The Limitations of Experimental Design:
A Case Study Involving Monetary Incentive
Effects in Laboratory Markets

STEVEN J. KACHELMEIER*
McCombs School of Business, University of Texas at Austin, 1 University Station B6400, Austin, TX 78712, USA
email: kach@mail.utexas.edu

KRISTY L. TOWRY
Goizueta Business School, Emory University, 1300 Clifton Road, Atlanta, GA 30322, USA
email: Kristy_Towry@bus.emory.edu

Abstract

We replicate an influential study of monetary incentive effects by Jamal and Sunder (1991) to illustrate the difficulties of drawing causal inferences from a treatment manipulation when other features of the experimental design vary simultaneously. We first show that the Jamal and Sunder (1991) conclusions hinge on one of their laboratory market sessions, conducted only within their fixed-pay condition, that is characterized by a thin market and asymmetric supply and demand curves. When we replicate this structure multiple times under both fixed pay and pay tied to performance, our findings do not support Jamal and Sunder's (1991) conclusion about the incremental effects of performance-based compensation, suggesting that other features varied in that study likely account for their observed difference. Our *ceteris paribus* replication leaves us unable to offer any generalized conclusions about the effects of monetary incentives in other market structures, but the broader point is to illustrate that experimental designs that attempt to generalize effects by varying multiple features simultaneously can jeopardize the ability to draw causal inferences about the primary treatment manipulation.

Keywords: experimental design, monetary incentives, market power

JEL Classification: C92

1. Introduction

The *sine qua non* of experimental design is the ability to manipulate one feature of interest while holding all other features constant. It is only through this *ceteris paribus* structure that laboratory experimentation can lay claim to a stronger sense of causal inference than that characterizing most empirical-archival studies (Shadish et al., 2002). Nevertheless, the goal of generalizing results as broadly as possible can tempt the experimental researcher to manipulate multiple features from one treatment to the next. This temptation is perhaps

---

*Author to whom correspondence should be addressed.

especially acute in experimental economics, due to the high costs involved in both time and money to conduct multiple sessions in which participants interact. Researchers can acknowledge the limitations of non-*ceteris-paribus* designs as opportunities for future research (e.g., Krahnen and Weber, 2001), but the fundamental problem remains that when a treatment feature and other contextual features vary at the same time, the latter can account for any treatment effect attributed to the former.

In this research note, we illustrate the point by replicating an influential study of monetary incentive effects by Jamal and Sunder (1991) under *ceteris paribus* conditions. We find that when we hold constant the laboratory market structure most responsible for their results, we are unable to duplicate their primary conclusion about the incremental effects of salient monetary incentives on laboratory market behavior. Our goal is not to single out Jamal and Sunder (1991), as there are many other examples of non-*ceteris paribus* designs in the literature. Nor is our goal to assert any general refutation of their conclusion about monetary incentive effects, as the same design features that enable *ceteris paribus* conclusions also limit our ability to generalize results for something as complex as incentive effects in market behavior. Rather, our more modest goal is simply to illustrate the limitations of experimental economic research in which more than one feature is varied across sessions, using Jamal and Sunder (1991) as a case in point.

In Section 2, we summarize Jamal and Sunder (1991), comment on its influence, and explain the features that limit its ability to offer causal inferences. Section 3 describes the refinement in our replication of Jamal and Sunder (1991) that enables a *ceteris paribus* comparison in the setting most responsible for their claimed findings. Section 4 contrasts our findings against those reported by Jamal and Sunder (1991), and Section 5 returns to our motivating theme about the limitations of non-*ceteris-paribus* designs, along with the limitations of our replication study.

2. Jamal and Sunder (1991)

Jamal and Sunder (1991) (henceforth JS) investigate the incremental influence of performance-contingent payments on equilibrium price convergence across six laboratory market sessions. Their primary manipulation is whether compensation to participants is *salient* (i.e., based on profits earned) or *fixed* (i.e., a constant stipend per period, irrespective of outcomes). In addition, JS manipulate several other market features, including *learning* (salient-pay periods occur either before or after the fixed-pay periods), *experience* (whether participants had participated in a previous session), number of trading *periods*, market *size* (four or eight participants), and market *structure* (i.e., the degree of market power by buyers vs. sellers). These features vary in a non-orthogonal manner across sessions, as detailed in Table 1.

In each trading period, JS calculate a "Coefficient of Convergence" (COC), defined as follows:

$$COC = \frac{\sqrt{\frac{\sum (Tradeprice - CE)^2}{n}}}{CE} \times 100$$

Table 1. Experimental design used by Jamal and Sunder (1991).

Market session	Number of periods and order of treatment	Experienced participants?	Number of traders/equilibrium quantity traded	Market structure
1	5 periods fixed followed by 4 periods salient	No	4/2–4	Symmetric
2	5 periods salient followed by 5 periods fixed	No	8/4–5	Symmetric
3	5 periods fixed followed by 3 periods salient followed by 4 periods fixed	Yes Yes	8/4–5	Symmetric
4	9 periods (all fixed)	No	8/4–5	Symmetric
5	10 periods (all fixed)	No	4/2–3	Asymmetric (seller surplus > buyer surplus)
6	6 periods fixed followed by 5 periods salient	Yes	4/2–3	Symmetric

where

$n = $ the number of trades completed during a period, and
$CE = $ the competitive equilibrium price.

This metric is a useful statistic for measuring equilibrium price convergence, because it reflects both the average distance from the competitive equilibrium price and the level of dispersion. In an analysis of covariance, JS regress the Coefficient of Convergence as defined above against indicator variables for payment type (salient vs. fixed), market size, and experience. All three variables are statistically significant, with the payment-type effect indicating significantly better convergence (i.e., lower *COC* measures) when payments to participants are tied to profit performance than when payments are fixed and invariant to outcomes. Others have cited this result as evidence of the influence of monetary incentives in experimental economics (e.g., Smith and Walker, 1993; Friedman and Sunder, 1994; Anderson and Sunder, 1995; Jamal and Sunder, 1996; Camerer and Hogarth, 1999; Tung and Marsden, 2000; Brandouy, 2001; Hertwig and Ortmann, 2001). However, we explain below why the non-orthogonal nature of the manipulations in the JS experimental design call this result into question.

A regression approach (e.g., analysis of covariance) presumably separates the incremental effects of the separate variables in the regression. However, there are two problems with the regression approach used by JS. First, JS treat *COC* measures from each individual trading period as statistically independent observations, but clearly multiple measures within any given laboratory market session are correlated. As discussed later, an appropriate control for this problem is to include the market period in the model as a separate, "within-subjects" (i.e., repeated-measures) experimental factor (e.g., see Glantz and Slinker, 1990, Ch. 9 or Kuehl, 1994, Ch. 15). While this problem can be fixed statistically, a second and more

fundamental problem is that statistics alone cannot disentangle effects that vary simultaneously in an experimental design. A closer scrutiny of the JS results indicates that of their six market sessions, weak and noisy convergence to the competitive equilibrium prediction occurs in only two cases: the first half of market session 1 (fixed payments) and the entirety of session 5 (also fixed payments). The fact that both of these cases involve fixed payments seems to lend credence to the JS conclusion that salient payments tied to performance improve the reliability of competitive price predictions. However, JS market sessions 1 and 5 are also the only two sessions conducted by JS with inexperienced participants in a thin market of only four traders (see Table 1).

More specifically, in JS market session 1, an erratic price pattern in the first few periods under fixed payments converges to a more stable price pattern nearer the competitive equilibrium price prediction in the last few periods under salient payments. JS themselves recognize the obvious confound with learning in this session (i.e., learning alone would facilitate convergence in the later periods), so they conduct market session 2 in which the salient-pay periods occur first, followed by the fixed-pay periods. However, market session 2 also involves twice as many traders in a design with twice as many supply and demand gradations and twice the equilibrium quantity—all features that would facilitate competitive equilibrium convergence. Thus, the quicker convergence pattern in session 2 as compared to session 1 may be attributable to these other market features as opposed to the use of salient pay.

A more telling example is market session 5, in which fixed pay is used throughout. This session exhibits a marked degree of price volatility below the competitive equilibrium price. But this is also the only session conducted by JS with an asymmetric supply and demand structure, generating far more profit at the competitive equilibrium price for sellers than for buyers. As explained by Holt et al. (1986), an asymmetric surplus in double auction markets leads to asymmetric market power for the side with the smaller surplus, because that side can restrict quantity at less opportunity cost than the side with the larger surplus. This phenomenon would tend to be especially pronounced in thin markets such as JS market session 5, with only two buyers and two sellers. The price pattern observed by JS in session 5 is consistent with the general patterns found by Holt et al. (1986) in various salient-pay double-auction markets with asymmetric supply and demand. Thus, it is not clear that the lack of price convergence observed by JS in session 5 can be attributed to their use of fixed pay in this session.

To corroborate our suspicion that the unique structure of JS market session 5 likely accounted for the bulk of the difference that the authors attribute to fixed vs. salient pay, we repeated the analysis of covariance conducted by JS using their data (JS, Table 2) and approach (i.e., disregarding the statistical independence problem across periods), but adding an indicator variable for market session 5. Adding this variable increases R^2 from .53 to .67. The session 5 indicator variable is statistically significant ($p < .001$), but unlike the original model, the indicator variable for salient vs. fixed pay is no longer significant ($p = .23$). Consistent with the original model, the variables for market size and trader experience remain significant.

Thus, based on a closer inspection of the JS design and data, we have reason to suspect their conclusion about the incremental effects of fixed vs. salient pay on market price

behavior. By varying multiple features in market session 5, a session that is critical to the reported results but is conducted only within the fixed-pay condition, the JS design leaves open the question of whether it is the pay structure or other features unique to this session that account for the observed difference. Below we describe a replication study we conducted to find out.

3. Replication of JS market session 5

We conducted eight laboratory market sessions, each identical in structure to JS market session 5 (see figure 1 for supply and demand curves). In four of these eight sessions, fixed pay of $1.00 per period (per participant) in the first twelve periods is followed by salient pay equal to actual trading profits in the final six periods. In the other four sessions, this pattern is reversed (twelve periods of salient pay followed by six periods of fixed pay). Thus, our replication affords both a between-markets comparison across sessions and a "within-subjects" comparison within sessions of the incremental effects of fixed vs. salient pay, under *ceteris paribus* conditions. We use the structure in JS market session 5 for this replication because that session is pivotal to the payment-type effect reported by JS, as explained above. However, as we explain in our concluding comments, the thinness and asymmetric supply and demand structure of this market structure limit our ability to generalize conclusions about the effects of salient vs. fixed pay from this single market design.

As depicted in figure 1, the structure of JS session 5 involves four cost levels on the supply curve ($1.15, $2.35, $3.50, and $4.00), and four redemption value levels on the demand curve ($4.65, $4.00, $3.50, and $2.35), with two of the costs and two of the redemption values assigned to each of two sellers and two buyers, respectively. We reassigned these
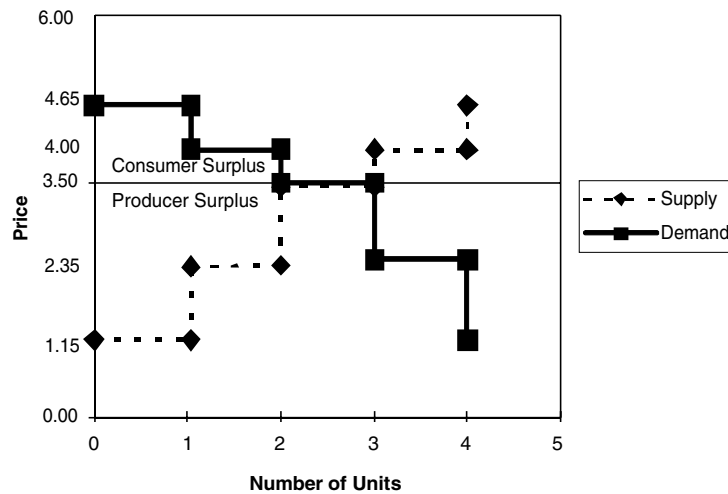

Figure 1. Market structure.

costs and redemption values each period, following a predetermined pattern that resulted in each seller and buyer experiencing all possible costs and redemption values an equal number of times, with the same pattern used in all sessions. The competitive equilibrium price is $3.50, at which either two or three units trade (with no marginal incentive for the third unit). At this equilibrium price, sellers earn a total surplus of $3.50 and buyers earn $1.65.

Participants were students recruited on a volunteer basis from business courses at a large university. Sessions were conducted on personal computers in a dedicated research laboratory, using double-auction software developed by Plott (1991). Instructions explained the payment scheme appropriate to each set of experimental periods, adhering to the JS payment manipulation as closely as possible.[1] In the salient-pay periods, participants calculated profits and were aware that these calculations determined their cash payoffs. In the fixed-pay periods, participants calculated their nominal profits from trading in a similar manner, but were informed that these profits were irrelevant to their compensation from participating.[2] Instead, as in the JS design, participants knew they would be paid $1.00 for each fixed-pay period. Similar to JS, this structure allowed us to vary the *nature* of pay (fixed vs. contingent on profit performance) without varying the mere *presence* of pay. In sum, our intent was to manipulate fixed vs. salient payments to participants in a manner identical to JS, but under *ceteris paribus* conditions. Participants earned an average of $19.80 each from combining the profits earned in the salient-pay periods and the $1.00 per period rewards in the fixed-pay periods, which when added to a fixed stipend for participating generated average payments of $25.30.

4.   Results

Table 2 reports trading prices and the coefficient of convergence (as defined in Section 2), averaged by period for the treatment with twelve periods of fixed pay followed by six periods of salient pay (i.e., the four fixed-to-salient sessions) and for the opposite treatment (i.e., the four salient-to-fixed sessions). Figure 2 plots average prices each period by session.

Focusing first on the coefficient of convergence, which is the dependent variable used by JS, Table 2 indicates that when we replicate the market structure primarily responsible for the JS results multiple times in both payment conditions, the average coefficient of convergence is identical in the first twelve periods (0.16 under both fixed and salient pay), and differs by only a negligible amount in the final six periods (0.13 vs. 0.14). The average convergence pattern across periods is depicted in figure 3 for each payment condition. The treatment lines cross at several points, suggesting no discernable differences between payment schemes, a conclusion that is corroborated in the statistical analyses we report below.

For statistical support, Table 3 reports a two-factor analysis of variance (ANOVA) on the coefficient of convergence metrics, with payment type as a between-sessions factor and trading period entering the model as a within-sessions factor. A within-sessions factor controls for interdependence across repeated observations within each session, and is equivalent to a regression approach with separate indicator variables for the individual market sessions (Glantz and Slinker, 1990, Ch. 9; Kuehl, 1994, Ch. 15). It is unnecessary to include

Table 2. Coefficients of price convergence and average trading prices.

	Coefficient of convergence[a]		Average trading price	
Trading period	Fixed-pay phase of the fixed-to-salient markets	Salient-pay phase of the salient-to-fixed markets	Fixed-pay phase of the fixed-to-salient markets	Salient-pay phase of the salient-to-fixed markets
Panel A. Periods before the reversal of payment type				
1	0.06	0.18	$3.46	$3.23
2	0.17	0.26	3.03	2.81
3	0.15	0.20	3.20	3.21
4	0.15	0.22	3.07	3.38
5	0.18	0.15	3.12	3.39
6	0.14	0.20	3.10	3.24
7	0.21	0.13	2.84	3.25
8	0.13	0.09	3.17	3.46
9	0.14	0.10	3.24	3.45
10	0.15	0.08	3.11	3.34
11	0.21	0.16	3.08	3.36
12	0.17	0.18	2.94	3.32
Average 1–12	0.16	0.16	$3.11	$3.29
Trading period	Salient-pay phase of the fixed-to-salient markets	Fixed-pay phase of the salient-to-fixed markets	Salient-pay phase of the fixed-to-salient markets	Fixed-pay phase of the salient-to-fixed markets
Panel B. Periods after the reversal of payment type[b]				
13	0.14	0.11	$3.11	$3.31
14	0.12	0.10	3.14	3.22
15	0.14	0.19	3.07	3.09
16	0.13	0.12	3.09	3.52
17	0.11	0.17	3.15	3.29
18	0.16	0.13	2.97	3.34
Average 13–18	0.13	0.14	$3.09	$3.30

[a]The coefficient of convergence is defined in Section 2. It captures both the distance from the competitive equilibrium price prediction and the level of price dispersion. Larger (smaller) coefficients indicate weaker (stronger) equilibrium price convergence.

[b]In periods 13–18, the payment manipulation is reversed. Thus, the columns in Panel B represent the opposite treatments as the columns in Panel A for the same market sessions.
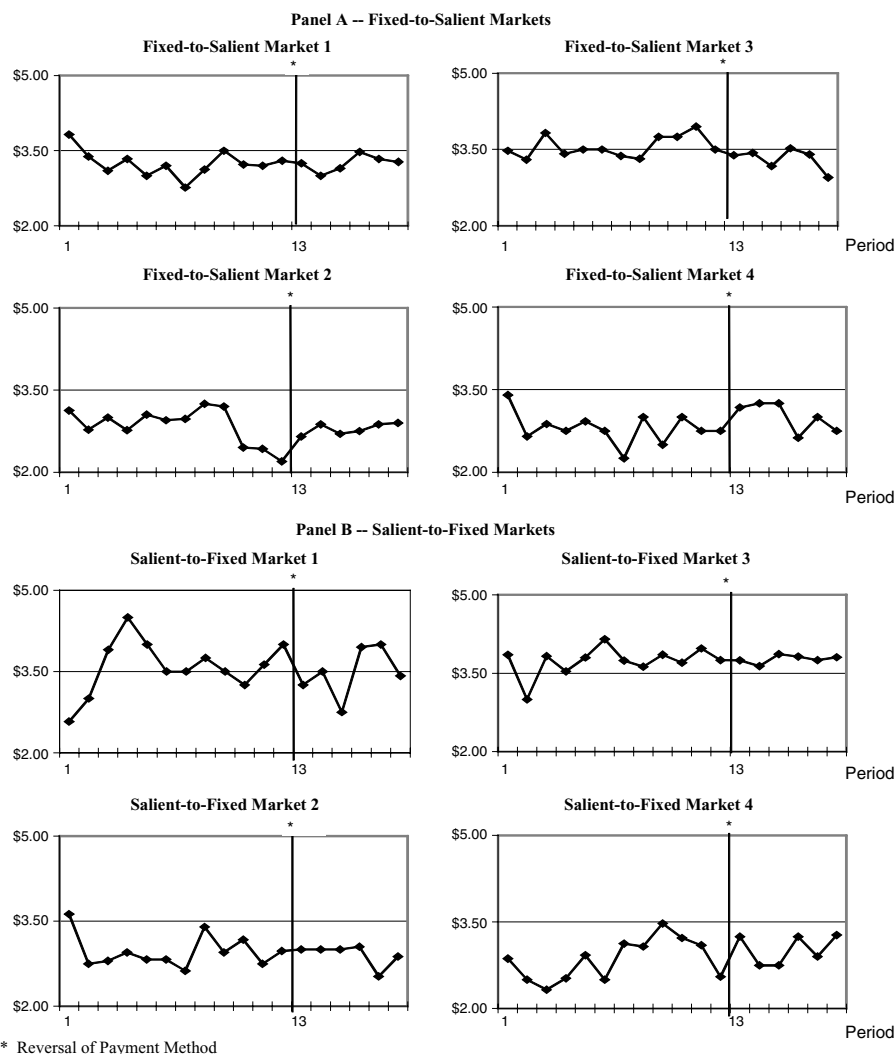
Panel A -- Fixed-to-Salient Markets

Figure 2. Trading prices.

other factors in the model, given that we hold all other features constant across sessions. We restrict this analysis to the first twelve periods, before the reversal of payment type. (Below we report a time-series analysis across all periods.) For the first twelve periods, the F-statistic for the effect of fixed vs. salient pay is negligible ($F = 0.04$; $p > .50$), corroborating the apparent similarity between treatment conditions in Table 2 and figure 3. Indeed, the ANOVA indicates no factors that are statistically significant at conventional levels.

Turning to a broader analysis of market prices, we conduct a time series analysis to determine if there is any "shock" in the price pattern attributable to the reversal of payment
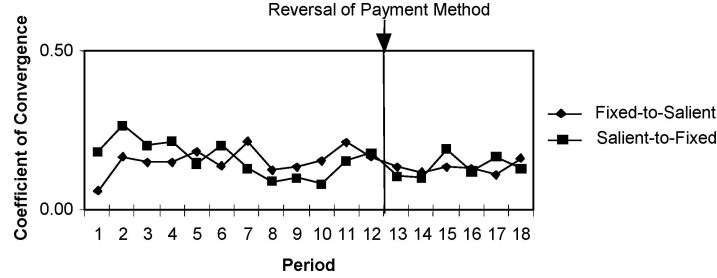
Figure 3. Coefficient of price convergence.

types after period 12. Specifically, following the advice of Glass et al. (1975), we fit an integrated, moving-average time-series regression model to the average price data, with a different model for the average of the four salient-to-fixed pay sessions and the four fixed-to-salient pay sessions.[3] The estimated model is as follows:

$$TRADEPRICE_p = \alpha + \beta SEG_p + (1 - \theta) \sum_{i=1}^{p-1} \eta_i + \eta_p$$

where

$TRADEPRICE_p =$ average trading price in period p,

$SEG_p = $ 0 in periods 1–12 and 1 in periods 13–18, to measure the incremental effect of reversing the payment scheme,

$\theta = $ Integrated moving average parameter, estimated via maximum-likehihood iteration to generate the θ with the smallest residual squared error, and

$\eta = $ error term.

In this model, the coefficient β estimates the incremental price effect of changing the payment scheme, controlling for trend and moving-average serial correlation. Results indicate

Table 3. Analysis of variance for coefficient of convergence, periods 1–12.

Source	df	Sum of squares	F	p-value (two-tailed)
		Panel A. Between-sessions factor		
Payment type (salient vs. fixed)	1	18	0.04	>.50
Error	6	2,573		
		Panel B. Within-sessions factors		
Period	11	1,006	1.45	.18
Period × Payment type	11	1,101	1.59	.13
Error	66	4,149		

Note: Reported p-values are two-tailed, after applying the Huynh and Feldt correction for variance heterogeneity.

that β is not significant in either the fixed-to-salient pay markets ($t = 1.24$; $p = .24$) or in the salient-to-fixed pay markets ($t = 1.69$; $p = .11$).

There is of course always some possibility that our failure to detect a statistically significant price effect between the fixed and salient pay conditions is due to low statistical power, rather than the true absence of an effect (McCloskey and Ziliak, 1996). A cursory review of the descriptive statistics in Table 2 does suggest that the average trading price of \$3.29 in the first twelve salient-pay periods of the salient-to-fixed markets is closer to the competitive equilibrium price prediction of \$3.50 than is the average trading price of \$3.11 in the first twelve fixed-pay periods of the fixed-to-salient markets. However, closer scrutiny of the price pattern after the reversal of payment schemes suggests that this apparent treatment difference is illusory. Specifically, in the last six periods, the average price difference between the salient-to-fixed and the fixed-to-salient markets persists and even widens slightly (\$3.30 vs. \$3.09), even though the payment schemes are reversed in these periods. More specifically, within the salient-to-fixed markets, the last observed average market price in the salient-pay phase (period 12) is \$3.29, and the last observed average price in the fixed-pay phase (period 18) is \$3.30. Similarly, the average period 12 price in the fixed-to-salient markets is \$3.11 and the average period 18 price is \$3.09. We conclude that any price differences between these markets likely reflect idiosyncratic price volatility across sessions, and not any systematic influence of the payment scheme.

Further evidence is apparent from the session-specific price plots in figure 2, in which we see price volatility that is largely consistent with observations by Holt et al. (1986) about market power in double auctions. For the most part, when prices diverge from the competitive equilibrium, they lie below the equilibrium price prediction of \$3.50, which is exactly what we would expect from the asymmetric market power of buyers in this setting. This below-equilibrium price behavior characterizes the second and fourth sessions of the fixed-to-salient pay markets as well as the second and fourth sessions of the salient-to-fixed pay markets. Other sessions seem to gravitate closer to the equilibrium price, with the exception of the first session of the salient-to-fixed pay condition, which shows relatively large price gyrations throughout. In sum, we observe price volatility in both treatments, with a tendency for prices to sometimes fall below the competitive equilibrium prediction. The volatility is consistent with what one would expect in a relatively thin market, and below-equilibrium prices are consistent with observations noted by Holt et al. (1986) when market surplus divisions are asymmetric. We conclude that there is no tenable basis for attributing an effect of fixed vs. salient payment incentives in this market structure, calling into question the different conclusions of JS from an experiment that implemented this structure only once, in their fixed-pay condition.

5.   Discussion and limitations

Evidence from an eight-session experimental design that varies the nature of compensation to participants but holds constant the features of a relatively thin market and an asymmetric division of market surplus indicates no significant effect of performance-based

payments on market behavior. This finding is noteworthy because the different conclu-
sions of Jamal and Sunder (1991) depend primarily on only one session conducted un-
der this structure, within their fixed-pay condition. Accordingly, our results suggest that
the effect of performance-based pay reported by JS is likely compromised by their con-
founding of market parameters (i.e., thinness and asymmetry) with the intended treatment
manipulation.

An important limitation of our conclusion is that we cannot offer any broad assertions
about the relevance or irrelevance of incentive-based pay in experimental economics based
on a design conducted only within this somewhat atypical market structure. The very features
of our design that enable a *ceteris paribus* comparison also limit our ability to generalize
results. It is possible if not likely that other market structures could exhibit significant
incentive effects, consistent with the general conclusion emerging from the incentives liter-
ature that incentive effects are context-specific and sensitive to a variety of factors such as
task difficulty, participant ability, and the extent to which accuracy depends on effort (for
reviews, see Arkes, 1991; Camerer and Hogarth, 1999; Bonner and Sprinkle, 2002). The
reader should not interpret our findings as evidence that incentives do not matter or as a
reason not to use incentive-based pay in experimental economic designs. Given the practi-
cal need to compensate participants as motivation to attend and complete an experimental
session, we see little reason not to pay participants based on performance in cases where it
is feasible to do so.

Broader conclusions about the effects of incentive-based pay in laboratory markets and
other experimental economic tasks await more ambitious studies that vary features that
theory suggests would lead to greater or lesser incentive effects, examining the interactions
between these features and the variation of payment schemes. Indeed, a desire to generalize
incentive effects across different settings is likely what led JS to conduct their study in six
different sessions with six different market structures. Our more modest objective in this
note is simply to point out that one cannot have it both ways. If the intent is to generalize,
the only defensible (and costly) approach is to vary the features of interest under *ceteris
paribus* conditions, with multiple sessions conducted within each unique experimental
cell.[4] Otherwise the researcher may be unable to disentangle a true treatment effect from
the effects of other features that vary simultaneously, as we have illustrated using JS as
a case in point. As mentioned in the introduction, our intent is neither to single out JS
nor to offer any general conclusions on incentive effects in experimental economics, but
rather is to encourage all experimentalists to adhere to the principles of sound experimental
design.

Notes

1. To avoid anticipatory behavior, the instructions for the last six periods with a new payment scheme (fixed-to-salient or salient-to-fixed) were not distributed until after the first twelve periods were completed.
2. Reiterating the instructions, the profit calculation sheets in the fixed-pay condition indicated that profits were "for our information" only, with explicit indication that each period's cash payoff was a fixed payment of $1.00, irrespective of profit results. Conversely, profit calculation sheets in the salient-pay condition noted explicitly that "Total profit = Your cash payoff for the period."
3. Ideally, the first stage of a time-series approach fits the data to the time series process most closely associated with the data, but we do not have enough time-series observations to do this. Accordingly, we apply the observation of Glass et al. (1975) that a simple first-order integrated moving average process (i.e., ARIMA 0,1,1) is typical for time-series data from experiments. For details of the maximum-likelihood estimation process, see Glass et al. (1975, pp. 134–140).
4. In a full factorial design, all unique treatment combinations are represented. Short of a full factorial, incomplete block and fractional factorial designs allow certain combinations to be omitted under specific circumstances, but also hinge on various assumptions of independence across treatment conditions (i.e., the absence of interactions). See Kuehl (1994, Chs. 9–12) or other research design texts for further detail.

References

Anderson, M.J. and Sunder, S. (1995). "Professional Traders as Intuitive Bayesians." *Organizational Behavior and Human Decision Processes*. 64, 185–202.

Arkes, H.R. (1991). "Costs and Benefits of Judgment Errors: Implications for Debiasing." *Psychological Bulletin*. 110, 486–498.

Bonner, S.E. and Sprinkle, G.B. (2002). "The Effects of Monetary Incentives on Effort and Task Performance: Theories, Evidence, and a Framework for Research." *Accounting, Organizations and Society*. 27, 303–345.

Brandouy, O. (2001). "Laboratory Incentive Structure and Control-Test Design in an Experimental Asset Market." *Journal of Economic Psychology*. 22, 1–26.

Camerer, C.F. and Hogarth, R.M. (1999). "The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework." *Journal of Risk and Uncertainty*. 19, 7–42.

Friedman, D. and Sunder, S. (1994). *Experimental Methods: A Primer for Economists.* Cambridge: Cambridge University Press.

Glantz, S.A. and Slinker, B.K. (1990). *Primer of Applied Regression and Analysis of Variance*. New York: McGraw-Hill.

Glass, G.V., Willson, V.L., and Gottman, J.M. (1975). *Design and Analysis of Time Series Experiments*. Boulder, CO: Colorado Associated University Press.

Hertwig, R. and Ortmann, A. (2001). "Experimental Practices in Economics: A Methodological Challenge for Psychologists?" *Behavioral and Brain Sciences*. 24, 383–451.

Holt, C.A., Langan, L.W., and Villamil, A.P. (1986). "Market Power in Oral Double Auctions." *Economic Inquiry*. 24(1), 107–123.

Huynh, H. and Feldt, L.S. (1976). "Estimation of the Box Correction for Degrees of Freedom from Sample Data in the Randomized Block and Split Plot Designs." *Journal of Educational Statistics.* 1, 69–72.

Jamal, K. and Sunder, S. (1991). "Money vs. Gaming: Effects of Salient Monetary Payments in Double Oral Auctions." *Organizational Behavior and Human Decision Processes.* 49, 151–166.

Jamal, K. and Sunder, S. (1996). "Bayesian Equilibrium in Double Auctions Populated by Biased Heuristic Traders." *Journal of Economic Behavior and Organization*. 31, 273–291.

Krahnen, J.P. and Weber, M. (2001). "Marketmaking in the Laboratory: Does Competition Matter?" *Experimental Economics*. 4, 55–85.

Kuehl, R.O. (1994). *Statistical Principles of Research Design and Analysis*. Belmont, CA: Duxbury Press.

McCloskey, D.N. and Ziliak, S.T. (1996). "The Standard Error of Regressions." *Journal of Economic Literature*. 34, 97–114.

Plott, C.R. (1991). "A Computerized Laboratory Market System and Research Support Systems for the Multiple Unit Double Auction." Social Science Working Paper 783. Pasadena: California Institute of Technology.

Shadish, W.R., Cook, T.D., and Campbell, D.T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* Boston: Houghton Mifflin.

Smith, V.L. and Walker, J.M. (1993). "Monetary Rewards and Decision Cost in Experimental Economics." *Economic Inquiry*. 21, 245–261.

Tung, Y.A. and Marsden, J.R. (2000). "Trading Volumes with and without Private Information: A Study Using Computerized Market Experiments." *Journal of Management Information Systems*. 17, 31–57.