrich the Setting?

As a thought exercise, one might speculate how the Bloomfield and Wilks (2000) and Dietrich et al. (2001) studies of disclosures involving uncertainty could be enriched with a more contextually realistic approach. For instance, Kennedy et al. (1998) used realistic case materials to examine the effects of minimum, maximum, range, and best estimate disclosures of contingent environmental liabilities on judgments made by financial executives, bankers, and M.B.A. students. Nevertheless, incorporating the full institutional complexity of forward-looking disclosures of risks and opportunities is just as daunting a challenge as addressing alternative comprehensive income reporting formats in a laboratory market. For example, Hodder et al. (2001) describe the considerable complexity associated with market risk disclosures recently mandated by the Securities and Exchange Commission (also see Linsmeier and Pearson 1997). While judgment and decision-making experiments can certainly help to disentangle the various behavioral phenomena associated with these disclosures, any attempt to be fully "realistic" can overwhelm the experimental participant. Moreover, from a theoretical perspective, there is value to a step-by-step approach that intentionally abstracts away from some aspects of reality in order to isolate a construct of theoretical interest. In this spirit, abstract laboratory market studies such as Bloomfield and Wilks (2000) and Dietrich et al. (2001) can be viewed as building policy-relevant theoretical insights from the ground up, with the added advantage of offering direct evidence on questions such as informational market efficiency.

STRATEGIC CONSIDERATIONS

Surrounding the interplay between individual-judgment behavioral studies and experimental economic studies of accounting policy issues is the fact that choices among alternatives set forth in accounting policies do not occur in a vacuum. While the researcher can manipulate accounting policy alternatives as an experimental treatment factor, in practice such alternatives are often chosen and interpreted with strategic incentives in mind. In the parlance of research, a key question is whether accounting alternatives are *exogenous*, as a manipulated treatment factor implies, or *endogenous*, when a strategic party is allowed to choose.

Experiments with Manipulated Disclosures

We illustrate the distinction between a view of accounting policy alternatives as exogenous or endogenous by returning to the earlier example of comprehensive income reporting format. While Hirst and Hopkins (1998) and Maines and McDaniel (2000) *manipulated* whether comprehensive income is reported in a separate performance statement or in the Statement of Stockholders' Equity, a substantial majority of companies in practice opt for the second alternative. This observation is a primary reason favoring experimentation over an empirical-archival study of these alternative disclosure formats. Nevertheless, the lack of instances in practice where comprehensive income is reported in a separate performance statement begs consideration of the decision-maker's strategic inferences upon observing this alternative. What motives could drive such a choice, and with what intended signal?

Similar questions surround the studies of risk-related disclosures by Bloomfield and Wilks (2000) and Dietrich et al. (2001). While these studies use laboratory market designs involving competitive market actions taken by strategic traders, the disclosures themselves are manipulated treatment variables, and hence are implicitly *non-strategic*. This approach enables the authors to test under *ceteris paribus* conditions the hypothesized information-processing effects of different ways to disclose risk. However, in cases where managers *choose* among a menu of acceptable risk-related disclosures, the additional possibility emerges that users can infer the private information motivating such choices.

Experiments with Strategic Disclosures

The strategic dimensions of accounting policy alternatives are just as amenable to experimentation as are the nonstrategic dimensions. For instance, Kennedy et al. (1998) compare their primary results to a supplemental condition in which case materials indicate that disclosures of contingent environmental liabilities are made by an independent engineering consultant rather than by management itself. This comparison enables the authors to separate out any incremental effects of management's opportunistic incentives to choose strategically among the disclosure alternatives. They find that although this manipulation has a significant effect on the perceived credibility of the disclosure, it does not significantly interact with the information-processing treatment effects of the different disclosure alternatives.

One might conjecture that Kennedy et al.'s (1998) follow-up comparison involving hypothetical strategic or nonstrategic sources, while helpful, is not as forceful as an interactive experiment with actual choices among accounting alternatives made by strategic human agents. There are some inroads in the use of experimental economic methods to investigate such choices and the corresponding strategic inferences made by those

who receive this information. For example, King and Wallin (1991) conduct a market experiment in which a seller could choose among a menu of truthful disclosures, revealing, at the seller's discretion, the precision with which buyers learned about the value of an asset on which they entered competitive bids.⁶ The authors vary both the number of disclosure alternatives available to the seller and the buyers' *ex ante* knowledge of these disclosure menus. Findings provide general support for strategic predictions in both disclosure choices and the corresponding inferences implied in buyers' bids, except in a experimental condition where buyers are unaware of the disclosure menu and the number of disclosure choices is large.

One possible policy implication from King and Wallin (1991) is that in cases where accounting regulators allow many acceptable alternatives, user awareness of the set of possibilities becomes more critical. For example, firms have a variety of alternatives for assigning costs to inventories, such as FIFO, LIFO, and weighted average, and for how they use and report special-purpose entities. Financial statement users are arguably far more familiar with the former set of choices than with the latter. Following King and Wallin (1991), these different levels of familiarity likely afford management with greater opportunities to influence market values by strategic use of special-purpose entities than by choosing FIFO over LIFO. Other strategic disclosure studies employing the tools of experimental economics differentiate among multiple strategic equilibria (King and Wallin 1995), investigate reputations for reliable reporting (King 1996), and examine the ability of decision makers to infer the precision of management's private information from its disclosure choices (Dickhaut et al. 2001).

While the studies noted above use experimental economics to address strategic disclosure issues in the laboratory, they do not consider the information-processing issues underlying much of the judgment and decision-making literature in accounting; and while laboratory market studies such as Bloomfield and Wilks (2000) and Dietrich et al. (2001) tackle information-processing issues, they do not incorporate strategic disclosure considerations. These different perspectives are not necessarily cause for concern if one views them as different pieces of a larger puzzle. Exogenous manipulation of accounting alternatives is helpful for some policy questions, whereas endogenous incorporation of such alternatives in participants' choices is helpful for other questions. There is also potential for experiments that simultaneously incorporate elements of both cognitive information-processing and strategic considerations. For example, future research could explore whether strategic agents anticipate cognitive information-processing biases in deciding among disclosure alternatives, and whether users learn over time to correct these biases when they are aware of management's disclosure motives.

CONCLUSIONS

We can summarize the theme of this commentary by observing that experiments can attack issues of *ex ante* relevance to accounting policymakers from multiple fronts. Echoing the similar sentiments expressed by Cuccia (1994) regarding tax policy research, the different perspectives implied by individual-judgment and experimental-economic approaches to accounting policy questions need not be viewed as competing alternatives. Some issues, such as the different formats used to report comprehensive

⁶ King and Wallin's (1991) experiment employs an antifraud rule, meaning that while sellers could choose to withhold truthful information, they could not willfully communicate erroneous information.

income, may best be approached using expert individual decision makers in a contextually rich setting. Other issues, such as the processing of information on future risks and opportunities, are amenable to experimental-economic approaches. Similar observations can be raised for the additional considerations that emerge when accounting alternatives are chosen strategically rather than assigned randomly.

Addressing the Lack of Realism

Another way to state this theme is to revisit perhaps the most frequently encountered criticism of accounting experimentation: *lack of realism* (Abdel-khalik 1994). Accounting experiments *are* unrealistic, but our observations suggest that we gain some synergy when different experimental approaches capture different elements of reality. Individual-judgment studies from the behavioral tradition of accounting experimentation offer realistic case materials and often involve experienced professionals as participants. However, they lack the richness of strategic interaction, often infer behavior from hypothetical judgments rather than actions, and might raise questions about incentives. Laboratory market or other interactive designs from the experimental economics tradition offer the discipline of a well-defined competitive incentive structure, measure actual behavior rather than hypothetical judgments, and can address the issue of market efficiency as an empirically testable question. Yet, they are often highly abstracted from real-world settings, and are often limited for practical reasons to student participants.

Similar trade-offs pertain to the strategic considerations of accounting choices. Both individual-judgment and experimental economic approaches often use designs in which accounting alternatives are manipulated exogenously, affording the control of a *ceteris paribus* design. However, such designs can suppress the strategic considerations of endogenous choices by opportunistic managers and the corresponding inferences of financial statement users who are aware of the potential for opportunism. Other designs incorporate explicit accounting choices by participants in a role analogous to managers of firms, but are limited in their ability to tease out more fundamental information-processing issues. For example, what could we learn about users' potential reactions to a disclosure alternative if it is never observed?

All of these approaches are unrealistic, but in different ways. Ideally, we can build on this diversity to capture the relative strengths and address the relative weaknesses of multiple approaches. The individual judgment researcher can point to laboratory markets as evidence that market competition does not necessarily weed out all instances of biased individual judgment. The laboratory market researcher can point to individual judgment studies as evidence that findings in a highly abstract setting often generalize to similar findings in studies with expert participants and more contextually rich experimental materials. We can also undertake a more sophisticated consideration of managers' strategic disclosure choices if we first understand how users react to information in a nonstrategic disclosure environment.

Integrating Multiple Perspectives

To be sure, there is also a critical role for "bridge-building" experiments that integrate multiple perspectives. For example, Dietrich et al. (2001) motivate their hypotheses from an individual decision-making perspective with theoretical roots in the psychology of risk perceptions, but test these hypotheses in a competitive laboratory

market. Support for the authors' hypotheses in market prices leads them to question why efficient-market arguments do not prevail. Follow-up analysis of market microstructure indicates that market trading in their setting is characterized more by a "follow-the-herd" process than by any systematic attempt by a subset of traders to arbitrage price inefficiencies. Price prediction data collected at the individual level before each trading period help to corroborate these conclusions.

In general, we encourage creativity over dogma in the design of accounting-policy experiments.⁷ Both psychological and economic theories and methods can be integrated in a multitude of ways to attack different accounting policy questions.⁸ To be sure, the distinctions between individual-judgment approaches and experimental-economic approaches are likely to remain even if researchers enthusiastically embrace the goal of integrating these approaches. As illustrated by our examples, different questions are better suited to some approaches than others. Nor would we recommend that all accounting policy experiments should or should not incorporate strategic reporting choices and corresponding user inferences. Nonstrategic experiments can build the information-processing groundwork for a more sophisticated understanding of strategic motives. There is value to a controlled experimental setting in which accounting alternatives are manipulated systematically under *ceteris paribus* conditions, enabling true random assignment and control over self-selection bias. Nevertheless, experimentalists should be aware of the strategic implications of their research questions, and should strive over time to interpret psychological information-processing phenomena within a strategic context.

Our conclusion is that accounting experiments are responding effectively to the call expressed a few years ago by Beresford and Johnson (1995), among others, for research with *ex ante* accounting policy insights. There are considerable challenges, and the age-old friction between the internal validity of experimental control and the external validity of generalization to realistic settings is unlikely to go away any time soon. In addition, Libby and Kinney (2000) voice the concern that the quality-control process of peer review in accounting journals can delay well-intended *ex ante* policy research to the extent that it is no longer *ex ante* upon publication. Nevertheless, we perceive a more open environment in contemporary accounting scholarship that is increasingly receptive to building the linkage from accounting research to accounting policy. Experiments from both individual judgment and from laboratory market perspectives addressing both nonstrategic information-processing issues and strategic inferences from choices among accounting alternatives offer different comparative advantages toward the overall goal of providing policy-relevant insights.

⁷ See Haynes and Kachelmeier (1998) for a similar conclusion on the role of rich vs. abstract contextual wording in accounting experiments.

⁸ The examples cited in this commentary focus on financial-accounting questions. There has also been progress integrating psychological and economic insights in other areas of accounting experimentation, including examples from auditing (Zimbelman and Waller 1999; King 2002), managerial accounting (Evans et al. 2001; Kachelmeier and Towry 2002), and taxation (Moser et al. 1995; Boylan and Sprinkle 2001).

REFERENCES

- Abdel-khalik, A. R. 1994. Factors limiting the role of behavioral research in standard setting. *Behavioral Research in Accounting* 6 (Supplement): 213–222.
- American Institute of Certified Public Accountants (AICPA). 1994. *Improving Business Reporting—A Customer Focus: Comprehensive Report of the Special Committee on Financial Reporting*. Jersey City, NJ: AICPA.
- Barth, M. E., W. H. Beaver, and W. R. Landsman. 2001. The relevance of the value relevance literature for financial accounting standard setting: Another view. *Journal of Accounting and Economics* 31 (September): 77–104.
- Beresford, D. R., and L. T. Johnson. 1995. Interactions between the FASB and the academic community. *Accounting Horizons* 9 (December): 108–117.
- Berg, J., J. Dickhaut, and K. McCabe. 1995. The individual versus the aggregate. In *Judgment and Decision-Making Research in Accounting and Auditing*, edited by R. H. Ashton, and A. H. Ashton, 102–134. New York, NY: Cambridge University Press.
- Bloomfield, R., and T. J. Wilks. 2000. Disclosure effects in the laboratory: Liquidity, depth, and the cost of capital. *The Accounting Review* 75 (January): 13–41.
- . 2002. The “incomplete revelation hypothesis” and financial reporting. *Accounting Horizons* 16 (September): 233–243.
- Botosan, C. A., and M. A. Plumlee. 2001. Estimating expected cost of equity capital: A theory-based approach. Working paper, University of Utah.
- Boylan, S. J., and G. B. Sprinkle. 2001. Experimental evidence on the relation between tax rates and compliance: The effect of earned vs. endowed income. *The Journal of the American Taxation Association* 23 (Spring): 75–90.
- Camerer, C. F., and R. M. Hogarth. 1999. The effects of incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty* 19: 7–42.
- Cuccia, A. D. 1994. The economics of tax compliance: What do we know and where do we go? *Journal of Accounting Literature* 13: 81–116.
- Dickhaut, J., M. Ledyard, A. Mukherji, and H. Saprà. 2001. Information management and valuation: An experimental investigation. Working paper, University of Minnesota.
- Dietrich, J. R., S. J. Kachelmeier, D. N. Kleinmuntz, and T. J. Linsmeier. 2001. Market efficiency, bounded rationality, and supplemental business reporting disclosures. *Journal of Accounting Research* 39 (September): 243–268.
- Evans, J. H. III, R. L. Hannan, R. Krishnan, and D. V. Moser. 2001. Honesty in managerial reporting. *The Accounting Review* 76 (October): 537–559.
- Financial Accounting Standards Board (FASB). 1982. *Disclosures about Oil and Gas Producing Activities*. Statement of Financial Accounting Standards No. 69. Stamford, CT: FASB.
- . 1997. *Reporting Comprehensive Income*. Statement of Financial Accounting Standards No. 130. Norwalk, CT: FASB.
- Gibbins, M., and K. Jamal. 1993. Problem-centered research and knowledge-based theory in the professional accounting setting. *Accounting, Organizations and Society* 18 (July): 451–466.
- Gonedes, N. J. 1972. Efficient markets and external accounting. *The Accounting Review* 47 (January): 11–21.
- , and N. Dopuch. 1974. Capital market equilibrium, information production, and selecting accounting techniques: Theoretical framework and review of empirical work. *Journal of Accounting Research* 12 (Supplement): 48–129.
- Haynes, C. M., and S. J. Kachelmeier. 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.
- Hirst, D. E., and P. E. Hopkins. 1998. Comprehensive income reporting and analysts' valuation judgments. *Journal of Accounting Research* (Supplement): 47–75.
- Hodder, L., L. Koonce, and M. L. McAnally. 2001. SEC market risk disclosures: Implications for judgment and decision making. *Accounting Horizons* 15 (March): 49–70.

- Hopkins, P. E. 1996. The effect of financial statement classification of hybrid financial instruments on financial analysts' stock price judgments. *Journal of Accounting Research* 34 (Supplement): 33–50.
- , R. W. Houston, and M. F. Peters. 2000. Purchase, pooling, and equity analysts' valuation judgments. *The Accounting Review* 75 (July): 257–281.
- Kachelmeier, S. J. 1996. Do cosmetic reporting variations affect market behavior? A laboratory study of the accounting emphasis on unavoidable costs. *Review of Accounting Studies* 1 (2): 115–140.
- , and K. L. Towry. 2002. Negotiated transfer pricing: Is fairness easier said than done? *The Accounting Review* 77 (July): 571–593.
- Kahneman, D., and A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* (March): 263–291.
- Kennedy, J., T. Michell, and S. E. Sefcik. 1998. Disclosure of contingent environmental liabilities: Some unintended consequences? *Journal of Accounting Research* 36 (Autumn): 257–277.
- King, R. R., and D. E. Wallin. 1991. Market-induced information disclosures: An experimental markets investigation. *Contemporary Accounting Research* 8 (Fall): 170–197.
- , and ———. 1995. Experimental tests of disclosure with an opponent. *Journal of Accounting and Economics* 19 (February): 139–167.
- . 1996. Reputation formation for reliable reporting: An experimental investigation. *The Accounting Review* 71 (July): 375–396.
- . 2002. An experimental investigation of self-serving biases in an auditing trust game: The effect of group affiliation. *The Accounting Review* 77 (April): 265–284.
- Lee, C. M. C. 2001. Market efficiency and accounting research: A discussion of: "Capital Market Research in Accounting" by S. P. Kothari. *Journal of Accounting and Economics* 31 (September): 233–253.
- Libby, R., and W. R. Kinney, Jr. 2000. Does mandated audit communication reduce opportunistic corrections to manage earnings to forecasts? *The Accounting Review* 75 (October): 383–404.
- , R. Bloomfield, and M. W. Nelson. 2002. Experimental research in financial accounting. *Accounting, Organizations and Society* (forthcoming).
- Linsmeier, T. J., and N. D. Pearson. 1997. Quantitative disclosures of market risk in the SEC release. *Accounting Horizons* 11 (March): 107–135.
- Maines, L. A. 1994. The role of behavioral accounting research in financial accounting standard setting. *Behavioral Research in Accounting* 6 (Supplement): 204–212.
- . 1995. Judgment and decision-making research in financial accounting: A review and analysis. In *Judgment and Decision-Making Research in Accounting and Auditing*, edited by R. H. Ashton, and A. H. Ashton, 76–101. New York, NY: Cambridge University Press.
- , L. S. McDaniel, and M. S. Harris. 1997. Implications of proposed segment reporting standards for financial analysts' investment decisions. *Journal of Accounting Research* 35 (Supplement): 1–24.
- , and ———. 2000. Effects of comprehensive-income characteristics on nonprofessional investors' judgments: The role of financial-statement presentation format. *The Accounting Review* 75 (April): 179–207.
- Moser, D. V., J. H. Evans III, and C. K. Kim. 1995. The effects of horizontal and exchange inequity on tax reporting decisions. *The Accounting Review* 70 (October): 619–634.
- Schipper, K. 1994. Academic accounting research and the standard setting process. *Accounting Horizons* 8 (December): 61–73.
- Solomon, I., and M. D. Shields. 1995. Judgment and decision-making research in auditing. In *Judgment and Decision-Making Research in Accounting and Auditing*, edited by R. H. Ashton, and A. H. Ashton, 137–175. New York, NY: Cambridge University Press.

- Sprinkle, G. B. 2002. Perspectives on experimental research in managerial accounting. *Accounting, Organizations and Society* (forthcoming).
- Swieringa, R. J., and K. E. Weick. 1982. An assessment of laboratory experiments in accounting. *Journal of Accounting Research* 20 (Supplement): 56-101.
- . 1996. Some comments about research and the FASB. *Behavioral Research in Accounting* 8: 1-7.
- . 1998. Accounting research and policy making. *Accounting and Finance* 38 (July): 29-49.
- Tversky, A., and D. Kahneman. 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5 (October): 297-323.
- Zimbelman, M. F., and W. S. Waller. 1999. An experimental investigation of auditor-auditee interaction under ambiguity. *Journal of Accounting Research* 37 (Supplement): 135-155.