Improving the Efficiency of Crowdsourcing Contests

Haifeng Xu
Computer Science Department,
University of Southern California,
Los Angeles, CA, US
haifengx@usc.edu

Kate Larson
Cheriton School of Computer Science,
University of Waterloo,
Waterloo, Ontario, Canada
klarson@uwaterloo.ca

ABSTRACT

While winner-take-all crowdsourcing contests are wide spread in practice, several researchers have observed that their social welfare can be poor due to effort exerted by contestants who are never rewarded. In this paper we study the problem of efficiency in winner-take-all crowdsourcing contests. Using a discrete choice model to capture contestants' production qualities, we introduce a mechanism which filters out low-expertise contestants, before they are asked to produce a solution. We show that under a set of natural assumptions, such a mechanism has desirable incentive properties, attracts high-quality contestants and can improve social welfare. We also provide insights into the problem of prize setting for such contests.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence

Keywords

Mechanims Design, Incentive Compatible, Winner-take-all Crowdsourcing, Prize Setting.

1. INTRODUCTION

The past few years has seen an explosion of interest in crowdsourcing, as companies and individuals try to leverage communities and networks of workers in order to perform certain tasks, and to take advantage of a global workforce. One of the more prevalent forms of crowdsourcing seen today is crowdsourcing competitions, where a call for solutions for some problem or task is placed and individuals or teams of individuals submit solutions in response to the call. Examples of such competitions range from the X Prize (*www.xprize.org*) and Netflix Prize (*www.netflixprize.com*) where awards to winning submissions have been worth millions of dollars, to numerous websites supporting general crowdsourcing competition platforms, for example, TopCoder (*www.topcoder.com*), CrowdFlower (*www.crowdflower.com*) and 99designs (*www.99designs.ca*), to name a few. However, with crowdsourcing there arrive new challenges. For example, it becomes harder to control workers' professional skills [2, 9] and the management focus becomes how to leverage the skills of potential workers instead of enhancing the skills of current workers [12].

Appears in: *Alessio Lomuscio, Paul Scerri, Ana Bazzan, and Michael Huhns (eds.), Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014), May 5-9, 2014, Paris, France.*

There has also been increasing interest in the research community, as researchers develop and analyze models of crowdsourcing competitions in order to better understand their properties and to study how to improve them. One popular model arising in the literature has been to view a crowdsourcing competition as a variant of an all-pay auction [1, 5, 6]. The insight is the observation that in a crowdsourcing competition, all contestants must exert (potentially costly) efforts to participate while only the competitor judged to have submitted the best entry is actually awarded the prize. This is similar to an all-pay auction where all bidders pay their bid amount, though only the winning bidder is allocated the item being auctioned. DiPalantino and Vojnovic were among the first to make this observation, and used the all-pay auction model in order to analyze multiple simultaneous contests where competitors were able to choose which contests to participate in [6]. Archak and Sundarajan also used the all-pay model when studying strategic behaviors of risk-averse contestants in single competitions [1]. They also had quite negative results with respect to social welfare in that they showed that the expected total utility of all contestants was asymptotically zero due to the (wasted) efforts of non-winning contestants. Chawla *et al* derived a theory of optimal crowdsourcing design (for optimizing the principal's utility) which mirrored that of optimal auction design [5]. They also commented on the problem of wasted efforts by non-winning contestants.

Recent work by Cavallo and Jain has explicitly studied the problem of efficiency in crowdsourcing contests [3, 4]. Differing from the works just discussed which all assumed that the quality of the product produced by contestants was deterministic, they used a stochastic production model to capture the uncertainty that might arise in the quality. Cavallo and Jain were able to design efficient (that is, social welfare maximizing) mechanisms for crowdsourcing but relied on being able to make payments to multiple contestants as opposed to the winner-take-all model which is commonly seen in practice [3]. Focusing on winner-take-all models, they studied their effectiveness when it comes to implementing efficient outcomes [4]. They were able to show that under certain assumptions on the quality distributions (*i.e.* uniform or exponential) it is possible to implement efficient outcomes most of the time.

In this work we also explicitly address the issue of efficiency in crowdsourcing contests. We introduce a discrete choice model to capture the inherent uncertainty in the quality of the solutions produced by contestants [7, 13, 14]. Using an insight from the work by Cavallo and Jain that social welfare is maximized when the top few contestants exert full efforts while all the others exert zero effort [4], we explicitly design a mechanism which aims to filter high-quality contestants and to allow them to compete. We show that under certain natural assumptions the mechanism is incentive compatible, and that the new contest form incentivizes high-quality contestants and increases social welfare. We are also able to provide insights for guiding the contest designer or principal to set the prize amount in order to encourage a certain level of product quality.

2. THE MODEL

In this paper we study winner-take-all crowdsourcing contests in which the contest principal is only interested in the best solution, and the winner takes the entire prize. These types of competitions are widespread in practice, including programming competition platforms such as TopCoder and Kaggle. We assume the contest principal launches a contest with a monetary prize M, that the contest only occurs once, and that the winner of the contest (*i.e.* the contestant who submits the best solution) receives the entire prize.

There are N rational and risk-neutral contestants who are interested in the contest. For simplicity we use $[N]$ to denote the set of N players. Each contestant $i \in [N]$ is characterized by an expertise type θ_i and an effort amount E_i that i can devote during the contest. We normalize the domain of θ_i to be the interval $[0, 1]$, and assume E_i is chosen from the interval $[0, E]$, in which the upper bound E could be given in the form of time requirement or deadline constraint.

Previous literature on crowdsourcing contests has typically assumed that contestants' types (expertise) are drawn from some common distribution [1, 5, 6]. We, instead, use a discrete choice model [14]. Given expertise profile $\{\theta_i\}_{\in[N]}$ and effort profile $\{E_i\}_{\in[N]}$, $E_i\theta_i$ is the idealized production quality provided by contestant i, at the beginning of the contest, before any work occurs. However, due to uncertainties during the task completion and in the contest principal's revision process, we assume that each contestant has some additional random uncertainty ϵ_i, meaning that the true quality submitted is $E_i\theta_i + \epsilon_i$. We assume that the ϵ_i are identically and independently distributed, with probability density function $f(\epsilon)$ and cummulative density function $F(\epsilon) = \int_{-\infty}^{\epsilon} f(t)dt$. We further assume that $f(\epsilon)$ is first-order differentiable in its support set, and the expectation and variance exist.

Given the model outlined above, the probability that a contestant i will win the contest is

$$\begin{aligned}\mathbf{Pr}(i \text{ win}) &= \mathbf{Pr}(E_i\theta_i + \epsilon_i > E_n\theta_n + \epsilon_n,\ \forall n \neq i)\\ &= \mathbf{Pr}(\epsilon_n < E_i\theta_i + \epsilon_i - E_n\theta_n,\ \forall n \neq i)\\ &\stackrel{I}{=} \mathrm{E}_{\epsilon_i}\Big(\Big(\prod_{n\in[N]\setminus\{i\}} F(E_i\theta_i + \epsilon_i - E_n\theta_n)\Big)\Big)\\ &= \int_{-\infty}^{\infty}\Big[\prod_{n\in[N]\setminus\{i\}} F(E_i\theta_i + \epsilon - E_n\theta_n)\Big]f(\epsilon)d\,\epsilon \end{aligned} \quad (1)$$

in which, step I follows from the *i.i.d* assumption of ϵ_i's and the following fact: by fixing ϵ_i, the continuity of f gives $\mathbf{Pr}(\epsilon_n < E_i\theta_i + \epsilon_i - E_n\theta_n) = \mathbf{Pr}(\epsilon_n \leq E_i\theta_i + \epsilon_i - E_n\theta_n) = F(E_i\theta_i + \epsilon_i - E_n\theta_n)$.

Denote the idealized production quality $E_i\theta_i$ as α_i, *i.e.*, $\alpha_i = E_i\theta_i$. Now we can simply denote the probability $\mathbf{Pr}(i \text{ win})$ as $p_i(\alpha_i, \alpha_{-i})$, where α_{-i} represents those idealized production qualities $\{\alpha_n\}_{n\in[N]\setminus\{i\}}$. So $p_i(\alpha_i, \alpha_{-i}) = \int_{-\infty}^{\infty}[\prod_{n\in[N]\setminus\{i\}} F(\alpha_i + \epsilon - \alpha_n)]f(\epsilon)d\,\epsilon$. The following properties of $p_i(\alpha_i, \alpha_{-i})$ following from Equation 1 shows that both better idealized quality and weaker competitors result in greater chance of winning and this chance is "fair" to every contestant.

PROPOSITION 1. *$p_i(\alpha_i, \alpha_{-i})$ is monotonically increasing on α_i and decreasing on α_n, $\forall n \neq i$. Furthermore, two contestants have the same probability to win if they have the same production quality and the same quality profile of competitors* i.e., *$p_i(\alpha_i, \alpha_{-i}) = p_j(\alpha_j, \alpha_{-j})$ if $\alpha_i = \alpha_j$ and $\alpha_{-i} = \alpha_{-j}$.*

In practice, $f(\epsilon)$ can be chosen according to the scenario at hand.[1] For example, a normal distribution results in a model that captures symmetric noise. To get a concrete understanding of the model, we show a particular example.

EXAMPLE 1. *(Logit Discrete Choice Model [10]) If the random uncertainty distribution f has the form*

$$f(\epsilon) = \lambda e^{(-e^{-\lambda\epsilon})}e^{-\lambda\epsilon},\ \epsilon \in R,$$

then the probability for contestant i to win, i.e., *$\alpha_i + \epsilon_i > \alpha_n + \epsilon_n$ $\forall n \neq i$, has the following closed form:*

$$\mathbf{Pr}(i \text{ win}) = p_i(\alpha_i, \alpha_{-i}) = \frac{e^{\lambda\alpha_i}}{\sum_{n\in[N]} e^{\lambda\alpha_n}}.$$

Here $\lambda > 0$ is a parameter to adjust the influence weight of the "representative" value α_i. As $\lambda \to \infty$, "representative" values tend to be decisive and the contestant who has the highest product quality wins the contest with a probability that tends to 1. When $\lambda \to 0$, uncertainties dominate the probabilities, therefore each contestant tends to have an equal chance to win.

3. A MECHANISM FOR FILTERING CONTESTANTS

As has been pointed out by previous researchers (see, for example [8, 15]), winner-take-all crowdsourcing contests can suffer from inefficiencies caused by potentially significant amounts of wasted effort by non-winning entrants. In this section we propose a mechanism that has contestants participate in a pre-qualification stage, before they exert any effort. This allows us to filter contestants based on the idealized production qualities they can produce and then run a contest with only the contestants deemed to provide high enough production qualities. We call the contestants which are allowed to participate in the final competition *competitors*.

DEFINITION 1 (K-COMPETITOR CONTEST). *A K-*Competitor Contest *is defined to be a pair $(\mathcal{C}(N, M), \Re_K)$, in which $\mathcal{C}(N, M)$ is a standard crowdsourcing contest with prize M and N potential contestants, while $\Re_K$ is a rule which selects exactly $K < N$ competitors from the N contestants.*

Generally, a K-competitor contest allows only K competitors to participate by applying some sifting rules. The rest of this section studies a particular sifting-rule coupled with incentives designed to ensure that contestants do not act so as to manipulate the process.

We use $[K]$ to denote the set of K competitors selected by the K-competitor contest. Furthermore, for any $i \in [K]$, let α_{-i}^K denote the idealized production quality profile for the $K - 1$ other competitors. Define $G_i(x) = \prod_{k\in[K]\setminus\{i\}} F(x - \alpha_k)$. Then, given Equation 1, the probability of competitor i winning when the idealized production qualities are α_i and α_{-i}^K is

$$p_i(\alpha_i, \alpha_{-i}^K) = \int_{-\infty}^{\infty} G_i(\epsilon + \alpha_i)f(\epsilon)d\epsilon.$$

Now we consider contestants' utilities. Note that $\alpha_i = \theta_i E_i$ can also be viewed as contestant i's cost amount during the contest. To

[1]The discrete choice model can also be viewed as a restricted version of stochastic production [3, 4], in which contestant i's production quality x obeys the distribution $f(x - \alpha_i)$.

get a monetary cost value, typically one needs to multiply the cost amount by a constant denoting the value per unit cost. Without loss of generality, we assume the interval $[0, E]$ is properly scaled, so that $E_i\theta_i$ represents directly the monetary cost. Following the definition introduced in previous literature [1], we define the *surplus* of i, S_i, to be the difference between the expected reward competitor i achieves by participating and the cost of the effort from participating. In particular

$$S_i = S_i(\alpha_i, \alpha_{-i}^K) = p_i(\alpha_i, \alpha_{-i}^K)M - \alpha_i.$$

We now design the rule $\Re_K$ as follows: we first run an auction in which each potential contestant i is required to submit its private expertise type θ_i and the effort amount E_i it can devote during the contest. Then any contestant with quality $\alpha_i = E_i\theta_i$ ranked in the top K is chosen to be a competitor. Each competitor i has to pay an entry fee, B_i, in order to actually participate. By participating, competitors provide solutions to the principal, who selects the one it judges as best, and rewards the prize M. Formally, we describe the mechanism in Algorithm 1, which we call *Top-K Rule*.

Algorithm 1 Top-K Rule

Input: N potential contestants

Output: K selected competitors for the contest

1. Each potential contestant i submits a bid pair (θ_i, E_i) indicating its expertise type (θ_i) and the amount of effort it can devote during the contest (E_i). Let $\alpha_i = E_i\theta_i$. Denote the ranked idealized quality profile as $\{\alpha_1, \alpha_2, ...\alpha_N\}$, $\alpha_i \geq \alpha_{i+1}$ $\forall i = 1, ..., N-1$. Ties are broken at random.
2. Bidders with top K qualities, $[K] = \{1, 2, ..., K\}$, are selected for the contest.
3. Each selected competitor $i \in [K]$ pays
$$B_i = S_i(\alpha_{K+1}, \alpha_{-i}^K) = p_i(\alpha_{K+1}, \alpha_{-i}^K)M - \alpha_{K+1}$$
as an entry fee for participation.
4. The selected competitors participate in the contest, with a winner being selected by the principal who awards the prize M.

In the mechanism presented in Algorithm 1, the potential contestants who rank $K+1$ to N do not partake in the actual contest and do not need to pay an entrance fee. Their utility is 0 since they incur no cost from exerting effort. For any top K ranked contestant, the expected utility is $S_i(\alpha_i, \alpha_{-i}^K) - B_i$. We note that the entry fees paid by the competitors differ in that competitors' with higher idealized production qualities pay higher fees.

PROPOSITION 2. *In Algorithm 1, the competitor with higher quality pays more. Specifically,* $\forall i, j \in [K]$, $i \neq j$,

$$B_j \geq B_i, \text{ if } \alpha_j \geq \alpha_i.$$

3.1 Truthfulness of the Top-K Rule

In this section we show the truthfulness (*ex post*) of Algorithm 1 for a broad family of distributions over the uncertainty ϵ. That is, each contestant i has incentive to truthfully reveal its expertise θ_i and the effort amount E_i it can devote during the contest. We require the following properties on the uncertainty distribution: **1)** $f(\epsilon)$ is symmetric and **2)** $f(\epsilon)$ is single-peaked at $\epsilon = 0$. That is $f(\epsilon) = f(-\epsilon)$, $\forall \epsilon \geq 0$ (symmetric) and $f'(\epsilon) \leq 0$, $\forall \epsilon \geq 0$ (single-peaked). With these two assumptions it is easy to show that $f'(-\epsilon) = -f'(\epsilon) \geq 0\ \forall \epsilon \geq 0$.

Before we prove the main result of this section, several lemmas are needed. We start by proving bounds on the function $G_i(x) = \prod_{k\in[K]\setminus\{i\}} F(x - \alpha_k)$. In the rest of the section we assume that $\alpha_1, \alpha_2, \ldots, \alpha_K$ are the ranked top K idealized production qualities, unless explicitly stated otherwise.

LEMMA 3.1. *If* $\alpha_K \geq \frac{1}{2}\alpha_1$, *then* $\forall \alpha \in [0, \alpha_1]$, $\forall x \geq 0$ *and* $\forall i \in [K]$, *we have*

$$G_i(x + \alpha) - G_i(-x + \alpha) \geq G_i(x) - G_i(-x). \quad (2)$$

Furthermore, for any $i, j \in [K]$, $\alpha_j \geq \alpha_i$ *and* $\alpha \in [0, \alpha_i]$, *we have*

$$G_j(x + \alpha) - G_j(-x + \alpha) \geq G_i(x + \alpha) - G_i(-x + \alpha). \quad (3)$$

The proof can be found in the Appendix.

Inequalities 2 and 3 together show that for any i in $[K]$ the difference $G_i(x + \alpha) - G_i(-x + \alpha)$ is minimal when $\alpha = 0$. Using this observation we can provide a bound for $\frac{\partial p_i(\alpha, \alpha_{-i}^K)}{\partial \alpha}$.

LEMMA 3.2. *If* $\alpha_K \geq \frac{1}{2}\alpha_1$, *then* $\forall \alpha \in [0, \alpha_1]$ *and* $\forall i \in [K]$, *we have* $\frac{\partial p_i(\alpha, \alpha_{-i}^K)}{\partial \alpha} \geq \frac{\partial p_i(\alpha, \alpha_{-i}^K)}{\partial \alpha}|_{\alpha=0}$.

Similarly, we also have the following bound:

LEMMA 3.3. *If* $\alpha_K \geq \frac{1}{2}\alpha_1$, *then* $\forall i, j \in [K]$ *and* $\alpha_j \geq \alpha_i$, *we have* $\frac{\partial p_j(\alpha, \alpha_{-j}^K)}{\partial \alpha} \geq \frac{\partial p_i(\alpha, \alpha_{-i}^K)}{\partial \alpha}$ *for any* $\alpha \in [0, \alpha_i]$. *Specifically,* $\frac{\partial p_j(\alpha, \alpha_{-j}^K)}{\partial \alpha}|_{\alpha=0} \geq \frac{\partial p_i(\alpha, \alpha_{-i}^K)}{\partial \alpha}|_{\alpha=0}$.

The proof of Lemma 3.2 and Lemma 3.3 can be found in the appendix.

Lemmas 3.2 and 3.3 show that the rate of increase in $p_i(\alpha_i, \alpha_{-i}^K)$ is minimal when $i = K$ and $\alpha = 0$. Using these results we are able to further show that under certain conditions, we can guarantee that surplus, $S_i(\alpha_i, \alpha_{-i}^K)$ is monotonically increasing on α_i in $[0, \alpha_1]$ (recall that $\alpha_1 = \max\{\alpha_j | j \in [N]\}$).

THEOREM 3.4. *Let* $f(\epsilon)$ *be symmetric and single peaked at* $\epsilon = 0$, *if* $\alpha_K \geq \frac{1}{2}\alpha_1$, *then* M *satisfies the following constraint:*

$$M \geq \frac{-1}{\int_{-\infty}^{\infty} G_K(\epsilon) f'(\epsilon) d\epsilon} \quad (4)$$

if and only if *for any* $i \in [K]$, *the surplus* $S_i(\alpha, \alpha_{-i}^K)$ *is monotonically increasing on* α *in* $[0, \alpha_1]$ *and* $S_i(\alpha, \alpha_{-i}^K) \geq 0, \forall \alpha \in [0, \alpha_1]$.

PROOF. We first prove the $\Rightarrow$ direction. Based on Lemmas 3.2 and 3.3, we have, for any $i \in [K]$ and $\alpha \in [0, \alpha_1]$,

$$\begin{aligned}
\frac{\partial S_i(\alpha, \alpha_{-i}^K)}{\partial \alpha} &= M\frac{\partial p_i(\alpha, \alpha_{-i}^K)}{\partial \alpha} - 1 \\
&\geq M\frac{\partial p_K(\alpha, \alpha_{-K}^K)}{\partial \alpha}|_{\alpha=0} - 1 \\
&= -M\int_{-\infty}^{\infty} G_K(\epsilon) f'(\epsilon) d\epsilon - 1 \\
&\geq 0
\end{aligned}$$

whenever Constraint 4 holds. This shows that $S_i(\alpha, \alpha_{-i}^K)$ is monotonically increasing on $\alpha \in [0, \alpha_1]$. Since $S_i(0, \alpha_{-i}^K) \geq 0$, so $S_i(\alpha, \alpha_{-i}^K) \geq 0$ for any $\alpha \in [0, \alpha_1]$.

The proof of $\Leftarrow$ needs to pick the special case where $i = K$ and $\alpha = 0$: the monotonicity of $S_K(\alpha, \alpha_{-i}^K)$ at $\alpha = 0$ gives the Constraint 4. $\square$

Since $\alpha = \theta_i E_i$, Theorem 3.4 means that competitors have no incentive to manipulate their expertise or effort during the actual contest. First, it is impossible for a competitor to exert more effort or show more expertise than their true type. On the other hand, if they exert less effort or show less expertise then their surplus decreases. This means competitors have incentives to behave "truthfully" during the actual contest.

We now prove the main result of this section, which says that potential contestants also have incentives to report their expertise and effort amounts truthfully in the *Top-K Rule* mechanism.

THEOREM 3.5. *Given a* K*-competitor contest* $(\mathcal{C}(N, M), \Re_K)$ *where* $\Re_K$ *is the* Top-K Rule, *assume* $f(\epsilon)$ *is symmetric and single peaked at* $\epsilon = 0$. *If* $\alpha_K \geq \frac{1}{2}\alpha_1$ *and* M *satisfies constraint 4, then* $\Re_K$ *is truthful ex-post. That is, it is a best response for any competitor* $i \in [K]$ *to report the true expertise type* θ_i *and effort amount* E_i *it can devote during the contest.*

PROOF. Denote the ranked true idealized qualities as $\alpha_1 \geq ... \geq \alpha_N$. Recalling that $\alpha_i = \theta_i E_i$, the proof shows that contestant i has no incentive to reveal anything other than its true expertise θ_i and effort amount E_i under the assumption that all other competitors are truthfully revealing their true expertise θ_{-i}^K and effort amounts E_{-i}^K.

First, assume that α_i is such that $\alpha_1 \geq \alpha_i \geq \alpha_K$. This means that i would have been selected for participation if it revealed its true expertise θ_i and effort amount E_i, and its expected utility would be

$$U_i(\alpha_i, \alpha_{-i}) = S_i(\alpha_i, \alpha_{-i}^K) - B_i = S_i(\alpha_i, \alpha_{-i}^K) - S_i(\alpha_{K+1}, \alpha_{-i}^K).$$

Note that $U_i(\alpha_i, \alpha_{-i}) \geq 0$ due to the monotonicity of $S_i(\alpha_i, \alpha_{-i}^K)$ on α_i in $[0, \alpha_1]$ (Theorem 3.4). Contestant i has no incentive to submit a pair (θ_i', E_i') to make $\alpha_i' = \theta_i' E_i' > \alpha_i$, as such a bid pair will still cause the contestant to qualify for the competition, would not change its entrance fee (which is independent of α_i) and would not improve the final chance that it would be allocated the prize M since E_i and θ_i are fixed. Contestant i has no incentive to submit a pair (θ_i', E_i') causing $\alpha_i' = \theta_i' E_i' < \alpha_i$ since either this will not change the non-negative expected utility (as long as $\alpha_i' => \alpha_{K+1}$) or it would cause the contestant to no longer qualify, resulting in utility of 0. Therefore, the pair (θ_i, E_i) is a best response.

Consider the case where $\alpha_i < \alpha_K$. Then, if contestant i truthfully reveals its expertise θ_i and effort amount E_i, it would not qualify and its utility would be 0. The similar situation would arise if it announced a pair (θ_i', E_i') making $\alpha_i' = \theta_i' E_i' < \alpha_K$. However, the contestant does not have incentive to bid a pair (θ_i', E_i') such that $\alpha_i^{'} = \theta_i' E_i' \geq \alpha_K$ either, since then the contestant would qualify as a competitor but with utility $U_i(\alpha_i, \alpha_{-i}) = S_i(\alpha_i, \alpha_{-K}^K) - S_i(\alpha_K, \alpha_{-K}^K) \leq 0$ due to the monotonicity of $S_i(\alpha, \alpha_{-K}^K)$ (note that if i enters top K, then α_K will becomes the $(K+1)$'th biggest idealized quality). $\square$

Theorem 3.5 relies on three main assumptions. First, the symmetric and single peaked assumption on f captures a broad family of uncertainty or noise distributions, *e.g.*, the commonly used normal distributions. Second, by having $\alpha_K \geq \frac{1}{2}\alpha_1$, the top K competitors are likely to be true competitors. That is, certain pathological cases where one competitor is significantly stronger than all others are not possible. Finally, the constraint required for M gives a lower bound for the prize, and Theorem 3.4 actually shows that this bound is also necessary to build up a proper contest setting.

3.2 Further Mechanism Properties

In this section, we study some further properties of the Top-K Rule. For ease of discussion, we assume that the assumptions required for Theorems 3.4 and 3.5 hold.

First, we note that if constraint 4 holds strictly, then the surplus $S_i(\alpha_i, \alpha_{-i})$ is strictly increasing on α_i, and the utility $U_i(\alpha_i, \alpha_{-i}) = S_i(\alpha_i, \alpha_{-i}^K) - S_i(\alpha_{K+1}, \alpha_{-i}^K)$ would be positive if $\alpha_i > \alpha_{K+1}$. Formally, the following proposition follows from Theorem 3.4.

PROPOSITION 3. *If constraint 4 holds strictly, that is,* $M > \frac{-1}{\int_{-\infty}^{\infty} G_K(\epsilon) f'(\epsilon) d\epsilon}$, *then the competitors' utilities would be positive if they really produce strictly better idealized qualities,* i.e., $U_i(\alpha_i, \alpha_{-i}) > 0$ *for any* $i \leq K$, *if* $\alpha_i > \alpha_{K+1}$.

Earlier, Proposition 2 showed that higher-quality competitors pay higher entrance fees in the Top-K mechanism. Now we show that, in spite of this, the high-quality competitors' expected utilities, even when accounting for the entrance fees, are still higher compared to those with lower qualities, therefore the Top-K mechanism will still attract high quality contestants (and thus competitors).

THEOREM 3.6. *The competitor with higher quality has higher utility. Specifically,* $\forall i, j \in [K]$, $i \neq j$,

$$U_j \geq U_i, \text{ if } \alpha_j \geq \alpha_i.$$

PROOF. Due to the monotonicity of $S_j(\alpha_j, \alpha_{-j})$ on $\alpha_j \in [0, \alpha_1]$ in Theorem 3.4, $\alpha_j \geq \alpha_i$ yields $S_j(\alpha_j, \alpha_{-j}) \geq S_j(\alpha_i, \alpha_{-j})$. So we have $U_j = S_j(\alpha_j, \alpha_{-j}) - S_j(\alpha_{K+1}, \alpha_{-j}) \geq S_j(\alpha_i, \alpha_{-j}) - S_j(\alpha_{K+1}, \alpha_{-j})$. To prove the theorem, we show the following stronger inequality:

$$S_j(\alpha_i, \alpha_{-j}) - S_j(\alpha_{K+1}, \alpha_{-j}) \geq S_i(\alpha_i, \alpha_{-i}) - S_i(\alpha_{K+1}, \alpha_{-i}). \tag{5}$$

Inequality 5 has the following equal descriptions:

$$\begin{aligned}
& S_j(\alpha_i,\alpha_{-j}) - S_j(\alpha_{K+1},\alpha_{-j}) \geq S_i(\alpha_i,\alpha_{-i}) - S_i(\alpha_{K+1},\alpha_{-i}) \\
\Leftrightarrow \quad & M(p_j(\alpha_i,\alpha_{-j}) - p_j(\alpha_{K+1},\alpha_{-j})) - (\alpha_i - \alpha_{K+1}) \\
& \geq M(p_i(\alpha_i,\alpha_{-i}) - p_i(\alpha_{K+1},\alpha_{-i})) - (\alpha_i - \alpha_{K+1}) \\
\Leftrightarrow \quad & p_j(\alpha_i,\alpha_{-j}) - p_j(\alpha_{K+1},\alpha_{-j}) \geq p_i(\alpha_i,\alpha_{-i}) - p_i(\alpha_{K+1},\alpha_{-i}) \\
\Leftrightarrow \quad & \int_{\alpha_{K+1}}^{\alpha_i} \left(\frac{\partial\, p_j(\alpha,\alpha_{-j})}{\partial\, \alpha} - \frac{\partial\, p_i(\alpha,\alpha_{-i})}{\partial\, \alpha}\right) d\alpha \geq 0
\end{aligned}$$

in which the last equivalence relation is based on the observation that $p_j(\alpha_i, \alpha_{-j}) - p_j(\alpha_{K+1}, \alpha_{-j})$ can be written in an integral form, *i.e.*,

$$p_j(\alpha_i, \alpha_{-j}) - p_j(\alpha_{K+1}, \alpha_{-j}) = \int_{\alpha_{K+1}}^{\alpha_i} \frac{\partial\, p_j(\alpha, \alpha_{-j})}{\partial\, \alpha} d\alpha$$

To show Inequality 5, we only need to prove $\frac{\partial\, p_j(\alpha,\alpha_{-j})}{\partial\, \alpha} \geq \frac{\partial\, p_i(\alpha,\alpha_{-i})}{\partial\, \alpha}$ for any $\alpha \in [\alpha_{K+1}, \alpha_i]$. This turns out to be true due to the conclusions in Lemma 3.3. $\square$

Theorem 3.6 shows that even though high-quality competitors pay more for participating, they actually don't pay "too much more" in the sense that they are still guaranteed higher expected utility, compared to lower quality competitors.

As a corollary of Theorem 3.6 and Proposition 2, the following relations holds for surplus $S_i = U_i + B_i$.

PROPOSITION 4. *For any* $i, j \in [K]$, *if* $\alpha_j \geq \alpha_i$, *then* $S_j \geq S_i$.

4. SOCIAL WELFARE AND PRIZE SETTING

The motivation of introducing the Top-K Rule was to improve social welfare by reducing the amount of effort induced on contestants who were unlikely to change the outcome of the contest. In this section we explicitly study the relationship between the social welfare which results when using the Top-K Rule and the social welfare that could have been achieved if no filtering took place.[2]

We first consider the maximum product quality under the discrete choice model, which we denote as $Q^{(K)}$ when there are K competitors participating. Let $G(x, K)$ and $g(x, K)$ be the CDF and PDF of $Q^{(K)}$, respectively. It's easy to see that

$$\begin{aligned} G(x,K) &= \mathbf{Pr}(Q^{(K)} \le x) \\ &= \prod_{k\in[K]} \mathbf{Pr}(\alpha_k + \epsilon_k \le x) \\ &= \prod_{k\in[K]} F(x-\alpha_k) \end{aligned}$$

To describe the best product's value to the principal, we follow the formulation in [3] and assume the principal has value v for each *unit* of product quality. So in a winner-take-all crowdsourcing contest, the principal's value from the product pool is $vQ^{(K)}$. The social welfare is then described by the following form [3]:

$$W_K(\alpha_1, ..., \alpha_K) = \mathrm{E}(Q^{(K)})v - \sum_{n\in[K]} \alpha_k$$

in which $\mathrm{E}(Q^{(K)}) = \int_{x\in R} g(x, K)xdx = \int_{x\in R} \frac{d\, G(x,K)}{dx} xdx$ is the expectation of maximum product quality.

Let $W_N(\alpha_1, ..., \alpha_N)$ and $W_K(\alpha_1, ..., \alpha_K)$ be the social welfare of a normal contest and its corresponding K-competitor contest form. We are interested in the gap between these two welfares, *i.e.*, $W_K - W_N$.

We first consider the gap between the expected maximum product qualities.

LEMMA 4.1. *The difference between the expected maximum product qualities satisfies $\mathrm{E}(Q^{(N)}) - \mathrm{E}(Q^{(K)}) \ge 0$. Furthermore, $\mathrm{E}(Q^{(N)}) - \mathrm{E}(Q^{(K)})$ is monotonically decreasing on α_k, $\forall k \le K$ and increasing on α_k, $\forall k > K$.*

The proof is found in the Appendix.

Lemma 4.1 is consistent with our intuition in that more workers result in a higher quality product being produced and the difference between the maximum qualities, $Q^{(N)} - Q^{(K)}$, is largest when all the contestants have the same skill type α_K.

We now prove the main result of this section, namely that it is possible to provide a *lower bound* to the difference between the social welfare from the Top-K mechanism and the social welfare when all contestants participate with no filtering.

THEOREM 4.2. *If $K \ge 2$ and $\alpha_K \ge \frac{1}{2}\alpha_1$, then we have $W_K - W_N \ge \sum_{n=K+1}^{N} (\alpha_n - v\bar{\epsilon}_n)$, where $\bar{\epsilon}_n = \frac{1}{n}\mathrm{E}(\max_{i\in[n],i.i.d} \epsilon_i) = \int_{-\infty}^{\infty} F^{n-1} f(\epsilon)\epsilon d\,\epsilon$ is decreasing on n.*

PROOF. Since

$$\begin{aligned} W_K - W_N &= \mathrm{E}(Q^{(K)})v - \sum_{k\in[K]} \alpha_k - \mathrm{E}(Q^{(N)})v + \sum_{n\in[N]} \alpha_n \\ &= [\mathrm{E}(Q^{(K)}) - \mathrm{E}(Q^{(N)})]v + \sum_{n=K+1}^{N} \alpha_n, \end{aligned}$$

we only need to show $\mathrm{E}(Q^{(K)}) - \mathrm{E}(Q^{(N)}) \ge -\sum_{n=K+1}^{N} \bar{\epsilon}_n$, i.e., $\mathrm{E}(Q^{(N)}) - \mathrm{E}(Q^{(K)}) \le \sum_{n=K+1}^{N} \bar{\epsilon}_n$.

From Lemma 4.1 we know that

$\mathrm{E}(Q^{(N)}) - \mathrm{E}(Q^{(K)})$ is monotonically decreasing on α_k, $\forall k \le K$ and increasing on α_k, $\forall k > K$. So we have

$$\begin{aligned} &\mathrm{E}(Q^{(N)}(\alpha_1, ..., \alpha_N)) - \mathrm{E}(Q^{(K)}(\alpha_1, ..., \alpha_K)) \\ \le\ &\mathrm{E}(Q^{(N)}(\alpha_K, ..., \alpha_K)) - \mathrm{E}(Q^{(K)}(\alpha_K, ..., \alpha_K)). \end{aligned}$$

Now we compute the upper bound for $\mathrm{E}(Q^{(N)}(\alpha_K, ..., \alpha_K)) - \mathrm{E}(Q^{(K)}(\alpha_K, ..., \alpha_K))$. For simplicity, we denote the common type α_K as α. So

$$\begin{aligned} &\mathrm{E}(Q^{(N)}(\alpha, ..., \alpha)) - \mathrm{E}(Q^{(K)}(\alpha, ..., \alpha)) \\ =\ &\int_{x\in R} \frac{d\, F^N(x-\alpha)}{dx} xdx - \int_{x\in R} \frac{d\, F^K(x-\alpha)}{dx} xdx \\ \overset{I}{=}\ &\int_{x\in R} \frac{d\, F^N(x)}{dx}(x+\alpha)dx - \int_{x\in R} \frac{d\, F^K(x)}{dx}(x+\alpha)dx \\ \overset{II}{=}\ &\int_{x\in R} \frac{d\, F^N(x)}{dx} xdx - \int_{x\in R} \frac{d\, F^K(x)}{dx} xdx \\ \overset{III}{=}\ &\int_{-\infty}^{\infty} [NF^{N-1}(x) - KF^{K-1}(x)] f(x) xdx \\ \overset{IV}{=}\ &\int_{0}^{\infty} [NF^{N-1}(x) - KF^{K-1}(x) - NF^{N-1}(-x) + \\ &KF^{K-1}(-x)] f(x) xdx \end{aligned}$$

in which: step I uses variable transform $x \to x + E\alpha$; step II uses the fact $\int_{x\in R} \frac{d\, F^N(x)}{dx} \alpha dx - \int_{x\in R} \frac{d\, F^K(x)}{dx} \alpha dx = \alpha - \alpha = 0$; step IV follows from variable transform $x \to -x$ when $x \le 0$ and the symmetry of $f(x)$.

We claim that:

CLAIM 1. *if $K \ge 2$, then $NF^{N-1}(x) - KF^{K-1}(x) - NF^{N-1}(-x) + KF^{K-1}(-x) \le \sum_{n=K+1}^{N} [F(x)^{n-1} - F(-x)^{n-1}]$ for any $x \ge 0$.*

The proof of the claim is in the Appendix. With this claim, we have

$$\begin{aligned} &\mathrm{E}(Q^{(N)}(\alpha, ..., \alpha)) - \mathrm{E}(Q^{(K)}(\alpha, ..., \alpha)) \\ \le\ &\int_0^{\infty} \sum_{n=K+1}^{N} [F(x)^{n-1} - F(-x)^{n-1}] f(x) xdx \\ =\ &\sum_{n=K+1}^{N} \int_{-\infty}^{\infty} F(x)^{n-1} f(x) xdx \end{aligned}$$

which proves the theorem. $\square$

Note that $\bar{\epsilon}_n$ is the expectation of the largest order statistics of n *i.i.d noise* divided by $n > K$. Given an appropriate K, $\bar{\epsilon}_n$ is small enough to ensure $\sum_{n=K+1}^{N} (\alpha_n - v\bar{\epsilon}_n) > 0$. In such a situation, the Top Kmechanism has higher social welfare even in the worst case (*i.e.* when all the contestants' types are identical). If $\bar{\epsilon}_n$ is negligible compared with α_n, K-competitor contest saves almost all the rank $K+1$ to N contestants' efforts.

[2] Note that we are taking a pessimistic perspective in that we make no claims that the social welfare we study in the general case is achieved in equilibrium. In particular, it is not possible to simply view the general contest as a Top N Rule which automatically inherits the results from Theorem 3.5, as some of the required conditions no longer necessarily apply.

4.1 Setting the Contest Prize

In this section we show how our model may be helpful to the principal when it comes to setting the proper contest size. This has been noted as being a challenging problem in practice when adopting and running crowdsourcing contests [11].

By using $G_K(x)$, the Inequality 4 of Theorem 3.4 gives the following lower bound of the prize M :

$$M \geq \frac{-1}{\int_{-\infty}^{\infty}[\prod_{k\in[K]\backslash\{K\}} F(\epsilon - \alpha_k)]f'(\epsilon)d\epsilon}. \tag{6}$$

This lower bound could be a good choice of M in practice – it guarantees $S_i(\alpha, \alpha_{-i})$ (≥ 0) is increasing on $\alpha \in [0, \alpha_1]$, so that the contest attracts any potential contestant, especially those skilled ones. The necessary and sufficient conditions in Theorem 3.4 shows that the bound is tight, *i.e.*, Inequality 6 is the best lower bound of M which still guarantees monotonicity of $S_i(\alpha, \alpha_{-i})$, $\forall i \in [K]$ and $\forall \alpha \in [0, \alpha_1]$.

It is also interesting to study the relationship between the lower bound on M and the contest parameters.

THEOREM 4.3. *Let*

$$L_M(\alpha_1, ..., \alpha_{K-1}) = \frac{-1}{\int_{-\infty}^{\infty}[\prod_{k\in[K]\backslash\{K\}} F(\epsilon - \alpha_k)]f'(\epsilon)d\epsilon} \tag{7}$$

be the lower bound of M, then L_M is increasing on α_k for any $k \in [K]\backslash\{K\}$.

We leave the proof to the Appendix. An interpretation of this theorem is that a higher prize amount is required if the principal wants to obtain higher idealized production qualities (*i.e.*, higher α_i) by attracting more skilled contestants (*i.e.*, higher θ_i) or expecting competitors to contribute more effort (*i.e.*, bigger E_i). This means that the principal should be able to balance the prize amount and the idealized production quality with targeting competitors with desired skill levels and effort amounts.

5. AN EMPIRICAL EXAMPLE

In this section we provide results from a simulation study in order to highlight certain properties of the mechanism. As an example, we set the number of competitors, K, to be 2 and the set of potential contestants to be $N = 10$. We assume E_i is in the interval $[0, 100]$ and $\theta_i \in [0, 1]$, resulting in $\alpha_i \in [0, 100]$ for any i. We drew α_i randomly from the truncated normal distribution $\frac{c}{20\sqrt{2\pi}}e^{-\frac{(x-50)^2}{400}}$, $x \in [0, 100]$, where c is a rescaling constant. We set $f(x) = \frac{1}{20\sqrt{2\pi}}e^{-\frac{x^2}{400}}$ and the value of each unit of product quality to be $v = 5$. Finally, we set $M = 814$, just slightly higher than the actual lower bound of the prize $L_M = 804$, which was computed using Equation 7.

Our first observation is that, as in line with Lemma 4.1, we saw a decrease in quality as we compared the expected value of unit of product quality obtained if all 10 contestants took part in the competition, compared to only the top 2. That is $Q^{10} = 95.7$ while $Q^2 = 72.6$. However, we also saw a significant increase in the social welfare when using the Top-K mechanism. In particular, the Top-2 mechanism resulted in social welfare equal to 240, while the social welfare if all competitors competed in the actual contest (*i.e.* with no filtering) was only 16. The principal's total utility from running the Top-2 mechanism was 170, which included both the total value of the best product quality produced and the entry fees collected from competitors 1 and 2, minus the actual prize amount.

Table 1 summarizes the statistics for the different contestants. Note that only the top two agents (in bold) actually participated in the competition once we ran the Top-2 rule. The table clearly shows the monotonicity property of S_i, B_i and U_i within top K competitors. The values for all other agents, $3 \leq i \leq 10$, are computed for the case where the contestant i misreported so as to become one of the top 2 candidates, and competed against contestant 1. That is, these values are the best the agents could do by misreporting. We emphasize that by being truthful, their utilities would have been zero, and thus this shows that honest reporting is in the agents' best interest.

Agent	α_i	$p_i(\alpha_i, \alpha_{-K}^K)$	$p_i(\alpha_i, \alpha_{-i})$	S_i	B_i	U_i
1	**62.4**	**0.53**	**0.23**	**368**	**322**	**46**
2	**60.2**	**0.47**	**0.20**	**323**	**298**	**25**
3	57.9	0.44	0.17	298	323	-25
4	55.8	0.41	0.14	277	323	-46
5	54.7	0.39	0.13	266	323	-57
6	41.0	0.23	0.04	142	323	-181
7	40.0	0.21	0.04	135	323	-188
8	33.3	0.15	0.02	91	323	-232
9	32.6	0.14	0.02	87	323	-236
10	24.5	0.09	0.01	49	323	-274

Table 1: Statistics of Contestants.

6. CONCLUSION

In this paper we looked at the problem of efficiency in winner-take-all crowdsourcing contests. Using a discrete choice model to capture contestants' production qualities, we defined a mechanism which filters out low-production-quality contestants, before they are asked to produce a solution. We showed that such a mechanism has desirable incentive properties under a set of natural assumptions, and does improve social welfare. We were also able to provide insights into the problem of prize setting.

We believe that this work complements the research of Cavallo and Jain [4]. While the discrete choice model we used can be seen as a special case of the stochastic production model used by them, it allows us to work with a much broader range of distributions. It would be interesting to further study the relationship between the two models, as well as their respective strengths and weaknesses when it comes to modeling these contests.

There are several other promising directions for future work. One parameter in the mechanism is K, and so a natural question which arises is how should K be set so as to optimize the outcome for the principal. A similar question is how the principal will behave rationally if she is also considered as a strategic player in this game. Our analysis has also only looked at the single-winner setting, and so it would be interesting to look at the problem where multiple winners are allowed in order to understand how that would change the mechanism, and what advantages or disadvantages might arise. Finally, since crowdsourcing contests are wide spread in practice, it would be interesting to both verify the realism of our discrete choice model assumptions, and see how our proposed mechanism would work in practice.

7. REFERENCES

[1] Nikolay Archak and Arun Sundararajan. Optimal design of crowdsourcing contests. In *International Conference on Information Systems (ICIS 2009)*, 2009.

[2] Daren C. Brabham. Crowdsourcing as a Model for Problem Solving. *Convergence: The International Journal of Research into New Media Technologies*, pages 75–90, 2008.

[3] Ruggiero Cavallo and Shaili Jain. Efficient crowdsourcing contests. In *Proceedings of AAMAS 2012*, pages 677–686, 2012.

[4] Ruggiero Cavallo and Shaili Jain. Winner-take-all crowdsourcing contests with stochastic production. In *Workshop on Social Computing and User Generated Content, EC-12*, 2012.

[5] Shuchi Chawla, Jason D. Hartline, and Balasubramanian Sivan. Optimal crowdsourcing contests. In *Proceedings of SODA 2012*, pages 856–868, 2012.

[6] Dominic DiPalantino and Milan Vojnovic. Crowdsourcing and all-pay auctions. In *Proceedings of EC 2009*, pages 119–128, 2009.

[7] Quang Duong and Sebastien Lahaie. Discrete choice models of bidder behavior in sponsored search. In *Proceedings of WINE 2011*, pages 134–145, 2011.

[8] Lav R. Varshney Gireeja Ranade. To crowdsource or not to crowdsource? In *HCOMP 2012*, 2012.

[9] David R. Karger, Sewoong Oh, and Devavrat Shah. Iterative learning for reliable crowdsourcing systems. In *Proceedings of NIPS 2011*, pages 1953–1961, 2011.

[10] Daniel McFadden. Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics*, pages 105–142. Academic Press, 1974.

[11] Benjamin Satzger, Harald Psaier, Daniel Schall, and Schahram Dustdar. Auction-based crowdsourcing supporting skill management. *Inf. Syst.*, pages 547–560, 2013.

[12] Gregory D. Saxton, Onook Oh, and Rajiv Kishore. Rules of crowdsourcing: Models, issues, and systems of control. *IS Management*, pages 2–20, 2013.

[13] Panagiotis G. Ipeirotis Siamak Faridani, Bjĺorn Hartmann. What's the right price? pricing tasks for finishing on time. In *HCOMP 2011*, 2011.

[14] Kenneth E. Train. *Discrete Choice Methods with Simulation*. Cambridge University Press, 2003.

[15] Lav R. Varshney, Joong Bum Rhim, Kush R. Varshney, and Vivek K. Goyal. Categorical decision making by people, committees, and crowds. In *ITA*, pages 425–434. IEEE, 2011.

APPENDIX

Proof of Lemma 3.1

The proof of Lemma 3.1 is as follows.

PROOF. Let $D(t) = F(x+t) - F(-x+t)$, so $D'(t) = f(x+t) - f(-x+t)$.

$f(\epsilon) = f(-\epsilon)$ implies $D'(t) = -D'(-t)$. Since $f(\epsilon)$ is decreasing on $\epsilon \geq 0$, $D'(t) \leq 0$ for any $x \geq 0$. Thus, $D(t)$ is also symmetric and single peaked at $t = 0$ when $x \geq 0$.

$\forall k \in [K]$, $\alpha_k \geq \alpha_K \geq \frac{1}{2}\alpha_1$, so $|0 - \alpha_k| \geq |\alpha - \alpha_k|$ $\forall \alpha \in [0, \alpha_1]$. The symmetry and single peaked property of $D(t)$ implies $D(\alpha - \alpha_k) \geq D_k(-\alpha_k)$. That is, $\forall \alpha \in [0, \alpha_1]$,

$$F(x+\alpha-\alpha_k) - F(-x+\alpha-\alpha_k) \geq F(x-\alpha_k) - F(-x-\alpha_k). \quad (8)$$

Let $\Delta_k(x) = F(x+\alpha-\alpha_k) - F(x-\alpha_k)$, Inequality 8 implies that $\Delta_k(x) \geq \Delta_k(-x) \geq 0$. Since $F(x+\alpha-\alpha_k) = \Delta_k(x) + F(x-\alpha_k)$, we have

$$\begin{aligned}
& G_i(x+\alpha) - G_i(-x+\alpha) \\
= & \prod_{k\in[K]\setminus\{i\}} F(x+\alpha-\alpha_k) - \prod_{k\in[K]\setminus\{i\}} F(-x+\alpha-\alpha_k) \\
= & \prod_{k\in[K]\setminus\{i\}} [F(x-\alpha_k) + \Delta_k(x)] \\
& - \prod_{k\in[K]\setminus\{i\}} [F(-x-\alpha_k) + \Delta_k(-x)] \\
\geq & \prod_{k\in[K]\setminus i} F(x-\alpha_k) - \prod_{k\in[K]\setminus i} F(-x-\alpha_k) \\
= & G_i(x) - G_i(-x),
\end{aligned}$$

in which the inequality " $\geq$ " uses the following fact: if $a_k \geq c_k \geq 0$ and $b_k \geq d_k \geq 0$, then $\prod_{k=1}^n (a_k + b_k) - \prod_{k=1}^n (c_k + d_k) \geq \prod_{k=1}^n a_k - \prod_{k=1}^n c_k$.

To show inequality 3, consider any $\alpha \in [0, \alpha_i]$, we have $|\alpha - \alpha_j| \geq |\alpha - \alpha_i|$. Therefore $D(\alpha - \alpha_i) \geq D(\alpha - \alpha_j)$ which gives $\Delta(x) = F(x+\alpha-\alpha_i) - F(x+\alpha-\alpha_j) \geq F(-x+\alpha-\alpha_i) - F(-x+\alpha-\alpha_j) = \Delta(-x)$. So,

$$\begin{aligned}
& G_j(x+\alpha) - G_j(-x+\alpha) \\
= & \frac{G_j(x+\alpha)}{F(x+\alpha-\alpha_i)} F(x+\alpha-\alpha_i) \\
& - \frac{G_j(-x+\alpha)}{F(-x+\alpha-\alpha_i)} F(-x+\alpha-\alpha_i) \\
\geq & \frac{G_j(x+\alpha)}{F(x+\alpha-\alpha_i)} (F(x+\alpha-\alpha_j) + \Delta(x)) \\
& - \frac{G_j(-x+\alpha)}{F(-x+\alpha-\alpha_i)} (F(-x+\alpha-\alpha_i) + \Delta(-x)) \\
= & G_i(x+\alpha) - G_i(-x+\alpha) \\
& + \frac{G_j(x+\alpha)}{F(x+\alpha-\alpha_i)} \Delta(x) - \frac{G_j(-x+\alpha)}{F(-x+\alpha-\alpha_i)} \Delta(-x) \\
\geq & G_i(x+\alpha) - G_i(-x+\alpha)
\end{aligned}$$

□

Proof of Lemma 3.2

The proof of Lemma 3.2 is as follows.

PROOF. We have the following derivation:

$$\begin{aligned}
& \frac{\partial\, p_i(\alpha, \alpha_{-i}^K)}{\partial \alpha} \\
= & \frac{\partial \int_{-\infty}^{\infty} G_i(\epsilon+\alpha) f(\epsilon) d\,\epsilon}{\partial\, \alpha} \\
\overset{I}{=} & \frac{\partial \int_{-\infty}^{\infty} G_i(\epsilon) f(\epsilon-\alpha) d\,\epsilon}{\partial\, \alpha} \\
\overset{II}{=} & -\int_{-\infty}^{\infty} G_i(\epsilon) f'(\epsilon-\alpha) d\,\epsilon \\
\overset{III}{=} & -\int_{-\infty}^{0} G_i(\epsilon+\alpha) f'(\epsilon) d\,\epsilon - \int_{0}^{\infty} G_i(\epsilon+\alpha) f'(\epsilon) d\,\epsilon \\
\overset{IV}{=} & -\int_{0}^{\infty} (G_i(\epsilon+\alpha) - G_i(-\epsilon+\alpha)) f'(\epsilon) d\epsilon \\
\overset{V}{\geq} & -\int_{0}^{\infty} (G_i(\epsilon) - G_i(-\epsilon)) f'(\epsilon) d\epsilon \\
= & \frac{\partial\, p_i(\alpha, \alpha_{-K}^K)}{\partial \alpha} |_{\alpha=0}
\end{aligned}$$

in which: step I uses variable transform $\epsilon \to \epsilon - \alpha$; step II uses the assumption that $f(\epsilon)$ is one-order differentiable within the support set; step III uses variable transform $\epsilon \to \epsilon + \alpha$; step IV uses symmetry property $f'(\epsilon) = -f'(-\epsilon) \le 0$; step V uses Inequality 2 from Lemma 3.1 and the fact that $f'(\epsilon) \le 0, \forall \epsilon \ge 0$. $\square$

Proof of Lemma 3.3

The proof uses Inequality 3 from Lemma 3.1 and the expression of $\frac{\partial p_i(\alpha, \alpha_{-i}^K)}{\partial \alpha}|_{\alpha=0}$ from the proof of Lemma 3.2.

Proof of Lemma 4.1

The proof of Lemma 4.1 is as follows.

PROOF. We have the following derivation:

$$\begin{aligned}
& \mathrm{E}(Q^{(N)}) - \mathrm{E}(Q^{(K)}) \\
= \; & \int_{x\in R} \frac{d\,G(x,N)}{dx} x dx - \int_{x\in R} \frac{d\,G(x,K)}{dx} x dx \\
= \; & x(G(x,N) - G(x,K))|_{-\infty}^{+\infty} - \int_{x\in R} (G(x,N) - G(x,K)) dx \\
\overset{I}{=} \; & \int_{x\in R} (\prod_{n\in[K]} F(x-\alpha_n) - \prod_{n\in[N]} F(x-\alpha_n)) dx \\
= \; & \int_{x\in R} \prod_{n\in[K]} F(x-\alpha_n)(1 - \prod_{n=K+1}^{N} F(x-\alpha_n)) dx \\
\ge \; & 0
\end{aligned}$$

in which step I follows from the following fact:

$$\lim_{x\to\infty} x(G(x,N) - G(x,K)) = \lim_{x\to\infty} \frac{G'(x,N) - G'(x,K)}{-1/x^2} = 0,$$

because of the existence of the variance of $f(\epsilon)$.

For any $n \in [K]$, α_n only shows up in $\prod_{n\in[K]} F(x - E\alpha_n)$ which is decreasing on α_n, so $\mathrm{E}(Q^{(K)}) - \mathrm{E}(Q^{(N)})$ is also decreasing on α_n. Similarly, for any $n > K$, α_n only shows up in $1 - \prod_{n\in K+1}^{N} F(x - E\alpha_n)$ which is increasing on α_n. Therefore, $\mathrm{E}(Q^{(K)}) - \mathrm{E}(Q^{(N)})$ is increasing on $\alpha_n, \forall n > K$. $\square$

Proof of the Claim in Theorem 4.2

The proof of the Claim used in Theorem 4.2 is as follows.

PROOF. For simplicity, let $a = F(-x) \le \frac{1}{2}, \forall x \ge 0$. According to the symmetry of $f(x)$, we have $F(x) = 1 - a$. So we only need to prove: $N(1-a)^{N-1} - K(1-a)^{K-1} - Na^{N-1} + Ka^{K-1} \le \sum_{n=K+1}^{N}[(1-a)^{n-1} - a^{n-1}]$. Let $B_n = n(1-a)^{n-1} - na^{n-1}$ be a series, we compute $B_{n+1} - B_n, \forall n \ge 2$:

$$\begin{aligned}
B_{n+1} - B_n &= (1-a)^n - a^n - [n(1-a)^{n-1}a - na^{n-1}(1-a)] \\
&\le (1-a)^n - a^n
\end{aligned}$$

in which the inequality "$\le$" follows from the facts that $a \le \frac{1}{2}$ and $n \ge 2$, so $(1-a)^{n-1}a \ge a^{n-1}(1-a)$.

So we have $B_N - B_K \le \sum_{n=K+1}^{N}[(1-a)^{n-1} - a^{n-1}] = \sum_{n=K+1}^{N}[F(x)^{n-1} - F(-x)^{n-1}]$ $\square$

Proof of Theorem 4.3

The proof of Theorem 4.3 is as follows.

PROOF. We prove that $-\int_{-\infty}^{\infty} G_K(\epsilon) f'(\epsilon) d\epsilon$ is decreasing on $\alpha_k, \forall k \in [K]\backslash\{K\}$.

Note that

$$-\int_{-\infty}^{\infty} G_K(\epsilon) f'(\epsilon) d\epsilon = -\int_{0}^{\infty} (G_K(\epsilon) - G_K(-\epsilon)) f'(\epsilon) d\epsilon.$$

Since $-f'(\epsilon) \ge 0, \forall \epsilon \ge 0$, it's sufficient to prove

$$G_K(x) - G_K(-x) = \prod_{k\in[K]\backslash\{K\}} F(x-\alpha_k) - \prod_{k\in[K]\backslash\{K\}} F(-x-\alpha_k)$$

is decreasing on α_k for any $x \ge 0$.

$\forall \delta \in [0, \alpha_1]$, let $\Delta_k(x) = F(x - (\alpha_k - \delta)) - F(x - \alpha_k)$. Similar to the argument in the proof of Lemma 3.1, one can show that $\Delta_k(x) \ge \Delta_k(-x) \ge 0$.

Then we have the following derivation:

$$\begin{aligned}
& G_K(x) \frac{F(x - (\alpha_k - \delta))}{F(x - \alpha_k)} - G_K(-x) \frac{F(-x - (\alpha_k - \delta))}{F(-x - \alpha_k)} \\
= \; & G_K(x) \frac{F(x-\alpha_k) + \Delta_k(x)}{F(x-\alpha_k)} - G_K(-x) \frac{F(-x-\alpha_k) + \Delta_k(-x)}{F(-x-\alpha_k)} \\
= \; & G_K(x) - G_K(-x) \\
& + [\frac{G_K(x)}{F(x-\alpha_k)} \Delta_k(x) - \frac{G_K(-x)}{F(-x-\alpha_k)} \Delta_k(-x)] \\
\ge \; & G_K(x) - G_K(-x)
\end{aligned}$$

This shows that $G_K(x) - G_K(-x)$ is decreasing on $\alpha_k, \forall k \in [K]\backslash\{K\}$. $\square$