NOAH LIM, MICHAEL J. AHEARNE, and SUNG H. HAM*

Sales contests are short-term incentives that managers use to raise sales effort. The extant marketing theory predicts that the optimal prize structure should have two characteristics: (1) The number of prizewinners should be greater than one, and (2) prize values should be unique and rank ordered. However, this theory has not been empirically examined. This article presents two empirical studies that examine whether the prize structure of a sales contest affects sales performance. In each study, the authors investigate the incremental effects of introducing multiple prizewinners and unique rank-ordered prizes into a sales contest. The first study consists of two laboratory experiments in which participants make decisions that closely reflect the decision trade-offs in the theoretical model of sales contests. The second study consists of two field economic experiments in which trained salespeople sell fundraising sponsorships to companies. The results across the experiments are remarkably consistent: The number of prizewinners in a sales contest should indeed be greater than one. However, introducing rank-ordered prizes into contests with multiple prizewinners does not boost sales effort and revenues.

Keywords: sales contests, sales management, experimental economics, field experiments

Designing Sales Contests: Does the Prize Structure Matter?

Sales contests are short-term incentives that managers use to motivate salespeople to meet sales targets. It is estimated that up to 90% of firms with a sales force regularly conduct sales contests (Murphy and Sohi 1995), and total expenditures on sales contests in the United States were reported to be more than $26 billion in 2000 (Center for Concept Development and Incentive Federation Inc. 2001). When designing a sales contest, a major question that managers confront is, What should be the prize structure of the contest? Specifically, how many prizewinners should there be, and what should be the value of each prize?

A survey of industry practices reveals that firms have used various prize structures in their sales contests. To name a few, Avis Rent A Car conducted a contest with the simplest possible prize structure—a one-winner (1W) contest, in which the salesperson with the highest dollar sales won a dinner for two (Keenan 1995). Phillips Foods used a grouped-winners (GW) format—that is, a contest with multiple prizewinners and identical prize values. In this instance, the top four salespeople each won an all-expenses-paid trip to Baltimore (Hartnett 2006). Webb Furniture implemented a ranked-winners (RW) contest, which has multiple prizewinners and unique rank-ordered prizes—the top-ranked seller won $10,000, and sellers ranked second through fifth won $8,000, $6,000, $4,000, and $2,000, respectively (*Furniture Today* 2005). NBC adopted a prize structure that is a hybrid of the GW and RW contests—there were three "gold prizes," in which the top three cable advertisement sellers won a trip to New York or Los Angeles; ten "silver prizes," in which the next

*Noah Lim is Assistant Professor of Marketing (e-mail: noahlim@uh.edu), Michael J. Ahearne is Associate Professor of Marketing and Executive Director of the Sales Excellence Institute (e-mail: mahearne@uh.edu), and Sung H. Ham is a doctoral candidate (e-mail: shham@uh.edu), Department of Marketing and Entrepreneurship, C.T. Bauer College of Business, University of Houston. The authors are grateful to the two anonymous *JMR* reviewers for their illuminating insights and helpful suggestions. They thank Ed Blair, Carl Herman, Jim Hess, and Eli Jones for their encouragement and support. They received superb research assistance from Hua Chen and Kihoon Park. They also benefited from seminar participants at the University of Houston and the University of Texas at Dallas Behavioral Research Seminar Series. This research is partly supported by the Institute for the Study of Business Markets Grant No. 0608. Jeffrey Inman served as associate editor for this article.

ISSN: 0022-2437 (print), 1547-7193 (electronic)

Journal of Marketing Research
Vol. XLVI (June 2009), 356–371

ten highest-ranked sellers won a trip to a conference; and hundreds of "bronze prizes," in which sellers received T-shirts and coffee mugs (Donohue 2004). Given the range of prize structures firms have adopted in practice and the plethora of alternative contest designs, managers must decide on the following questions: Does the prize structure of a sales contest matter at all? Is there an optimal prize structure, and if so, what are its characteristics?

In a significant contribution to marketing theory, Kalra and Shi (2001) address these questions and show that the prize structure in a sales contest is an important determinant of sales effort.[1] Using a game-theoretic model, they demonstrate that under the regular assumptions of risk-averse salespeople and logistically distributed sales outcomes, the optimal prize structure is the RW contest, which has the following two characteristics: (1) The number of prizewinners is greater than one, and (2) prizes are rank ordered and unique to that rank—that is, salespeople ranked higher should receive prizes with greater monetary values, and salespeople with different sales ranks should not receive prizes with identical values. The intuition for their result is as follows: First, because risk-averse salespeople place less incremental value on larger prize spreads, it is better to allocate the total prize money across more ranks. Second, having unique rank-ordered prizes motivates salespeople to increase effort to achieve the highest sales ranks possible because every rank that is higher (above the lowest rank that qualifies them as a prizewinner) carries a greater monetary reward. The upshot is that when managers conduct sales contests, they should increase the design complexity of a prize structure by introducing both multiple prizewinners and unique rank-ordered prizes to reap the greatest boost in sales effort and revenues.

However, the usefulness of the theory has been limited because its predictions have not been empirically examined. The need for empirical validation is especially strong given that the theoretical predictions, which are based on comparing the symmetric multiperson Nash equilibrium effort levels across different prize structures, rely on two assumptions that may not be met even in the best-controlled empirical settings. First, salespeople are assumed to be completely homogeneous in their preferences and sales effectiveness. Second, they are assumed to be motivated only by monetary incentives and can make perfectly accurate strategic trade-offs between the costs of expending sales effort and the expected rewards given by a prize structure in a sales contest. We note that it is virtually impossible (and probably not useful) to replicate the perfect homogeneity assumption, particularly in salespeople's preferences over monetary outcomes, in any empirical test. Moreover, for the theory to be useful, it must accommodate bounded-rational behavior, such as different levels of strategic sophistication among salespeople (Camerer, Ho, and Chong 2004; Ho, Lim, and Camerer 2006). Given that both heterogeneity and bounded rationality are unavoidable when theory is translated into an empirical setting, the predictions of the marketing theory of sales contests are by no means assured, even if the primary goal of the empirical test is to examine whether the predictions can be supported by the aggregate behavior across the contestants.

Furthermore, the criteria for a valid empirical test cannot be met by data that are typically available to marketing researchers. First, to construct a causal test of whether the prize structure of a sales contest affects sales effort and revenues, both exogenous variation in the prize structures and the random assignment of salespeople into different contest formats are required. Second, because sales effort is a strategic variable in a sales contest and can vary with the number of contestants, it is crucial to ensure that the number of salespeople in each contest is identical. Third, an accurate test of theory must include the theoretically optimal (or effort-maximizing) prize distribution for the RW contest because there are many ways to distribute prize values across ranks under this prize structure. To ensure this, the researcher must have *ex ante* control over the parameters that influence the effort-maximizing design, such as the aggregate degree of risk aversion of the salespeople and the effects of other monetary incentives that are available to salespeople. The only empirical methodology that meets these criteria is an economic experiment—that is, an experimental study in which decisions are aligned with monetary incentives (for successful applications in marketing, see Amaldoss and Jain 2002, 2005; Amaldoss et al. 2000; Ding et al. 2005; Krishna and Ünver 2008; Lim and Ho 2007).

This article presents both laboratory and field economic experiments that address the empirical question whether the prize structure in a sales contest matters. We examine the performance of three different prize structures by adopting the following experimental design: First, we investigate the incremental impact of introducing multiple winners into a contest by comparing the simplest 1W contest with the GW and RW contests. Second, we study the impact of introducing unique rank-ordered prizes in contests with multiple prizewinners by comparing the GW contest with the theoretically optimal RW contest. In the laboratory experiments, participants make decisions that closely reflect the trade-offs salespeople face when choosing sales effort. The results from the laboratory experiments represent the most accurate empirical outcomes of the theoretical model because the researcher wields the greatest degree of control over the factors that affect whether the model's assumptions are satisfied. The field economic experiments consist of a pair of studies; the first compares the 1W contest with the GW contest in a setting in which salespeople sell fundraising sponsorships through a golf tournament, and the second compares the GW and RW contests in an environment in which salespeople raise funds for a university educational program. The main purpose of the field experiments is to provide researchers with insights into the usefulness of the theoretical model in environments in which the theory is expected to apply. This article is also one of the first in

[1]There are other studies in marketing that examine sales contests, but only Kalra and Shi (2001) address the question of the optimal prize structure from the perspective of a profit-maximizing firm. For example, Murphy, Dacin, and Ford (2004) survey salespeople about their preferences for monetary awards, the duration of a contest, and the proportion of winners in a contest. On the last issue, they show that the most-preferred contest is one in which prizes are awarded to the top 40% of contestants. Gaba and Kalra (1999) study the degree of risk-taking behavior by varying the proportion of "winners" (i.e., participants who receive a prize) in a contest. They show that more risky prospects were chosen when the proportion of winners was "low." They consider only the GW contest in their study.

marketing to employ the methodology of field economic experiments (Harrison and List 2004; Krishna and Ünver 2008).

The results of the laboratory and field experiments are remarkably consistent: The prize structure of a sales contest matters, but not exactly in the way the prevailing wisdom predicts. Our findings strongly indicate that introducing multiple prizewinners into a sales contest indeed improves sales performance. However, introducing unique rank-ordered prizes into a multiple-winner sales contest does not boost sales effort and revenues. In other words, the GW contest performs as well as the RW contest. Thus, designing the optimal sales contest is less complex than what existing theory prescribes because managers need not solve for the prize values for each prizewinner; they simply need to fix the prize values to be identical across winners and to focus on determining the optimal number of winners, which can be calculated using a simple formula.

We organize the remainder of the article as follows: The next section describes the economic model underlying the marketing theory of sales contests and presents the hypotheses. Then, we detail the design and results of the laboratory experiments. This is followed by the design and the results of the field experiments. The final section concludes with a discussion of the managerial implications, as well as the limitations of the research and directions for further research.

MODEL AND THEORETICAL PREDICTIONS

Model Overview

In a sales contest, there are N salespeople contending for prizes based on their individual sales ranking, which is determined by a sales metric (usually dollar sales). The prize structure of a sales contest is given by $P_1 \geq P_2 \geq \ldots \geq P_N$, where P_1 is the monetary value received by the salesperson with the highest rank, and so on. Let $P_j \geq 0$ denote the monetary value received by the salesperson with the jth rank (j = 1, 2, 3, ..., N). The manager is given a budget, B, for the total prize money in the sales contest and must decide on a prize structure that will motivate salespeople to expend the greatest sales effort. We define the terms "prizewinner," "winner," and "prize" to refer only to cases in which $P_j > 0$.

The marketing theory of sales contests was first proposed by Kalra and Shi (2001) and draws on tournament theory in economics (Green and Stokey 1983; Lazear and Rosen 1981; Nalebuff and Stiglitz 1983) and the marketing literature on sales compensation (e.g., Basu et al. 1985). The optimal prize structure for a sales contest is derived through the following steps: (1) The firm must determine how salespeople make decisions about sales effort given a prize structure, and (2) the manager then incorporates knowledge about salespeople's behavior into his or her decision calculus and chooses the prize structure that maximizes the firm's profits. In this article, we highlight only the features of the model that are critical for understanding the experimental tests in the following sections and refer the reader to the first section of the Web Appendix (http://www.marketingpower.com/jmrjune09) for other important mathematical details of the salesperson's behavior and the manager's decisions in the model.

The theory of sales contests assumes that all the N salespeople in the contest are homogeneous in their preferences and sales abilities. Given a prize structure, each salesperson, i, chooses an effort level, e_i, which maximizes his or her utility. The salesperson's utility is assumed to be separable in monetary rewards and sales effort and is given by $U(R_i, e_i) = u(R_i) - c(e_i)$. The salesperson is assumed to be risk averse over the monetary outcomes ($u'[R_i] > 0$ and $u''[R_i] < 0$). The most common specification for $u(R_i)$ is the constant relative risk aversion utility function R_i^α/α, where $0 < \alpha < 1$. The parameter, α, captures the salesperson's degree of risk aversion, with a lower α indicating greater aversion to risk. Sales effort is costly (i.e., time for other activities needs to be sacrificed), and the cost of effort $c(e_i)$ is assumed to be strictly increasing and convex. The sales metric according to which the salesperson is ranked is $s(e_i) = h(e_i) + \varepsilon_i$, where $h(e_i)$ is the deterministic sales effort function that represents the salesperson's ability to convert effort into dollar sales and ε_i is a stochastic component that reflects uncertainty in sales that each salesperson faces (e.g., changes in customer demand). The assumptions of the model are that $h'(e_i) > 0$, $h''(e_i) \leq 0$, and $\varepsilon_i \sim \text{Logistic}(0, \pi^2\beta^2/3)$.

Each salesperson evaluates the trade-off between expending sales effort and the probability of attaining monetary rewards for a given prize structure by choosing effort to maximize his or her expected utility. Given this setup and applying the common specifications of $h(e_i) = e_i$ and $c(e_i) = ke_i^2$, Kalra and Shi (2001) show that the closed-form equilibrium effort level of each salesperson for a given prize structure is as follows:

$$(1) \qquad e^* = \sum_{j=1}^{N} \frac{P_j^\alpha (N - 2j + 1)}{2\alpha k\beta N(N+1)}.$$

The manager rationally anticipates the salesperson's behavior and takes this into account when choosing the prize structure that maximizes the firm's profits. To design the optimal contest, the manager must also ensure that the resultant expected utility of each salesperson is above a reservation level (i.e., satisfying the salesperson's participation constraint) and that the total value of the prizes must be equal to B.

Theory Prediction: The Prize Structure of a Sales Contest Matters

Kalra and Shi (2001) show that the optimal prize structure is a RW contest that has multiple winners (i.e., $P_{w^*} > 0$, where $w^* > 1$) and unique rank-ordered prizes. They also examine two other prize structures we described previously, the 1W contest and the GW contest, and explain why they do not yield as much sales effort as the RW contest.[2]

First, the optimal sales contest must have more than one prizewinner. This is because when salespeople are risk averse, their marginal valuations for larger prize values are lower than their marginal valuations for smaller prizes. This

[2]There are many ways to distribute prizes within the RW and GW prize structures. In this article, we consider only the optimal, or effort-maximizing, design for each prize structure.

leads to smaller optimal prize spreads, which means that it is better to have prizes with smaller values across ranks than to have a single large prize. This principle from the marketing theory of sales contests can be formally described as follows:

> H_1: When salespeople are risk averse, sales contests with multiple prizewinners are superior to a sales contest with only one prizewinner. Specifically, the pattern of sales effort in the following prize structures is predicted to be (a) $e^*(RW) > e^*(1W)$ and (b) $e^*(GW) > e^*(1W)$.

Note that the optimal number of prizewinners in a sales contest increases with the salesperson's level of risk aversion. In the GW prize structure, it can be shown that the optimal number of prizewinners, w_{GW}^*, is equal to $\text{Max}\{1, [N(1 - \alpha)/(2 - \alpha)]\}$, where $[N(1 - \alpha)/(2 - \alpha)]$ is rounded off to the nearest integer.

Second, the marketing theory of sales contests states that prize values should be unique and rank ordered because having prize spreads motivates salespeople to obtain the highest ranks possible. If all prize values across the "winning ranks" (the first w* ranks in which prizes are awarded) are identical, the salesperson has no incentive to expend effort to achieve ranks that are higher than the lowest rank that qualifies for a prize. The same logic applies to any pair of consecutive winning ranks, and thus no two prize values should be identical. Naturally, larger prizes should be awarded for higher ranks so that salespeople will expend effort to achieve the highest possible sales rankings. Formally, we predict the following:

> H_2: In sales contests with multiple prizewinners, sales effort is higher in the prize structure with unique rank-ordered prize values than in the one in which prize values are identical. That is, $e^*(RW) > e^*(GW)$.

Before describing the experimental tests of our hypotheses, we extend the existing theory by providing insights into the following question: Given that the optimal prize structure of a sales contest must have the two characteristics of multiple winners and unique rank-ordered prizes, what is the relative importance of these two characteristics? In other words, what are the effort losses if managers implement the simpler 1W or GW designs instead of the optimal RW prize structure?

Table 1 provides the outcome of our analysis for parameter values of B = \$100, $\beta = 1$, and k = 10.[3] Columns 1 and 2 show the respective effort-maximizing contest designs for the RW and GW prize structures for different contest sizes (N = 8, 15) and the levels of risk aversion (α = .25, .5, .75). Columns 3 and 4 show the relative sales effort in the 1W and GW contests, respectively, as a percentage of the sales effort in the optimal RW contest. First, Column 3 indicates that the greatest boost to sales effort from adopting the more complex RW prize structure over the simplest 1W design occurs when salespeople are more risk averse. Second, sales effort under the GW prize structure ranges from 94.5% to 96% of that of the RW contest. That is, the incentive effect of achieving higher ranks, conditional on being a prizewinner, is small. These observations imply that the benefit of having multiple prizewinners (while keeping prize values identical) is greater than the marginal effect of introducing unique rank-ordered prizes into contests with multiple prizewinners.

LABORATORY EXPERIMENTS

We conducted a pair of laboratory experiments designed to test H_1 and H_2. For the experiments to be considered valid empirical tests of the marketing model of sales contests, we must consider several other criteria. Table 2 outlines these criteria and how well they are met by the laboratory experiments in this article.

First, as we mentioned previously, the prize distributions in the GW and RW contests must be theoretically optimal. To obtain these, the researcher must first specify the level of the contest size, N, and the total prize money, B; estimate the degree of risk aversion, α, of the salespeople; and ensure that the participation constraint is satisfied. More-

[3]The parameter values of B = 100, $\beta = 1$, and k = 10 are chosen both for simplicity and to ensure that the participation constraint of the salesperson is satisfied in all three prize structures given the levels of N and α. We do not vary β and k because the effort-maximizing designs of the RW and GW prize structures do not depend on these parameters as long as the salesperson's participation constraint is met. Our analyses using other parameter values yield similar results. We did not find any cases in which the relative effort of GW was lower than 93%.

Table 1
OPTIMAL CONTEST DESIGNS AND RELATIVE EFFORT IN THE 1W AND GW CONTESTS

	1: RW Contest $\{P_1, P_2, \ldots, P_{w^*}\}$ (\$)	*2: GW Contest* $\{P_1, P_2, \ldots, P_{w^*}\}$ (\$)	*3: Relative Effort of the 1W Contest*[a] (%)	*4: Relative Effort of the GW Contest*[a] (%)
8 Contestants				
α = .25	{49, 31, 16, 4}	{33.3, 33.3, 33.3}	58.7	95.5
α = .50	{58, 30, 11, 1}	{33.3, 33.3, 33.3}	76.4	94.5
α = .75	{77, 20, 3}	{50, 50}	93.8	95.6
15 Contestants				
α = .25	{30, 23, 18, 13, 9, 5, 2}	{16.7, 16.7, 16.7, 16.7, 16.7, 16.7}	39.0	96.0
α = .50	{35, 26, 18, 11, 6, 3, 1}	{20, 20, 20, 20, 20}	59.1	94.5
α = .75	{52, 29, 13, 5, 1}	{33.3, 33.3, 33.3}	84.8	95.6

[a]The parameters are B = \$100, $\beta = 1$, and k = 10. Relative effort refers to sales effort relative to that in the optimal RW contest.

Table 2
CRITERIA FOR A VALID EMPIRICAL TEST AND THE RELATIVE STRENGTHS OF THE LABORATORY AND FIELD EXPERIMENTS

1: Criteria	*2: Laboratory Experiments*	*3: Field Experiments*
Optimal prize distributions for the GW and RW contests	Yes Specifications: N = 8 α = .50 (measured for Experiment 1) α = .55 (measured for Experiment 2) B = $15 per round in Experiment 1, $6.75 per round in Experiment 2 Participation constraint satisfied: a. No other monetary incentives b. Expected utility is positive	Yes Specifications: N = 15 α = .50 (measured) B = $300 in Experiment 1, $200 in Experiment 2 Participation constraint satisfied: a. No other monetary incentives
Effort directly observable	Yes The following must be specified: $h(e) = e$ $\varepsilon_i \sim \text{Logistic}(0, \pi^2\beta^2/3)$, $\beta = 30$ $c(e) = ke^2$, $k = .0002$	No This is because h(e), β, and c(e) cannot be accurately determined
Homogeneity of Salespeople		
Sales productivity	Close Control Control mechanism: h(e), c(e), and β are identical and common knowledge.	Partial Control There may be heterogeneity in sales productivity across contestants. Control mechanism: Each salesperson undergoes the same sales training courses, and range of sales experience (number of years) is small.
Preferences over monetary prizes	Not Controlled Preferences may be heterogeneous. Risk aversion parameter (α) used is an aggregate measure.	Not Controlled Preferences may be heterogeneous. Risk aversion parameter (α) used is an aggregate measure.
Trained salespeople performing real selling task	No	Yes

over, both the contest size and the total prize money must be identical across different contests. Second, though not a necessary condition for a valid test, it is useful if the precise level of effort (or a decision proxy) can be predicted and directly observed. This can be accurately done only in a laboratory setting, in which the researcher can explicitly specify the sales effort function, the cost of effort function, and the variance parameter of the logistically distributed sales outcomes (h[e], c[e], and β, respectively). Third, the model assumes that the salespeople are homogeneous in sales productivity and their degree of risk aversion. In the laboratory experiments, we are able to meet the homogeneity assumption for sales productivity because h(e), c(e), and β are identical and common knowledge to all salespeople. However, we do not control for homogeneity of risk aversion for two reasons: First, we are unaware of any reliable method of inducing an identical level of risk aversion across contestants. Second, and more important, because managers do not have control over salespeople's risk preferences (and do not attempt to homogenize levels of risk aversion when designing sales contests), we believe that it is important for the theory to be empirically validated under this condition.[4] In this case, the procedure closest to satisfying the assumption is to estimate an aggregate level of risk aversion across the contestants and to use this measure to design the theoretically optimal sales contest. Finally, we recognize that there may be concerns about external validity issues because the laboratory experiments do not involve trained salespeople undertaking real selling tasks. We address this concern with the field experiments.

The structure of our laboratory experiments is similar to previous research that tests tournament theory in economics and management (Bull, Schotter, and Weigelt 1987; Orrison, Schotter, and Weigelt 2004; Schotter and Weigelt 1992), except for two major innovations that we introduce. First, we study complex prize structures, such as the RW contests with as many as eight contestants, whereas the focus of prior research has been on simpler prize structures with a smaller group of contestants. Second, we endogenously determine the prize distributions in the RW and GW contests after estimating the aggregate risk aversion parameter, so they correspond to the theoretically optimal prize distributions for each prize structure.[5]

[4]In contrast, managers often try to meet the homogeneity assumptions for sales productivity to create a contest that is perceived by contestants as being fair, by adjusting the sales metric by which salespeople are ranked. For example, instead of gross sales, the metric might be sales dollars over quota and/or sales for a new line of products.

[5]Prior research has either studied the 1W prize structure with two contestants (Bull, Schotter, and Weigelt 1987; Schotter and Weigelt 1992) or compared the 1W and GW prize structures while allowing the number of contestants (two, four, or six) and the total prize money B to vary (Orrison, Schotter, and Weigelt 2004). In the latter, the authors exogenously determined the prize distributions in the GW contests, and they assumed that the contestants were risk neutral. Moreover, they assumed that sales outcomes were uniformly distributed.

Experiment 1

The first laboratory experiment compares the 1W, GW, and RW contests using a between-subjects design. Before we specify the prize values in the contests with multiple prizewinners, we must estimate the aggregate level of risk aversion of the salespeople.

Measuring risk aversion. Holt and Laury (2002) introduced the procedure we use. One hundred twenty participants were randomly recruited from students in an undergraduate sales management class who had indicated their availability to participate in our experiments. Each student was given a questionnaire containing six lottery-choice questions, and their task was to select one of two lotteries (Option A or Option B) for each of the six questions. The six questions for one of the two payoff conditions we employed appear in Columns 1 and 2 of Table 3. Note that across the six decisions, the stakes in Options A (safer choice) and B (riskier choice) are fixed, though the probabilities vary, with the differences in the expected value of the two options (which were not provided to the participants) increasing in favor of Option B as participants answer the questions. The basic idea of this procedure is that each participant's level of risk aversion can be inferred from the number of safe choices (Option A) he or she makes before crossing over to Option B because the choice proportion corresponds to an interval for the risk aversion parameter, α, in the constant relative risk aversion utility function. For example, if a participant chooses Option A for the first two decisions, followed by Option B for the next four, it implies that $.59 < \alpha < .85$ for that participant. Sixty participants were given the exact lotteries in Table 3, and the other 60 made choices in another payoff condition in which the stakes in Table 3 were increased by 1.5 times. We chose the lottery stakes so that they would span the range of monetary outcomes from participating in the experimental contests. Before the participants made their decisions, they were told that six of them would be randomly selected (we ultimately selected three from each payoff condition) to participate in one of the six questions, and they would play the lottery (Option A or B) they had selected for that question. All payments were made in cash, and all six participants received their payments within a week of the study.

Using the choice data, we estimated an aggregate risk aversion parameter for the participants using a probabilistic choice rule first derived by Luce (1959), in which a decision maker's probability of choosing Option A, when faced with two options, A and B, with utilities U_A and U_B, respectively, is given by $Pr(A) = U_A^{1/\lambda}/[U_A^{1/\lambda} + U_B^{1/\lambda}]$, where λ is a rationality parameter that captures the degree of sensitivity of choice probabilities to utilities (with smaller values representing greater sensitivity). This choice rule also has the property that choice probabilities remain unchanged when payoffs are scaled by a multiplicative constant, so that participants who select the same number of safe choices across the two payoff conditions have the same degree of risk aversion (see also Goeree, Holt, and Palfrey 2005). The maximum-likelihood estimate of the aggregate risk aversion parameter across the two payoff conditions is $\alpha = .476$ (log-likelihood = –376.22 and $\lambda = .132$), which is not significantly different from $\alpha = .5$ ($\chi^2 = .628$, d.f. = 1, $p = .427$). This estimated value of α also falls within the range of .33 to .71, as reported by Holt and Laury (2002).

Design parameters. The values of the other parameters we selected are B = \$15, N = 8, h(e) = e, $c(e) = ke^2$, k = .0002, and $\beta = 30$. Together with the risk aversion coefficient estimate of $\alpha = .5$, these parameter values yield the effort-maximizing designs of {\$5.00, \$5.00, \$5.00, \$.00, ..., \$.00} and {\$8.75, \$4.46, \$1.61, \$.18, \$.00, ..., \$.00} for the GW and RW contests, respectively. The theoretical predictions of sales effort are 62.76, 77.64, and 82.17 for the 1W, GW, and RW prize structures, respectively. The contest designs and predictions also appear in the top half of Table 4. We selected the parameter values so that they would meet all the following conditions: First, we chose

Table 3

LOTTERY CHOICE QUESTIONS FOR MEASURING RISK AVERSION

1: Option A	*2: Option B*	*3: Difference in Expected Value (A–B) (not Provided to Participants) (\$)*
1. 40% chance of \$20 and 60% chance of \$16	40% chance of \$38.50 and 60% chance of \$1.00	1.60
2. 50% chance of \$20 and 50% chance of \$16	50% chance of \$38.50 and 50% chance of \$1.00	–1.75
3. 60% chance of \$20 and 40% chance of \$16	60% chance of \$38.50 and 40% chance of \$1.00	–5.10
4. 70% chance of \$20 and 30% chance of \$16	70% chance of \$38.50 and 30% chance of \$1.00	–8.45
5. 80% chance of \$20 and 20% chance of \$16	80% chance of \$38.50 and 20% chance of \$1.00	–11.80
6. 90% chance of \$20 and 10% chance of \$16	90% chance of \$38.50 and 10% chance of \$1.00	–15.15

Table 4

THEORETICAL PREDICTIONS FOR THE LABORATORY EXPERIMENTS

Prize Structure	*Optimal Design* $\{P_1, P_2, ..., P_{w*}\}$	*Effort* (e^*)	*Relative Effort (%)*	*Expected Utility*
Experiment 1				
1W	{15.00}	62.76	76.4	.18
GW	{5.00, 5.00, 5.00}	77.64	94.5	.47
RW	{8.75, 4.46, 1.61, .18}	82.17	100.0	.34
Experiment 2				
GW	{5.00, 5.00, 5.00}	76.50	93.8	.48
RW	{9.16, 4.33, 1.39, .12}	81.59	100.0	.29

B = \$15 (for each contest in each decision round) so that the monetary incentives would be sufficiently motivating. Second, we wanted N to be large enough so that the optimal number of prizewinners in each of the GW and RW contests would be greater than two; thus, we selected N = 8. Third, we used the simple functions of h(e) = e and $c(e) = .0002e^2$ to reduce the mathematical complexity of the decision tasks. Fourth, because we knew that the predicted effort levels for the GW and RW can be close, we chose parameters that separate the point predictions as much as possible. Fifth, we did not want the predictions to be anchored on any focal numbers. Finally, we chose the parameter combination so that the participation constraints would be satisfied in all the contests.

Procedure. A total of 72 participants were randomly selected from the 120 participants who completed the risk aversion survey to participate in the experiment. We further randomly divided these participants into three sessions of 24 participants each, with each session assigned to one of three prize structures studied (1W, GW and RW contests). Participants received course credit for arriving on time and were paid in cash privately at the end of the experiment. Participants earned an average of US\$19, with the minimum and maximum at \$2 and \$57, respectively. Each experimental session consisted of 15 decision rounds and lasted about an hour. Because there were 24 participants in a session and N = 8, there were three contest groups in every round. In each round, participants were randomly and anonymously matched with seven other contestants. In addition, participants were reminded that each of them would encounter the same set of contestants only once during the 15 rounds. We did this to minimize any reputation effects, so that the one-shot nature of the theoretical model is preserved. After participants entered the laboratory, they were seated at separate computer terminals, and the experimenter read the instructions aloud. The experiment was implemented using z-Tree software (Fischbacher 1999).

We designed the instructions and the decision task to be as simple as possible. In line with prior experimental tests of tournament theory (Bull, Schotter, and Weigelt 1987; Orrison, Schotter, and Weigelt 2004; Schotter and Weigelt 1992; Weigelt, Dukerich, and Schotter 1989), the instructions did not contain the words "contest," "prize," "winning," or "losing," so that no nonpecuniary incentives were introduced. Participants were told that their decision task was to select a decision number (e_i) from 30 to 100. Each decision number is associated with a decision cost ($.0002e_i^2$), which is greater the higher the decision number they choose. The decision cost corresponding to each decision number was provided to each participant on an information sheet. Participants were then told that after they entered their decision number into the computer program, the program would generate a random number (ε_i) ranging from –172 to 172.[6] To explain the probability distribution of ε_i, we did not use the term "logistic distribution" but rather communicated the concept to the participants in three complementary ways using another information sheet: We provided (1) a histogram showing the relative frequencies of the numbers; (2) a table showing the probabilities for every interval of ten numbers (e.g., –169 to –160, –159 to –150) except for the endpoints, so that the interval size was three (i.e., –172 to –170 and 170 to 172); and (3) a "full table" displaying the probability that each integer would be drawn. Participants were next told that the computer would add the decision number with the random number to form the final number ($s[e_i]$), by which they would be ranked from highest (Rank 1) to lowest (Rank 8). They would then earn a fixed payment (P_j) that corresponded to their rank. The fixed payment for each rank varied with the contest treatment of the participant. The cash earning for each participant in each round was equal to the fixed payment less the decision cost. We also gave every participant a start-up payment of \$5 in all three treatment sessions. We did this so that the cash earnings of the participants would remain positive, especially in the earlier rounds. After completing a practice round to familiarize themselves with the software, participants proceeded to choose their decision number for 15 rounds.

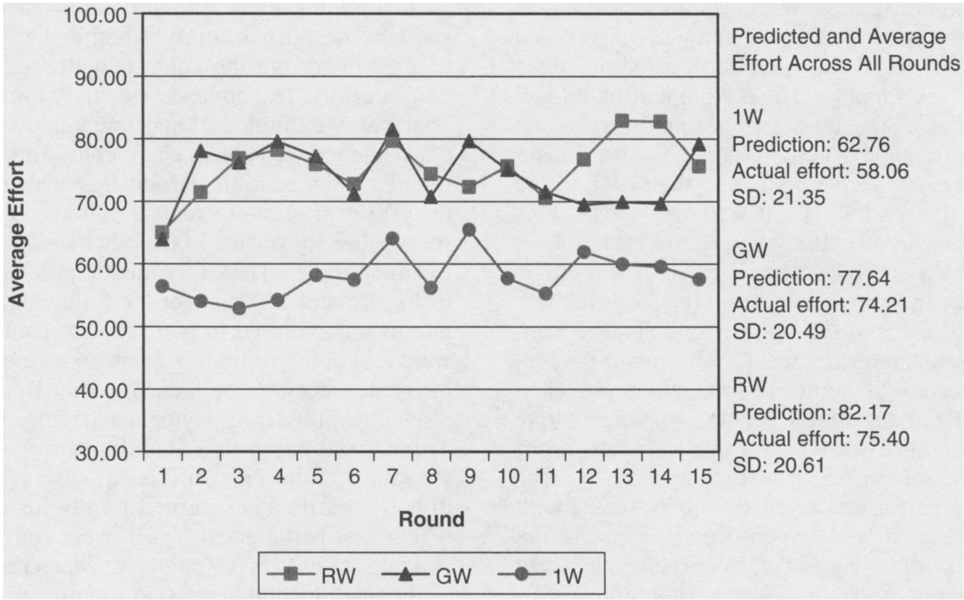
Results. Figure 1 reports the actual average effort level (decision number) for the three prize structures. Across the 15 rounds, average effort is 58.06, 74.21, and 75.40 in the 1W, GW, and RW contests, respectively. The key observation in the graph is that average effort in the 1W contest is directionally lower than the contests with multiple prizewinners in every round. Moreover, the average effort level in the RW contest is not always directionally higher relative to the GW contest.

We now proceed to conduct formal statistical tests of the hypotheses. Because participants made multiple decisions in the experiments, we account for potential within-subject correlation by clustering the standard errors at the participant level in all the statistical tests we report. We begin by comparing the average effort across all rounds with the theoretical point predictions in each of the contests using one-sample t-tests. In the 1W and RW contests, average effort is lower than the respective predicted levels of 62.76 ($t = -2.00$, $p = .046$) and 82.17 ($t = -2.93$, $p = .004$). The underinvestment in effort is small because the hypothesis that average effort is equal to 96% of the predicted effort levels cannot be rejected for the 1W ($t = -.93$, $p = .35$) and RW ($t = -1.51$, $p = .13$) contests. In the GW contest, average effort of 74.21 is not significantly different from the theoretical prediction of 77.64 ($t = -1.58$, $p = .116$). Next, we check whether there are dynamics in the data in each of the three contests by grouping the data into five blocks of three rounds and performing ordinary least squares (OLS) regressions of effort on the blocks (with one of the blocks as the base). For the 1W contest, the only difference in effort across blocks arises between Block 1 (Rounds 1–3) and Block 3 (Rounds 7–9), with effort being higher in the latter ($p = .012$). There are no differences in effort across the blocks of rounds in the GW contest. In the RW contest,

[6]Because the logistic distribution is continuous with an infinite range, we needed to "discretize" and truncate the distribution to simplify the instructions. To do so, we first calculated the probability that each integer x would be drawn using $F(X < x + .5) - F(X < x - .5)$. Next, we selected the endpoints with a view to balance the trade-off between overloading the participants with too much information and achieving an accurate representation of a logistic density with $\beta = 30$. With endpoints of –172 and 172, we found that using the expected frequencies for each integer for 5000 draws produces a maximum-likelihood estimate of $\beta = 29.37$ (log-likelihood = –26,730). This is not significantly different from $\beta = 30$ ($\chi^2 = 3.0$, d.f. = 1, $p = .083$).

Figure 1

PRIZE STRUCTURES WITH MULTIPLE WINNERS LEAD TO GREATER EFFORT



the only difference across blocks is that effort in Block 5 (Rounds 13–15) is significantly higher than effort in Blocks 1 and 4 (*ps* = .01 and .011, respectively). However, effort in Block 5 is no different than effort in Blocks 2 and 3. Overall, although there are no consistent directional trends in the data, there are some differences in effort across blocks. Consequently, we report the results of the hypotheses tests for the data pooled across all rounds and also on a block-by-block basis.

Result 1: Sales contests with multiple prizewinners lead to greater effort. The top half of Table 5 displays the results of OLS regressions of effort on prize structures for the pooled data and for each of the five blocks. In each of the regressions, the 1W contest was the base, with treatment dummies for the prize structures with multiple prizewinners. The results of the regression for the pooled data in the first column confirm the predictions of H_1; that is, effort is higher in the GW contest (t = 5.04, $p = .000$) and in the RW

Table 5

HAVING MULTIPLE PRIZEWINNERS BOOSTS EFFORT, BUT UNIQUE RANK-ORDERED PRIZES DO NOT

	Overall	*Block 1*	*Block 2*	*Block 3*	*Block 4*	*Block 5*
Test of H_1						
Constant (Base = 1W)	58.06*	54.46*	56.64*	61.90*	58.31*	59.01*
	(2.35)	(2.75)	(3.11)	(3.00)	(3.08)	(3.74)
GW	16.15*	18.28*	19.39*	15.42*	13.74*	13.92*
	(3.20)	(3.42)	(4.27)	(4.56)	(4.19)	(5.01)
RW	17.33*	16.68*	19.00*	13.56*	16.00*	21.43*
	(3.30)	(3.95)	(4.68)	(3.93)	(4.75)	(4.81)
Test of H_2						
Constant (Base = GW)	74.21*	72.74*	76.03*	77.32*	72.04*	72.93*
	(2.17)	(2.03)	(2.92)	(3.44)	(2.84)	(3.33)
1W	−16.15*	−18.28*	−19.39*	−15.42*	−13.74*	−13.92*
	(3.20)	(3.42)	(4.27)	(4.56)	(4.19)	(5.01)
RW	1.19	−1.60	−.39	−1.86	2.26	7.51
	(3.17)	(3.49)	(4.55)	(4.27)	(4.60)	(4.49)
Observations	1080	216	216	216	216	216
Clusters	72	72	72	72	72	72
R^2	.1263	.1608	.1610	.0992	.0980	.1471

*Significant at the 5% level.

Notes: Standard errors are in parentheses.

contest (t = 5.26, p = .000) than in the 1W contest. The other columns in Table 5 show that this result holds when the data are analyzed on a block-by-block basis.

Result 2: Having unique rank-ordered prizes does not lead to greater effort. We now test H_2, which predicted that effort would be higher in the RW contest than in the GW contest because the prize spreads of unique rank-ordered prizes in the former motivate salespeople to aim for the highest possible sales rankings. The bottom half of Table 5 displays the results of the OLS regressions of effort on prize structures with the GW contest as the base treatment. There is no difference in effort between the GW and the RW contests for the pooled data (t = .37, p = .71). This finding also holds when we analyze the data at the block level; in each of the five blocks, there is no statistical difference in effort between the two contests at the 5% level.

Discussion. The results of the OLS regressions with standard errors clustered at the participant level support the prediction that contests with multiple prizewinners lead to greater effort but not the prediction that having unique rank-ordered prizes further boosts effort. We also conducted two additional statistical tests that confirm these findings. First, we performed an analysis of variance with the prize structure as a between-subjects factor and the decision round as a within-subjects factor. The results show that the prize structure of a sales contest does affect effort ($F_{2, 69}$ = 17.57, p = .000), and the tests of effort between contests show that effort in both the RW and the GW contests is greater than that in the 1W contest (both ps = .000), though there is no difference in effort between the RW and the GW contests (p = .718). Next, we used the average effort of each participant as the unit of analysis. The nonparametric Mann–Whitney tests reveal that compared with the 1W contest, effort is higher in the RW (Z = 4.27, p = .000) and the GW (Z = 4.43, p = .000) contests. Again, effort levels are not different across the two multiple prizewinners contests (Z = .536, p = .592).

We were also able to estimate the implied risk aversion coefficient of the participants using the experimental data and the expression for equilibrium effort in Equation 1. The nonlinear least squares estimate of the aggregate risk aversion coefficient is .505 (t = 11.60, p = .000), and the maximum-likelihood estimate (assuming that effort decisions are normally distributed with the equilibrium effort as the mean) is .507. These estimates provide additional evidence that our measure of .476 using Holt and Laury's (2002) procedure tracks the aggregate risk aversion of the participants well.

Finally, we computed the power of the tests in the OLS regressions, assuming a Type I error rate of 5% and using the theoretical predictions of the differences across the contests as the alternative hypothesis. For the tests of H_{1a} and H_{1b}, the power of each of the tests is above .99. However, the power of the test of H_2 is .30, which is below the usual benchmark of .8. This is mainly because the predicted effort levels in the GW and RW contests are small. The power levels reported are conservative because we accounted for each participant making repeated observations (so that effort decisions by the same participant are not independent); had we not done so, we would have obtained an erroneous power level of .84 for the test of H_2. Because of the low power of the test, we designed a second laboratory experiment that tests H_2 with sufficient statistical power.

Experiment 2

In this experiment, we employ a within-subject design (which yields more statistical power than a between-subjects design) to study whether having unique rank-ordered prizes in sales contests with multiple prizewinners boosts effort. The contests we study are the GW and RW contests. We divided 48 participants evenly between two experimental sessions, each consisting of 30 decision rounds. In one session, participants made effort decisions in the GW contest for 15 rounds, followed by the RW contest for another 15 rounds. In the other session, we reversed the treatment order. The experimental procedure was identical to Experiment 1 except for the following: Because participation was required in two separate contests, participants were told at the beginning of the experiment that the fixed-payment schedule would change after Round 15 (they were also shown the fixed-payment schedules of both contests). These fixed-payment schedules correspond to the prize structure of the contests. Next, instead of earning cash directly, participants earned points for the 30 rounds in which they participated, which were converted into dollars at a rate .45 dollars per point earned, so that the dollar value of the total prize money, B, in each round was .45 times 15 points, which is equal to $6.75.

To design the optimal prize distributions for the GW and RW contests, we used Holt and Laury's (2002) procedure to estimate the aggregate risk aversion coefficient for the participants. Seventy potential participants (including the 48 participants who actually participated) answered the lottery choice questions given in Table 3, and the estimated risk aversion coefficient, α, is .55 (log-likelihood = –225.3, λ = .155). Given this estimate and using the parameter values of B = 15, N = 8, h(e) = e, c(e) = $.0002e^2$, and β = 30 from Experiment 1, we obtain the effort-maximizing designs (in points) of {5.00, 5.00, 5.00, .00, ..., .00} and {9.16, 4.33, 1.39, .12, .00, ..., .00} for the GW and RW contests, respectively, with corresponding effort predictions of 81.59 and 76.50 (see bottom half of Table 4).

Results. As in Experiment 1, we account for within-subject correlation by clustering the standard errors at the participant level in all the statistical tests we report. Figure 2 shows the average effort levels across all 15 rounds for the two prize structures. In the GW contest, average effort is 76.34, which is not significantly different from the predicted level of 76.50 (t = –.07, p = .944). Average effort is 76.31 in the RW contest, which is lower than the predicted level of 81.59 (t = –2.29, p = .022). Again, the level of underinvestment in effort is small because the hypothesis that average effort is equal to 97% of the predicted effort level cannot be rejected (t = –1.23, p = .22). Furthermore, when we divide the data into five blocks of three rounds each, OLS regressions of effort on blocks reveal no statistical differences in effort among blocks in each of the contests.

Table 6 displays the results of the OLS regressions of effort on prize structures with the GW contest as the base treatment. We also control for any possible carryover effects from the first contest treatment with an order effect dummy. The pattern of results is identical to that in Experi-

Figure 2
SUMMARY RESULTS OF EXPERIMENT 2

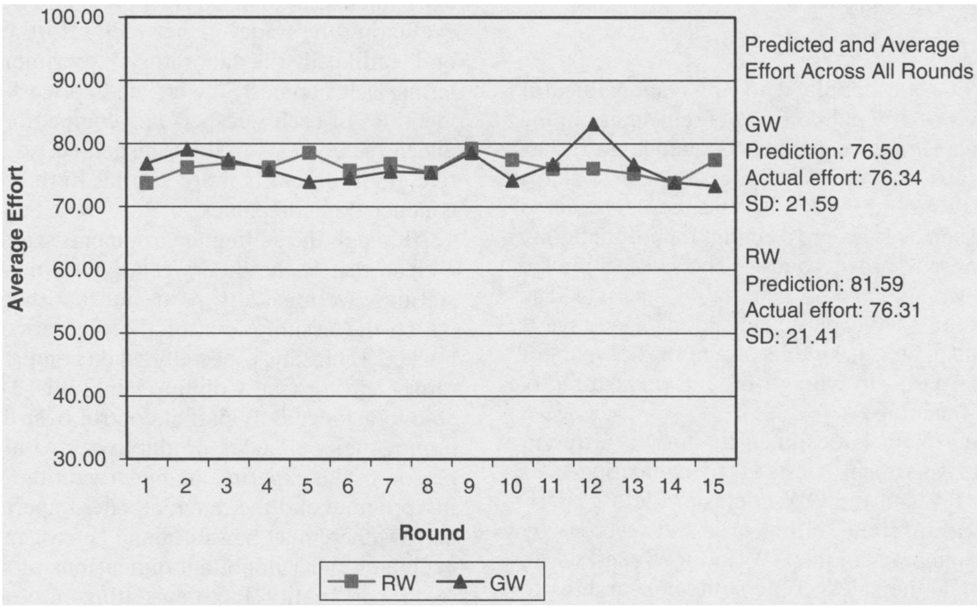


Table 6
EXPERIMENT 2 CONFIRMS THAT UNIQUE RANK-ORDERED PRIZES DO NOT MATTER

	Overall	*Block 1*	*Block 2*	*Block 3*	*Block 4*	*Block 5*
Constant (Base = GW)	74.95*	76.42*	74.57*	74.20*	75.99*	73.54*
	(2.30)	(2.37)	(2.56)	(2.54)	(2.63)	(3.36)
RW	–.03	–2.19	2.05	.65	–1.54	.85
	(1.47)	(2.18)	(2.41)	(2.14)	(2.47)	(2.80)
Carryover effect (from	2.80	2.74	.41	4.60*	4.07	2.17
first contest treatment)	(1.47)	(2.18)	(2.41)	(2.14)	(2.47)	(2.80)
Observations	1440	288	288	288	288	288
Clusters	48	48	48	48	48	48
R^2	.0042	.0075	.0024	.0123	.0103	.0025

*Significant at the 5% level.
Notes: Standard errors are in parentheses.

ment 1. There is no difference in effort between the GW and the RW contests for the pooled data (t = –.02, p = .981).[7] Moreover, this result is consistent across each of the five blocks. More important, the power of the test is .92, which is sufficiently high. Next, a repeated measures analysis of variance shows no difference in effort between the two contests ($F_{1, 47}$ = .001, p = .982). Using the average effort of each participant as the dependent variable, we conducted nonparametric Wilcoxon signed-rank tests, which reveal no difference in effort between the RW and the GW contests (Z = –.164, p = .870). Thus, the results of Experiment 2 confirm that there is no difference in effort levels between the GW and the RW contests. Finally, the nonlinear least squares estimate of the implied aggregate risk aversion coefficient from the effort data is .558 (t = 3.31, p = .002), which is close to our measure of .55 using the lottery procedure, suggesting that our measure is robust.

Overall, the two laboratory experiments show that the prize structure of a sales contest matters but not exactly in the way existing marketing theory predicts. The optimal sales contest should indeed have multiple prizewinners, but having unique rank-ordered prizes does not boost sales effort.

FIELD EXPERIMENTS

We conducted two field experiments designed to test H_1 and H_2 further. These experiments deliver value to our research by serving as an external validity check because

[7]The results of the OLS regressions used to test the hypotheses in both laboratory experiments remain unchanged when we include the participants' level of risk aversion as a covariate. Details are available on request.

the environmental settings in the field experiments correspond to those to which the marketing theory of sales contests should apply. However, as we discuss subsequently, the trade-off is that the field experiments do not meet the assumptions of the theoretical model of sales contests as well as the laboratory experiments.

Participants and Field Environment

We recruited 60 salespeople who are responsible for fund-raising activities at a public research university in the United States. These fund-raising events include golf tournaments, career fairs, media advertising, and sponsorships for educational programs. All the salespeople received training in areas such as sales prospecting, rapport building, sales presentations, and order requests. The salespeople are diverse: 51% are women, and 63% are non-Caucasian. The average salesperson is 23 years of age, and the age range is 20–29 years. Salespeople have an average of 4.5 years of work experience, with an average of 2.3 years directly related to selling functions.

Each field experiment is designed to study a different hypothesis. Field Experiment 1 tests H_1 by comparing sales outcomes in the 1W and the GW contests. Field Experiment 2 involves a different selling task and tests H_2 by comparing sales outcomes in the GW and RW contests. In Field Experiment 1, the salespeople sold sponsorships to the business community through a golf tournament. The sponsorship products include individual player and team sign-ups for the golf tournament. The full list of products and prices appears in Table 7. In Field Experiment 2, salespeople solicited donations from companies to raise funds for the university, and they were provided a list of 200 potential sponsors they could contact. The contest cycle in each experiment is six weeks. In all the sales contests, the salespeople received only one formal update of the sales dollars of the other contestants from the researchers (at the end of the fourth week into the contests). The results of the contests were announced only when both experiments were concluded. Other than the prizes in the sales contests, there were no other monetary incentives. The same group of 60 salespeople performed the two selling tasks, which overlapped for four weeks, albeit with different sales peaks. To our knowledge, performing multiple selling tasks (each with its own incentives) is common to salespeople. Note that both the selling tasks and their sales cycles are features of the fund-raising tasks that the salespeople are supposed to undertake and were not constructed specially for this research.

Table 7
PRODUCT LINE AND PRICES IN FIELD EXPERIMENT 1

Product Description	*Price ($)*
Individual player	150
Driving range sponsor	250
Golf towel sponsor	500
Foursome team	500
Dinner sponsor	500
Snack/drink sponsor	500
Sign sponsor	1,000

The field experiments have two advantages over the laboratory experiments because the empirical setting is more "natural" along two dimensions. First, they involve real selling tasks by salespeople who received formal sales training. Compared with the participants in the laboratory experiments, trained salespeople could be more skillful in evaluating the trade-off between effort and rewards. Second, unlike in the laboratory experiments, we used the terms "sales contest," "win," and "prizes" and revealed the identities of each salesperson's competitors when we introduced the contests to the salespeople, so that the competitive environment is more similar to that of sales contests conducted in companies.

Although the selling environments we described provide a clean test of the theory relative to most other company settings, we must still point out that there are features in which the assumptions of the theoretical model are not tracked as closely, especially when compared with the laboratory setting (see Column 3 in Table 2). First, we were able to achieve only partial control over the assumption of homogeneity in sales productivity. Although each salesperson in the experiments underwent the same sales training program and the range of sales experience is relatively narrow, there may be substantial heterogeneity across salespeople on other important dimensions of sales productivity (e.g., personality factors, selling approach). Moreover, salespeople may believe that they are competing against others who may be more or less skilled in selling. Second, although sales effort is the primary outcome measure in the theoretical model, we are unable to use this measure, because it is extremely difficult to observe and quantify sales effort in the field environment. This means that only sales outcomes can be used to determine whether there are differences between prize structures. Furthermore, a feature of the field test is that we cannot generate point predictions for dollar sales for the field experiments. For these, it would be necessary to have accurate measures of the sales effort function and its cost, which in turn would require a good measure of sales effort. We note that the lack of point predictions is not specific to our experimental design but is likely to be a feature of virtually all field settings. However, the consequence is that our hypotheses and the statistical analyses are limited in that they are only directional in nature (because even if they are not observed, the point predictions exist). Finally, the sample size in our study is relatively small, which limits the power of the statistical tests. Thus, our results must be interpreted with these caveats in mind.

Contest Design

To design prize distributions for the GW and the RW prize structures that are theoretically optimal, we estimated the salespeople's aggregate level of risk aversion using Holt and Laury's (2002) procedure. Of the 60 salespeople, 56 made the six lottery decisions with the payoffs given in Table 3 multiplied by 1.5. We then randomly selected 3 salespeople and paid them according to their lottery decisions for one of the six questions. The estimated aggregate risk aversion parameter, α, is .47 (log-likelihood = –171.9, λ = .13), which is not significantly different from α = .50 (χ^2 = .6, d.f. = 1, p = .439).

In Field Experiment 1, which compares the 1W and GW prize structures, the 60 salespeople were randomly divided into four contest groups, with two contest groups for each prize structure. The number of contestants, N, in each contest group was 15. We deliberately chose N to be larger than 8, the contest size in the laboratory experiments, because we wanted to determine whether the empirical outcomes would be similar for a larger sales force. The total prize money in each contest group was B = \$300. In the 1W contest groups, the salesperson with the highest dollar sales received \$300, and the other 14 received no monetary rewards. In the GW contest groups, $w_{GW}^* = 5$, so the salespeople who were ranked from first to fifth in each contest group received \$60 each. Table 8 displays the prize distributions for the contests.

In Field Experiment 2, which compares the RW and GW prize structures, 60 salespeople were again randomly divided into four contest groups of 15 each (salespeople do not compete against the same set of contestants in the two selling tasks). Again, there were two contest groups for each prize structure. The lower half of Table 8 displays the prize distributions for the contests. The total prize money in each contest group was B = \$200, which is lower than in Field Experiment 1 because prior data indicate that dollar sales are higher in the golf tournament. The optimal design for the RW contest has six prizewinners, and the prize distribution is {\$70, \$51, \$36, \$23, \$13, \$7}.[8] For the GW prize structure, because the optimal number of winners is $w_{GW}^* = 5$, salespeople ranked first to fifth in each contest group received \$40 each.

Results

For every contest in each experiment, there are 30 observations (two contest groups, N = 15 each). Table 8 also reports the average dollar sales for each contest in the two experiments. In Field Experiment 1, the average dollar sales in the GW contest are \$1,391, which is directionally higher than the average of \$897 in the 1W contest. In Field Experiment 2, average sales in the RW and GW contests are \$333.33 and \$358.33, respectively.

We performed likelihood-ratio tests to check for statistical differences in average sales using the theoretical assumption of $s_i \sim \text{logistic}(\mu[\cdot], \pi^2\beta^2[\cdot])$, where $(\cdot)$ represents a prize structure. First, we freely estimate the means and variances for each prize structure. Second, we compare the log-likelihood in the unrestricted model with that of a restricted model in which the average sales of the prize structures are constrained to be equal. The results appear in Table 9. In Field Experiment 1, the likelihood-ratio test reveals that average dollar sales in the GW contest is higher than those in the 1W contest. Specifically, the log-likelihood drops from –483.0 in the unrestricted model to –486.5 under the constraint of $\mu_s(GW) = \mu_s(1W)$ ($\chi^2 = 7.0$, d.f. = 1, $p = .008$). In Field Experiment 2, the likelihood-ratio test indicates that there is no difference in the average sales between the RW and the GW contests. The log-likelihood values in the full and restricted models (the latter with the constraint $\mu_s[RW] = \mu_s[GW]$) are –398.955 and –398.972, respectively, and the chi-square statistic is .034 (d.f. = 1, $p = .854$).

This pattern of results is also supported by Wilcoxon rank-sum tests. In Field Experiment 1, sales in the GW contests are indeed higher than when there is only one prizewinner (Z = 3.00, $p = .003$). In Field Experiment 2, there are no differences in sales between the RW and the GW contests (Z = –.44, $p = .660$).[9] Finally, in Field Experiment 1, the bootstrapped 95% confidence interval of the difference in average sales between the GW and the 1W contests is [72, 1048]. The same measure for the RW and GW contests in Field Experiment 2 is [–75, 124], which includes 0.

Because all the salespeople performed the two selling tasks, we can use a correlation analysis of their ranks in the two contests to examine whether there is strong evidence of heterogeneity in sales productivity. The Spearman rank correlation coefficient in our experiments is .246 and is not significantly different from zero ($p = .063$), indicating that heterogeneity might not be a key driver of the results.

Overall, the findings of the field experiments are consistent with those of the laboratory experiments: Introducing multiple prizewinners into sales contests increases sales

[8]Strictly speaking, the optimal RW design prescribes awarding \$6 to the salesperson ranked sixth and \$1 to the salesperson ranked seventh. We believed that the \$1 prize was too low and awarded \$7 to the sixth place.

[9]The results do not change if we use t-tests instead of the Wilcoxon rank-sum tests: Sales are higher in the GW contest than in the 1W contest (t = 2.01, $p = .049$), and there are no differences in sales between the RW and GW contest (t = –.491, $p = .625$). The levels of "observed power" of the t-tests—that is, power measures computed using the differences in the empirical means—are .51 and .08 in Field Experiments 1 and 2, respectively. These measures are not as accurate as those we computed for the laboratory experiments, in which we have the theoretical predictions of effort (these predictions exists but are unobservable in the field context). Furthermore, Hoenig and Heisey (2001) highlight the problem of drawing inferences from "observed power"; because there is a direct relationship between the p-value and observed power, an observed null effect will rarely have high power.

Table 8
CONTEST DESIGNS AND SUMMARY OUTCOMES OF FIELD EXPERIMENTS

Design	*Optimal Design* $\{P_1, P_2, ..., P_{w^*}\}$	*Average Sales (\$)*	*SD*
Field Experiment 1			
1W	{300}	897.07	778.31
GW	{60, 60, 60, 60, 60}	1,391.00	1096.97
Field Experiment 2			
GW	{40, 40, 40, 40, 40}	358.33	181.98
RW	{70, 51, 36, 23, 13, 7}	333.33	211.05

Table 9
RESULTS OF FIELD EXPERIMENTS CONFIRM FINDINGS OF LABORATORY EXPERIMENTS

Models	*Log-Likelihood*	*Chi-Square Statistic*	p-*Value*
Field Experiment 1			
Full model	–483.0	—	—
Restriction:			
$\mu_s(GW) = \mu_s(1W)$	–486.5*	7.0	.008
Field Experiment 2			
Full model	–398.955	—	—
Restriction:			
$\mu_s(RW) = \mu_s(GW)$	–398.972	.034	.854

*Significant at the .1% level.

revenue. However, introducing unique rank-ordered prizes into multiple-winner contests does not boost sales revenues.

DISCUSSION AND CONCLUSION

Sales contests are implemented regularly in companies, but there has been little theoretical and no empirical research on the question of what should be the optimal prize structure in a sales contest. In practice, sales managers have been calling for help on how to formulate and measure the effectiveness of their sales contests (Litsikas 2006). However, as demonstrated in the 2001 Federation Study (see Center for Concept Development and Incentive Federation Inc. 2001), these "black-box" characteristics have not stopped sales managers from spending even more money on sales contests. Meanwhile, managers are left with limited means of evaluating the impact of their contests (Brodsky 2003). As questions regarding contest design increase along with corporate spending on sales contests, the need for specific direction and quality measurement methods becomes even more pressing.

This article contributes to marketing by presenting the first set of empirical studies that address these issues. The empirical studies consist of two laboratory experiments, in which participants make decisions that closely parallel the effort–reward trade-offs salespeople in sales contests typically face, and two field experiments, in which trained salespeople are randomly assigned to participate in sales contests with different prize structures. We believe that this empirical strategy is novel, and taken together, the laboratory and field experiments create an empirical test that is both internally and externally valid. The results of the laboratory and field experiments are remarkably consistent and indicate that the prize structure of a sales contest is indeed an important determinant of sales effort. Specifically, sales contests with multiple prizewinners motivate salespeople to expend more effort. However, contrary to the prediction of the existing theory of sales contests, there is no need to have unique rank-ordered prizes in contests with multiple winners; that is, keeping the prize values identical across prizewinners does not result in a significant drop in sales effort.

Our findings imply that designing the optimal sales contest is a less complex task than prescribed; managers need only to determine the optimal number of winners in the GW contest. This can be done when the number of salespeople in a contest is known and the aggregate level of risk aversion of the contestants is measured using the lottery choice procedure we used herein. Moreover, because calculating optimal prize spreads among prizes is less important than theory suggests, managers can opt more readily for noncash prizes, in which the perceived differences in monetary values of the prizes may be less precise, without worrying about a reduction in effort.

Many authors in the practitioner literature have observed that the same few salespeople tend to win all of a firm's sales contests and have posed the question of how to motivate more than just these few top performers (Bishop 1999; Farber 1994). The current research suggests that having multiple winners with identical prizes in sales contests can enhance sales performance by raising overall morale in the sales force. The overall morale in a 1W contest is likely to be low because while the winner is ecstatic, the rest of the sales force may wallow in defeat. Creating multiple prizes allows more salespeople to be "a winner," and having prizes of identical values can raise the morale of salespeople who are not the top performers, because they have the opportunity to obtain equal recognition. Managing these morale levels is important because they can affect sales effort in subsequent periods. Researchers should examine this further by building dynamics into their models to test the long-term impact of sales contests on sales effort and performance.

As cooperation and teamwork become increasingly important in sales organizations (Cummings 2007), it is also important that the prize structure of a sales contest does not endanger cooperation within the sales force. Because of the large differences in prize values across ranks, the 1W and RW contests can create a hypercompetitive and cutthroat environment. The GW contest may create the most cooperative contest environment because salespeople who know that they will be a prizewinner no longer face incremental pressure to boost sales. Therefore, they may be more willing to spend time helping other salespeople without adversely affecting their own prize earnings.

We now make several observations about the experimental design and findings of this article that, in turn, should stimulate further research. First, the results of both laboratory experiments suggest a small but consistent underinvestment in effort. This underinvestment may be explained by a recent finding that decision makers overweight the costs of up-front payments (Ho and Zhang 2008); similarly, in sales contests, sales effort can be viewed as an up-front cost because it must be expended before the contest outcomes are known. Indeed, if we allow the cost of effort parameter k to be scaled by a multiplicative term μ (with $\mu > 1$ because effort costs are overweighted), in the first laboratory experiment, the maximum-likelihood estimate of μ, assuming that effort is normally distributed with the expression for equilibrium effort in Equation 1 as the mean, is 1.074. A major finding in our research is that the degree of underinvestment is higher in the RW contest, so effort levels are not different from those of the GW contest. This result is surprising from an economics perspective because the rank-ordered prizes provide an incentive effect to achieve the highest ranks possible. However, two behavioral forces can cancel this effect. The first is a reference-dependent effect based on research by Medvec, Madey, and Gilovich (1995), who find that Olympic silver medalists appeared more unhappy than bronze medalists because the former compare their outcome with the possibility that they could have won the gold medal; if we apply this to sales contests, differentiating prizewinners through rankings may reduce effort because the utilities from prizes that are below the highest rank may be perceived as lower if salespeople compare their rankings and prizes with those at higher ranks. Conversely, in the GW contest, contestants know that there will be no such utility losses, because all prizewinners will experience the same level of achievement. The second force is that salespeople may withhold effort in a RW contest if they are not confident about achieving the top rank because the prizes associated with lower ranks are comparatively smaller. This effect will be greater if the prize spreads are convex (i.e., absolute prize values are smaller when moving from the top rank to the lower ranks) because a larger proportion of the total prize money is concentrated in the highest ranks. As we noted previously, the GW contest may motivate those who are not confident about getting first place to increase their efforts because coming in second or third will be equally rewarding. We believe that the study of psychological and social effects of sales contests is a fruitful area for further research. For example, would a contest with more winners than losers motivate contestants to put in more effort so that they would not be in the minority? If contests are repeated for the same set of salespeople, could there be pressure on winners to avoid losing subsequently, or would salespeople who did not win in previous contests be even more motivated to win to avoid embarrassment? These effects may be moderated by the degree to which salespeople make social comparisons in the organization.

Second, note that the participants in our research exhibited a moderate amount of risk aversion over the lottery stakes, which is the theoretical basis for the empirical finding that prize structures with multiple prizewinners yield greater effort. It is important to point out that though the cash prizes in the contests were the sole monetary reward for expending sales effort in the experiments, this is not the case in companies, because sales contests are used along with existing compensation schemes, such as salary and commissions. Thus, in the measurement of risk aversion using lottery stakes that span the range of monetary outcomes of the contest, the degree of risk aversion should be lower if salespeople rationally integrate their major sources of income into their utility for monetary wealth. Conversely, there is empirical evidence that decision makers engage in "narrow framing"; that is, they evaluate the attractiveness of uncertain outcomes in isolation and do not integrate other sources of wealth into their decision calculus, thus exhibiting moderate levels of risk aversion over relatively small monetary stakes (Barberis, Huang, and Thaler 2006). Similarly, salespeople's risk preferences over prize outcomes may be separately evaluated from other sources of compensation.[10] Further research should be directed to study the factors that affect the degree to which salespeople engage in narrow framing when evaluating the rewards of supplementary compensation, such as prizes from sales contests or bonuses for meeting sales targets, because the optimal design of these compensation schemes depends critically on the level of risk aversion managers assume. In sales contests, the degree of narrow framing (and the measured level of risk aversion) may depend on how motivated salespeople are by the contests, which is a function not only of the total prize money but also of how the contest is implemented, such as whether salespeople will receive frequent reminders about the contest and updates on rankings.

An important qualification of our findings is that they are based on empirical tests that attempt to control for heterogeneity in the sales productivity of salespeople; in the laboratory experiments, the participants are completely homogeneous in their sales productivity, and in the field experiments, they are relatively homogeneous in terms of their narrow range of sales experience and lack of territories. Because the typical sales force in companies is likely to be more heterogeneous in sales productivity, we discuss how our research findings can extend to these environments, particularly when it is difficult for managers to create more homogeneous contests by either developing a metric that adjusts for territorial differences when ranking the salespeople or designing separate contests for different groups of salespeople. First, as long as salespeople are risk averse, the optimal number of winners in a sales contest still exceeds one. Moreover, we expect the number of prizewinners to increase if the sales productivities of the contestants are different because managers need to motivate the "disadvantaged" salespeople to expend effort, which in turn will induce the "advantaged" salespeople to do so. Second, we expect the GW contest to perform as well as, if not better than, the RW contest, so the finding that introducing unique rank-ordered prizes into contests with multiple prizewinners does not boost effort should continue to hold. A comparison of the model that assumes homogeneity in sales productivity with tournament models in the economics literature that allow for heterogeneity (see the second section of the Web Appendix at http://www.marketingpower.com/jmrjune09) shows that though economic theory still predicts a positive marginal effect of prize spreads on effort, this effect is weaker when salespeople are heterogeneous in sales abilities or territorial advantages. In other words, the incentive effect of having prize spreads through implementing a RW contest is predicted to be even smaller if there are differences in sales productivity among the contestants.

As we mentioned previously, our empirical tests do not control for heterogeneity in risk preferences among contestants and bounded-rational behavior because these factors are present and are difficult to control in virtually any empirical setting, so the theory must survive these to be useful. The substantial dispersion in effort choices in the

[10]In both laboratory experiments, there is also no evidence that the degree of risk aversion is lower with an increase in income earned as the experiment progressed. For example, in Experiment 1, the implied aggregate risk aversion from the effort data in the first eight rounds is .498, and the same measure is .514 in the last seven rounds. The two estimates are not significantly different ($F_{1, 71}$ = .025, p = .875). This result also holds from different partitions of the data. Moreover, it can be argued that the participants in our studies engaged in narrow framing because they should also be close to risk neutrality if they evaluated the options according to their overall wealth.

laboratory experiments suggests that heterogeneity in risk preferences and bounded-rational behavior are indeed present. Although the focus of this article was to study aggregate behavior across different prize structures, future models of sales contests should incorporate these factors so that individual behavior can be explained better. Approaches that relax the Nash equilibrium solution concept, such as incorporating noisy best responses (Lim and Ho 2007) or heterogeneity in levels of strategic thinking (Camerer, Ho, and Chong 2004), are particularly promising because they can predict dispersion in effort decisions even when salespeople are homogeneous in sales productivity.

Finally, in the field experiments, we provided all the contestants with only one informational update about the rankings and dollar sales of other contestants during the contest cycle. We did not study whether the provision of feedback or how the frequency and type of feedback affect sales effort. This question is one that managers confront when implementing sales contests and remains an open area of research.

REFERENCES

Amaldoss, Wilfred and Sanjay Jain (2002), "David vs. Goliath: An Analysis of Asymmetric Mixed Strategy Games and Experimental Evidence," *Management Science*, 48 (8), 972–91.

——— and ——— (2005), "Conspicuous Consumption and Sophisticated Thinking," *Management Science*, 51 (10), 1449–66.

———, Robert Meyer, Jagmohan Raju, and Amnon Rapoport (2000), "Collaborating to Compete," *Marketing Science*, 19 (2), 105–126.

Barberis, Nicholas, Ming Huang, and Richard Thaler (2006), "Individual Preferences, Monetary Gambles and Stock Market Participation: A Case for Narrow Framing," *American Economic Review*, 96 (4), 1069–1089.

Basu, Amiya K., Lal Rajiv, V. Srinivasan, and Richard Staelin (1985), "Salesforce Compensation Plans: An Agency Theoretic Perspective," *Marketing Science*, 4 (4), 267–91.

Bishop, Bill (1999), "Sales Contests Need Reworking to Motivate All Agents," *National Underwriter/Life & Health Financial Services*, 103 (13), 7.

Brodsky, Norm (2003), "The Sales Commission Dilemma," *Inc.*, 25 (5), 49.

Bull, Clive, Andrew Schotter, and Keith Weigelt (1987), "Tournaments and Piece Rates: An Experimental Study," *Journal of Political Economy*, 95 (1), 1–33.

Camerer, Colin F., Teck H. Ho, and Juin-Kuan Chong (2004), "A Cognitive Hierarchy Model of Games," *Quarterly Journal of Economics*, 119 (3), 861–98.

Center for Concept Development and Incentive Federation Inc. (2001), "Federation Study 2001: A Study of the Incentive Merchandise and Travel Marketplace," (February), (accessed March 2, 2009), [available at http://www.incentivecentral.org/business_motivation/whitepapers/federation_study_2001.1926.html].

Cummings, Betsy (2007), "Group Dynamics," *Sales & Marketing Management*, 159 (1), 8.

Ding, Min, Jehoshua Eliashberg, Joel Huber, and Ritesh Saini (2005), "Emotional Bidders: An Analytical and Experimental Examination of Consumers' Behavior in a Priceline-Like Reverse Auction," *Management Science*, 51 (3), 352–64.

Donohue, Steve (2004), "Sales Reps Go for Olympic Gold," *Multichannel News*, 25 (21), 22.

Farber, Barry J. (1994), "Sales Managers: Get Real!" *Sales & Marketing Management*, 146 (10), 25.

Fischbacher, Urs (1999), "Z-tree. Zurich Toolbox for Readymade Economic Experiments: Experimenter's Manual," Working Paper No. 21, Institute for Empirical Research in Economics, University of Zurich.

Furniture Today (2005), "Webb Dealer Wins $10K in Drawing, 30 (9), 20.

Gaba, Anil and Ajay Kalra (1999), "Risk Behavior in Response to Quotas and Contests," *Marketing Science*, 18 (3), 417–34.

Goeree, Jacob, Charles Holt, and Thomas Palfrey (2005), "Regular Quantal Response Equilibrium," *Experimental Economics*, 8 (4), 347–67.

Green, Jerry R. and Nancy L. Stokey (1983), "A Comparison of Tournaments and Contracts," *Journal of Political Economy*, 91 (3), 349–64.

Harrison, Glenn and John List (2004), "Field Experiments," *Journal of Economic Literature*, 42 (4), 1013–1055.

Hartnett, Michael (2006), "The Seafood Seasons: With the Lenten Season Becoming More and More Important to Frozen and Fresh Seafood Programs, Winning Products and Programs Are Already Being Carefully Scrutinized," *Frozen Food Age*, 55 (1), 18–20.

Ho, Teck H., Noah Lim, and Colin F. Camerer (2006), "Modeling the Psychology of Consumers and Firms Using Behavioral Economics," *Journal of Marketing Research*, 43 (August), 307–331.

——— and Juanjuan Zhang (2008), "Designing Price Contracts for Boundedly Rational Customers: Does the Framing of the Fixed Fee Matter?" *Management Science*, 54 (4), 686–700.

Hoenig, John M. and Dennis M. Heisey (2001), "The Abuse of Power: The Pervasive Fallacy of Power Calculations for Data Analysis," *American Statistician*, 55 (1), 19–24.

Holt, Charles and Susan Laury (2002), "Risk Aversion and Incentive Effects," *American Economic Review*, 92 (5), 1644–55.

Kalra, Ajay and Mengze Shi (2001), "Designing Optimal Sales Contests: A Theoretical Perspective," *Marketing Science*, 20 (2), 170–93.

Keenan, William, Jr. (1995), "Selling with Sales Contests," *Sales & Marketing Management*, 147 (6), 35.

Krishna, Aradhna and Untku Ünver (2008), "Improving the Efficiency of Course Bidding at Business Schools: Field and Laboratory Studies," *Marketing Science*, 27 (2), 262–82.

Lazear, Edward P. and Sherwin Rosen (1981), "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89 (5), 841–64.

Lim, Noah and Teck H. Ho (2007), "Designing Price Contracts for Boundedly Rational Customers: Does the Number of Blocks Matter?" *Marketing Science*, 26 (3), 312–26.

Litsikas, Mary (2006) "Maximize Incentive ROI," *Sales & Marketing Management*, 158 (1), 16.

Luce, Duncan (1959), *Individual Choice Behavior*. New York: John Wiley & Sons.

Medvec, Victoria, Scott Madey, and Thomas Gilovich (1995), "When Less Is More: Counterfactual Thinking and Satisfaction Among Olympic Medalists," *Journal of Personality and Social Psychology*, 69 (4), 603–610.

Murphy, William H., Peter A. Dacin, and Neil M. Ford (2004), "Sales Contests Effectiveness: An Examination of Sales Contests Design Preferences on Field Sales Force," *Journal of the Academy of Marketing Science*, 32 (2), 127–43.

——— and Ravipreet S. Sohi (1995), "Salespersons' Perceptions about Sales Contests: Towards a Greater Understanding," *European Journal of Marketing*, 29 (13), 42–66.

Nalebuff, Barry J. and Joseph E. Stiglitz (1983), "Prizes and Incentives: Toward a General Theory of Compensation and Competition," *Bell Journal of Economics*, 14 (Spring), 21–43.

Orrison, Alannah, Andrew Schotter, and Keith Weigelt (2004), "Multiperson Tournaments: An Experimental Examination," *Management Science*, 50 (2), 268–79.

Schotter, Andrew and Keith Weigelt (1992), "Asymmetric Tournaments, Equal Opportunity Laws, and Affirmative Action: Some Experimental Results," *Quarterly Journal of Economics*, 107 (2), 511–39.

Weigelt, Keith, Janet Dukerich, and Andrew Schotter (1989), "Reactions to Discrimination in an Incentive Pay Compensation Scheme: A Game-Theoretic Approach," *Organizational Behavior and Human Decision Processes*, 44 (1), 26–44.