EXPERIMENTAL EVIDENCE FOR AGENCY MODELS OF SALESFORCE COMPENSATION

Mrinal Ghosh
University of Michigan Business School
University of Michigan
Ann Arbor, MI 48109

George John
Carlson School of Management
University of Minnesota
Minneapolis, MN 55455

Forthcoming in *Marketing Science* (2001)

The authors gratefully acknowledge the valuable comments provided by the current and previous Editors, and the reviewers.

ABSTRACT

EXPERIMENTAL EVIDENCE FOR AGENCY MODELS OF SALESFORCE COMPENSATION

Academic work on sales compensation plans feature agency models prominently, and these models have also been used to build decision aids for managers. However, empirical support remains sketchy. We conducted three experiments to investigate three unresolved predictions involving the incentive-insurance trade-off posited in the model. First, compensation should be less incentive-loaded with greater effort-output uncertainty so as to provide additional insurance to a risk-averse agent. Second, flat wages should be used for verifiable effort so as to avoid unnecessary incentives. Third, less incentive-loaded plans should be used with more risk-averse agents so as to provide additional insurance.

Our design implemented explicit solutions from a specific agency model, which offers greater internal validity compared to extant laboratory designs that either did not implement explicit solutions or excluded certain parameters. In Experiment I, data from working manager subjects supported the first prediction, but only when risk-averse agents undertook non-verifiable effort. We interpret this as disclosing the model's "core" circumstance wherein it orders the data when the incentive-insurance trade-off is relevant. Thus, when verifiable effort made incentives moot, as is the case for the second prediction, the model failed to order the data.

Building on these results, we reasoned that the third prediction should find support among risk-averse agents, but not among risk–neutral agents, since insurance is a moot point with the latter agents. To this end, we added risk-neutral utility functions for agents in Experiment II. Data from MBA student subjects supported the predictions, but only when risk-averse agents

undertook non-verifiable effort. In those cells where the incentive–insurance trade–off was moot (either due to risk–neutrality or else due to verifiability), the data did not support the predictions.

We confronted several validity threats to these results. To begin, Experiment I used the standard agency solution which equalizes an agent's expected utility from the predicted plan with his expected utility from rejecting it. Subjects might have broken these ties on grounds like fairness. In order to assess whether this confounded the results, we derived new solutions in Experiment II that broke ties in favor of the predicted plan (by a 10% margin in the expected utility). Our results were robust to this change.

Second, our agents' behavior in Experiments I and II were much more consistent with predictions compared to the principals' behavior, which brought up task comprehension as a validity threat since our principals faced a more complex experimental task than the agents. To address this threat, we used three decision rounds in Experiment III to reduce the principals' task comprehension problems. A related validity threat arose from the relatively small gap in some cells between a principal's predicted expected utility and her next best choice. To address this threat, we derived new solutions with larger gaps to make her choices "easier." The results were again robust to these changes, which removes these validity threats.

We also addressed two alternative explanations. Might principals be predisposed to pick salary plus commission plans regardless of the model's predictions? If so, we should find such plans chosen uniformly across different experimental conditions. Pooling the data from our three experiments, we rejected this predisposition explanation by finding variation that was more consistent with treatment differences across cells. Second, might agents choose higher effort levels because of a demand bias? If so, we should find agents picking high effort regardless of the plan actually offered to them. Using pooled data, we rejected this explanation by finding variation that was more consistent with a utility-maximizing reaction to the plan actually offered to them.

Finally, we included manipulation checks to assess whether principals and agents perceived experimental stimuli identically as per the “common knowledge” assumption in game theory. These data showed no differences between agents and principals’ perceptions of stimuli.

Our experiments move the literature from simply asking whether the model works to pinpointing the circumstances where it orders behavior. The primary stylized fact we uncovered is the persistent and striking lack of support for the agency model outside the circumstance where risk-averse agents undertake non-verifiable effort. The model’s failure when there is no material insurance-incentive trade-off deserves scrutiny in future work.

1. INTRODUCTION

Agency Theory (AT) is arguably the most prominent economics-based approach to studying salesforce compensation issues. Derived from more general "mechanism design" principles, Holmstrom's initial model (1979) has been refined and applied to sales compensation in a series of papers (Lal, 1982; Basu, 1983; Basu, Lal, Srinivasan and Staelin (BLSS), 1985; and Lal and Staelin, 1986). Managerial decision support systems that design and implement compensation plans based on this approach have also been developed (e.g., Mantrala, Sinha and Zoltners 1994).

The intuition behind these models is as follows. A principal (the firm's owner) wishes to induce an agent (the salesperson) to undertake costly effort on her[1] behalf. Her goal is to derive the least costly compensation contract that will induce the agent to take the desired level of effort (from the standpoint of the principal). Compensation contracts written over *verifiable*[2] events are the only feasible contracts. AT's fundamental insight is that optimal contracts balance the stronger incentive effects of performance-based compensation (e.g., commissions) against the better insurance properties of flat-wage compensation.

To fix the intuition, consider the extremes. Suppose she designs a flat-wage contract over a non-verifiable level of effort. A utility-maximizing agent would shirk and deliver only the minimum level of effort. On the other hand, a commission-only contract tied to verifiable output (like sales) induces the agent to deliver more effort. However, he is now subject to variability in income from factors outside his control (e.g., product quality, brand advertising, etc.). Such contracts are unappealing to a risk-averse agent, unless they includes a sufficient risk premium (which then makes it more costly to the principal). Observe that if effort were verifiable, then the flat wage would be the least costly contract as it avoids the risk premium. More generally, AT models hold that the optimal combination of salary and incentive compensation required to evoke the desired level of effort from the salesperson is a function of several variables. These include the uncertainty of sales response to effort, the risk aversion of the agent, the outside opportunities available to the agent, and the set of verifiable events.

Despite its prominence, the empirical validity of AT compensation models remains sketchy with equivocal results found in the relatively small number of studies conducted to date. For example, results from studies testing the prediction that greater variability in sales leads to more salary-weighted plans have been supportive (Lal, Outland, and Staelin, 1994), non-supportive (John and Weitz, 1989), and opposite (Coughlan and Narasimhan, 1992). Similarly, the only test of the AT prediction that higher risk-aversion leads to salary-weighted plans found support only among a sub-set of relatively more risk-averse agents (Joseph and Kalwani, 1995).

2. GOALS AND CONTRIBUTIONS

Our experiments focus on three fundamental aspects of AT that remain unresolved in the literature; viz. a) the effort–output uncertainty of sales response, b) the verifiability of the agent's effort and b) the agent's risk preferences on the form of compensation contract. We proceed with the BLSS model as the point of departure in our discussion, and employ their functional forms and solutions. In some instances, we are required to depart from the BLSS predictions. For instance, our verifiable effort predictions are derived from our own specification and solutions as the BLSS model does not provides solutions for this case.

2.1 Results

Our data suggests that the basic AT prediction of the incentive-insurance trade-off holds to some degree in specific circumstances. In settings where *risk–neutral principals deal with risk–averse agents whose actions are non-verifiable*, we find that higher levels of effort–output uncertainty evoke more salary–weighted compensation plans. Outside this circumstance, the posited trade-off fails to find *any* support in our experiments. For instance, with risk-neutral agents (where insurance is moot) the expected commission-only plans were *not* selected. Likewise, with verifiable actions (where output-based incentives are moot) the expected salary-only plans were *not* selected.

2.2 Contributions

Our study extends the empirical AT literature in a number of ways. From a substantive standpoint, our results respond to Lafontaine and Slade's conclusion from their review that risk effects in AT are a "puzzle (that) deserves further attention" (1998). We uncover the conditions under which the posited risk effects do hold up. From a methodological standpoint, although the stronger internal validity of experiments over surveys is well known, our designs go further than extant experiments. By implementing a fully specified AT model in each experimental condition, we are able to generate stronger probes of the theory. In contrast, extant experimental protocols either fail to implement an explicit AT model (e.g., Umanath et al 1993, 1996) or else fail to implement one or more factors of an explicitly specified AT model (e.g., Berg et al 1992).

2.3 Organization of Paper

The empirical literature is reviewed in Section 3. Section 4 describes the generic experimental protocol used in our study. In Section 5, we report the specific methodology and results of Experiment I followed by a discussion of its results and potential artifacts. Experiments II and III, which are designed to address potential threats to the validity of the results from Experiment I, are presented in Section 6 and 7 respectively. Alternative behavioral explanations are considered in Section 8 using data pooled across the three experiments. Section 9 closes the paper with a discussion about the insights and the limitations of our work.

3. LITERATURE REVIEW

Empirical studies in AT can be usefully classified into questionnaire surveys and experiments. The primary dependent variable is the form of the compensation plan, particularly the split between incentive compensation and flat wages. We organize our literature review by the three relevant factors identified above.

3.1 Studies of Effort-Output Uncertainty

Greater uncertainty about the sales volume, conditional on an effort level, exposes an agent to higher risk. Here, AT predicts that salary–weighted plans are required to provide insurance to the risk-averse agent. This prediction has been studied in four survey-based studies. John and Weitz (1988) find *no support* for it in their primary data collected from 161 salesforces. Coughlan and Narasimhan (1992) found marginally *negative support* in their secondary database of 286 companies. Finally, Lal, Outland and Staelin (1994) collected data about 157 salespeople from 4 salesforces, and found *positive support* for the prediction.

Three internal validity threats pose problems for these results. First, the three studies employing salesforce level data to study the individual-level AT model suffer from aggregation biases (John and Weitz, 1988,1989; Coughlan and Narasimhan, 1992). Second, the measures employed in these studies suffer from face validity problems. John and Weitz's (1988, 1989) measure of uncertainty as the difficulty of forecasting sales omits the conditional variability aspect of the construct's definition. Coughlan and Narasimhan's (1992) measure of uncertainty as the number of calls required to close a sale appears to track the k parameter of the BLSS model more closely rather than the n parameter (representing the number of prospects) that BLSS use as their uncertainty parameter.[3] Lal, Outland and Staelin's (1994) measure of uncertainty as the difficulty in forecasting an individual salesperson's output *conditional* on his effort level comes closest to the AT conceptualization. Interestingly, this study is also the only one to find support for the prediction.

Amershi and Butterworth (1988) describe succinctly the most problematic internal validity threat faced by survey-based studies. In order "…to make exact predictions (for testing), *both* technology and (risk) preferences must be known" (emphasis added). Technology refers to the effort–output function, while risk preferences refer to the utility function. AT models depend non-trivially on the assumptions made about these functions. Hence, without controlling (knowing) both functions, internal validity is seriously compromised. Unfortunately, such control is impossible in a survey.

In principle, this defect could be resolved by switching to an experimental setup where we can explicitly manipulate risk preferences and effort–output uncertainty. Unfortunately, these manipulations have not been implemented simultaneously in extant work. For instance, two experimental studies (Umanath et. al 1993, 1996) have investigated the impact of effort–output uncertainty on compensation plans, but neither one manipulated utility functions and effort–output functions. Instead of deriving solutions form an explicitly specified AT model, they constructed verbal scenarios based on unspecified forms of these functions. Thus, we cannot match up the results from their experiments against the theoretically expected utility maximizing choice.

3.2 Studies of Verifiability of Effort

The AT prediction that verifiable effort evokes flat wages has been examined in three studies. John and Weitz's (1989) survey work discovered that increased difficulty in monitoring salespeople leads to *increased* salary–weighted compensation which is contrary to the expectation.[4].

Berg et al (1992) did find support for the AT prediction in their laboratory experiment, but their design is seriously compromised by not implementing the reservation utility parameter. This omission should bias the results because it changes the game type from an ultimatum game to a dictator game.[5] Epstein (1992) demonstrates this bias in his replication of the Berg et al study. The only difference in the two studies is Epstein's introduction of the reservation wage parameter (albeit at an arbitrary level). This single change produced dramatically different results with fewer than 50% of Epstein's subjects acting according to AT compared with 90% of the Berg et al subjects. Unfortunately, he did not incorporate a verifiable effort condition, so his study is not dispositive about this AT prediction.

3.3 Studies of Risk Preferences

Joseph and Kalwani (1995) investigate the agent's risk preferences in a survey-based study. As per AT, they find that the proportion of salary to total compensation increases with risk aversion, but this holds only for relatively high levels of risk aversion.[6] Turning to the experimental studies on this topic, we find that risk preferences have never been manipulated as a factor at multiple levels in any single study, so the AT prediction remains unresolved.

3.4 Summary of Gaps

The survey-based work has made headway in uncovering some support for the broad insights of AT models, yet they suffer from a generic inability to probe the theoretical mechanisms rigorously. Difficulties in measuring unobserved constructs like risk preferences and controlling for technology and risk preferences simultaneously make it difficult to sort out the inconsistent survey-based findings. Although experimental designs are stronger per se, the extant experiments have been compromised by various design problems. Addressing these gaps is our primary design challenge.

4. IMPLEMENTING AN AGENCY MODEL

We employ the BLSS model of salesforce compensation as our starting point. To the extent possible, we use the functional forms specified in their model. All the compensation contracts from this model are of the general form $(A + Bx)^2$ with A and B representing fixed and incentive (commission rate) parameters respectively. Appendix 1 summarizes the parameters and the functional forms used.[7]

4.1 General Procedure

Subjects are told that they will be participating in a marketing simulation, and will earn rewards in proportion to their individual profits in the simulation. No other incentive is offered.

Setting. Each subject is randomly assigned to a cell to act as a principal or as an agent. Neither party knows the identity of the other. Each principal is told that she has to hire a rep for the sole task of selling computer systems to automobile dealers at a fixed price. All revenues and costs in the experiment are denominated in 'pesos'. All communication is limited to the decisions at hand, and is delivered via the experimenter. Figure 1 shows the sequence of moves.

Inducing Utility Functions. We use the risk induction procedure developed and validated by Berg et al (1986) to implement specific utility functions for each subject. This procedure only requires that subjects prefer more money to less and that they understand elementary arithmetic. Essentially, the procedure allows subjects to play a prize wheel using the pesos that they earned during the main phase of the experiment. The probability of earning a reward with these wheels increases with more pesos. The win area on a wheel is calculated from the desired utility function. For instance, all the wheels in Figure 2 show a risk-averse utility function where the winning arc area increases at a decreasing rate. In contrast, it would increase at a constant rate for a risk-neutral function. The point values at 0 and 360 degrees correspond to the minimum and maximum pesos that subjects can earn in that experimental condition.

Our principals' wheels always mapped a risk–neutral function, $U_P(Q)$, where Q denotes the principal's payoffs *net* of the agent's compensation. Our agents' wheel always mapped $U_A(q)–V(t)$ where $U_A(q)$ is the agent's utility for compensation q and $V(t)$ is his disutility for effort level t. We require multiple wheels for agents because each wheel maps the difference of two distinct functions. Each wheel corresponds to one of the multiple courses of action available to an agent in a particular experimental condition. For instance, an agent facing three options (reject the pay plan; accept the pay plan and take high effort; accept the pay plan and take low effort) faces three prize wheels. Figure 2 shows an illustrative set of wheels for an agent.[8]

Principal's decision. We show each principal her prize wheel as well as her agent's prize wheels. Each principal is asked to select a compensation plan from a menu of three plans. Each plan is described in tabular and graphical formats showing each party's payoffs given various levels of realized sales (Figure 3). We also provide them with sales forecast graphs depicting the likelihood of attaining each level of sales (Figure 4).

Agent's decision. We show each agent his principal's offer using the same graphical and tabular formats described above. Each agent is asked to accept or reject the offered plan. We report this choice to the principal. If the agent accepts the plan, he is then asked to choose an effort level. This latter decision is *not* reported back to the principal. Each agent also sees one prize wheel for each of his courses of action.

Disutility for Effort. The agent's disutility for effort, $V(t)$, is operationalized using prospect list(s)[9]. This list(s) is available for purchase from a market research firm (played by a confederate). In some cells, agents could choose between two prospect lists; a good (costlier) list and a poor (cheaper) list. In other cells, agents have only one (good) list available. The prospects on the good list are more likely to buy than the prospects on the poor list. Being more expensive, buying a good list induces higher disutility than does a poor list. The prospect list information is shown to all subjects as sales forecast graphs (Figure 4).

Reservation utility. We operationalize outside market opportunities by crediting the agent with an amount of pesos that he may keep if (and only if) he rejects the pay plan offered by his principal.

Effort–output uncertainty. We operationalize the stochastic effort–output function by using a binomial function simulator (Appendix 1). The number of prospects/trials (*n*) indexes the variability in sales output conditional on the effort level (*t*) of the agents. Subjects are shown sales forecast graphs for specific values of *n* (4, 8 in the low and high uncertainty cells respectively).

Verifiability of agent's effort. Verifiability is operationalized as the number of prospect lists available for purchase. Two lists implement non-verifiable effort, while a single list implements verifiable effort. To understand this, recall that the agent's list purchase behavior is *not* reported back to the principal. Nevertheless, she can infer his effort level in a single list cell from knowing whether her pay plan was accepted or rejected. Suppose that her agent accepts the plan. She knows his effort level since there is only one list (with a known disutility) available to him. In contrast, with two lists, she cannot infer which list was purchased because of the stochastic effort–output function. In the event that the agent rejects the plan, the effort level is always known to be zero.

Payoffs. Once the sales volumes are simulated, the principal and the agent receive their respective payoffs. Each of them places their earnings on the appropriate prize wheel. The wheels are spun, and the rewards are dealt out.

5. EXPERIMENT I

This design tests the posited effect of a) effort–output uncertainty, and b) verifiability of agent's effort on the choice of pay plans. We construct 4 cells, each involving one of two levels of effort-output uncertainty (n = 4, 8) and one of two levels of verifiability (either one or two prospect lists). All other model parameters are held constant across the cells. In each cell the principal chooses from a menu of three compensation contracts. To guard against demand artifacts from non-comparable choice sets, we develop an identical menu of three contracts in each cell. Using the particular parameter values and specific functions documented in Appendix 1, we derive the predicted contracts for each cell. These contracts provide each principal her highest net expected utility while meeting the participation constraint and incentive compatibility conditions for her agent.[10]

Table 1 summarizes the four cells and the predicted contracts. Table 2 shows the expected utilities for principals and agents for all possible courses of action and the predicted equilibrium (in bold)[11] in each cell. We offer an intuitive description of the predictions below.

5.1 Predictions

The mixed pay plan is predicted in each of the two non-verifiable effort conditions (Cells 1 and 2). However, the commission rate shrinks from 24 (Cell 1) to 12 (Cell 2) in line with the rise in uncertainty. The salary-only plan is predicted in both of the two verifiable effort conditions (Cells 2 and 4).[12]

To be more precise, AT predicts *contract-action pairs*[13]. As such, “Plan X–high effort” is the predicted pair in Cell 1, “Plan Y–high effort” is predicted in Cell 3, and “Plan Z–high effort” is predicted in Cells 2 and 4.[14].

5.2 Subjects

Fifty participants in an executive development program at a large mid–western university were randomly assigned into 25 principal-agent dyads across the four cells. On average, our subjects had over 10 years of work experience and they came from a variety of backgrounds and functional areas. To motivate them, the experimenter (instructor) announced that winners would get a nominal amount of money ($5) and a T–shirt imprinted with the program logo and the names of all the winners. Losers would get nothing. This prize structure is consistent with the binary lottery format required for our utility induction procedure (Roth and Malouf, 1979; Berg et al (1986). This announcement was displayed prominently on flip boards using mock winner lists and a sample T–shirt. These actions generated a high degree of interest. No other cover story was used to disguise the study. Post-study briefings revealed no problems with hypothesis guessing.

5.3 Results

Table 3 reports the observed outcomes. Strictly speaking, AT predicts that everyone will behave as per the theory. There is no ‘error’ theory. As a practical matter, however, we cannot test the data against a 100% conformity benchmark, as it would be too “easy” to reject the theory. We require a more realistic baseline whereby AT can be said to order the data better than a reasonable null hypothesis. To this end, we use the null hypothesis that all the possible outcomes in a cell are equally likely. Behaviorally, this reduces to assuming that our subjects are choosing random courses of action. For instance, in Cell 1, there are 9 possible contract-action outcomes (three choices for the principal each paired with three choices for the agent), and the predicted contract-action pair would be randomly chosen 1/9 of the time. The observed choice should exceed this baseline for us to favor the theory over the null hypothesis.

Contract–action pairs. Table 3 shows that 11 of 25 dyads (44%) chose the predicted contract–action pair across the 4 cells. There are 9 possible outcomes in two cells and 6 outcomes in the other two cells. Using a more conservative null hypothesis of 6 equally likely outcomes (i.e., 17%), the estimated confidence interval of 0.44 +/– 0.20 excludes this baseline, so we can reject the null hypothesis at the 95% confidence level.[15]

We also observe a decided pattern of support for the model. In the two non-verifiable effort cells, 10 of 13 dyads (77%) selected the hypothesized contract–action pair, which is significantly greater than our baseline of 9 equally likely outcomes (0.77 +/– 0.24 excludes 0.11). These data also support the predicted trade–off between incentives and risk–sharing. Recall that the more incentive loaded plan (X) is predicted in the lower variance condition (Cell 1) and the more salary loaded plan (Y) is predicted in the higher variance condition (Cell 3). Indeed, this is the case in the data.

However, we find no support at all for AT in the other two cells. Only 1 of 12 dyads (8%) in these verifiable effort cells acted consistent with the theory. This is not significantly different from the null hypothesis of 6 equally likely outcomes (0.08 +/– 0.16 includes 0.17). To gain more insight, we decompose the contract–action pairs into principal's choices and agent's reactions.

Principals' choices. Overall, 13 of 25 principals (52%) offered the predicted contract, but our null hypothesis of 3 equally likely choices cannot be rejected (0.52 +/– 0.20 includes 0.33). Incidentally, these observed levels are very close to the corresponding data reported in Epstein (1992). Further analysis on our data reveals a decided pattern. In the two non-verifiable effort conditions, 12 of 13 principals (92%) chose the predicted plan, which is significantly greater than our baseline (0.92 +/– 0.15 excludes 0.33).

However, in the verifiable effort cells, only 1 of 12 principals (8%) chose the predicted plan. While this is statistically different from the baseline (0.08 +/– 0.16 excludes 0.33), it is in the *wrong* direction. In sum, when agents' effort levels cannot be verified, principals appear to make the incentive–insurance trade–off as per AT in offering compensation plans. However, when agents' effort levels *can* be verified, virtually all the principals ignored the AT position that there is no need to offer the (more expensive) incentive plans.

Agents' reactions. Strictly construed, AT does not speak to off–equilibrium observations. Nevertheless, for each agent, we can calculate his expected utility maximizing reaction *conditional* on the offered plan (Table 2). In order to construct a baseline for these data, we have to contend with agents facing three options (reject; low effort; high effort) in two cells but with only two options in the other two cells. Additionally, a reject reaction is indistinguishable from high effort reactions because both actions yield the same expected utility. We constructed a conservative baseline wherein a 'reject' reaction is included as an outcome that is consistent with the theory. As such, 0.50 is our baseline for these data. Across the four cells, 22 of 25 agents (88%) reacted in the fashion consistent with AT, which is significantly higher than our baseline (0.88 +/– 0.13 excludes 0.50).

Our agents' reactions appear to be more consistent with theory than our principals' choices. A likely reason is that an agent's task is less complex compared to his principal. We shall return to this issue later. In addition, unlike the contract-action data, or even the principals' choices, these reaction data are identical across the verifiability manipulations. To sum up, agents appear to act in their own self–interest given the plans offered to them by the principals. There is no evidence of "collusion" with their principals.

5.4 Discussion

AT fares better than our baseline in the two cells that implement the 'core' circumstance of the model, viz. when a risk–averse agent's actions are non-verifiable. Here, the contract–action pairs, principals' choices, and agents' reactions are all more consistent with AT than our baselines. Most importantly, we find that incentive plans selected are sensitive to the risk faced by the agents. Cells with greater variability in expected outcomes given effort levels evoked pay plans with a lower incentive component as per AT. To the best of our knowledge, this is the first experimental study to explicitly disclose the trade–off predicted by AT between incentives and insurance.

In contrast, AT does not seem to hold up at all outside these 'core' conditions. Recall that only 1 of 12 principals chose the predicted salary–only plan (Z) that was expected in Cells 2 and 4. Why are salary–only plans so unpopular among our principals?

One possible reason is that each agent's expected utility from the predicted option in each cell (i.e. high effort) is exactly equal to his expected utility from the reject option. This equality in expected utility is particularly transparent with salary–only plans since there is no uncertainty about their expected compensation. Recall that our subjects see their peso payoffs as win areas on their prize wheels. The agent's win area for his salary–only plan and the corresponding area for his reject option are immediately and obviously equal. In contrast, there is a range of possible win areas for an incentive plan. Thus, it is plausible that agents perceive salary-only plans as "unfair". Anticipating this, principals might have resorted to the more generous incentive plans. This type of strategic bargaining is not accommodated in AT, which assumes take-it-or-leave-it ultimatum offers by principals. If this "tie" problem exists, it constitutes a serious confound, as it parsimoniously orders our data. Experiment II is designed to tackle this validity threat.

6 EXPERIMENT II

Experiment II introduces changes intended to rule out the tie-break validity threat, and to specify more closely the boundaries of the 'core' circumstance of AT. Cells 1 and 2 involve risk–averse and risk–neutral agents respectively undertaking non-verifiable effort. Cell 3 involves risk–averse agents undertaking verifiable effort. As before, principals are risk–neutral. Table 4 summarizes the three cells and the predicted pay plans.

6.1 Differences from Experiment I

Tie–breaking. We derived new solutions such that the reject option now yields 10% *less* expected utility to an agent than his utility from the predicted contract-action pair for that cell.

Risk preferences. We induced a risk–neutral utility function for agents in Cell 2. The trade–off between incentives and insurance is moot when agents are risk–neutral. Thus, if AT fails whenever this trade–off is not relevant, it should fail here just as it did with verifiable effort in the first experiment.

Other differences. Sixty-six MBA students were randomly assigned into 33 principal-agent dyads across the 3 cells. On average, these subjects had 2-3 years of work experience, and they came from a variety of backgrounds. Unlike the previous experiment, cash was the sole reward offered. Winners earned $10 and losers earned $1.

6.2 Predictions

The salary plus incentive plan (X), the incentive–only plan (Y), and the salary–only plan (Z) are the principals' predicted choices in Cells 1, 2, and 3 respectively. Contingent on being offered these plans, the agents' predicted reaction is high effort in all cells. Table 5 shows the expected utilities for each possible course of action.

6.3 Results

Contract–action pairs. Table 6 shows that 8 of 33 dyads (24%) chose the predicted contract–action pair. Overall, we cannot reject our null hypothesis of equally likely random choices (0.24 +/– 0.15 includes 0.17). However, there is support in certain cells. As we saw previously, given risk-averse agents undertaking non-verifiable effort (Cell 1), 8 of 12 dyads (75%) chose the predicted contract-action pair. This is significantly higher than our baseline (0.75 +/– 0.27 excludes 0.11). However, given risk-neutral agents undertaking non-verifiable effort (Cell 2), *none* of the 10 dyads chose the predicted contract–action pair. Finally, despite the changes we introduced to break the utility tie, none of the 11 dyads in Cell 3 chose the predicted salary-only plan. These data dispose of the validity threats discussed in connection with the results of Experiment 1. They also strengthen our notion that AT orders the data support only when the insurance-incentive trade-off is relevant.

Principals' choices. Overall, 12 of 33 principals (36%) acted consistent with the theory, which does not reject our null hypothesis (0.36 +/– 0.17 includes 0.33). However, there is support in one cell. Fully 92%, of the choices in Cell 1 were consistent with AT, as compared to 10%, and 0% in Cells 2 and 3 respectively. This selective pattern reaffirms Experiment I. Support is found in the 'core' circumstance (Cell 1), while Cell 2 shows choices made in the *opposite* direction to AT. Finally, the Cell 3 data rule out the tie-break validity threat.

Agents' reactions. Overall, 23 of 33 agents (70%) undertook reactions consistent with AT, which is significantly higher than our baseline. These data affirm the robustness of our conclusion from Experiment I that the agents' reactions are more consistent with AT than are the principals' choices.

6.4 Discussion

We have found that the design changes in Experiment II effectively dispose of the tie-break validity threat, while uncovering robustness in the pattern of support. AT is supported only when risk-averse agents undertake non-verifiable effort. However, there are three unresolved validity threats. First, we find consistently greater support in the agent reaction data. Why? Given the simpler task facing the agents compared to principals, this raises the possibility that principals might have shown stronger support had their task comprehension been higher. One way to test this is to allow our subjects to play repeated trials. A drift towards the predicted equilibrium over the trials would confirm that task comprehension is an artifact that explains the non–supportive results in our data.

The second unresolved threat concerns the manipulations themselves. Do our subjects see these treatments as intended? In particular, perceptual differences of identical stimuli across principals and agents would be a validity threat given the common knowledge assumption built into AT. Manipulation checks in our protocol can be used to assess this threat.

The third unresolved threat is the matter of 'near ties' faced by principals. In some cells, the principal's expected utility from the predicted contract-action pair is very close to her next best alternative.[16] If principals are indifferent to small differences, they might choose randomly across the two alternatives, which would then dampen support for AT. We can address this threat by choosing parameter values that create a larger gap in expected utility between the predicted action and the closest alternative.

7 EXPERIMENT III

Experiment III consists of the 3 cells in Table 7. Here, 52 participants in an executive development program were randomly assigned into 26 principal-agent dyads across the 3 cells.

7.1 Differences from Experiment II

Repeated trials. In order to rule out task comprehension as a possible validity threat, each subject makes three rounds of decisions. This necessitated large changes in the protocol. Given that the model being tested is a single–period model, we need to be careful that the multiple rounds do not inadvertently create a multi–period (repeated) game. Hence, we announced that the first two rounds would be trial rounds, to be followed by one 'real' round. We followed each practice round by debriefing the subjects about the procedure. This served to reinforce the role of the practice rounds as opportunities to learn the structure of the game. We also announced that principals and agents would be randomly reassigned to different anonymous partners after each round. This rules out any effect of future anticipated interaction with the same partner.

Near ties. New parameter values were chosen to create larger gaps (greater than 10%) in the expected utility between the predicted option and the next best action for the principal.

Manipulation checks. In order to assess possible differences between the principals and the agents on their perceptions of identical stimuli, we included a post-experimental questionnaire with items for two constructs: disutility of effort and the uncertainty in outcomes. The items consisted of 5-point Likert-type scales.

7.2 Predictions

Table 8 shows the expected utilities in each cell. "X-high effort", "Y-high effort", and "Z-high effort" are the predicted contract-action pairs in Cells 1, 2, and 3 respectively.

7.3 Results

Manipulation checks. The mean response to the item measuring the sacrifice associated with purchasing a prospect list significantly greater than the mid-point of the scale (3.47 vs. 3). More importantly, the principals' ratings are not significantly different from the agents' ratings. The average response to the question about uncertainty of sales outcomes was significantly higher for the 8-prospect list compared to the response for the 4–prospect list (3.44 vs. 2.85). Again, these ratings did not differ across principals and agents. These data ease concerns about the validity threat posed by manipulation failures.

Contract–action pairs. Table 9 shows outcomes that are remarkably similar to the data from our previous experiments. Significant support for AT is confined to Cell 1 (risk-aversion with non-verifiable effort). Here, 20 of 27 dyads (74%) behaved as per the prediction. However, there is no significant increase in support across the repeated trials, which effectively disposes of the task comprehension issue.

The outcomes in Cell 2 (risk-neutral agents) are not significantly different from our baseline. Only 5 of 27 dyads (19%) chose the predicted pair. In Cell 3 (with verifiable effort), only 1 of 24 dyads (4%) chose the predicted contract–action pair. This is significantly lower than our baseline expectation. As with Cell 1, Cells 2 and 3 show no discernable effect of task comprehension, as the outcomes are uniform across the three rounds. In sum, neither task comprehension or manipulation failure seems to pose viable validity threats to the results.

Principals' choices. These data are also consistent with the previous experiments. Across the three cells, 81%, 26%, and 8% of the observations in Cells 1, 2, and 3 were consistent with AT. Again, support is confined to the 'core' circumstance (Cell 1), while Cell 2 is not different from the baseline, and Cell 3 shows outcomes that are significantly lower than the baseline. No trend occurs in the data across the trials, which rules out task comprehension as a validity threat.

Agents' reactions. Again, the agents' reactions are consistent with the theory in all the three cells (88%, 76% and 83% in Cells 1, 2 and 3 respectively). These outcomes are all significantly higher than our baseline.

8 ALTERNATIVE BEHAVIORAL EXPLANATIONS

Principals' Choices. The most popular choice of pay plan across the three studies is the mixed plan. Might principals choose this plan because of a decision bias or predisposition? There are reasons to expect such a bias. The mixed plan looks less extreme and there is a well-documented decision-making bias towards "picking the middle" (e.g., Simonson, 1989). Alternatively, subjects might assume that mixed plans are more common in practice than either extreme, and follow this norm in their role as principals. If true, then our results for those cells where such plans were indeed picked (and attributed to the model) might well be explained without invoking AT. Furthermore, this explanation is particularly problematic since it parsimoniously explains the persistent lack of support for both salary-only and incentive-only plans that we find in the data.

In order to discriminate AT's predictions about mixed plans from a general bias towards such plans, we reasoned as follows. If our principals have a general bias toward picking the mixed plan, such plans should be chosen at the same rate across the cells regardless of the inductions. Conversely, if the experimental inductions are responsible for the results, the proportion of mixed plans being chosen should vary depending on the inductions. We pooled the data from the three studies to test this alternative explanation. Specifically, we sorted the 10 different experimental conditions from the three experiments into two categories. Let the *Mixed* category consist of those 4 cells where AT predicts a mixed plan, while *Pure* consist of the 6 cells where AT predicts either a salary–only or a commission–only plan. If our subjects were predisposed toward the mixed plan, the observed proportion of mixed plans should be equal across the two categories. Conversely, if AT is correct, the mixed plans should be chosen more frequently in the *Mixed* category.

In fact, the observed proportion of mixed plans is significantly higher in the *Mixed* category than in the *Pure* category ($z=3.84$; $p < 0.05$). This argues in favor of the AT explanation and rules out the validity threat posed by a predisposition towards the mixed plan. Further support for the AT position over the predisposition bias comes from the 4 cells in the *Mixed* category where the mixed plan was the theoretically predicted choice. Here, subjects chose various mixed plans in a discriminating fashion. For instance, subjects in Experiment I chose *different* mixed plans for *different* levels of uncertainty. This provides further evidence that principals did not naively choose mixed plans.

Agents' Reactions. It might be argued that agents are socially predisposed toward high effort. After all, agents see that higher effort yields more sales, and the social desirability pressures within a classroom simulation of a selling situation might guide them to choose the more "appropriate" effort level. We constructed the following test to discriminate between a predisposition explanation for higher effort and the utility maximizing explanation in AT.

We capitalized on the fact that not all principals picked the predicted plans in our data. In total, there are 135 contract offers across the 10 cells in the 3 experiments. Of these, 19 were plans where an expected utility-maximizing agent should *not* have reacted with high effort.

These data allow us to sort between the two explanations. Let *High* represent those 116 observations where high effort is the expected utility maximizing reaction, while *Not High* represent the 19 observations where low effort, or rejecting the contract, is the expected utility maximizing reaction. Under the predisposition interpretation, the proportion of high effort reactions should be *equal* across the two categories. Conversely, under the AT interpretation, the proportion of high effort choices should be greater in the *High* category compared to the *Not High* category. In fact, high effort choices were significantly more frequent in the *High* category than in the *Not High* category ($z=6.39$; $p < 0.05$), which rejects the null hypotheses. This rules out the alternative explanation that agents might have responded with high effort because of a social desirability bias.

9 DISCUSSION

9.1 Limitations

Although we have ruled out several threats to validity as well as alternative explanations, a few limitations should be underscored. The first concern is the appropriate way to translate as-if mathematical models into realistic experimental predictions and statistical tests. In our tests, we scored an observation as supporting AT if the predicted contract is selected over the other two plans. Clearly, this is not the only possible test. For instance, a far more 'difficult' test would be to allow the principal to design a contract *de novo*. In all likelihood, very few principals would actually come close to the mathematically correct formulation. In contrast, a far 'easier' test would be to credit the theory if the principal chose any contract that was 'vaguely right' in the sense of offering some incentive for non-verifiable effort. Our test lies somewhere in between these extremes. All our conclusions are predicated on our null hypothesis.

A second limitation is the Berg et al procedure used to induce specific utility functions. Although its validity has been established previously (Berg et al 1986; Prasnikar, 1993; Rietz, 1993), and is regarded as one of the standard means of controlling for individual risk preferences (Roth 1995), we did no manipulation checks of the procedure itself. In order to assess whether the desired utility functions were actually induced, we would have had to include a lottery selection task. This would have interfered with the main purpose, and risked introducing other demand artifacts.

A third limitation of the study involves the construct validity of our effort manipulation and the verifiability manipulation. Instead of exerting actual physical effort, our subjects had pesos deducted from their account corresponding to the cost of the prospect list. The psychological impact of physical effort and monetary sacrifice are not likely to be identical in utility. The robustness of our results to other forms of effort manipulation remains unknown. As for effort verifiability, we did not include a manipulation check on our subjects' perception that a single prospect list rendered the effort verifiable *de facto*. To the extent that subjects did not realize this, they may have felt it necessary to provide incentives.

The final limitation pertains to the impact of our institutional setting. Traditionally, work in experimental economics abstracts away from the institutional context as much as possible. In contrast, we believe that the institutional context matters, and developed a much richer task. Despite our effort, some aspects of the institutional setting could not be created in a lab. Specifically, we were unable to generate a lab setting with a fixed fee contract for risk–neutral agents. Although this is the correct contract as per theory, real salespeople are never compensated in this fashion (although independent retailers are) and subjects perceive such fees as unrealistic.

Thus, we implemented the mathematically equivalent, but institutionally different commission only plans in place of fixed-fee plans.

9.2 Implications

What do we learn about AT from our experiments? Our fundamental result is that AT finds some support in lab settings where risk–neutral principals deal with risk–averse agents whose actions are non-verifiable. To put it into perspective, consider that across all the three experiments, 87% of the plans offered in this setting were ‘correct’ choices in that they traded off insurance against incentives in the predicted fashion.

In stark contrast, only 6% of the plans were correct choices in those settings offered where effort was verifiable. Likewise, only 21% of the plans were correct choices in those settings where agents were risk–neutral. This selective support is very robust. Recall that we incorporated a series of design changes intended to dispose of different validity threats and/or alternative explanations. We used different utility functions, broke exact ties in expected utility for agents, separated near ties in expected utility for principals, allowed for both single and multiple rounds of decisions, used two different subject populations, and checked for manipulation failures. Finally, we analyzed the data to test the possibility that our principals might have been predisposed to offer the ‘popular’ (i.e. mixed compensation) plans and that our agents were predisposed to undertake the ‘right’ (i.e., high) effort. We were able to rule out all these validity threats and alternative explanations.

We are led to a specific conclusion. The common element in all the settings where support is not forthcoming is the lack of relevance of the theoretical mechanism highlighted by AT. Specifically, when verifiable effort renders incentives moot, or when risk–neutrality among agents renders insurance moot, the basic insurance–incentive trade–off espoused by the theory is not applicable. However support is forthcoming in those conditions where this trade-off is present (risk-averse agents with non-verifiable effort). To concretize our contribution, re-consider our understanding of the three specific issues that we studied, in the light of our current results.

Effort-Output Uncertainty. Recall that the contradictory findings in the extant survey work could be attributed to the difficulty of measuring uncertainty conditional on effort. While the protocols used in prior experimental work (Umanath et al, 1993, 1996) obviated the measurement problem, they never implemented explicit functions. Hence, their results were not dispositive of the effort-output uncertainty effect. Our findings provide some closure on this issue. We find support for the need to insure risk–averse agents against uncertainty when they undertake non-verifiable effort. Furthermore, the extent of insurance provided increased as these agents faced more effort–output uncertainty. To the best of our knowledge, this is the only empirical evidence linking effort-output uncertainty to pay plans in an unambiguous fashion.

Verifiability of agent's effort. Recall that there is scant evidence about this effect. Berg et al (1992) studied the issue, but their design was compromised because agents were not free to reject the offered plan. Likewise, the John and Weitz (1989) survey is inconclusive because they dealt only with varying degrees of performance auditability instead of the contrast between verifiable and non-verifiable effort. Our findings provide no comfort for AT. Verifiable effort simply does not evoke salary–only plans. We speculate that salary–only plans are relevant only within multi–period relationships (e.g., between employer and employee) and not for single–period interactions as posited in AT.

Risk preferences of the agent. Only one study Joseph and Kalwani (1995) provides results relevant to this issue. They find that the proportion of salary to total compensation increases, but only for high levels of risk aversion. Our results clarify and extend our knowledge about this effect. Changes in risk-averse agents' risk preferences do have the predicted effect, but only as long as these changes do not extend to the point where the agents become risk-neutral. The insurance-incentive trade-off is absent when agents are risk-neutral.

Agenda for Future Research. We believe our study refines the scope of agency models of salesforce compensation and provides direction for future research in AT modeling. First, multi–period agency models that include reputation effects should be developed further because they are more likely to account for the data, particularly the circumstances under which salary-only plans are used. Furthermore, these models could be made richer by prominently expanding the set of variables included in compensation contracts to include partially verifiable elements. While reliance on all or nothing notions of contractability makes for model clarity, the real world offers effort as contractible to some greater or lesser degree. Along these lines, Dutta et al (1994) develop an agency model of exclusive territory dealers whose bootlegging efforts (shirking) are only partially enforceable (viz., verifiable with a lag). Hopefully, future designs will develop and examine similar models of compensation directly.

A second suggested line of inquiry is the development of models that include other mechanisms that could restrain agents from shirking on non-verifiable effort. Recent experimental work (e.g., Berg, Dickhaut, and McCabe, 1994) shows that norms of reciprocity can restrain shirking even in anonymous, one-shot interactions. Extrapolating, we argue that salary–only plans might be feasible if supported by such social control mechanisms.

A third line of inquiry concerns ownership of relevant assets as per the incomplete contracting models pioneered by Grossman and Hart (1986). Their model implies that incentive compensation is not very effective when used with employees, since the terms of the incentive contract can be changed by employers ex post. For example, high-potential territories can be reassigned or redistricted. Yet, we often observe incentive compensation paid to company salespeople. The support for incentive compensation that is implied by de facto employee control over assets such as the customer base needs to be developed further.

A final line of inquiry consists of efforts to model incentives and monitoring simultaneously (e.g., Joseph and Thevaranjan 1998). They offer the prospect of a better understanding of the impact of effort verifiability.

Agenda for Managers. AT is a useful framework for thinking about incentive compensation plans, particularly the implications of risk–sharing and shirking. Our results assist practitioners in applying AT logic more closely. First, incentive plans are more useful when salespeople have significant discretion about setting their effort levels. Here, the conflicting objectives of the firm and the salesperson must be dealt with squarely and financial incentives based on verifiable output are very useful in balancing these goals.

Second, our data alerts managers to those circumstances that require them to soften the incentive component of pay plans when they are used to motivate non-verifiable effort. The principal factor here is the income variability faced by employees. By cushioning risk, they can reduce compensation costs. Finally, our data provide managers a better appreciation of AT's limitations, especially for employee compensation design. Employee relationships are inherently multi–period in nature, and a rich variety of control mechanisms are simultaneously being used. AT does not help managers to incorporate these multiple mechanisms into compensation design.

REFERENCES

Amershi, Amin H. and Butterworth, James (1988), "Explorations in the Theory of Single and Multiple Agent Agencies," in Economic Analysis of Information and Contracts, G. Felthan, A. Amershi, and W. Ziemba (Ed.), Boston: Kluwer-Nijhoff Publishing.

Basu, Amiya K. (1983), "A Study on Optimal Salesperson Compensation Plans," Unpublished Doctoral Dissertation, Stanford, CA: Graduate School of Business, Stanford University.

Basu, Amiya K., Rajiv Lal, V. Srinivasan, and Richard Staelin (1985), "A Theory of Salesforce Compensation Plans", Marketing Science, 4, Fall, 267-91.

Berg, Joyce E., Lane Daley, John Dickhaut and John O'Brien (1986), "Controlling Preferences for Lotteries on Units of Experimental Exchange", Quarterly Journal of Economics, 101, 1, 281-306.

Berg, Joyce E., Lane Daley, John Dickhaut and John O'Brien (1992), "Moral Hazard and Risk Sharing: Experimental Evidence", in Research in Experimental Economics, Vol. 5, R. M. Isaac (Ed.), Greenwich CT: JAI Press, 1-34.

Berg, Joyce E., John Dickhaut and Kevin McCabe (1994), "Trust, Reciprocity, and Social History", Working Paper, Accounting Department, University of Minnesota.

Coughlan, Anne T. and Chakravarti Narasimhan (1992), "An Empirical Analysis of Salesforce Compensation Plans", Journal of Business, 65, 1, 93-121.

Dutta, Shantanu, Mark E. Bergen, and George John (1994), "The Governance of Exclusive Territories when Dealers can Bootleg", Marketing Science, 13, Winter, 83-99.

Epstein, Seth (1992), "Testing Principal-Agency Theory", Research in Experimental Economics, Vol. 5, R. M. Isaac (Ed.), Greenwich, CT: JAI Press, 35-60.

Forsythe, Robert, Joel L. Horowitz, N. E. Savin, and Martin Sefton (1994), "Fairness in Simple Bargaining Experiments," Games and Economic Behavior, 6, 347-369.

Grossman, Sanford J. and Oliver D. Hart (1986), "The Cost and Benefits of Ownership: A Theory of Vertical and Lateral Integration", Journal of Political Economy, 94, 691-719.

Holmstrom, Bengt (1979), "Moral Hazard and Observability", Bell Journal of Economics, 10, 180-191.

-------------------and Paul Milgrom (1991), " Multi-Task Principal-Agent Analyses: Incentive Contracts, Asset Ownership and Job Design," Journal of Law, Economics And Organization, 7, p. 24-52.

John, George and Barton A. Weitz (1988), "Explaining Variation in Sales Compensation Plans: Empirical Evidence for the Basu et al Model", Working Paper, Marketing Department, University of Minnesota.

--------------------------------- (1989), "Salesforce Compensation: An Empirical Investigation of Factors Related to the Use of Salary Versus Incentive Compensation", Journal of Marketing Research, 26, 1, 1-14.

Joseph, Kissan and Manohar U. Kalwani (1995), "The Impact of Environmental Uncertainty on the Design of Salesforce Compensation Plans", Marketing Letters, 6, 3, 183-197.

----------------- and Alex Thevaranjan (1998), "Monitoring and Incentives in Sales Organizations: An Agency-Theoretic Perspective", Marketing Science, 17, Spring, 107-123.

Lafontaine, Francine and Margaret Slade (1998), "Incentive Contracting and the Franchise Decision", in K. Chatterjee and W. Samuelson (Eds.) Advances in Business Applications of Game Theory, Kluwer Academic Press (forthcoming).

Lal, Rajiv (1982), "A Theory of Salesforce Compensation Plans", Unpublished Doctoral Dissertation, Graduate School of Industrial Administration, Carnegie-Mellon University.

------------ Donald Outland, and Richard Staelin (1994), "Salesforce Compensation Plans: An Individual-Level Analysis", Marketing Letters, 5, 2, 117-130.

------------ and Richard Staelin (1986), "Salesforce Compensation Plans in Environments with Asymmetric Information", Marketing Science, 5, Summer, 179-198.

Mantrala, Murali K., Prabhakant Sinha and Andris A. Zoltners (1994), "Structuring a Multi-product Quota Bonus Plan for a Heterogeneous Sales Force: A Practical Model-Based Approach", Marketing Science, 13, Spring, 122-144.

Oliver, Richard and Barton A. Weitz (1989), "The Effects of Risk Preference, Perceived Uncertainty, and Incentive Compensation on Salesperson Motivation," Working Paper, Marketing Department, University of Florida.

Padmanabhan, V. and Ram C. Rao (1993), "Warranty Policy and Extended Service Contracts: Theory and an Application to Automobiles", Marketing Science, 12, Summer, 230-247.

Peck, Charles (1982), Compensating Field Sales Representatives, New York: The Conference Board.

Prasnikar, Vesna (1993), "Binary Lottery Payoffs: Do they Control Risk Aversion?" Discussion Paper, Center for Mathematical Studies in Economics and Management Science, Northwestern University.

Rietz, Thomas A. (1993), "Implementing and Testing Risk-Preference-Induction Mechanisms in Experimental Sealed-Bid Auctions", Journal of Risk and Uncertainty, 7, 199-213.

Roth, Alvin E. (1995), "Introduction to Experimental Economics", in J. H. Kagel and A. E. Roth (Eds.), The Handbook of Experimental Economics, Princeton: Princeton University Press, 3-98.

Roth, Alvin E. and Michael W. K. Malouf (1979), "Game-Theoretic Models and the Role of Information in Bargaining", Psychological Review, 574-594.

Simonson, Itamar (1989), "Choice Based on Reasons: The Case of Attraction and Compromise Effects," Journal of Consumer Research, 16, 158-174.

Umanath, Narayan S., M. R. Ray and T. L. Campbell (1993), "The Impact of Perceived Environmental Uncertainty and Perceived Agent Effectiveness on the Composition of Compensation Contracts", Management Science, 39, January, 32-45.

Umanath, Narayan S., M. R. Ray and T. L. Campbell (1996), "The Effect of Uncertainty and Information Asymmetry on the Structure of Compensation Contracts: A Test of Competing Models", Management Science, 42, June, 868-874.

Figure 1: Sequence of Moves in the Experiment

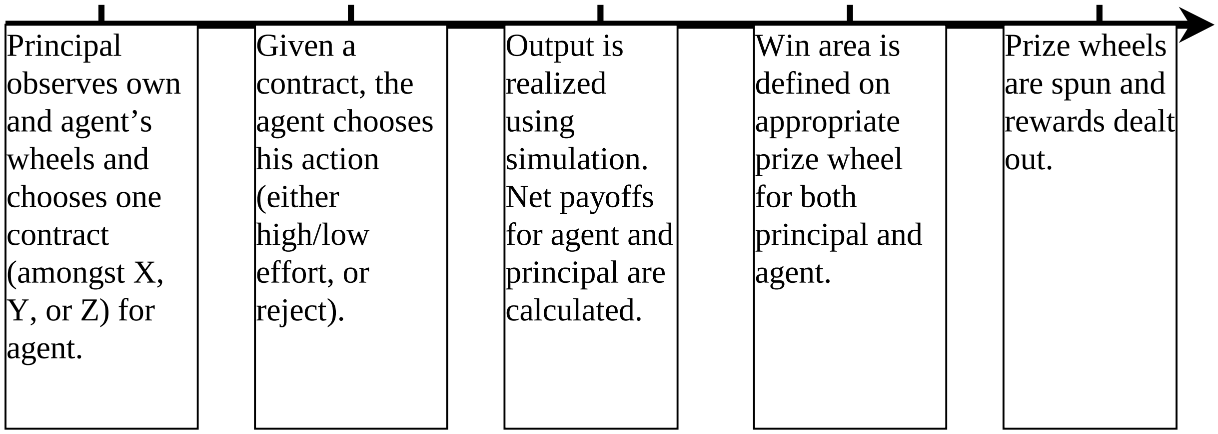

Figure 2: Sample of Multiple Prize Wheels for Agent

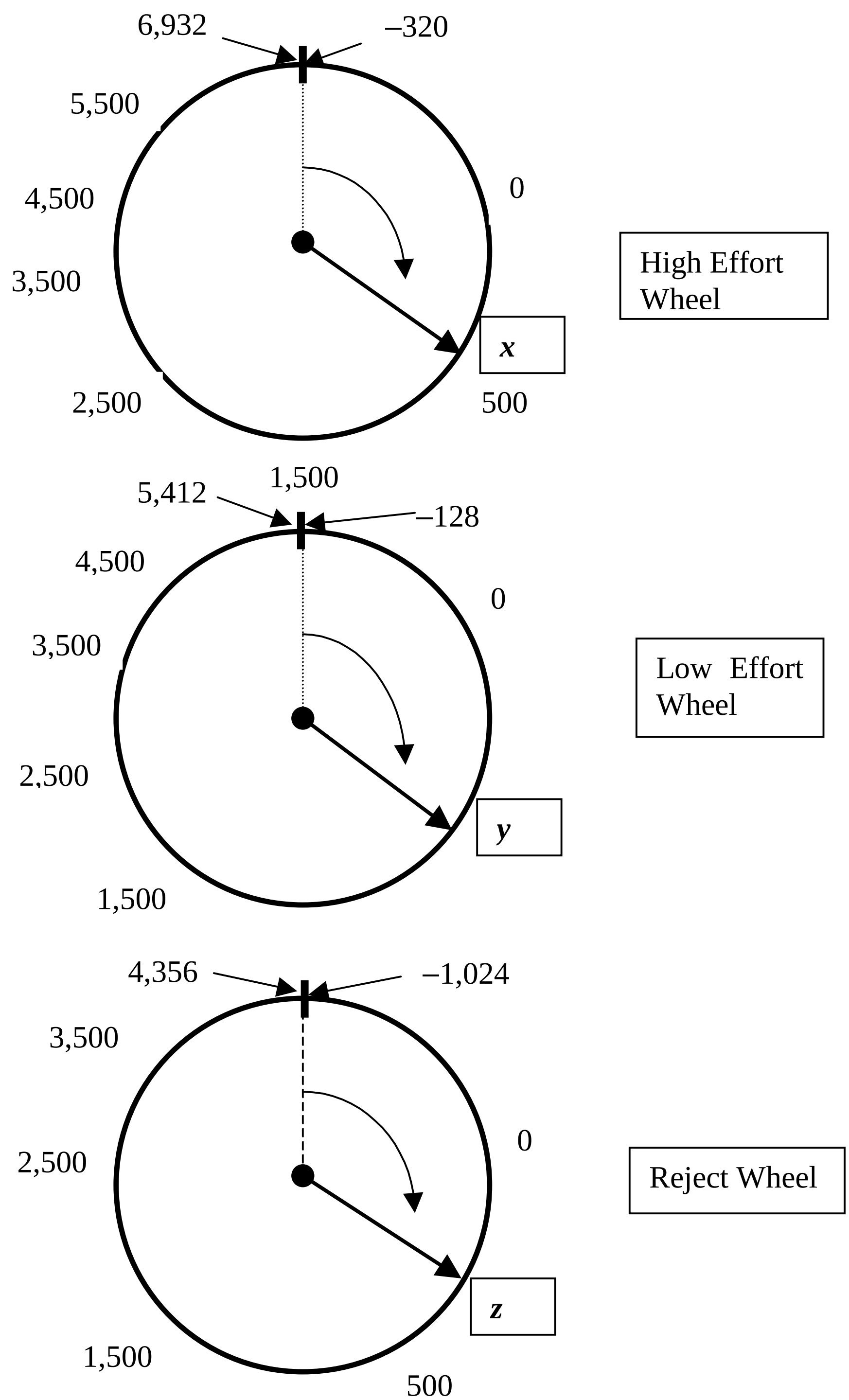

Figure 3: Sample Graphs of Principal's and Agent's

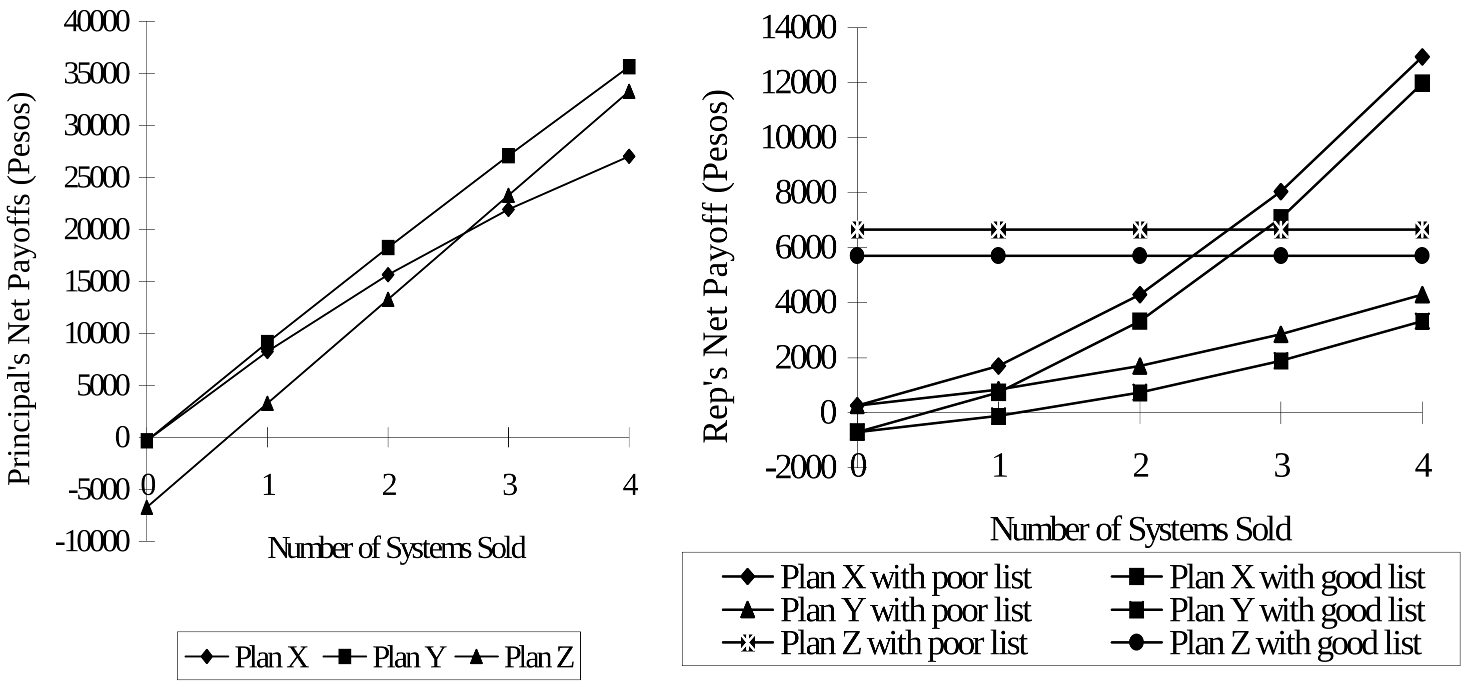

Figure 4: Sample Graphs for Sales

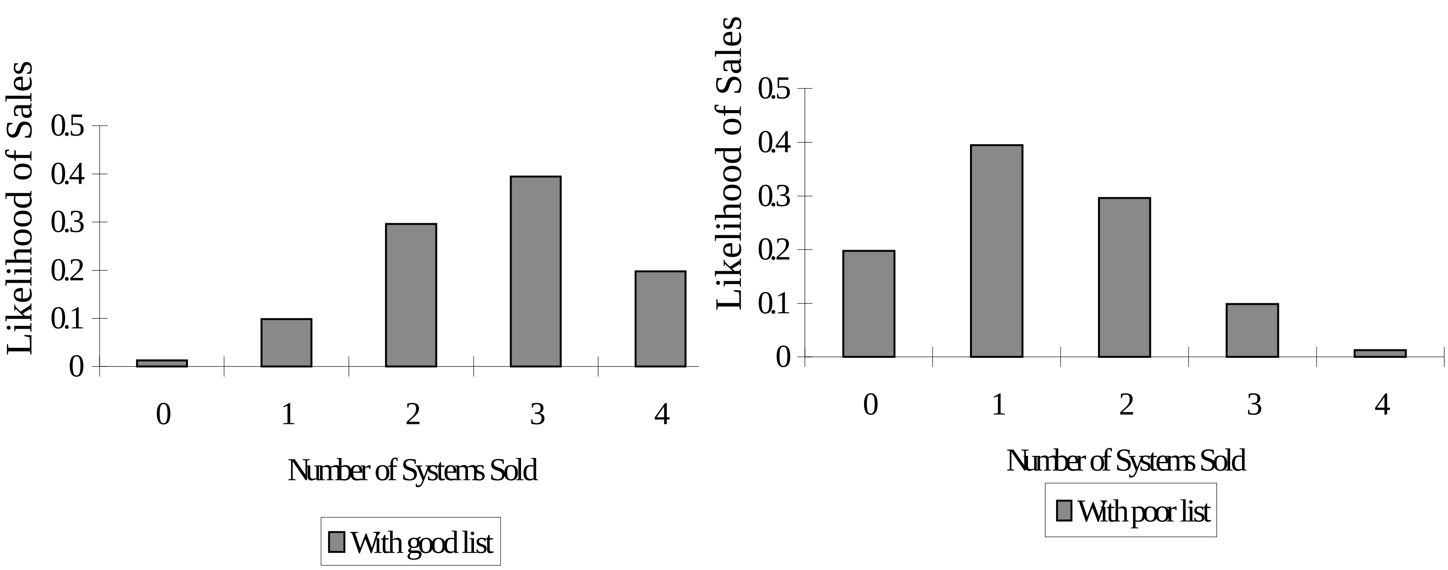

TABLE 1. DESIGN OF EXPERIMENT I

Cell Inductions	Cell 1 (n=7) Unobservable Effort 4 prospects	Cell 2 (n=6) Observable Effort 4 prospects	Cell 3 (n=6) Unobservable Effort 8 prospects	Cell 4 (n=6) Observable Effort 8 prospects
Predicted Plan	Mixed Plan X: $(18 + 24x)^2$	Salary-Only Plan Z: $(82)^2$	Mixed Plan Y: $(18 + 12x)^2$	Salary-Only Plan Z: $(82)^2$

TABLE 2: EXPECTED UTILITIES FOR EXPERIMENT I

Cell	Principal's Offer	Agent's Reaction					
		High Effort		Low Effort		Reject Contract	
		Principal	Agent	Principal	Agent	Principa l	Agent
Cell 1	X	**19431**	**100**	10321	84	0	100
	Y	24039	36	12049	52	0	100
	Z	19943	100	6609	148	0	100
Cell 2	X	19431	100	N/A	N/A	0	100
	Y	24039	36	N/A	N/A	0	100
	Z	**19943**	**100**	N/A	N/A	0	100
Cell 3	X	30993	228	18919	148	0	100
	Y	**46353**	**100**	23911	94	0	100
	Z	46609	100	19943	148	0	100
Cell 4	X	30993	228	N/A	N/A	0	100
	Y	46353	100	N/A	N/A	0	100
	Z	**46609**	**100**	N/A	N/A	0	100

Cell entries are expected utility

Cell values in bold indicate predicted equilibrium

TABLE 3:RESULTS OF EXPERIMENT I

Outcome	Cell 1	Cell 2	Cell 3	Cell 4
Contract–Action Pair	7/7	1/6	3/6	0/6
Principal's Choice	7/7	1/6	5/6	0/6
Agent's Reaction*	7/7	6/6	4/6	5/6

Cell entries indicate m/n observations are consistent with prediction

*: Agents' reactions are conditional on principal's offer

TABLE 4: DESIGN OF EXPERIMENT II

Cell Inductions	Cell 1 (n=12) Unobservable Effort Risk–Averse Agent	Cell 2 (n=10) Unobservable Effort Risk–Neutral Agent	Cell 3 (n=11) Observable Effort Risk–Averse Agent
Predicted Plan	Mixed Plan X: $(23 + 12x)^2$	Incentive-Only Plan Y: 815x	Salary Plan Z: $(87)^2$

TABLE 5: EXPECTED UTILITIES FOR EXPERIMENT II

Cell	Principal's Offer	Agent's Reaction					
		High Effort		Low Effort		Reject Contract	
		Principal	Agent	Principal	Agent	Principal	Agent
Cell 1	X	**45508**	**110**	23386	94	0	100
	Y	48987	66.7	24493	73	0	100
	Z	45764	110	19098	158	0	100
Cell 2	X	45508	249	23386	115.2	0	100
	Y	**48987**	**110**	24493	70.9	0	100
	Z	45764	238.8	19098	286.8	0	100
Cell 3	X	45508	110	N/A	N/A	0	100
	Y	48987	66.7	N/A	N/A	0	100
	Z	**45764**	**110**	N/A	N/A	0	100

Cell entries are expected utility

Cell values in bold indicate predicted equilibrium

TABLE 6: RESULTS OF EXPERIMENT II

Outcome	Cell 1	Cell 2	Cell 3
Contract–Action Pair	8/12	0/10	0/11
Principal's Choice	11/12	1/10	0/11
Agent's Reaction*	9/12	9/10	5/11

Cell entries indicate m/n observations are consistent with prediction

*: Agents' reactions are conditional on principal's offer

TABLE 7: DESIGN OF EXPERIMENT III

Cell Inductions (Three Rounds)	Cell 1 (n=9) Unobservable Effort Risk–Averse Agent	Cell 2 (n=9) Unobservable Effort Risk–Neutral Agent	Cell 3 (n=8) Observable Effort Risk–Averse Agent
Predicted Plan	Mixed Plan X: $(23 + 12x)^2$	Incentive-Only Plan Y: 815x	Salary-Only Plan Z: $(58)^2$

TABLE 8: EXPECTED UTILITIES FOR EXPERIMENT III

Cell	**Principal's Offer**	**Agent's Reaction**					
		High Effort		Low Effort		Reject Contract	
		Principal	Agent	Principal	Agent	Principal	Agent
Cell 1	X	**8175**	**110**	4719	94	0	100
	Y	11354	66.7	5287	73	0	100
	Z	12636	52	4636	100	0	100
Cell 2	X	8175	249	4636	115	0	100
	Y	**11354**	**110**	5287	71	0	100
	Z	12636	70.6	4636	118.6	0	100
Cell 3	X	8175	197	N/A	N/A	0	100
	Y	11354	132	N/A	N/A	0	100
	Z	**12636**	**110**	N/A	N/A	0	100

Cell entries are expected utility.

Cell values in bold indicate predicted equilibrium

TABLE 9: RESULTS OF EXPERIMENT III

Outcome	Cell 1			Cell 2			Cell 3		
	Decision Round			Decision Round			Decision Round		
	T1	T2	T3	T1	T2	T3	T1	T2	T3
Contract–Action Pair	5/9	8/9	7/9	3/9	2/9	0/9	1/8	0/8	0/8
Principal's Choice	6/9	8/9	8/9	3/9	3/9	1/9	1/8	1/8	0/8
Agent's Reaction*	8/9	8/9	8/9	5/9	8/9	9/9	6/8	7/8	7/8

Cell entries indicate m/n observations are consistent with prediction

*: Agents' reactions are conditional on principal's offer

APPENDIX 1

Experiment I

1) Utility function for principal: $U_p(q)=q$ where q is in pesos.

2) Utility function for risk averse agent : $U_a(q)=2q^{\left(\frac{1}{2}\right)}$, where q is in pesos.

3) Disutility of effort for agent: $V(t)=t^2$, where t is the level of effort; t_{high}=8, and t_{low}=4.

4) Sales given effort is the following binomial function:

   $f(x=i/t)= {}^nC_i\left[p(t)\right]^{(i)}\left[1-p(t)\right]^{(n-i)} = \pi_i(t)$, where probability of sale p(t) = t/12.

5) Price of good prospect list: $U_a^{-1}(8^2)=1{,}024$ pesos

6) Price of poor prospect list: $U_a^{-1}(4^2)=64$ pesos

7) Reservation utility of the agent: $m = 100$

8) Reservation wage: $U_a^{-1}(100)=2{,}500$ pesos

9) Selling price of system: 10,000 pesos.

Experiment II

The differences from Experiment I are noted below. Otherwise, all parameters and functions remain the same.

1) Utility function for risk neutral agent: $U_a(q)=\dfrac{q}{25}$

2) Price of prospect lists: 1600, 400 pesos for good, poor list respectively

3) Contracts designed to break:ties: $m = 110$.

Experiment III

The changes and additions are noted below. Otherwise, all other parameters remain the same.

1) Utility function for risk–averse agents in Cell 1 are same as Expt. I

2) Utility function for risk–neutral agents in Cell 2 are same as Expt. II

3) Utility function for risk–averse agents in Cell 3: $U_a(q) = 3q^{\left(\frac{1}{2}\right)}$

4) Price of good prospect list (Cell 1, 2, 3): 1024, 1600, and 455 pesos respectively

5) Price of poor prospect list (Cell 1, 2): 64 and 400 pesos respectively

6) Reservation wage (Cell 1, 2, 3): 2500, 2500, 1110 pesos respectively

7) Selling price per unit system: 3000 pesos

8) Contracts designed to break ties: $m = 110$.

FOOTNOTES

---

1 To minimize confusion, we shall refer to the principal as "she", and to the agent as "he" throughout the paper.

2 Verifiable means that a third-party like a court can enforce the contract relatively easily. Although observability and verifiability often track each other, they are conceptually distinct. In many circumstances, the two parties themselves may be able to observe an event, but a third party is unable to verify it. Slacking off on the job would be an example.

3 We thank a reviewer for pointing this out to us.

4 We note that the multi-task agency model of Holmstrom and Milgrom (1991) does accommodate this finding.

5 In the ultimatum game, the responder can reject the proposer's offer. In contrast, in a dictator game, the responder cannot reject the proposer's offer (see Forsythe et al 1994). Thus, in a dictator game, the proposer need not take into account the responder's preference while making the offer. The dictator game structure is inconsistent with the structure of AT where the principal's offer has to explicitly take into consideration the agent's preferences and reservation utility.

6 This paucity of work might be due to the difficulty of crafting survey questions to measure risk–preferences. Indeed, very few survey-based studies in any stream have measured risk preferences (exceptions include Oliver and Weitz 1989, Padmanabhan and Rao 1993). Evidently, the Lal, Outland and Staelin (1994) study also included such a measure, but it was not presented in the published paper (personal communication with Richard Staelin, 1998).

7 Technical appendices including derivations of our predicted contracts are .available on the Marketing Science web site, which is currently located at www.smeal.psu.edu/mktg/MktgSciJournal.

8 It is important to note that even though the three wheels in Figure 2 have different values in pesos for the same corresponding angle, the underlying net utility is the same for all the three wheels.

9 Although our. experimental setup operationalizes only two levels of effort, the underlying disutility function for effort is convex.

10 The literature does not provide explicit global solutions for contract-action pairs that solve the agent's optimal effort level and the principal's optimal compensation plan. As such, our solutions are not globally optimal either. Instead, the predicted pair in each cell is the best one from the set of pairs available in each cell. One cannot compare these predicted pairs across the cells, which rules out testing main effects or cell contrasts in a classic ANOVA analysis. All tests must be done against a specified null hypothesis on a cell-by-cell basis. Also observe that in all our cells, the predicted contract-action pair involves the high effort level. It is possible to select parameter values judiciously such that the predicted contract-action pair involves low effort. This would be an interesting extension of this experimental setup.

11 To understand the intuition behind the equilibrium contract–action pair in each cell, consider the principal's alternatives in Cell 1. Plan Y offers her the highest net expected utility *if* the agent undertakes high effort. However, if Y were offered, the agent's best reaction would be to reject the contract altogether. Plan Z offers the next highest level of expected utility

to the principal if the agent takes high effort. However, if Z were offered, the agent's best reaction would be to undertake low effort. Finally, if the principal were to offer Plan X, the agent would be indifferent between undertaking high effort and rejecting the contract. Hence the principal's offer of contract X and the agent's reaction of high effort is the predicted contract–action pair in Cell 1.

12 The BLSS model does not provide an explicit prediction when effort is verifiable. Even though this prediction is a logical consequence of general agency models, observe that it does not rely directly on the incentive–insurance trade–off, which is moot here. It might be construed as a set of circumstances at the 'boundaries' of the theory's domain.

13 A contract action pair is one combination of a pay plan chosen by a principal and the subsequent action of the agent (accept, reject, etc.).

14 It should be noted that while the principal earns zero net utility if the agent rejects her offer, her win area on the prize wheel is not zero. In fact, the principals realize a win area of about 16% on their own prize wheels if the agent rejects the pay plan.

15 We opt for a more conservative two-tailed test throughout the analysis.

16 We thank a reviewer for pointing this out to us.