Crowdsourcing and Optimal Market Design

B. Pakzad-Hurson†

†Department of Economics, Brown University

BROWN

Mechanism Design is Hard with Interdependent Values

The term **Interdependent Values (IDV)** refers to situations where an agent's valuations depend on the information/type of other's. One natural cause of this is imperfect observation of characteristics of goods to be allocated.

Unfortunately, many positive results with independent values do not extend to interdependent values. Agents have incentives to misreport information to the mechanism to prevent helping out their competitors. In fact, it is often impossible to achieve desirable allocations in IDV settings [1], [2], [5], [6].

To get around impossibility results, I leverage the information of "the crowd" to approximate desirable allocations. Agents are punished to ensure truthful reporting. As the market grows large, my methodology is **accurate** in that it nearly always replicates the full-information optimal allocation, and **non wasteful** in that the size aggregate punishments converges to zero.

Extended Example

The results of this project are general to a wide class of allocation settings. To ground the analysis, consider a labor market in which each firm can accommodate only one worker. Workers and firms have arbitrary preferences over partners, and no monetary transfers are allowed (or equivalently, salaries are fixed).

A *matching* is a function from the set of workers and firms into itself such that each firm is either matched to a single worker or is unmatched, each worker is either matched to a single firm or is unmatched, and firm i is matched to worker n if and only if n is matched to i.

A market designer wishes to create a *(full-information) stable* matching: no firm or worker prefers to be unmatched than to be matched to its assigned partner, and there is no worker-firm blocking pair who prefer one another to their assigned partners.

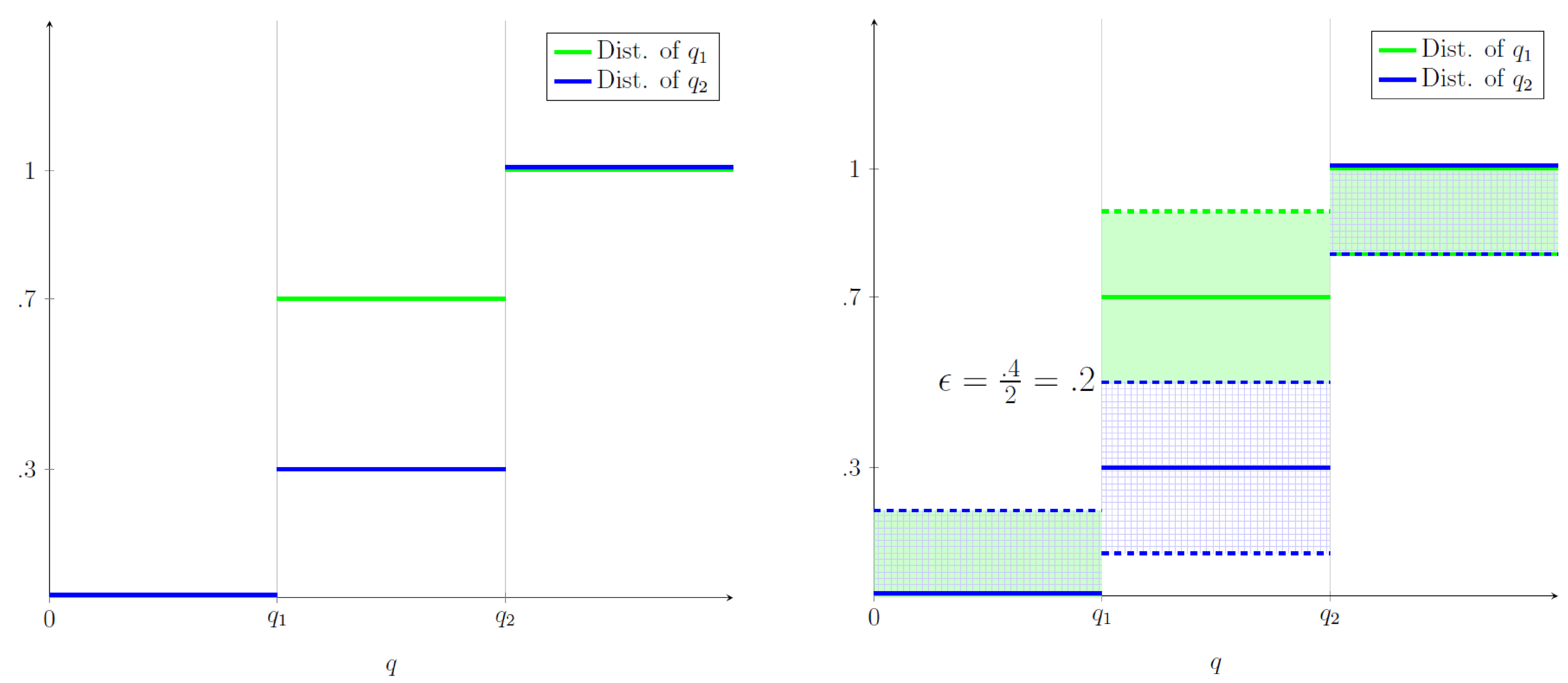
The deferred acceptance mechanism of [4] produces such a matching. However, if worker qualities are imperfectly observed by firms, there exists no mechanism that can guarantee a stable matching [1]. This poster discusses how to generate one with high probability.

(Simplified) Model

- N workers and firms. $I \leq N$ "expert firms."
- Possible qualities of each worker: $q_1, ..., q_Q$.
- Quality for each worker is drawn iid according to some known distribution.
- Each of the I expert firms receives a signal of each worker which is independent (of signals of other firms) conditional on quality of worker. Other firms do not receive signals/are unable to communicate signals to market designer.
- All utilities of matching with a potential partner (other than the quality) common knowledge.

Step 1: Identifying Qualities

Suppose all of the firms' signals are public. The market designer labels a worker as having a particular quality if the empirical distribution of her signals has a pointwise distance from the expected signal distribution less than some ϵ. The first graph illustrates the expected signal distributions when there are two possible qualities. The second graph demonstrates how to calculate ϵ: a worker is assigned quality q_1 iff the distribution of her signals is in the green region, and she is assigned quality q_2 iff the distribution of her signals is in the blue region.



- Need I to be large (many firms with signals) in order to accurately assess qualities.
- Law of Large Numbers implies fraction of misidentified workers goes to zero, but not *number* of misidentified workers: if N grows fast, we have many opportunities to make mistakes.
- Need to identify the qualities of **all** workers. Misidentifying small fraction but large number of workers can lead to a large amount of instability. I provide an example in which nearly all firms and nearly half of workers are involved in (infinitely many) blocking pairs in the limit as the market grows large.

Proposition: *As the market grows large, a designer observing I signals correctly identifies qualities of all workers with probability converging to one if N grows "slower than" $e^{2I\epsilon^2}$.*

The designer can use the crowdsourced qualities in the full-information mechanism to create a stable matching. The bound on relative growth rates is nearly tight:

Proposition: *There is some $\varphi > 2$ such that if N does not grow "slower than" $e^{\varphi I\epsilon^2}$, then the designer will misidentify arbitrarily many workers with probability converging to one.*

$\Rightarrow$

Possible to replicate full-information markets in the presence of interdependent values if and almost only if growth-rate condition satisfied

Step 2: Truthful Reporting

If the designer does not observe firms' signals, can she still approximate a stable matching?

Firms may have incentives to lie to game the mechanism, even in the limit. One form of manipulation is related to the main result of [3]: firms are better off by saying unattainable workers are of highest quality in order to increase chances of matching with a (higher quality) attainable worker.

Crowdsourcing Mechanism

1. Each firm submits a report on the quality of each worker.
2. Assign quality to each worker given reports.
3. Run full-information mechanism (DA) using the assigned qualities.
4. Charge each firm i a penalty proportional to $\frac{2e^{-2I\epsilon^2}}{C-(C+1)2e^{-2I\epsilon^2}}$ where C is a known constant each time its report for a worker differs from the assessed quality of that worker.

Theorem: *The game induced by this mechanism has a Bayes Nash equilibrium in which all firms truthfully report their signals. If N grows "slower than" $\frac{e^{2I\epsilon^2}}{I}$ then the worst-case aggregate punishment converges to zero and a stable matching is created with probability converging to one.*

Note that this growth-rate condition is close to the previous one.

$\Rightarrow$

By using punishment, a designer can crowdsource information almost as effectively as if she had the same information herself

How Big is a Large Market?

The convergence rates in these results are exponential. The accuracy and waste of this mechanism are appealing in "normally" sized markets.

Consider the following approximation of the academic market for economists:

- 100 expert firms/signals.
- ≈75% chance of each signal is correct.
- Firms value workers between \$0 and \$10 million.

Then

- Probability of mis-identifying worker quality ≈ 1 in 1 million.
- Penalty for each incorrect report ≈ \$1 needed for truthtelling.

References

[1] Archishman Chakraborty, Alessandro Citanna, and Michael Ostrovsky. "Two-sided matching with interdependent values". In: *Journal of Economic Theory* 145.1 (Jan. 2010), pp. 85–105.

[2] Yeon-Koo Che, Jinwoo Kim, and Fuhito Kojima. "Efficient Assignment with Interdependent Values". In: *Journal of Economic Theory* 158.A (2015), pp. 54–86.

[3] Vincent P. Crawford. "Comparative Statics in Matching Markets". In: *Journal of Economic Theory* 54 (1991), pp. 389–400.

[4] David Gale and Lloyd S. Shapley. "College Admissions and the Stability of Marriage". In: *The American Mathematical Monthly* 69.1 (1962), pp. 9–15.

[5] Philippe Jehiel and Benny Moldovanu. "Efficient Design with Interdependent Valuations". In: *Econometrica* 69.5 (2001), pp. 1237–1259.

[6] Philippe Jehiel et al. "The Limits of ex post Implementation". In: *Econometrica* 74.3 (2006), pp. 585–610.