1983 Competitive Manuscript Award: The Effects of Job Standard Tightness and Compensation Scheme on Performance: An Exploration of Linkages
Author(s): Chee W. Chow
Reviewed work(s):
Published by: American Accounting Association
Stable URL: http://www.jstor.org/stable/247062
Accessed: 28/12/2011 03:29

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.
1983 Competitive Manuscript Award

The Effects of Job Standard Tightness and Compensation Scheme on Performance: An Exploration of Linkages

Chee W. Chow

ABSTRACT: This study explores the linkages among job standard tightness, type of compensation scheme, and performance. It postulates that job standard tightness and type of compensation scheme affect not only workers' effort, but also their self-selection among employment contracts, and through these, job performance. A laboratory experiment yielded the following results: among subjects with assigned treatments, job standard tightness and type of compensation scheme had significant independent, but insignificant interactive effects on performance. Subjects who were permitted to choose their own compensation schemes (given an assigned job standard) self-selected among these by skill. There was also some indication that being able to select one's own compensation scheme, per se, enhanced performance. If supported by future studies, these results suggest that job standards and compensation schemes may affect performance not just by motivating a given set of employees, but also by affecting the type of employees an organization attracts from the labor market.

I. INTRODUCTION

This study explores the ways that job standard tightness and type of compensation scheme affect job performance. That both job standard tightness and compensation schemes affect the motivation to exert effort has been abundantly documented in accounting and organizational behavior research [Hopwood, 1974; Ronen and Livingstone, 1975; Locke, et al., 1981; Mitchell, 1979, 1981]. However, the potential interactive effects between these factors have received far less scrutiny [Locke, et al., 1980]. Empirical evidence on the magnitude of such effects can have great practical significance because control system design encompasses choosing both the performance standard and compensation scheme. An exploration of

The author is indebted to Jean Cooper, Terence Mitchell, Michael Shields, William Waller, the anonymous reviewers, the participants of the University of North Carolina-Chapel Hill Accounting Workshop and the Columbia University Accounting Research Forum and the Manuscript Award Committee for numerous helpful suggestions. A special note of thanks is due James Jiambalvo, whose input significantly enhanced the entire project. Judy Woo's assistance in conducting the experiment is also appreciated.

Chee W. Chow is Associate Professor of Accounting, University of North Carolina at Chapel Hill.

Manuscript received May 1982.
Revisions received January 1983 and April 1983.
Accepted May 1983.
these interactive effects is included in this study.

In addition to studying the motivational effects of job standards and compensation schemes in a given employment setting, this investigation probes whether these factors affect the matching-up of employers and employees. Analytical results by Stiglitz [1975], Salop and Salop [1976] and Demski and Feltham [1978], among others, suggest that workers will self-select among employment contracts based on expected compensation. If different job standard tightness/compensation scheme packages attract workers of different skill, the design of such packages takes on an added significance since both effort and skill affect performance [Campbell, et al., 1970; Locke, Mento and Katcher, 1978].

Four propositions are developed in Section II for empirical investigation. The remainder of this paper is organized as follows: Section III describes the behavioral experiment used for data collection; Section IV presents the results; and a summary and conclusions are supplied in Section V.

II. DISCUSSION OF RESEARCH PROPOSITIONS

This section derives four propositions for empirical investigation. First, the related theoretical and empirical literature is reviewed. The propositions are then discussed in light of this background.

A. THEORETICAL BACKGROUND

Numerous studies have examined how job standard tightness and contingent compensation affect performance. Reviews of this literature are provided by Locke, et al. [1981] and Mitchell [1979]. Both reports note the preponderance of support for the following findings: first, that specific and challenging standards induce higher performance than easy standards; and second, that performance-contingent rewards (especially financial ones) appear to have an independent motivating effect on performance. Mitchell [1974, 1979] observes that much of this empirical evidence is based on the expectancy theory framework. Ronen and Livingstone [1975] likewise have used expectancy theory to integrate the findings on budget and behavior in the accounting literature. Many alternate formulations of the expectancy model have been used [Ferris, 1977; Rockness, 1977; Jiambalvo, 1979; Ferris, Dillard and Nethercott, 1980]; all of them postulate that an individual’s motivation to exert a given level of effort depends on: (1) the individual’s probability estimates (expectancies) that specific outcomes will follow from the exerted effort, and (2) the perceived desirabilities (valences) of the outcomes associated with performing at that level of effort.

As it stands, the expectancy model is a rather general statement about individuals’ purposeful behavior. It has provided structure to investigations of motivation and performance in numerous field and laboratory settings (see Ferris [1977, p. 605] for an extensive listing). However, the richness of this framework does not seem to have been fully realized. Besides explaining individuals’ choice of effort level in a given employment situation, expectancy theory also offers a way to view how individuals select their employment settings in the first place: each individual entering the employment market (whether for the first time or as part of an ongoing re-evaluation process) brings to it his particular mix of characteristics (skill, work ethic, beliefs, physical appearance, preferences, etc.). The process of employment selection can be thought of as each individual seeking the best ‘fit’ between his own characteristics and the requirements and rewards of the job such that his expected valence is
maximized. As an illustration, suppose that two employment alternatives exist: the first demands a high level of performance and compensates abundantly for attaining this standard; the second requires a far more modest level of performance, but the remuneration for achieving this standard is also much less. Which alternative an individual will select depends on, among other things, the valence he places on financial rewards versus job effort. However, other things being equal, his skill level may also significantly affect this decision. For a given amount of effort, the expectancy of achieving the desired performance level will tend to increase with skill. As a result, a highly skilled individual may prefer alternative one, while a less skilled person may prefer alternative two—even though the latter pays less for standard attainment, his expectancy of reaching the standard in this case may far exceed that for the former alternative.

By and large, empirical research on performance evaluation and compensation has focused on individuals' effort choices to the exclusion of employment contract selection. Even within this restricted domain, a key interdependence between performance standards and compensation is often overlooked. Demski and Feltham [1978, p. 356] observe:

In many discussions of standards, the compensation aspects are largely ignored. Alternative standards will be associated with different compensation functions, and the preferences for one standard over another must reflect the changes in those functions. A _ceteris paribus_ comparison is incomplete.

Their analysis further indicates that performance standards and compensation schemes can have interactive effects on performance. In view of this result and the interdependence between performance standard and compensation scheme in employment contract design, evidence on the interactive effects between these factors can have important practical implications; since any amount of expected compensation can be derived from innumerable combinations of performance standard and compensation scheme (e.g., tight standard/high compensation, easy standard/low compensation, etc.), evidence on the individual and interactive motivational effects of these factors can indicate which combination is likely to induce the highest performance for a given amount of pay.

The study by Demski and Feltham [1978] belongs to a rapidly expanding analytical literature, popularly referred to as "agency theory." (See Jennergren [1980] and Baiman [1982] for reviews.) This literature focuses on the design of optimal contractual relationships among self-interested individuals; it analyzes both individuals' action choices under alternate contracts and their contract selections. Demski and Feltham [1978] present an application of this analytical approach to the budgetary control setting.

Demski and Feltham [1978] focus on the issues of divergent incentives and information asymmetry between controller (or employer) and controllee (or employee). They assume that, other things being equal, the employee prefers to exert less effort while the employer prefers the opposite. Demski and Feltham [1978] observe that information costs typically deter the employer from collecting perfect information about the employee's effort level. Similarly, cost considerations often preclude perfect differentiation among job applicants of different skill levels. The result of imperfect information is that problems of excessive shirking by employees and adverse selection may arise. In an employment setting, the latter refers to the case
when an employment contract attracts respondees whose skill level critically differs from that of the population at large.\footnote{Extensive analysis of the adverse selection problem in a variety of settings is provided by Akerlof [1970], Harris and Raviv [1978], Rothschild and Stiglitz [1976], Salop and Salop [1976], and Wilson [1977].}

Demski and Feltham [1978] show that if employee skill or effort is costly to observe, a budget-based employment contract (i.e., employee compensation is contingent on meeting the performance standard) can be Pareto superior\footnote{A contract is Pareto superior to another if it makes at least one contractee better off and no one worse off.} to fixed pay or linear sharing rules (where the employer and employee split the output). They indicate that budget-based compensation not only can motivate more effort from the employee, but it can also help to screen out potential employees of lower-than-desired skill (since such individuals are less likely to meet the standard). However, a potential drawback of budget-based schemes is that they impose risk on the employee (since job performance may be affected by factors outside the employee’s control, such as weather and economic conditions). As a result, employees’ risk preferences can affect their employment contract selection.

Analytical results such as those by Demski and Feltham are valuable because they bring a different perspective on the nature of the employment relationship. However, this line of research is not without limitations of its own. Baiman [1982] observes that agency research typically invokes highly restrictive assumptions about the analytical setting, including homogeneous beliefs among individuals, a one-period world, one employer and one employee, costless decision making, and the central importance of financial rewards (to the exclusion of psychological considerations like peer group esteem and feeling good about a job well done).\footnote{On the role of non-financial incentives, Demski and Feltham [1978, p. 356] point out that “... discussions of standards often stress motivational aspects which we have ignored. In particular, we have ignored the possibility that the outcome relative to standard provides some form of non-pecuniary income to the worker.”} This prompts him to caution that “... (i)t is premature to translate the results of agency research into normative guidelines for the choice and design of information systems within decentralized firms” [p. 206]. At the same time, agency research has raised issues that have been neglected in the organizational behavior literature. Thus, combining the insights from both literatures may yield improved understanding of the key linkages in the employment relationship. Below, four propositions are generated by taking this approach.

B. Research Propositions

The first proposition is as follows:

\textit{Proposition One = Ceteris paribus, job standard tightness and type of compensation scheme affect job performance independently as well as interactively.}

The basis for this proposition has been indicated in the review of agency theory and the expectancy model. Inclusion of Proposition One in this study serves three major functions—first, it motivates a replication of past research, the result of which may help to gauge the reliability of this study. Second, it incorporates an examination of the interactive effects of job standard tightness and type of compensation scheme on performance. Finally, evidence on this proposition provides a base of comparison for the other parts of this study.

The second and third propositions focus on the employer-employee matching-up process:
Proposition Two = Ceteris paribus, individuals will self-select among performance-based compensation schemes based on their skill.

Proposition Three = Ceteris paribus, individuals’ risk preferences will moderate their self-selection among performance-based compensation schemes.

Proposition Two is suggested by the extensive agency research on adverse selection [Akerlof, 1970; Harris and Raviv, 1978; Rothschild and Stiglitz, 1976; Salop and Salop, 1976; Wilson, 1977], and specifically by Demski and Feltham’s [1978] analysis of employment contracting. The preceding section also showed that self-selection by skill is consistent with the expectancy framework. Proposition Three introduces risk preference as a moderating factor in employment contract selection. This factor is not an explicit part of the expectancy model, but it plays a key role in agency analysis [Baiman, 1982; Demski and Feltham, 1978]. The basic idea here is that individuals are not indifferent among alternatives with the same expected payoff. Consider two compensation schemes: scheme one pays a fixed wage; scheme two carries the same expected pay as scheme one, but actual pay is uncertain. If performance meets some predetermined (tight) standard, the pay exceeds the fixed wage; otherwise it is much lower. When faced with these alternatives, equally skilled individuals may not make the same choice. More risk-averse individuals may prefer the fixed wage arrangement, while less risk-averse ones may select the budget-based scheme.

From a budgetary control standpoint, the implications of Propositions Two and Three can be quite different. If risk preference is not a major determinant of performance, then motivating individuals to self-select on this characteristic would serve little useful purpose. On the other hand, providing incentives for job applicants to self-select by skill may greatly enhance firm productivity.

If some individuals are able to select their own employment packages (i.e., job standard tightness and type of compensation scheme), will their performance differ from individuals with identical skills and risk preferences, but who cannot choose their own packages? Evidence from equity theory [Carrell and Dittrich, 1978; Goodman, 1977] suggests that a difference may exist; when people feel that they are not being rewarded fairly and can do little to change the situation, their performance tends to decline. Thus, if self-selection increases the feeling of fairness, it may enhance performance.

Self-selection may also be viewed as a form of participation in the budgetary process. In this regard, considerable evidence exists that participation improves worker satisfaction, though its impact on performance is more equivocal [Mitchell, 1979; Brownell, 1982]. Still, in view of the equity theory findings, this proposition seems worthy of investigation:

Proposition Four = Ceteris paribus, individuals who can select their own job standard tightness/compensation scheme packages will out-perform individuals who are assigned to the same packages.

Evidence bearing on Proposition Four can have important implications for how changes in budgetary control systems should be introduced. For instance, if this proposition is supported, it may

---

4 Ross [1973] and Itami [1975] show that in some cases, divergence of risk preferences between employee and employer may induce the former to make resource allocations contrary to the latter’s best interests. This problem is probably more severe at the managerial level than at the worker level.
imply that changes should not simply be imposed. Rather, the employees should be offered a selection of alternatives.

Section III will describe how evidence was gathered to evaluate the propositions presented above.

III. The Experiment

A laboratory experiment was used to facilitate control over omitted variables. In a budgetary control setting, these variables include leadership style [De-Coster and Fertakis, 1968; Hopwood, 1972; Otley, 1978] and organizational structure [Bruns and Waterhouse, 1975; Swieringa and Moncur, 1975; Merchant, 1981]. Controlling for these factors would be costly in a field study.5

The experiment and its analysis are comprised of several steps. Part of the data collected in an early step was used both there and again in a later step. Because of this complication, an overview of the entire experiment is provided prior to discussing its administration.

A. Overview of the Experiment

The subjects consisted of 86 volunteers from undergraduate business classes. The experimental task was adapted from Foreward [1969] and Rockness [1977]. It involved decoding randomly sequenced alphabets which had been pre-punched (16 each) on computer cards. This task was used to simulate an assembly line setting, where workers verify the location of integrated circuits on circuit boards for small business computers.

The experimental manipulations are comprised of two major components, as follows:

1) The first component consisted of six cells (denoted cells one through six). These were derived from two levels of job standard tightness ("average" and "tight") times three types of compensation schemes (fixed pay, piece-rate, and budget-based). These schemes cover the

continuum of compensation contracts analyzed by Demski and Feltham [1978]. Fixed pay and piece-rate represent the two extremes of linear sharing contracts; the budget-based scheme is Demski and Feltham's "bang-bang" contract, where attainment of the job standard brings a large reward, while failure to achieve the standard is paid a much smaller sum. Of the three schemes studied, this is the only one which explicitly ties compensation to job standard attainment.

All subjects in cells one through six were assigned to their job standard and compensation scheme combinations. Their task performance was used to assess the independent and interactive effects of these treatments.

2) The second component consisted of a different set of subjects who were initially assigned to two cells: average and tight job standard (denoted cells seven and eight, respectively). Each subject then chose between the fixed pay and budget-based schemes. (Piece-rate was eliminated in order to preserve adequate sample sizes in the remaining cells).6 This self-selection process sorted the subjects into four cells (two levels each of job standard tightness and compensation scheme). Subject skill and risk preference were compared across these cells. In addition, performance was compared between these cells and the four corre-

5 The obvious drawback of a laboratory experiment is that it may lack external validity. Discussions of this issue are available in Ashton and Kramer [1980], Fromkin and Streufert [1975], Kruglanski [1975], and Tunnell [1977]. Its advantage is a better control for omitted variables. Since the purpose of this study is to evaluate theoretical propositions rather than to generalize results to other settings, the advantages of a laboratory experiment probably exceed its disadvantages [Calder, Phillips and Tybout. 1981].

6 In the "real world," the sequence of choices may be opposite to the one used here—employees may first select a compensation scheme, and then participate in setting the standard. The effects of sequencing on the experimental results are worthy of investigation. The particular sequence used in the study was chosen because it made the treatment manipulations more manageable.
sponding cells from the first experimental component (i.e., those assigned to fixed pay and budget-based compensation schemes).

**B. Administration of the Experiment**

The experiment was administered in two one-hour sessions, scheduled one week apart. The procedures that were followed are described below.

1) **The first (pre-test) session.** Each subject was given a packet of pre-punched computer cards and a decoding key. The experimental task was explained, followed by a five-minute practice period to ensure that the instructions were understood (a ten-card practice deck was supplied for this purpose). After the practice deck was set aside, the subjects were told to put their names on the first card of the 50-card experimental deck, and to work for 30 minutes at their normal speed.

At the end of 30 minutes the card decks were collected. Each subject then completed an instrument aimed at assessing his risk preference. This instrument contained a set of instructions and 17 fictitious simple lotteries, patterned after those in Hilton, Swieringa and Hoskin [1981]. The expected values of these lotteries ranged from $10 to $90, with a mean of $50. These expected values were provided in the instrument, and the subject’s task was to indicate the certainty equivalent for each lottery (i.e., the amount of certain payment that would make him just willing to forego the lottery). On completing this instrument, the subjects were thanked for their cooperation and reminded to return the following week.

2) **The Interim.** Between the first and second sessions, the specific parameters for the experimental treatments were established, and the subjects were assigned to the eight experimental cells as follows.

a) Specifying the Treatment Parameters.

Across all subjects, an average of 23.976 cards were correctly decoded in the pre-test. The standard deviation was 5.409 cards. The Kolmogorov-Smirnov goodness-of-fit test indicated that this distribution did not depart significantly from normality (Z = .673, significance = .791).

The average standard was set at 24 cards per half-hour so that it would be achievable with an average probability of 50 percent across all subjects. The “tight but reasonably attainable” standard was set at .67 standard deviations (or 3.5 cards) above the average standard.

According to the pre-test distribution, this 27.5-card standard was achievable with a 25 percent probability, on average. (When the subjects were ranked by pre-test performance, it turned out that 21 of them, or 24.5 percent of the sample, had surpassed this standard.)

All three compensation schemes were “make believe” in that no actual payment was involved. Expected pay was held constant at $6 (per half-hour) to approximate actual wages for this line of work.

---

7 Each lottery had a probability $P_A$ of paying $A$, and a probability $1 - P_A$ of paying $B$. $A$ was either $0$ or $50$, and $B$ was either $50$ or $100$. Probability $P_A$ was varied in multiples of 1 between .1 and .9, with corresponding adjustments in $1 - P_A$.

8 The subjects were told that errors in decoding would be offset against correct ones. No decoding error was found.

9 This standard was set by rounding upwards the pre-test average of 23.976. It would have been desirable to incorporate anticipated learning effects. This was not done because data were not collected for estimating this effect. As a result, ex post, both job standards were probably easier than they were meant to be.

10 A major concern with using fictitious payments is that the subjects may fail to internalize this treatment. However, there is evidence that hypothetical payoffs can work as well as real payoffs [Slovic, 1969; Kogan and Wallach, 1974]. They apparently did in this experiment.

11 This relatively high (but not excessive) wage was used because previous research has found that small monetary rewards do not yield detectable effects even if such effects exist [Pritchard and Curtis, 1973; Rosswork, 1977].
Thus, fixed pay was $6 regardless of performance. Piece-rate was $.25 per card (prorated for part of a card) so that $6 would be earned at the 24-card average performance level. The budget-based scheme was dichotomous: if performance equaled or surpassed the standard, compensation exceeded $6; otherwise it was lower. Since the probabilities of standard-attainment were unequal between the average and tight standards, their associated dichotomous pay also differed so as to keep expected pay at $6. For subjects facing the average (24-card) standard, pay was $9 for matching or exceeding the standard, and $3 otherwise. The corresponding figures for the tight (27.5-card) standard were $15 and $3, respectively.\(^\text{12}\)

b) Assigning Subjects to Treatments.

Stratified random assignment was used to equalize subject skill and risk preference across treatments. Pre-test performance (denoted PRE) proxied for skill; the sum of each subject’s 17 certainty equivalents (denoted RISKSC) was used to gauge his risk preference (or aversion).\(^\text{13}\) The entire sample was first ranked by RISKSC and assigned to top, middle, and bottom groups of 30, 30, and 26 members, respectively. Each group was in turn divided into ten-member subgroups based on PRE ranking (the lowest group in the bottom RISKSC class had only six members). Finally, the subjects in each ten-member subgroup were randomly assigned to the eight experimental cells. Because cells seven and eight will each become two cells after the self-selection process, they received two subjects in each assignment round. The six-member subgroup was randomly assigned among cells one through six. Thus, cells one through six had nine subjects each, while cells seven and eight each contained 16 subjects.

An analysis of variance (ANOVA) indicated that PRE and RISKSC were not significantly different across the eight cells.\(^\text{14}\) This result indicates that the objective of the subject assignment procedure was achieved.\(^\text{15}\)

3) The experimental session. Four subjects did not return for the second session. This reduced the sample sizes for cells one, two, seven, and eight by one each.

Each returning subject was given a fresh 50-card deck, a decoding key, and an individualized instruction sheet. This sheet contained information on the subject’s own performance in the pre-test. For the subjects in cells one through six, it also specified their performance standards and compensation schemes for the current session. The subjects in cells seven and eight received an assigned performance standard, but were per-

\(^\text{12}\) Since the average standard was achievable with a 50 percent probability, on average, expected pay under it was $3(.5) + $9(.5), or $6. The tight standard was achievable with a 25 percent probability, on average. So its associated expected pay was also $6 [3(.75) + 15(.25)].

\(^\text{13}\) RISKSC may be a potentially flawed proxy for risk preference/aversion. Libby and Fishburn [1977] observe that risk preference measures tend to be situation/task-specific. Since the fictitious lotteries were not directly tied to the experimental task in this study, they may fail to elicit the risk attitudes applicable to this setting. Also, simply adding up the 17 certainty equivalents is unlikely to adequately capture differences in risk attitudes. Despite these limitations, this crude approach was taken because the main objective was only to segregate the sample into several major categories. Perhaps a comforting fact is that RISKSC was significantly and positively correlated with the subjects’ self-assessed risk attitude. Of course, this may only indicate that both measures are equally inadequate proxies for risk attitudes. Thus, results based on RISKSC can be no more than suggestive.

\(^\text{14}\) The Kolmogorov-Smirnov goodness-of-fit test indicated that RISKSC was not significantly non-normally distributed (Z = 1.07, significance = .147). RISKSC was also not significantly correlated with pre-test performance.

\(^\text{15}\) The reason for randomizing on PRE and RISKSC is to help ensure that the subjects in each cell have about the same characteristics of interest. Later on, the effect of PRE will also be subject to statistical control in order to focus more sharply on the experimental treatments.
mitted to choose between the fixed pay and budget-based schemes.

The decoding task was reviewed and the subjects were reminded that it was important for them to behave as if in a real job situation. Each subject then put his name on the first card of the deck and was given 30 minutes for the task. This was followed by a 5-item exit questionnaire which asked each subject to indicate, on a 7-point Likert-type scale, his subjective assessment of the following: the tightness of his assigned standard (DIFF), the fairness of his compensation scheme (FAIR), his interest in the experimental task (TASK), his attitude towards risk (RISKAT), and his interest in the experiment as a whole (WHOLE). Upon completing this questionnaire, the experiment was concluded.

IV. EMPIRICAL RESULTS

Table 1 gives the sample sizes for each experimental cell as well as the cell means and standard deviations for eight variables: pre-test performance (PRE); experimental performance (PERF); sum of 17 certainty equivalents (RISKSC); and the five subjective assessments—tightness of assigned job standard (DIFF), fairness of the compensation scheme (FAIR), interest in the decoding task (TASK), interest in the experiment as a whole (WHOLE), and risk attitude (RISKAT). Kolmogorov-Smirnov goodness-of-fit tests indicated that all five subjective assessments deviated significantly from normality.17 Also note in Table 1 that cells 7A, 7B, 8A, and 8B are the outcome of self-selection by the subjects initially assigned to cells 7 and 8, respectively.

Empirical examination of the four propositions used different subsets of the cells in Table 1. These results will be presented in the same order as the propositions.

**Proposition One**

It was postulated that job standard tightness and type of compensation scheme affect performance both independently and interactively. This proposition was evaluated using cells 1 through 6. Table 1 shows that in every cell, mean experimental performance (PERF) exceeded average pre-test performance (PRE). This may be a reflection of learning effects between the two experimental sessions. Of more interest is the dispersion in mean PERF across cells, ranging from a low of 26.58 in cell 1 (average standard and fixed pay) to a high of 34.56 in cell 4 (tight standard and piece-rate). Figure 1 suggests that this variation in PERF is related to both job standard tightness and type of compensation scheme. With compensation scheme held constant, mean PERF differs between the average and tight job standards. Similarly, when job standard is held constant, mean PERF varies across the three compensation schemes.

The statistical significance of treatment effects was assessed by applying analysis of covariance (ANCOVA) to the 52 observations in cells 1 through 6. The dependent variable was experimental performance. Pretest performance was used as a covariate to control for the effects of subject skill. The results are

16 The end points of the scales were as follows: DIFF, 1 = "very easy," 7 = "very tough"; TASK and WHOLE, 1 = "very interesting"; 7 = "very boring"; FAIR, 1 = "very fair"; 7 = "very unfair"; RISKAT, 1 = "very risk averse"; 7 = "very risk seeking ."

17 The respective Z scores and their significance levels were as follows: PRE, Z = 641, significance = .805; PERF, .459 and .984; RISKSC, 1.158 and .137; DIFF, 1.394 and .041; FAIR, 2.004 and .001; TASK, 2.059 and .001; RISKAT, 1.996 and .001; WHOLE, 1.411 and .037.

18 These differences may also have resulted from introducing the "make believe" incentive schemes. This would not cause a problem since the analysis focuses on differences across schemes. Likewise, learning is unlikely to confound the results since it should apply to all (randomly assigned) subjects.
### Table 1

**Means and Standard Deviations of Measured Characteristics for Each Experimental Group**

(Standard deviations in parentheses)

<table>
<thead>
<tr>
<th></th>
<th>Average Performance Standard (Assigned)</th>
<th>Tight Performance Standard (Assigned)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed Pay</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell 1, N=8</td>
<td>PRE 24.60 (7.22)</td>
<td>PRE 23.46 (2.95)</td>
</tr>
<tr>
<td></td>
<td>PERF 26.58 (6.10)</td>
<td>PERF 28.78 (4.71)</td>
</tr>
<tr>
<td></td>
<td>RISKSC 837.0 (135.1)</td>
<td>RISKSC 811.5 (93.0)</td>
</tr>
<tr>
<td></td>
<td>DIFF 2.75 (1.28)</td>
<td>DIFF 3.50 (1.07)</td>
</tr>
<tr>
<td></td>
<td>FAIR 2.87 (.99)</td>
<td>FAIR 3.50 (1.93)</td>
</tr>
<tr>
<td></td>
<td>TASK 5.62 (1.51)</td>
<td>TASK 5.75 (1.28)</td>
</tr>
<tr>
<td></td>
<td>RISKAT 4.0 (1.19)</td>
<td>RISKAT 3.62 (1.19)</td>
</tr>
<tr>
<td></td>
<td>WHOLE 5.37 (1.30)</td>
<td>WHOLE 4.87 (1.73)</td>
</tr>
<tr>
<td><strong>Assigned Compensation Scheme</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Piece Rate</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell 3, N=9</td>
<td>PRE 23.44 (6.01)</td>
<td>PRE 22.92 (6.24)</td>
</tr>
<tr>
<td></td>
<td>PERF 29.22 (5.36)</td>
<td>PERF 34.56 (5.68)</td>
</tr>
<tr>
<td></td>
<td>RISKSC 799.4 (97.4)</td>
<td>RISKSC 745.6 (344.5)</td>
</tr>
<tr>
<td></td>
<td>DIFF 2.89 (1.62)</td>
<td>DIFF 4.22 (1.39)</td>
</tr>
<tr>
<td></td>
<td>FAIR 3.44 (1.59)</td>
<td>FAIR 2.89 (1.36)</td>
</tr>
<tr>
<td></td>
<td>TASK 5.78 (.83)</td>
<td>TASK 4.22 (1.79)</td>
</tr>
<tr>
<td></td>
<td>RISKAT 3.89 (1.17)</td>
<td>RISKAT 3.44 (1.24)</td>
</tr>
<tr>
<td></td>
<td>WHOLE 4.28 (1.30)</td>
<td>WHOLE 4.56 (1.24)</td>
</tr>
<tr>
<td><strong>Budget-Based Pay</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell 5, N=9</td>
<td>PRE 23.58 (5.77)</td>
<td>PRE 24.75 (6.85)</td>
</tr>
<tr>
<td></td>
<td>PERF 27.44 (6.47)</td>
<td>PERF 28.69 (4.68)</td>
</tr>
<tr>
<td></td>
<td>RISKSC 791.8 (135.7)</td>
<td>RISKSC 761.1 (200.8)</td>
</tr>
<tr>
<td></td>
<td>DIFF 3.67 (1.73)</td>
<td>DIFF 4.78 (1.56)</td>
</tr>
<tr>
<td></td>
<td>FAIR 2.56 (1.33)</td>
<td>FAIR 3.33 (1.87)</td>
</tr>
<tr>
<td></td>
<td>TASK 4.78 (1.99)</td>
<td>TASK 6.11 (1.17)</td>
</tr>
<tr>
<td></td>
<td>RISKAT 3.33 (1.32)</td>
<td>RISKAT 4.0 (1.50)</td>
</tr>
<tr>
<td></td>
<td>WHOLE 3.78 (1.39)</td>
<td>WHOLE 4.78 (1.56)</td>
</tr>
<tr>
<td><strong>Self-Selected Compensation Scheme</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fixed Pay</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell 7A, N=10</td>
<td>PRE 22.73 (5.38)</td>
<td>PRE 21.51 (1.55)</td>
</tr>
<tr>
<td></td>
<td>PERF 25.46 (6.33)</td>
<td>PERF 26.18 (2.66)</td>
</tr>
<tr>
<td></td>
<td>RISKSC 805.3 (105.7)</td>
<td>RISKSC 811.8 (69.7)</td>
</tr>
<tr>
<td></td>
<td>DIFF 3.75 (1.04)</td>
<td>DIFF 3.88 (1.73)</td>
</tr>
<tr>
<td></td>
<td>FAIR 2.75 (1.67)</td>
<td>FAIR 3.25 (2.12)</td>
</tr>
<tr>
<td></td>
<td>TASK 5.88 (.84)</td>
<td>TASK 5.63 (1.19)</td>
</tr>
<tr>
<td></td>
<td>RISKAT 3.13 (1.13)</td>
<td>RISKAT 3.63 (1.41)</td>
</tr>
<tr>
<td></td>
<td>WHOLE 5.25 (1.04)</td>
<td>WHOLE 4.63 (1.59)</td>
</tr>
<tr>
<td><strong>Budget-Based Pay</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell 7B, N=5</td>
<td>PRE 26.79 (5.83)</td>
<td>PRE 27.16 (2.52)</td>
</tr>
<tr>
<td></td>
<td>PERF 33.88 (7.30)</td>
<td>PERF 34.42 (5.48)</td>
</tr>
<tr>
<td></td>
<td>RISKSC 830.6 (178.8)</td>
<td>RISKSC 824.1 (73.3)</td>
</tr>
<tr>
<td></td>
<td>DIFF 3.75 (1.67)</td>
<td>DIFF 3.67 (2.07)</td>
</tr>
<tr>
<td></td>
<td>FAIR 3.13 (1.89)</td>
<td>FAIR 3.33 (2.42)</td>
</tr>
<tr>
<td></td>
<td>TASK 5.88 (1.36)</td>
<td>TASK 4.83 (2.32)</td>
</tr>
<tr>
<td></td>
<td>RISKAT 3.75 (1.17)</td>
<td>RISKAT 4.17 (1.47)</td>
</tr>
<tr>
<td></td>
<td>WHOLE 5.0 (1.69)</td>
<td>WHOLE 4.5 (2.17)</td>
</tr>
</tbody>
</table>

Note: PRE = Number of cards decoded in the pre-test session  
PERF = Number of cards decoded in the experimental session  
RISKSC = Sum of certainty equivalents to the 17 lotteries  
DIFF = Self-assessment of job standard difficulty  
FAIR = Self-assessment of fairness of compensation scheme  
TASK = Self-assessment of interest in the experimental task  
RISKAT = Self-assessment of risk preference/aversion  
WHOLE = Self-assessment of interest in the whole experiment
Subjects who were assigned to a compensation scheme and an average performance standard.

Subjects who were assigned to a compensation scheme and a tight performance standard.

reported in Panel A of Table 2. As may be expected, a significant relationship exists between experimental performance and pre-test performance. Holding the latter constant, there are significant main effects due to job standard tightness and type of compensation scheme, but the interactive effect between these factors is not statistically significant.19

A closer inspection of Figure 1 suggests that the effects of job standard and compensation scheme may have been “driven” by the subjects operating under piece-rate. An exploration of this possibility is warranted, as only fixed pay and budget-based compensation are re-

19 Since the subjects were randomly assigned based on skill, controlling for pre-test performance is not strictly necessary. An analysis of variance (ANOVA) was performed by excluding this variable. The entire ANOVA was significant at the .066 level. The main effects due to type of compensation scheme and job standard were significant at the .055 and .060 levels, respectively, and their interaction effect was only significant at the .520 level.
Table 2
ANCOVA of Experimental Job Performance for Subjects with Assigned Performance Standards and Compensation Schemes

Panel A
All subjects with Assigned Performance Standards and Compensation Schemes (N=52)

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>DF</th>
<th>Mean Square</th>
<th>F</th>
<th>Significance of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test Performance</td>
<td>1</td>
<td>1085.756</td>
<td>114.079</td>
<td>.001</td>
</tr>
<tr>
<td>Type of Compensation Scheme</td>
<td>2</td>
<td>74.844</td>
<td>7.864</td>
<td>.001</td>
</tr>
<tr>
<td>Performance Standard Tightness</td>
<td>1</td>
<td>94.987</td>
<td>9.980</td>
<td>.003</td>
</tr>
<tr>
<td>Type of Compensation* Standard</td>
<td>2</td>
<td>.513</td>
<td>.054</td>
<td>.948</td>
</tr>
<tr>
<td>Scheme Tightness</td>
<td>6</td>
<td>221.863</td>
<td>23.311</td>
<td>.001</td>
</tr>
<tr>
<td>Explained</td>
<td>45</td>
<td>9.518</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residual</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel B
The Subjects with Assigned Performance Standards and Fixed Pay or Budget-based Compensation Schemes (N=34)

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>DF</th>
<th>Mean Square</th>
<th>F</th>
<th>Significance of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Test Performance</td>
<td>1</td>
<td>584.318</td>
<td>56.467</td>
<td>.001</td>
</tr>
<tr>
<td>Type of Compensation Scheme</td>
<td>1</td>
<td>10.434</td>
<td>1.008</td>
<td>.324</td>
</tr>
<tr>
<td>Performance Standard Tightness</td>
<td>1</td>
<td>58.612</td>
<td>5.664</td>
<td>.024</td>
</tr>
<tr>
<td>Type of Compensation* Standard</td>
<td>1</td>
<td>.931</td>
<td>.090</td>
<td>.766</td>
</tr>
<tr>
<td>Scheme Tightness</td>
<td>4</td>
<td>163.462</td>
<td>15.797</td>
<td>.001</td>
</tr>
<tr>
<td>Explained</td>
<td>29</td>
<td>10.348</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residual</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The absence of significant interactive effects for job standard tightness and type of compensation scheme is contrary to expectations based on the agency literature. This finding may imply that such interactive effects either do not exist or are insignificant in magnitude. Alternatively, it may be an artifact of the levels of experimental treatments used in this study. Extensive replications are needed before one can differentiate between these alternatives.

---

20 Differences in the self-assessed measures were also analyzed. From Table 1, the self-assessed job difficulty measure (DIFF) was higher for the subjects assigned the difficult standard. However, comparing DIFF between the subjects facing the average and tight standards (with the nonparametric median test) yielded an insignificant result. This may indicate that the two levels of job standards, at least as they were perceived, were not as far apart as they were designed to be. Also, it appears that feelings towards the job were not significantly different across experimental treatments: TASK, WHOLE, and FAIR were all insignificantly different across job standards and types of compensation scheme.
Proposition Two

This proposition states that individuals will self-select among employment contracts based on skill. Cells 7A, 7B, 8A, and 8B were used to examine this postulate.

Proposition Two is motivated by the assumption that individuals seek to maximize their expected compensation. For a given performance standard and effort level, individuals with greater skill are more likely to achieve the standard (with its attendant higher compensation than fixed pay). Thus, other things equal, more highly skilled individuals can be expected to select budget-based pay, while less skilled individuals will tend towards fixed pay. This self-selection effect was explored by comparing pre-test performance between cells 7A and 8A (subjects who selected fixed pay) and 7B and 8B (subjects who selected budget-based pay). The mean and standard deviation of this proxy for skill were 22.227 and 4.378, respectively, for the fixed pay group. The budget-based group had a higher mean (27.017) and about the same standard deviation (4.173). The t-statistic for the difference in means was 2.93, which is significant at the .0035 (one-tailed) level. This result is consistent with Proposition Two. It indicates that skill played a significant role in the subjects' compensation scheme selection.\(^{21}\)

Proposition Three

It is postulated that individuals' risk attitudes will influence their choice of compensation scheme. This proposition was tested using two alternate proxies for risk attitude—the sum of certainty equivalents (RISKSC) and the subjective risk attitude assessment (RISKAT). Cells 7A, 7B, 8A and 8B were used in these tests.

First, RISKSC was compared between the fixed pay (cells 7A and 8A) and budget-based (cells 7B and 8B) groups. Mean RISKSC for these groups were 807.9 and 826.6, respectively. The standard deviations were 125.36 and 130.52, respectively. The t-statistic for the difference in means was .72, which is only significant at the .48 level. An important consideration in interpreting this result is that RISKSC may be a flawed proxy for risk attitude (see footnote 13). This concern motivated additional analysis using RISKAT. Even though this self-assessed measure also suffers from not being specific to the present experimental task, it does avoid the linear aggregation issue. Hence, evidence based on this proxy can strengthen the overall results.

The RISKAT comparison employed a nonparametric test because this variable is non-normally distributed. Applying the median test [Hollander and Wolfe, 1973] yielded a Chi-square statistic of .136, which has a .713 significance level. This result is consistent with the RISKSC finding, and may be a reflection of the significant (at the .008 level) positive correlation of .341 between RISKAT and RISKSC.

It is premature to interpret the evidence here as implying that risk preferences did not affect the subjects' self-selection process. The potential limitations of RISKSC and RISKAT have already been noted. In addition, risk may not have varied much across experimental treatments. Every subject faced the same task and performance was measured in terms of observed output. Given this setting, the only risk in the self-selection process related to how accurately each subject could assess his own skill. Since feedback was provided on pre-test performance, this source of risk was probably minor; even if an effect existed, it was likely to

\(^{21}\) This conclusion assumes that pre-test performance is a valid proxy for skill. This seems to be a reasonable assumption.
have been too small for detection. The considerations above indicate that more extensive tests of Proposition Three are called for, both by using more refined risk attitude measures, and by incorporating other sources of risk.

**Proposition Four**

This proposition states that self-selection per se will enhance performance. Evaluation of this postulate entailed comparing experimental performance between cells 1, 2, 5, and 6 (not free to self-select) and cells 7A, 7B, 8A, and 8B (free to self-select). However, using one ANCOVA for the entire comparison would be inappropriate. Given the evidence on self-selection (Proposition Two), job standard tightness and type of compensation scheme are not independent for the subjects in the latter four cells. In order to control for this collinearity, separate ANCOVA’s were performed for subjects who faced the same assigned performance standard. Thus, one ANCOVA used only the subjects facing the average standard (cells 1, 5, 7A, and 7B), and another utilized the subjects with the tight assigned standard (cells 2, 6, 8A, and 8B). In both cases, pre-test performance was used as a covariate to control for subject skill, and the treatments were availability of choice of compensation scheme and type of compensation scheme. The results are reported in Table 3 as Panels A and B, respectively. In addition, Figure 2 plots mean experimental performance for the subjects with the average standard. Figure 3 does this for those facing the tight standard.

First examine the average standard case. Figure 2 shows that mean experimental performance was slightly higher
FIGURE 2
THE RELATION OF EXPERIMENTAL PERFORMANCE TO COMPENSATION SCHEME AND AVAILABILITY OF CHOICE FOR THE SUBJECTS OPERATING UNDER THE AVERAGE PERFORMANCE STANDARD

Subjects who were assigned to a compensation scheme and an average performance standard.

Subjects who were assigned an average performance standard, and then selected their compensation scheme.

for the subjects who had been assigned to, as opposed to self-selected, fixed pay (26.58 versus 25.46). The reverse holds for the subjects under the budget-based scheme (27.44 versus 33.88). A likely cause of this pattern of differences is self-selection by skill: earlier it was shown that more skilled (i.e., higher pre-test performance) subjects tended to choose the budget-based scheme, and that pre-test performance was positively related to experimental performance.

Panel A of Table 3 also indicates a significant main effect due to pre-test performance. Given self-selection by skill, controlling for this factor is crucial if the other treatment effects are to be correctly identified. With this control in effect, availability of choice of compensation scheme has a significant main effect at the .062 level. This finding moderately supports the proposition that being per-
FIGURE 3
THE RELATION OF EXPERIMENTAL PERFORMANCE TO COMPENSATION SCHEME AND AVAILABILITY OF CHOICE FOR THE SUBJECTS OPERATING UNDER THE TIGHT PERFORMANCE STANDARD

<table>
<thead>
<tr>
<th>Performance (Number of cards decoded in half an hour)</th>
</tr>
</thead>
<tbody>
<tr>
<td>40.0</td>
</tr>
<tr>
<td>30.0</td>
</tr>
<tr>
<td>20.0</td>
</tr>
<tr>
<td>27.5</td>
</tr>
<tr>
<td>24.0</td>
</tr>
</tbody>
</table>

---

Subjects who were assigned to a compensation scheme and a tight performance standard.

Subjects who were assigned a tight performance standard, and then selected their compensation scheme.

mittent to choose one's own compensation scheme enhances performance. Contrary to an earlier section (Table 2, Panel B), type of compensation scheme now has a main effect which is significant (at the .009 level), while availability of choice and type of compensation scheme have an interactive effect which is significant at the .054 level. Together, these results indicate that performance not only depends on type of compensation scheme, but also on whether the scheme is imposed or self-selected.22

22 Did availability of choice affect performance through altering attitudes towards the job? This question was addressed by comparing the assigned and self-selected subjects on the following self-reported measures: interest in the experimental task (TASK), interest in the experiment as a whole (WHOLE), and fairness of the compensation scheme (FAIR). The median test was used because all three measures are nonnormally distributed. The Chi-square statistics were .063, .063, and .969, with significance levels of .802, .802 and .325, respectively.
Now turn to the subjects with the tight job standard. Figure 3 shows that under fixed pay, mean experimental performance was higher for the assigned subjects than for the self-selected ones (28.78 versus 26.18). The reverse is true under the budget-based scheme (28.69 versus 34.42).

Panel B of Table 3 shows that pre-test performance and type of compensation scheme have significant main effects at the .001 and .063 levels, respectively. Availability of choice is only significant at the .784 level, and its interactive effect with compensation scheme only has a .381 significance level. Thus, self-selection did not significantly affect performance when the job standard was tight.²³

Given that job attitudes were apparently similar between the two levels of job standard tightness, what accounts for the different results regarding availability of choice? The collected data precluded an exploration of this issue. Such an investigation is desirable, as it may yield useful insights into the relative impacts of job standards, type of compensation scheme, and availability of choice.

V. SUMMARY AND CONCLUSIONS

This study postulates that job standard tightness and type of compensation scheme affect both worker effort and self-selection and, through these, job performance. A laboratory experiment with business students yielded the following results: among subjects with assigned treatments, job standard tightness and type of compensation scheme had significant independent, but insignificant interactive effects on performance. Subjects who were permitted a choice of compensation schemes (given an assigned job standard) self-selected among these by skill. Risk attitude, as measured in this study, did not significantly affect the self-selection process. Also, there was some indication that being able to select one’s own compensation scheme, per se, enhanced performance. However, this effect was only observed in the average standard and not the tight standard case.

The results of this study indicate that job standards and compensation schemes may play both significant screening and motivational roles. However, generalizing these findings to practice will have to await extensive replications. Future studies need both to broaden and deepen the scope of analysis. The former refers to testing for effects with other tasks (e.g., managerial as opposed to assembly line), settings (e.g., cross-sectional or field study rather than a laboratory experiment), and types of employment contracts (e.g., whether noncontrollable factors are excluded from performance measurement). The latter alludes to refining both the theoretical and empirical specification of the linkages in the employment relationship. On the theoretical front, a tighter integration of the organizational behavior and agency literatures holds promise of additional insights. Operationally, more refined and comprehensive measures of risk attitudes, job attitudes, and individual characteristics should be used. The role of non-financial motivations (e.g., desire for peer esteem) also needs to be explicitly addressed. Finally, expanding the investigation to a multi-period setting can shed light on whether the incentive and performance effects are transitory or permanent.

This evidence suggests that self-selection did not significantly affect job attitudes. It should be noted, though, that the measures used here were relatively crude. Future research should incorporate more refined measures as well as more aspects of job attitudes to explore more fully the linkage between availability of choice and job performance.²³ For this subset of the sample, TASK, WHOLE, and FAIR did not differ significantly between the assigned and self-selected subjects. Their respective Chi-square statistics were only significant at the .368, .825, and .50 levels.
REFERENCES


