International Conference on Information Systems (ICIS)

ICIS 2009 Proceedings

Association for Information Systems *Year* 2009

Optimal Design of Crowdsourcing Contests

Nikolay Archak* Arun Sundararajan†

*New York University, narchak@stern.nyu.edu
†New York University, asundara@stern.nyu.edu

OPTIMAL DESIGN OF CROWDSOURCING CONTESTS

Completed Research Paper

Nikolay Archak
New York University,
Leonard N. Stern School of Business,
44 West 4th Street, New York, NY
narchak@stern.nyu.edu

Arun Sundararajan
New York University,
Leonard N. Stern School of Business,
44 West 4th Street, New York, NY
asundara@stern.nyu.edu

Abstract

This paper provides a game theoretic model of a crowdsourcing contest. Special attention is given to the asymptotic behavior of the contest outcome. We show that all significant outcomes of crowdsourcing contests will be determined by contestants in a small neighborhood (core) of the most efficient contestant type; in particular, the asymptotic structure of the crowdsourcing contests is distribution-free. Our formal analysis yields a managerially implementable and easily understood rule of thumb for the optimal division of the contest budget among multiple prizes. When agents are risk-neutral, the principal should optimally allocate all of its budget to the top prize even if it values multiple submissions. In contrast, if agents are sufficiently risk-averse, the principal may optimally offer more prizes than the number of submissions it desires. Our paper represents the first general analysis of the economics of crowdsourcing contests, provides a simple rule of thumb for determining the optimal prize structure for practitioners who are considering designing such contests, and also discusses the welfare implications of organizing production or R&D as a Web-based contest of this kind.

Keywords: auction theory, all-pay auction, all-pay contest, contest, contest design, crowdsourcing, electronic markets, incomplete information, Bayesian game

Introduction

Markets have been around for more than people can remember, yet the structure of market interactions is not "cast in stone" and markets evolve as other social institutions. Technological progress brought significant reduction in search (Bakos 1997) and coordination (Malone, Yates and Benjamin 1987) costs by placing information system as the foundation of the electronic market. Associated increase in diversity of the goods offered (Anderson 2004) was shown to have major impact on consumer welfare: Brynjolfsson, Hu and Smith (2003) analysis of on-line bookstores demonstrates that the increased book variety alone enhanced consumer welfare by 731 million USD to 1.03 billion USD in the year 2000, which is between 7 and 10 times as large as the consumer welfare gain from increased competition and lower prices in this market. Increase in diversity came with a price of higher informational asymmetries than in the traditional markets: old economic concepts of moral hazard (Holmstrom 1979), adverse selection (Akerlof 1970) and reputation (Kreps and Wilson 1982) became particularly salient in electronic markets.

Effects of electronic markets on consumer welfare and market efficiency are well studied in empirical and theoretical literature (Brown and Goolsbee 2002, Brynjolfsson and Smith 2000, Clemons, Hann and Hitt 2002**),** but electronic markets are still evolving, periodically giving birth to new amazing market mechanisms like Amazon's product review system and Google's sponsored search engine (Edelman, Ostrovsky and Schwarz 2007). A recent, prominent and quite controversial example of such new mechanism is "crowdsourcing". The term "crowdsourcing" was first used by Jeff Howe in a Wired magazine article (Howe 2006):

Simply defined, crowdsourcing represents the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call. This can take the form of peer-production (when the job is performed collaboratively), but is also often undertaken by sole individuals. The crucial prerequisite is the use of the open call format and the large network of potential laborers.

An important distinguishing feature of crowdsourcing, in addition to open call format and large network of contributors, is that it blurs boundaries between consumption and production creating a new consumer type: the "working consumer" (Howe 2006). The proactive nature of "working consumers" and their direct involvement in the production and innovation processes give new meaning to the "long tail" effect (Anderson 2004): while the original definition of the "long tail" referred to diversity of consumer tastes causing changes to the demand function and consequential adjustments on the supply side, crowdsourcing brings diversity of consumer experiences, skills, backgrounds and tastes straight to the supply side by allowing crowds to participate directly in the production processes. As it happened before with electronic markets, reduction in matching costs and increase in diversity (now on the production side) comes at a cost of potentially higher informational asymmetries in the market. The major goal of this paper is in understanding interplay between diversity of expertise and informational asymmetries in the crowdsourcing setting and studying its implications for the optimal design of the crowdsourcing mechanisms.

This traditional definition of crowdsourcing covers a broad range of activities that were originally performed in-house but now, due to reduction in production, search and coordination costs, can be outsourced to a crowd. Thus, crowdsourcing can involve experts, amateurs or any mix of those, the participation incentives can be monetary, intrinsic or mixed and it can be used to produce goods, services, ideas or obtain information. In this paper, we focus on a single, most popular form of crowdsourcing: **a crowdsourcing contest**. For a long time, organizations used contests to procure goods and services for which spot markets do not exist, contracts are incomplete and outcomes are not verifiable in the court. Although in the past most procurement contests were a privilege of the government and large corporations, the Web now allows almost any firm or individual to organize its own minor competition with just a few mouse clicks. Typical prizes vary from a million dollars for improving the performance of a movie recommender system[1] (Bennett and Lanning 2007) and thousands of dollars for minor pharmaceutical innovations[2] (Lakhani and Panetta 2007) to a few hundred dollars for designing a software component.[3] Most of these contests do not have an entry fee or restrict participation in any way except for legal restrictions where applicable. In fact, many

---

1 http://www.netflixprize.com

2 http://www.innocentive.com

3 http://www.topcoder.com

contests intentionally target large crowds, as emphasized by the term “crowdsourcing contest”. Very little is known about welfare properties of such contests or about their optimal design.

Although crowdsourcing contests attracted significant attention in popular press, there are only few empirical and theoretical studies of economic incentives and strategic behavior of individuals in crowdsourcing contests. Yang, Adamic and Ackerman (2008) examine behavior of users of the web-knowledge sharing market Taskn.com. They find significant variation in the expertise and productivity of the participating users: a very small **core** of successful users contributes nearly 20% of the winning solutions on the site. They also provide evidence of strategic behavior of the core participants, in particular, picking tasks with lesser expected level of competition. Motivated by this study, DiPalantino and Vojnovic (2009) provide a game-theoretic model of multiple simultaneous crowdsourcing contests in which agents select among, and subsequently compete in, a collection of crowdsourcing contests offering various rewards. They model crowdsourcing contests as all-pay auctions with incomplete information about contestant skills, the approach we also adopt in this paper. This paper complements results of DiPalantino and Vojnovic (2009): while they focus on strategic trade-off that risk-neutral contestants face when choosing between multiple simultaneous single-prize contests, we investigate strategic behavior of risk-averse contestants in a single contest with multiple prizes. We provide existence and uniqueness result for symmetric Bayes-Nash equilibria of the crowdsourcing contest with multiple prizes, investigate the asymptotic properties of this equilibrium as the number of contestants grows and characterize the asymptotically profit-maximizing contest design. Our results indicate that the optimal allocation of the prize budget among contest winners can vary quite widely depending on the risk aversion of contestants, however it is essentially distribution-free. In particular, while the principal facing risk-neutral agents should place all her budget on a single prize, the same principal facing sufficiently risk-averse contestant pool will optimally offer more prizes than the number of submissions she values, although the optimal prize amounts will exhibit exponentially decreasing marginal utility pattern.

The “procurement by contest” mechanism studied in this paper is largely new to the IS literature, and, although related, distinguished both contextually and analytically from those of basic auctions for Web-based markets (Kambil and van Heck 2004, van Heck and Ribbers 1997) and myriad reverse auctions that deal with e-sourcing (Zhong and Wu 2006) and underlie the procurement of products ranging from oil leases (Saidi and Marsden 1992) and timber harvest contracts (Athey and Levin 2001) to FCC spectrum (Cramton and Schwartz 2000) and sponsored search slots (Edelman, Ostrovsky and Schwarz 2007). These reverse auctions require signing a contract with the winning bidder and therefore may not be viable when contracts are incomplete and outcomes are not verifiable in a court. In the latter situations, a contest might be a more feasible option (Taylor 1995). The major difference of a contest from a reverse auction is that goods or services are delivered before the winner is announced, however only the contest winner or winners get paid. Asset-specificity of the efforts implies that the outside option for the goods and services produced is zero. Thus, although contests have strong auction flavor, the “bids” made by contestants are sunk and the corresponding game of the incomplete information is best described as an all-pay auction. For example, our result that a winner-take-all contest is optimal for risk-neutral agents can be compared with a well-known result that a single share structure is optimal for keyword packaging auctions with risk-neutral bidders assuming that the IHR (increasing hazard rate) condition holds (Chen, Liu and Whinston 2006): our result is obtained in a different setting (an “all-pay” auction rather than a regular auction), does not require the IHR condition on the distribution of valuations but holds only asymptotically, i.e., for contests with sufficiently large number of participants.

As we study the crowdsourcing contests, special emphasis is given to the asymptotic analysis of the equilibrium. To the best of our knowledge, this is the first paper to construct and describe distribution-free asymptotics of a large scale all-pay auction. The results we obtain replicate behavior of actual crowdsourcing contests, for instance, we show that, asymptotically, only small neighborhood of the most efficient contestant type determines the contest outcome; such neighborhood is essentially identical to the contest **core** in the Taskcn contests (Yang, Adamic and Ackerman 2008). By focusing our analysis on behavior of users in the **core,** we can provide simple characterization of the asymptotically optimal prize structure, in particular, managerially implementable and easily understood rule of thumb for the optimal division of a budget among multiple prizes.

In this paper, we extend traditional contest models by modeling a situation where the contest sponsor is interested in the quality of one **or several** best solutions. Our optimality results are most relevant for research/innovation oriented contests rather than labor tournaments like sales force contests where sales managers compete for bonuses and the firm is interested in the overall level of sales (Kalra and Shi 2001). In particular, our paper can be contrasted with the well-known paper of Moldovanu and Sela (2001) that studies optimal allocation of prizes in contests. The main difference is that Moldovanu and Sela (2001) consider a contest in which the sponsor is interested in maximizing the sum of all solutions, but in our setting the outcome of interest is the quality of the top K solutions.

While acknowledging that research/innovation contests have been extensively studied in the economic literature, we argue that our model represents a new contribution to this stream of research that has direct managerial implications for Web-based contest design while not compromising on rigor or generality in any way. Taylor (1995) investigates optimal design of "golden carrot" contests in which ex-ante identical participants compete to find the solution of the highest value to the organizer and the winner receives the specified prize. He finds that free entry in the contest is not optimal and the organizer should restrict participation by imposing an entry fee, one that extracts all participant surplus. Similar conclusions are obtained by Fullerton and McAfee (1999) in a complete information model of innovation contests with heterogeneous participants. They show that the optimal contest should include only the two "most skilled" competitors and propose using a "contestant selection all-pay auction" before the actual contest to select them. There are a few key assumptions that these prior research results rely on. For example, the result of Fullerton and McAfee (1999) requires that agents are risk-neutral; Taylor (1995) requires that it is feasible for the sponsor to appropriate part of the surplus via entry fees; Fullerton and McAfee (1999) require viability of a "contestant selection auction" which reveals the contestants' ability prior to the actual development. Although such assumptions are adequate for multi-billion dollar military R&D competitions, they seem less suitable for Web-based crowdsourcing contests where contestant anonymity and heterogeneity makes complete information unlikely and in which contestants are often budget-constrained individuals (rather than firms) who may display significant risk aversion especially with respect to initial losses such as entry fees.[4] We argue that these considerations justify the need for a very different game-theoretic model of Web contests, one that incorporates risk-averse and budget-constrained individuals, heterogeneity of expertise in the population, asymmetry of information on the Web and that can focus on investigating structure of the equilibrium when the number of participants is large.

We conclude the introduction by highlighting some of our key modeling features and salient results. Any good model of open Web contests needs to capture heterogeneity of "skills" or "expertise" across the pool of potential contestants. We use the word "expertise" in a very broad sense here. It could be, among other things, a proxy for raw abilities (talent), or for overall experience in a particular area, or for possession of some rare knowledge or skill, or a mix of these. Our modeling focus is not on the nature of this expertise but rather that it is distributed across the population. In a recent paper, Terwiesch and Xu (2008) show that, for expertise-based contests, a free entry in the equilibrium may or may not be optimal depending on the parameters of the expertise distribution. Our paper investigates a different but very related question: "if entry in an expertise-based contest is free and the contest attracts a lot of participants, what is the optimal number of prizes and prize amounts that should be awarded by the profit-maximizing sponsor?" Strikingly, the answer is distribution free and depends only on the utility function and marginal cost of effort.

Keeping in mind scenario of a Web contest, the natural modeling approach is to represent the agent's expertise as the agent's type in a game of incomplete information. We emphasize that incompleteness of information on opponents' expertise is an essential aspect of our model: a very different set of theoretical results will be obtained in a complete information setting (Siegel 2009). Note that, in contrast with some prior articles on the contest design, we do not limit ourselves to studying a single prize, instead allowing the sponsor to endogenously choose the number of prizes as well. We show that when participants are risk-averse, the optimal number of prizes can be strictly greater than the number of solutions desired by the sponsor and show that the optimal prize amounts exhibit the exponentially decreasing marginal utility pattern.

The rest of the paper is organized as follows. The Model section introduces the crowdsourcing contest model and provides the basic set of results related to existence and uniqueness of a symmetric Bayesian equilibrium of this kind of game. The next section presents the asymptotic results for the equilibrium. In the last part of the paper, we use asymptotic techniques to derive the optimal prize structure. The paper concludes with discussion of the results.

Due to space limitations, only sketches of the proofs are provided in the paper. Full proofs of all propositions are provided in the online Appendix to the paper which can be downloaded from the following URL: http://pages.stern.nyu.edu/~narchak/optimal_contest_design.pdf

---

[4] For development or research contests with prizes in the thousands or tens of thousands of dollars, running the "contestant selection auction" on the Web where the number of potential candidates is huge but willingness to bid of an average candidate is very small may also involve substantial transaction costs that can outweigh the surplus extracted from such auction.

Model

We model a continuum of problem solvers (agents) where agents have hidden type $\theta \in \left[\underline{\theta}, \overline{\theta}\right]$ that characterizes their expertise, with type $\underline{\theta} > 0$ representing the most skilled agent and type $\overline{\theta}$ representing the least skilled agent. More precisely, agents are assumed to have a linear cost of effort $C(\theta, q) = \theta q$, and thus θ is the constant marginal cost of the agent of the corresponding type. The higher the agent's type, the more effort is needed to realize a particular quality level.

N agents are randomly and independently chosen from the pool to compete on a single project. Agent types are described by the distribution function[5] Φ with the continuous probability density function ϕ. Each competing agent can submit at most one solution. The agent chooses her effort level, which determines the quality q of the submitted solution. Given the agent's type, there is a deterministic mapping from the effort to the realized quality, and the agent's choice can thus be represented in terms of their delivered quality q.[6] The contest sponsor awards monetary prizes $M_1 \ldots M_L$ to the top L (< N) submissions: the agent with the winning submission receives M_1, the first runner-up receives M_2, and so forth. If there are several submissions of the same quality, ties will be broken randomly. All other agents receive no prize, so we adopt a convention that $M_{L+1} = \ldots = M_N = 0$. The winner-takes-all contest is a particular case of this scenario when L = 1.

We model possible risk aversion of agents with respect to prizes by introducing a von Neumann-Morgenstern utility function $V(M)$. When choosing quality level q that induces a monetary lottery with prizes M_i and winning probabilities $p_i(q)$, an agent of type θ will choose to maximize the expected utility

$$\boxed{\sum_i p_i(q) V(M_i) - \theta q.}$$

We assume that V is twice continuously differentiable, strictly increasing with respect to prize M and concave. We also adopt the normalization condition V(0)=0.

For any $M_l \geq M_{l+1}$ define $\Delta V(M_l, M_{l+1}) = V(M_l) - V(M_{l+1})$, i.e., ΔV represents the gain in agent's utility from placing one spot higher.

As nonobservance of opponent skills is a natural assumption on the Web, our contest is a game of incomplete information, in which the distribution of types Φ as well as other parameters of the game are common knowledge, however agents do not know types of their competitors. Owing to the similarity of our game to an all-pay auction, we will refer to the quality q of an agent's submission as their "bid".

Our first proposition establishes that in "fair" crowdsourcing contests, **more skilled** contestants exert **higher** effort levels.

Proposition 1 (Monotonicity of Best Responses): *If prizes are "fair" ($M_1 \geq M_2 \geq \ldots \geq M_L$), the best response function of every agent in our game is non-increasing in her type.*

Why?

Proof: Consider an agent i of type θ who assumes that the other agents play strategies $b_2^i, \ldots, b_N^i$. Without loss of generality, we can take i = 1 and drop the superscript in the rest of the proof. The opponents strategies define a family of functions $P_j(b)$, j=1..N, where $P_j(b)$ represents the probability of the agent i winning the j-th prize after

---

5 Note that we disallow "atoms" in the skills distribution. This is not an essential requirement for the asymptotic analysis unless there is an atom at $\underline{\theta}$, however it simplifies presentation considerably by ensuring that the equilibrium does not involve mixed strategies.

6 While it is possible to extend the model to scenarios where solution quality is a stochastic function of the effort level, it is beyond the scope of the current paper, as our primary focus is on interplay between diversity of skills, private information and risk aversion in the crowdsourcing contests.

bidding b. For each l, define $Q_l(b)=\sum_{i=1}^{l}P_i(b)$: the probability of getting **at least** the l-th prize, where $Q_0(b)\equiv 0$. The expected utility of the agent i when bidding b can be written as $EU(\theta,b)=\sum_{l=1}^{L}P_l(b)V(M_l)-\theta b.$ Consider any $\theta_1>\theta_2$ and assume that the best response b_1 if the agent is of type θ_1 is larger than the best response b_2 of the same agent when the agent is of type θ_2. As b_2 is the best response of θ_2, it follows that

$$EU(\theta_2,b_2)=\sum_{l=1}^{L}P_l(b_2)V(M_l)-\theta_2 b_2 \geq \sum_{l=1}^{L}P_l(b_1)V(M_l)-\theta_2 b_1 = EU(\theta_2,b_1),$$

or

$$\sum_{l=1}^{L}P_l(b_2)V(M_l)-\sum_{l=1}^{L}P_l(b_1)V(M_l)\geq \theta_2(b_2-b_1)$$

Because (b_2-b_1) is negative, replacing θ_2 by a larger value of θ_1 on the right side of the equation above will not violate the inequality (the right side will go down, the left side will not be affected). But that means that type θ_1 also strictly prefers playing b_2 to playing b_1, and we have a contradiction. Q.E.D.

Proposition 1 is reassuring in that it ensures that in fair contests, rather than relying on ability and substituting away from effort, skill and effort are "complements", loosely speaking. It also shows that our game satisfies the single crossing condition for games of incomplete information (Athey 2001). One might be tempted to apply Corollary 2.1 from Theorem 2 of (Athey 2001) and conclude that there exists a pure strategy Bayes-Nash equilibrium. Unfortunately, this corollary requires the ex-post agent's payoff to be continuous with respect to bid, a condition that fails to hold for our contest. Although Section 4 of (Athey 2001) considers auctions and other games with discontinuities, her results do not directly cover all-pay auctions with multiple non-identical prizes, although it might be possible to extend them to such scenario. Nevertheless, a much simpler proof of existence follows from the symmetry of our game as one can derive an ordinary differential equation describing the symmetric equilibrium bid function. This approach is similar to analysis of auctions with risk-averse buyers (Maskin and Riley 1984). Because we are in a somewhat different setting (multiple prizes), we provide an independent proof.

Proposition 2 (Existence and Uniqueness): *The crowdsourcing contest has a unique symmetric pure strategy Bayesian Equilibrium. The equilibrium bid function* $b^*(\theta)$ *is strictly decreasing in type* θ *and satisfies the following equation*

$$b^*(\theta)=-\sum_{l=1}^{L}\int_{\theta}^{\bar{\theta}}\frac{Q_l'(s)\Delta V(M_l,M_{l+1})}{s}ds=\sum_{l=1}^{L}\frac{1}{B(l,N-l)}\int_{\theta}^{\bar{\theta}}\frac{\phi(s)\Phi(s)^{l-1}(1-\Phi(s))^{N-1-l}\Delta V(M_l,M_{l+1})}{s}ds$$

where $Q_l'(\theta)=-\frac{1}{B(l,N-l)}\phi(\theta)\Phi(\theta)^{l-1}(1-\Phi(\theta))^{N-1-l}$ *is the marginal probability that an agent of type* θ *places at or above the l-th spot and B(x,y) is the Beta function.*

Proof: The full proof is provided in online Appendix. The basic proof strategy is as follows. First, we know that the best response function of each agent is weakly monotone in her type. This fact can be used to show that, in any symmetric pure-strategy equilibrium, the bidding function must be **strictly** decreasing in θ, continuous and differentiable (almost everywhere). The first-order condition for the profit-maximization problem together with the boundary condition that type $\bar{\theta}$ bids zero gives integral representation for the bidding function. Uniqueness follows from the fact that any symmetric equilibrium bid function must satisfy the first-order condition and the boundary condition. Q.E.D.

Although the bid equation from Proposition 2 can be used to reconstruct the bid function numerically, it does not say much about strategic behavior of the agents. It is unclear, for example, how this behavior would change if a new prize is added, the number of competitors grows or the distribution of skills changes. Fortunately, the Envelope

Theorem provides some insight here. Note that

$$\frac{d}{d\theta}EU(\theta,b^*(\theta)) = \frac{\partial}{\partial\theta}EU(\theta,b^*(\theta)) + b^*{}'(\theta)\frac{\partial}{\partial b}EU(\theta,b^*(\theta)).$$

The second term is zero; from the proof of previous Proposition, $b^*(\theta)$ solves the first-order condition for the expected utility maximization. It follows that

$$\frac{d}{d\theta}EU(\theta,b^*(\theta)) = \frac{\partial}{\partial\theta}EU(\theta,b^*(\theta)) = -b^*(\theta).$$

Together with the boundary condition that type $\overline{\theta}$ bids zero and therefore has zero expected utility, this yields

$$EU(\theta,b^*(\theta)) = \int_{\theta}^{\overline{\theta}} b^*(s)ds.$$

This in turn implies that $\forall\theta_1,\theta_2 \in [\underline{\theta},\overline{\theta}] : \theta_1 < \theta_2 \Rightarrow EU(\theta_1,b^*(\theta_1)) > EU(\theta_2,b^*(\theta_2))$. In particular, we can see that in the symmetric equilibrium agents with more expertise get more surplus than less skilled ones, which leads to our next theorem:

Proposition 3 (Surplus Monotonicity): *Crowdsourcing contests are distributionally efficient: contestants with greater ability receive a higher expected surplus in the equilibrium.*

An intuitive foreshadowing of our asymptotic result is that, as N grows, competition intensifies, and $EU(\theta,b^*(\theta)) \to 0$, where the convergence will be shown to be uniform on $\theta \in [\theta_0,\overline{\theta}]$ for any $\theta_0 > \underline{\theta}$. Therefore, the asymptotic behavior of the bid function and other interesting statistics can be deduced from the following approximation

$$EU(\theta,b^*(\theta)) = \sum_{l=1}^{L} P_l(\theta)V(M_l) - \theta b^*(\theta) \approx 0.$$

In particular, the approximation above will later be used to derive the asymptotically optimal prize structure for a profit-maximizing sponsor interested in the expected quality of the top K solutions. We conclude this section with a couple of related definitions.

Definition 1: *For any k = 1..N, define the k-th place quality $Q^B_{k,L}$ as the expected quality of the k-th best solution, $Q^B_{k,L} = E\{b^*(\theta_k)\}$, where θ_k is the k-th order statistic (the k-th smallest value) for a random sample of size N from Φ. Define Q^A_L as the expected average quality, $Q^A_L = E\{b^*(\theta)\}$, where θ is randomly drawn from Φ.*

Crowdsourcing with a Large Pool of Contestants: Asymptotic Results

The bulk of our remaining analysis examine contest outcomes and optimal design of contests when N is large. We start by characterizing the asymptotic behavior of the bid function as the number of agents grows. For future analytical convenience, we also study the inverse bid function.

Definition 2: *The inverse bid function $I^*(b):[0,\infty) \to \left[\underline{\theta},\overline{\theta}\right]$ is defined by $I^*(b) = \inf\left\{\theta \mid b^*(\theta) \le b\right\}$.*

Note that, by construction and strict monotonicity of b*, $I^*(b^*(\theta)) \equiv \theta.$

The following Lemma establishes some important general properties of the inverse bid function.

Proposition 4 (Asymptotic behavior of the bid and the inverse bid functions in the symmetric equilibrium):

If L is fixed and $N \to \infty$, then

1. *The bid function of the most skilled type* $\underline{\theta}$ *converges to the socially optimal level of effort* $b^*(\underline{\theta}) \to \frac{V(M_1)}{\underline{\theta}}$.
2. *The bid function of any other type converges to the zero level of effort:* $b^*(\theta) \to 0$ *uniformly on* $\theta \in [\theta_0, \bar{\theta}]$ *for any* $\theta_0 > \underline{\theta}$.
3. *The inverse bid function* $I^*(b) \to \underline{\theta}$ *uniformly on* $b \in [b_1, b_2]$ *for any* $b_1, b_2 \in \left(0, \frac{V(M_1)}{\underline{\theta}}\right)$.
4. *The product* $\Phi(I^*(b))N \to \alpha(b)$ *uniformly on* $b \in [b_1, b_2]$ *for any* $b_1, b_2 \in \left(0, \frac{V(M_1)}{\underline{\theta}}\right)$, *where* $\alpha(b)$ *is defined as a solution of the equation* $e^{-\alpha(b)} \sum_{l=1}^{L} V(M_l) \frac{\alpha(b)^{l-1}}{(l-1)!} = \underline{\theta} b$.

All convergence results are uniform with respect to prizes $\underline{M} \le M_1 \le \ldots \le M_L \le \overline{M}$ *for any* $0 < \underline{M} \le \overline{M}$.

Proof: A complete proof is provided in online Appendix, a sketch of the proof is given below. The expected surplus of the type θ can be written as

$$EU(\theta, b^*(\theta)) = \sum_{l=1}^{L} P_l(\theta) V(M_l) - \theta b^*(\theta) \le \sum_{l=1}^{L} P_l(\theta) V(M_l).$$

For any $l \le L$, $P_l(\theta) \to 0$ uniformly on $\theta \in [\theta_0, \bar{\theta}]$ for any $\theta_0 > \underline{\theta}$, i.e., the probability of winning any non-zero prize for type θ converges to zero unless θ represents the most skilled agent type possible. Thus,

$0 \le EU(\theta, b^*(\theta)) = \sum_{l=1}^{L} P_l(\theta) V(M_l) - \theta b^*(\theta) \to 0$ uniformly, so it must be that both $EU(\theta, b^*(\theta))$ and $b^*(\theta)$ converge to zero uniformly.

Now, using this fact one can also prove that $EU(\underline{\theta}, b^*(\underline{\theta})) \to 0$ because $EU(\underline{\theta}, b^*(\underline{\theta})) = \int_{\underline{\theta}}^{\bar{\theta}} b^*(s) ds$ and the integral on the right side converges to zero because of the uniform convergence of the bid function. It remains to note that the best type always wins, so $EU(\underline{\theta}, b^*(\underline{\theta})) = V(M_1) - \underline{\theta} b^*(\underline{\theta}) \to 0$ or $b^*(\underline{\theta}) \to \frac{V(M_1)}{\underline{\theta}}$.

Convergence of $I^*(b) \to \underline{\theta}$ for $b < \frac{V(M_1)}{\underline{\theta}}$ follows from the fact that as any type different from $\underline{\theta}$ will sooner or later bid almost zero.

The final result on convergence of $\Phi(I^*(b))N$ uses the following asymptotic equivalences:

$$0 \leftarrow EU(I^*(b), b) = \sum_{j=1}^{L} P_j(I^*(b)) V(M_j) - bI^*(b) \approx$$

$$\sum_{j=1}^{L} \frac{1}{(j-1)!} \left\{ \left(1 - \frac{1}{\frac{1}{I^*(b)}} \right)^{\frac{1}{I^*(b)}} \right\}^{NI^*(b)} (NI^*(b))^{j-1} V(M_j) - bI^*(b) \approx \sum_{j=1}^{L} \frac{V(M_j)}{(j-1)!} e^{-NI^*(b)} (NI^*(b))^{j-1} - b\underline{\theta}.$$

Q.E.D.

The most important result of the previous Lemma says that, asymptotically, $I^*(b) \approx \Phi^{-1}\left(\frac{\alpha(b)}{N}\right)$ where $\alpha(b)$ solves $e^{-\alpha(b)} \sum_{l=1}^{L} V(M_l) \frac{\alpha(b)^{l-1}}{(l-1)!} = \underline{\theta} b$. One can see that bids that are significantly different from zero will be produced by agents with types "close" to the most efficient type $\underline{\theta}$ and this convergence is of order O(1/N). In other words, *all significant outcomes will be determined by contestants in a small neighborhood (**core**) of the most efficient possible contestant* and this neighborhood shrinks with speed of 1/N as the number of agents grows. This is a critical insight into the design of crowdsourcing contests because it suggests that the driver of better outcomes as one's pool of contestants grows on account of the Web is entirely provided by the greater likelihood of having access to a small group of highly qualified individuals competing against each other.

Moreover, the size of the **core** neighborhood is determined by the equation for $\alpha(b)$ which involves **only** the prize amounts and the utility function of the most efficient agent type. To give the readers some feeling of the shape of the first-order term as a function of b, we have plotted the function $\alpha(b)$ for different number of identical prizes ($V(M_1) = ... = V(M_L)$), $\underline{\theta} = 1$ on Figure 1.

In this particular example, $b^*(\underline{\theta}) \to \frac{V(M_1)}{\underline{\theta}} = 1$ so 1.0 is the upper support point of the bid distribution as can be seen from the figure. Note that the $\alpha(b)$ curve becomes steeper as the number of prizes grows reflecting that additional prizes give higher incentives for less skilled types to deliver higher quality. *In other words, contests with more prizes will have larger core neighborhood.*

The way to read the curves on Figure 1 is as follows. Assume, for example, that we have a thousand contestants (N = 1,000), which is not an unreasonable number for a Web contest. In the symmetric equilibrium for a single prize contest (L = 1), submission with quality of 0.9 (90% of the quality produced by the most experienced contestant) corresponds to $\alpha(b) \approx 0.1$. The type θ that produces this quality in the equilibrium will be located at

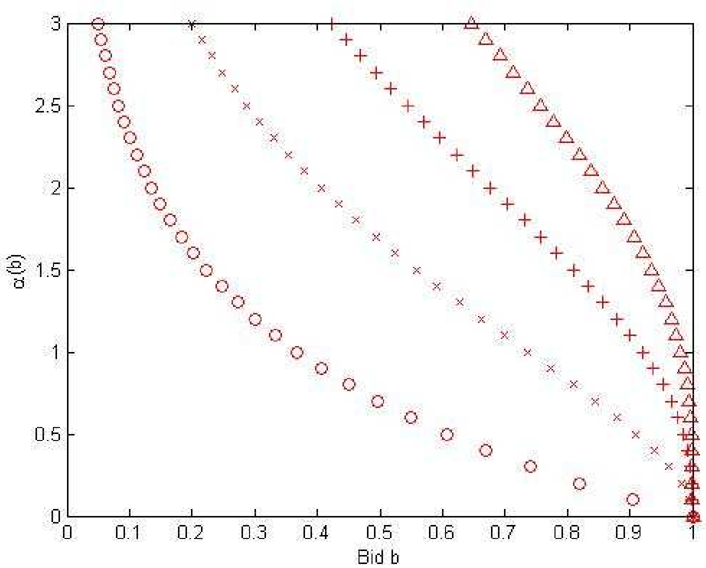

Figure 1: Asymptotic behavior of the inverse bid function I*(b) for different number of prizes (L = 1 - circle, L = 2 - cross, L = 3 - plus, L = 4 - triangle)

$\frac{\alpha(b)}{N} \approx 10^{-5}$ quantile of the distribution of expertise. Thus, the probability that no contestants will produce a solution of at least this quality is $\left(1-10^{-5}\right)^{1000} \approx e^{-0.1} \approx 0.9$ At the same time, if the contest also awards the second place prize (L=2), we have $\alpha(b) \approx 0.5$ and the probability that no contestants will produce a solution of at least this quality is

$e^{-0.5} \approx 0.6$. Thus, there is a four-fold increase in the probability of getting a good outcome that is induced simply by introducing a second prize. This is a simple example but it produces very useful insight on the value of an additional prize to the contest sponsor.

We further extend this idea by showing that the first-order approximation allows us to calculate the asymptotic expected quality of the k-th submission for any fixed k: if one fixes k and lets N grow, the type of the agent producing the k-th submission will start approaching $\underline{\theta}$, thus behaving as if $I^*(b) = \Phi^{-1}\left(\frac{\alpha(b)}{N}\right)$. This insight is formalized by the next Proposition. In order to state the results in a convenient form, we need to introduce the inverse of the function $\alpha(b)$ which will play important role in the following analysis.

Definition 3: *Define the core kernel function* $\mathrm{H}(\alpha):[0,\infty) \to \left(0, \frac{V(M_1)}{\underline{\theta}}\right]$ *as the unique solution of the following equation* $\mathrm{H}(\alpha(b)) \equiv b$, *i.e.,* $\mathrm{H}(\alpha) = \frac{1}{\underline{\theta}} e^{-\alpha} \sum_{l=1}^{L} V(M_l) \frac{\alpha^{l-1}}{(l-1)!}$.

Note that we call function H the *core kernel function* because it describes the bidding behavior of all agents in the contest core. Given some value α representing distance of an agent from the type $\underline{\theta}$, function $\mathrm{H}(\alpha)$ produces the bid value b such that, asymptotically, type $\Phi^{-1}\left(\frac{\alpha}{N}\right)$ will bid exactly b. Yet another way to think about function $\mathrm{H}(\alpha)$ is to imagine the core agent facing the following "asymptotic lottery": win the first prize with probability $e^{-\alpha}$, win the second prize with probability $\alpha e^{-\alpha}$, ..., win the L-th prize with probability $e^{-\alpha} \frac{\alpha^{L-1}}{(L-1)!}$. In such a scenario, the agent will be exactly indifferent between getting the lottery and delivering a solution of quality $\mathrm{H}(\alpha)$ and not participating. This lottery analogy is not occasional: it is exactly this lottery that the type $\Phi^{-1}\left(\frac{\alpha}{N}\right)$ will face asymptotically. The next Lemma develops this idea by showing how the function $\mathrm{H}(\alpha)$ determines the asymptotic game outcome.

Proposition 5 (Asymptotic outcomes of the symmetric equilibrium):

If L is fixed and $N \to \infty$, then convergence of the expected quality of the k-th best solution is given by the following expression

$$Q_{k,L}^{B} \to \int_0^\infty e^{-\alpha} \frac{\alpha^{k-1}}{(k-1)!} \mathrm{H}(\alpha) d\alpha = \frac{1}{\underline{\theta}} \sum_{j=1}^{L} V(M_j) \frac{1}{2^{k+j-1}} \binom{k+j-2}{k-1}.$$

Also, the expected total quality of all solutions converges as $NQ_L^A \to \int_0^\infty \mathrm{H}(\alpha) d\alpha = \frac{1}{\underline{\theta}} \sum_{j=1}^{L} V(M_j)$ *and the expected total utility of all contestants converges to zero:* $NEU(\theta, b^*(\theta)) \to 0$.

All convergence results are uniform with respect to prizes $\underline{M} \le M_1 \le ... \le M_L \le \overline{M}$ *for any* $0 < \underline{M} \le \overline{M}$.

Proof: Strict proof is provided in online Appendix, informal argument is as follows. The probability density function of the k-th best agent type is the probability density function of the k-th smallest element out of the sample of size N, which is given by $\frac{1}{B(l, N-k+1)} \phi(\theta) \Phi(\theta)^{k-1} (1-\Phi(\theta))^{N-1-k}$. This expression is asymptotically equivalent to $\frac{N^k}{(k-1)!} \phi(\theta) \Phi(\theta)^{k-1} (1-\Phi(\theta))^{N-1-k}$ (note that we are interested in the area where $\Phi(\theta)$ is small). Perform substitution $\theta = I^*(b) = \Phi^{-1}\left(\frac{\alpha(b)}{N}\right)$. The density will transform to $\frac{\alpha(b)'}{(k-1)!}\alpha(b)^{k-1}\left(1-\frac{\alpha(b)}{N}\right)^{N-1-k} \approx \frac{\alpha(b)'}{(k-1)!}\alpha(b)^{k-1}e^{-\alpha(b)}$

Expression for $Q^B_{k,L}$ can be obtained by integrating $\mathrm{H}(\alpha(b)) \equiv b$ against this density (and using α as the integration variable so that $\alpha'(b)$ disappears).

To prove that $\int_0^\infty e^{-\alpha}\frac{\alpha^{k-1}}{(k-1)!}\mathrm{H}(\alpha)d\alpha = \frac{1}{\underline{\theta}}\sum_{j=1}^{L} V(M_j)\frac{1}{2^{k+j-1}}\binom{k+j-2}{k-1}$ just substitute $\mathrm{H}(\alpha) = \frac{1}{\underline{\theta}}e^{-\alpha}\sum_{l=1}^{L} V(M_l)\frac{\alpha^{l-1}}{(l-1)!}$ and perform integration. Q.E.D.

Optimal Prize Structure for Crowdsourcing Contests

We now proceed to using our asymptotic results to derive the optimal prize structure. Consider a risk-neutral budget constrained tournament sponsor whose utility function is $U(Q_1,...,Q_K,M) = \sum_{k=1}^{K}\mu_k Q_k - M$, where K >= 1 represents the number of solutions the sponsor is interested in, $\mu_k > 0$ is the constant marginal utility of the quality of the k-th best solution ($\mu_1 \ge \mu_2 \ge ...\mu_k$) and M is the amount of money spent. A utility-maximizing sponsor should choose the number of prizes L, the total prize pot M and the allocation of prizes $M_1 + M_2 + ...M_L = M$ which maximizes the sponsor's utility. We will only consider ``fair'' allocations of prizes: $M_1 \ge M_2 \ge ...M_L \ge m$, where m is the minimum allowed payment prize.[7] Define the choice set

$$C = \left\{(L, M_1, M_2, ...) \mid M_1 \ge M_2 \ge ...M_L \ge m, M_l = 0 \text{ when } l > L\right\} \subset \aleph \times \Re^\infty .$$

The sponsor's maximization problem is

$$\max_C \sum_{k=1}^{K}\mu_k Q^B_{k,L}(M_1,...,M_L) - \sum_{l=1}^{L} M_l .$$

To make our definition valid in the sense that the maximum exists, we will require that either the total prize budget $\sum_{l=1}^{L} M_l$ is fixed or $\lim_{M\to\infty} V'(M) = 0$. In this latter case, the maximum utility an agent can get is V(M$_1$) and therefore the maximum utility the sponsor can achieve is $\sum_{k=1}^{K}\mu_k\frac{V(M_1)}{\underline{\theta}}$, which asymptotically grows slower than $\sum_{l=1}^{L} M_l$ so there is an upper bound on the size of optimal prizes. Denote this upper bound as $\overline{M}$. Because there is a minimum prize value, there is also a bound on the optimal number of prizes $\overline{L}$: $\overline{L}m \le \sum_{k=1}^{K}\mu_k\frac{V\left(\overline{M}\right)}{\underline{\theta}}$. Existence of the maximum for a fixed L immediately follows from continuity of the target function with respect to prize amounts and compactness of its domain when L is bounded. To find the actual maximum, one only needs to check all L from 0 up to $\overline{L}$.

As the number of participants grows, the expected k-th place quality $Q^B_{k,L}$ converges to its asymptotic limit $\int_0^\infty e^{-\alpha}\frac{\alpha^{k-1}}{(k-1)!}\mathrm{H}(\alpha)d\alpha = \frac{1}{\underline{\theta}}\sum_{j=1}^{L} V(M_j)\frac{1}{2^{k+j-1}}\binom{k+j-2}{k-1}$. One may also consider asymptotic version of the sponsor's

[7] The assumption of a minimum prize amount simplifies the analysis by ruling out optimal pricing schemes with “infinite” (function of N) number of prizes and infinitely decreasing prize amounts. Whereas mathematicians might be interested in infinitely decreasing prize amounts, economists are not, and so the assumption is economically justified. One can also think of *m* as the minimum payment justifying the transfer of intellectual property as required by law.

maximization problem given by

$$\max_C \sum_{k=1}^{K}\mu_k\left[\frac{1}{\underline{\theta}}\sum_{j=1}^{L}V(M_j)\frac{1}{2^{k+j-1}}\binom{k+j-2}{k-1}\right]-\sum_{l=1}^{L}M_l=\max_C \sum_{j=1}^{L}\left\{\frac{1}{\underline{\theta}}\left[\sum_{k=1}^{K}\mu_k\frac{1}{2^{k+j-1}}\binom{k+j-2}{k-1}\right]V(M_j)-M_j\right\}$$

Our next proposition establishes that for sufficiently large N solving this approximation problem *leads to outcomes that are sufficiently close to the finite-N optimal outcomes so that one can use these asymptotic results to guide the actual contest design.*

Proposition 6 (Optimal prize structure for the asymptotic optimization problem is asymptotically optimal):

Assume that the optimal solution of the asymptotic optimization problem exists and given by $(L^*,M_1^*,M_2^*...)$. *For each* $\varepsilon>0$ *there exists* $N^*>0$ *such that for any* $N>N^*$, *any* $(L,M_1,M_2,...)\in C$ *:*

$$\sum_{k=1}^{K}\mu_k Q_{k,L}^{B}(M_1,...,M_L)-\sum_{l=1}^{L}M_l<\sum_{k=1}^{K}\mu_k Q_{k,L}^{B}(M_1^*,...,M_{L^*}^*)-\sum_{l=1}^{L^*}M_l^*+\varepsilon$$

Proof: See online Appendix.

Fortunately, the asymptotic optimization problem is easy to solve as it achieves nice separation of prize amounts: the optimization condition for any prize amount M_j is independent of M_i for $i\neq j$ if one ignores the prize monotonicity constraints. Define $W_{kj}=\frac{1}{2^{k+j-1}}\binom{k+j-2}{k-1}$ and $X_j=\sum_{k=1}^{K}\mu_k W_{kj}$. Also, define the inverse marginal utility function $J:(0,V'(0)]\to[0,\infty)$ by $J(V'(m))\equiv m$. The asymptotic optimization problem transforms to $\max_C \sum_{j=1}^{L}\left\{\frac{X_j}{\underline{\theta}}V(M_j)-M_j\right\}$. Next, relax the monotonicity constraint for prizes ($M_j\geq M_{j+1}$). and solve for the optimal prize amount M_j. There are three possible cases:

- The first order condition (FOC) with respect to M_j is $\frac{1}{\underline{\theta}}X_jV'(M_j)=1$ gives the solution $M_j^*=J\left(\underline{\theta}X_j^{-1}\right)$ that is larger than the minimum prize amount m. Note that by concavity of V that will be the optimal solution.
- The FOC solution is smaller than the minimum prize amount m but the minimum prize amount is feasible, i.e., $\frac{1}{\underline{\theta}}X_jV(m)\geq m$. In this case, the optimal prize amount will be $M_j^*=m$.
- The minimum prize amount is not feasible, i.e., $\frac{1}{\underline{\theta}}X_jV(m)\leq m$. In this case the optimal prize amount will be zero.

We can summarize the optimal solution to the relaxed problem by $M_j^*=\max\left(m,J\left(\underline{\theta}\left[\sum_{k=1}^{K}\mu_kW_{kj}\right]^{-1}\right)\right)$ as long as this value satisfies $\frac{1}{\underline{\theta}}X_jV(M_j^*)\leq M_j^*$ and zero otherwise. It remains to show that this solution will not violate the fairness constraint $M_1\geq M_2\geq ...M_L$. As function V is concave, i.e., its derivative is non-increasing, the monotonicity of prize amounts will be immediate if values X_j are monotone and this is guaranteed by the following proposition.

Proposition 7 (Monotonicity Constraint is Satisfied)*: Sequence* $X_j = \sum_{k=1}^{K} \mu_k W_{kj}$ *is decreasing in j if* $\mu_1 \geq \mu_2 \geq ... \mu_k$*, therefore* $M_j^* = \max\left(m, J\left(\underline{\theta}\left[\sum_{k=1}^{K} \mu_k W_{kj}\right]^{-1}\right)\right)$ *is a non-increasing function of j.*

Proof: See online Appendix.

We can now describe the optimal solution of the asymptotic optimization problem.

Proposition 8 (Optimal Solution of the Asymptotic Optimization Problem)

The optimal solution of the asymptotic optimization problem will award $L^* = \inf\left\{L \mid \sum_{k=1}^{K} \mu_k W_{kL} \geq \frac{m\underline{\theta}}{V(m)}\right\}$ *prizes and the optimal prize amounts will be given by* $M_j^* = \max\left(m, J\left(\underline{\theta}\left[\sum_{k=1}^{K} \mu_k W_{kj}\right]^{-1}\right)\right)$ *where* J *is the inverse of the marginal utility of money.*

Proof: Immediate corollary of Proposition 7. Q.E.D.

We can analyze a number of examples based on this main design result.

Example 1: *When contestants are risk-neutral, the optimal design of the crowdsourcing contest with a fixed budget involves placing all of the budget on the top prize even if the sponsor values multiple submissions.* While the risk-neutral scenario is not directly covered by Proposition 8, one can approximate it by considering a sequence of optimization problems O_ε with $V_\varepsilon(M) = \frac{M^{1-\varepsilon}}{1-\varepsilon}$. Easy to see that the fractional prize holdings for the optimal prize allocation for the optimization problems O_ε converge to the optimal prize allocation for the risk-neutral scenario V(M) = M *with fixed prize budget*. From Proposition 8, we know that, in the optimal solution of the optimization problem O_ε, all non-zero prizes will be given by $M_j^*(\varepsilon) = \max\left(m, X_j^{\varepsilon} \underline{\theta}^{-\frac{1}{\varepsilon}}\right)$, therefore, when ε goes to zero, $\frac{M_j^*(\varepsilon)}{M_{j+1}^*(\varepsilon)} \geq \min\left(\frac{X_j^{\frac{1}{\varepsilon}} \underline{\theta}^{-\frac{1}{\varepsilon}}}{X_{j+1}^{\frac{1}{\varepsilon}} \underline{\theta}^{-\frac{1}{\varepsilon}}}, \frac{m^{\frac{1}{\varepsilon}}}{m}\right) = \min\left(\left[\frac{X_j}{X_{j+1}}\right]^{\frac{1}{\varepsilon}}, m^{\frac{1}{\varepsilon}-1}\right) \to \infty$. In other words, as ε goes to zero, the share of the optimal holdings of two consecutive prizes goes to infinity, i.e., all of the prize budget is shifted to the first place prize.

Example 2: *When participants are sufficiently risk-averse, the number of prizes is often more than the number of desired submissions.* Assume the sponsor is interested only in the best submission, $\mu_1 = 5$. The minimum prize amount is m = 1, agents have utility V(M) = log(M + 1) and the distribution of θ has support [0.5,1.5]. First, $\frac{m\underline{\theta}}{V(m)}$ is equal to 0.5/log(2)=0.721... Consulting with first row of the Table 1, we see that $\mu_1 W_{11}$ and $\mu_1 W_{12}$ are greater than 0.721, however $\mu_1 W_{13}$ = 0.625 < 0.721, therefore the optimal number of prizes is two. Although the sponsor is only interested in the value of the best submission, the optimal design will involve the first and the second place prizes. Moreover, the optimal number of prizes grows with sponsor's valuation μ_1. For example, if $\mu_1 = 10$, it would be optimal to use the third place prize as well.

Example 3: *There is an elegant alignment between the relative importance of each submission to the sponsor and the relative magnitude of each prize*. An interesting insight is obtained by examining a candidate set of W_{kj} values for small k,j=1..6 which is given in Table 1. These document the marginal contribution of the n-th prize to the k-th outcome (k-th place quality). What is notable is (a) that while each prize contributes to quality on all the outcomes, the contribution to of the n-th prize to the n-th outcome is (weakly) the highest. *Thus, the relative importance of the ranked contributions is aligned with the relative magnitude of the prize amounts.*

Table 1. Asymptotic weighting matrix W_{kj}

	1st Prize	2nd Prize	3rd Prize	4th Prize	5th Prize	6th Prize
1st Quality	0.5	0.25	0.125	0.0625	0.03125	0.015625
2nd Quality	0.25	0.25	0.1875	0.125	0.078125	0.046875
3rd Quality	0.125	0.1875	0.1875	0.15625	0.11719	0.082031
4th Quality	0.0625	0.125	0.15625	0.15625	0.13672	0.10938
5th Quality	0.03125	0.078125	0.11719	0.13672	0.13672	0.12305
6th Quality	0.015625	0.046875	0.082031	0.10938	0.12305	0.12305

Several important notes can be made about the results we derived above. At first, while the optimal prize amounts as well as the optimal number of prizes depend on the utility function of contestants as well as the contest sponsor, they, **asymptotically,** do not depend on the distribution of expertise in the population. At second, the previous example shows that the optimal number of prizes with risk averse contestants can sometimes be larger than the number of solutions desired by the sponsor. At third, the optimal prize amounts are directly related to the weight values W_{kj} which are exponentially decreasing as a function of the prize number j: $W_{kj} = \frac{1}{2^{k+j-1}}\binom{k+j-2}{k-1} = O\left(\frac{j^{k-1}}{2^{j-1}}\right)$. In other words, having everything else fixed, starting at some point every new prize will have twice less contribution to the contest outcome than the prize immediately above it. That will translate to the first order conditions, so one will have $V'(M_j) = \frac{\theta}{X_j} \approx \frac{\theta}{2X_{j+1}} = \frac{1}{2}V'(M_{j+1})$ as long as the minimum prize constraint is not binding. What this equation says is that the optimal prizes awarded should follow the following rule of thumb: each new prize should have approximately twice higher marginal utility than the prize immediately above it.[8] Once we reach a point at which it is not possible to satisfy this relationship, no new prizes should be awarded. Finally, one may compare the optimal prize scheme of the profit maximizing sponsor with the optimal solution of the social planner that maximizes the expectation of the total social surplus. As we know that, asymptotically, the principal expects to extract the whole surplus, it follows that the optimal prize scheme of the profit maximizing sponsor is identical to the optimal prize scheme of the social planner.

Conclusion

This paper presents a theoretical model of an all-pay contest with heterogeneous risk-averse wealth-constrained contestants and multiple prizes. As the model was driven primarily by the new phenomenon of Web tournaments that are known to attract significant participation, major attention was given to the asymptotic properties of the equilibrium. Considering asymptotics simplifies the model significantly and allows us to obtain a number of interesting results.

Consistent with prior empirical studies of crowdsourcing contests (Yang, Adamic and Ackerman 2008), we obtain that *all significant contest outcomes will be determined by contestants in a small neighborhood (**core**) of the most*

[8] Note that the marginal utility goes **down** with the prize amount: the higher is the prize amount, the less one is willing to do for an extra dollar.

efficient possible contestant and this neighborhood shrinks with speed of 1/N as the number of agents grows. In particular, asymptotically, only the support of the type distribution and the marginal cost at the support determine the contest outcome, therefore limiting behavior of the game has a particularly simple structure that can be summarized by a distribution free *core kernel function* H.

Additionally, we employ our asymptotic results to define the asymptotic optimization problem for the profit-maximizing tournament sponsor and show that the optimal solution of the finite-sample allocation problem converges to the asymptotic one as the number of participants grows. We derive the optimal solution of the asymptotic optimization problem. When the contestants are sufficiently risk-averse, the firm may optimally offer more prizes than there are the desired submissions, thus awarding prizes even to submissions it does not eventually want. Moreover, the optimal prize amounts exhibit the exponentially decreasing marginal utility pattern: each new prize should have approximately twice higher marginal utility than the prize immediately above it. This is a simple but very useful rule of thumb that can be used in the design of real crowdsourcing contests.

We also show that all-pay contests with incomplete information have a number of interesting welfare properties. First, the expected overall contestant surplus (the expected total utility of all contestants) is asymptotically zero. That is, asymptotically, the sum of the expected utilities of the eventual winners (the highest K types that receive prizes) has exactly the same magnitude as but the opposite sign of the sum of the expected utilities of the remaining participants. As one might expect, there is a welfare loss from using these contests relative to the first-best outcome under which only the best K participants produce at the socially optimal levels of quality. However, we also show that, conditional on using an all-pay contest of the kind we model, the menu of prizes that maximizes the sponsor's profits also, asymptotically, maximizes total welfare. Thus, the efficiency news is not all bad.

References

Akerlof, G. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics* (84), 1970, pp. 488-500.

Anderson, C. "The Long Tail," in *Wired Magazine,* 2004.

Athey, S. "Single Crossing Properties and the Existence of Pure Strategy Equilibria in Games of Incomplete Information," *Econometrica* (69:4), 2001, pp. 861-889.

Athey, S. and Levin, J. "Information and Competition in U.S. Forest Service Timber Auctions," *Journal of Political Economy* (109:2), 2001, pp. 375-417.

Bakos, Y. "Reducing Buyer Search Costs: Implications for Electronic Marketplaces," *Management Science* (43), 1997, pp. 1676-1692.

Bennett, J., and Lanning, S. "The Netflix Prize," in *Proceedings of KDD Cup and Workshop*, San Jose, California, August 2007.

Brown, J.R., and Goolsbee, A. "Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry," *Journal of Political Economy* (110), 2002, pp. 481-507.

Brynjolfsson, E., Hu, Y.J., and Smith, M.D. "Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers," *Management Science* (49), 2003, pp. 1580-1596.

Brynjolfsson, E., and Smith, M.D. "Frictionless Commerce? A Comparison of Internet and Conventional Retailers," *Management Science* (46), 2000, pp. 563-585.

Chen, J., Liu, D., and Whinston, A. "Resource Packaging in Keyword Auctions," *Proceedings of the 27th International Conference on Information Systems,* Milwaukee, WI, 2006.

Clemons, E.K., Hann, I.H., and Hitt, L.M. "Price Dispersion and Differentiation in Online Travel: an Empirical Investigation," *Management Science* (48), 2002, pp. 534-549.

Cramton, P. and Schwartz, J. "Collusive Bidding: Lessons from the FCC Spectrum Auctions," *Journal of Regulatory Economics* (17:3), 2000, pp. 229-252.

DiPalantino, D., and Vojnovic, M. "Crowdsourcing and All-Pay Auctions," *Proceedings of the 10th ACM International Conference on Electronic Commerce,* 2009, pp. 119-128.

Edelman, B., Ostrovsky, M., and Schwarz, M. "Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords," *The American Economic Review* (97:18), 2007, pp. 242-259.

Fullerton, R., and McAfee, P. "Auctioning Entry into Tournaments," *Journal of Political Economy* (107:3), 1999, pp. 573-605.

Holmstrom, B. "Moral Hazard and Observability," *The Bell Journal of Economics* (10), 1979, pp. 74-91.

Howe, J. "The Rise of Crowdsourcing," in *Wired Magazine,* June 14th 2006.

Kalra, A. and Mengze, S. "Designing Optimal Sales Contests: A Theoretical Perspective," *Marketing Science* (20:2), 2001, pp. 170-193.

Kambil, A. and van Heck, E. "Introduction to 'Innovative Auction Markets'" Special Issue. *Electronic Markets* (14:3), 2004, pp. 166-169.

Kreps, D.M. and Wilson, R. "Reputation and Imperfect Information" *Journal of Economic Theory* (27), 1982, pp. 253-279.

Lakhani, K. and Panetta J. "The Principles of Distributed Innovation," in *Innovations: Technology, Governance, Globalization* (2:3), 2007, pp. 97-112.

Malone, T.W., Yates, J., and Benjamin, R.I. "Electronic Markets and Electronic Hierarchies," in *Communications of the ACM* (30), 1987, pp. 484-497.

Maskin, E. and Riley, J. "Optimal Auctions with Risk Averse Buyers," *Econometrica* (52:6), 1984, pp. 1473-1518.

Moldovanu, B. and Sela, A. "The Optimal Allocation of Prizes in Contests," *The American Economic Review* (91:3), 2001, pp. 542-558.

Saidi, R. and Marsden, J. "Number of Bids, Number of Bidders and Bidding Behavior in Outer-Continental Shelf Oil Lease Auction Markets," *European Journal of Operational Research* (58:3), 1992, pp. 335-343

Siegel, R. "All-Pay Contests," *Econometrica* (77:1), 2009, pp. 71-92.

Terwiesch, C. and Xu, Y. "Innovation Contests, Open Innovation, and Multiagent Problem Solving," *Management Science* (54:9), 2008, pp. 1529-1543.

Taylor, C. "Digging for Golden Carrots: an Analysis of Research Tournaments," *The American Economic Review* (85:4), 1995, pp. 872-890.

van Heck, E. and Ribbers, P., "Experiences with Electronic Auctions in the Dutch Flower Industry," *Electronic Markets* (7:4), 1997.

Wu, D.J., P. Kleindorfer, and J. E. Zhang, "Optimal Bidding and Contracting Strategies for Capital-Intensive Goods," *European Journal of Operational Research,* Vol. 137, No. 3, pp. 657-676, 2002.

Yang, J., Adamic, L., and Ackerman, M. "Crowdsourcing and Knowledge Sharing: Strategic User Behavior on taskcn," *Proceedings of the 9th ACM International Conference on Electronic Commerce,* 2008, pp. 246-255.

Zhong, F., and D.J. Wu, "E-Sourcing: Impact of Bidding Behavior and Non-Price Attributes," *Proceedings of the 27th International Conference on Information Systems,* Milwaukee, WI, 2006.