Adverse Selection in Adaptive Settings

Carlos Gonzalez

University of Oxford, Department of Economics

November, 2022

Multi-Armed Bandits

- What is a (multi-armed) bandit?
  - Slot machines example. Exploration vs Exploitation
  - Let's be a bit more rigorous...
- Some important concepts
  - History H_t
  - Policy $\pi_t : H_t \mapsto A_t$
  - Regret $R(\pi, \epsilon)$ with ϵ the competitor class
- Two kinds of Bandits
  - Stochastic Bandits: $P_a : a \in \mathcal{A}$. Learner chooses action A_t
  - Adversarial Bandits: Arbitrary sequences $\{x_t\}_1^T$. Learner chooses P_t
  - Differences across $\mathbb{E}[R]$

Mutli-armed bandits

- Two main objects of interest in a bandit problem
- An Upper Bound in the Regret of an Algorithm
  - For a given algorithm α, what is the (order of the) regret of the worst bandit I can give you?
- A Lower Bound in the Regret of the Problem
  - Which is the regret of the algorithm with the lowest Upper Bound among all possible (reasonable) algorithms.
- Usual goal is to obtain sublinear regrets $R_\pi < \mathcal{O}(T)$

- Bandits are amazing, but how are they policy relevant?
- Rather than T periods, think about N agents. Every period, I select a policy parameter (a wage, a tax rate, a price) and I observe the behaviour of agent i
- If I understand policy parameters as (continuous) arms and rewards as particular realizations (either stochastic, either adversarial) of some unknown distribution (or sequence of rewards), then we are back to normal!
- Three foundational papers to my work
  - [Kleinberg and Leighton, 2003] A monopolist problem
  - [Cesa-Bianchi et al., 2021] A bilateral trade problem
  - [Cesa-Bianchi et al., 2022] A policy parameter problem
- More on their models and results later

Classic Problem, Classic Solutions

- [Akerlof, 1978], [Mas-Colell et al., 1995], [Cowell, 2018]
- Consider the following setting
  - N agents, with ability (type) $u_i \in \mathcal{U}$ and reservation value $v_i \in \mathcal{V}$. $\mathcal{U}, \mathcal{V}$ closed intervals in $\mathbb{R}^+$
  - Formally, consider a measure space $(\Omega, \mathcal{F}, \mu)$ and define two rv U, V such that $U : \mathcal{F} \mapsto \mathcal{B}(\mathcal{U})$, $V : \mathcal{F} \mapsto \mathcal{B}(\mathcal{V})$, where $\mathcal{B}(\cdot)$ is the Borel σ-algebra.
  - You may think of $F_{U,V}$ as the product measure $F_U \otimes F_V$ where F_Z is the induced measure of μ on $\mathcal{Z}$ defined via $(\mu \circ Z^{-1})(B)$ for every $B \in \mathcal{B}(\mathcal{U})$
  - u_i and v_i are simply the ith realizations of such variables
  - Agent i observes wage x_i and makes decision $J_i = \mathbb{1}(x_i > v_i)$
  - Define $J^j = \{i : J_i = j\}$

1. Bandits
2. Bandits as Policy
3. Classic Adverse Selection
   - Competitive Equilibrium
   - Monopolistic Competition
4. Adaptive Adverse Selection
5. Conclusion

Competitive Equilibrium

- Consider now two classic problems in Adverse Selection
- **Problem 1: Competitive Equilibrium**
  - Competitive market with 2 firms (wlog) where the policy planner wants to maximize welfare S^{CE} defined via

$$S^{\mathrm{CE}} = \int_{\mathcal{U},\mathcal{V}} [Jx + (1 - J)v] \ \ dF_{U,V} \tag{1}$$

  - Profits don't show up because in equilibrium they are driven down to zero through Bertrand-like competition

Competitive Equilibrium under Full Information

- Under full-information (i.e. $u_i \in H_i$) solution is given by $x_i = u_i$ (and $J^1 = \{i : x_i \geq v_i\}$).
- This result can be characterized under Competitive Equilibrium (CE) and Perfect Bayesian Equilibrium (PBE)
- Observe that social welfare (1) is maximized. Thus equilibrium is socially optimal $x_i = x_i^*, J^1 = J^{1*}$.

Competitive Equilibrium under Partial Information

- But life is not always as beautiful...
- Imagine only $F_{u,v} \in H_i$ for all i
- CE and PBE solutions to this game are characterized via

$$x = u : \{\mathbb{E}[u] = \mathbb{E}[u_i \mid i : v_i < x_i]\} \tag{2}$$

- (In most cases) the set of solutions is not empty
- Observe that there is no room for price discrimination under partial information
- Is any of these equilibria socially optimal? (In most cases) absolutely NOT!

Market Unraveling and Adverse Selection under CE

- Consider $v = r(u)$ with $r(\cdot)$ strictly increasing AND $r(u_i) < u_i$ for all i
- Under partial information ($r(\cdot)$ known, $F_{u,v}$ known) the market may (completely) unravel driven by Adverse Selection considerations
- Example?

- **Problem 2: Monopolistic Competition**
  - 1 monopolistic firm maximizes profits (Π) such that

$$\Pi = \int_{\mathcal{U},\mathcal{V}} [J(u_i - x_i)] \;\; dF_{U,V} \tag{3}$$

  - A social welfare S^{MC} can be defined as

$$S^{\mathrm{MC}} = \int_{\mathcal{U},\mathcal{V}} J\big((u_i - x_i) + \lambda(x_i - v_i)\big) \;\; dF_{U,V} \tag{4}$$

  - With $\lambda < 1$

Monopolistic Equilibrium under Full Information

- Under full-information (i.e. $u_i \in H_i$) solution is given by $x_i = \mathbb{1}(u_i \geq v_i)v_i$ (and $J^1 = \{i : u_i \geq v_i\}$).
- Workers' revenue is driven down to 0
- For $\lambda < 1$, social welfare is maximized (although, possibly, as policymakers we are not very happy with this result).

Monopolistic Equilibrium under Partial Information

- Define Partial Information like in the Competitive Equilibrium case
- Solution to this game is characterized via

$$x^{\mathrm{MC}} = \arg\max_{x} \mathbb{E}_v\left[J(v)\left(\mathbb{E}_u[u \mid v] - x + \lambda(x - v)\right)\right] \tag{5}$$

- An object not as fancy as the one in equation (2), but well-defined

Contributions of the Paper

- I derive **adaptive analogs** of the models above
- **Scope**: For the first time, I characterize models which focus on maximizing consumer surplus (not firm's revenue). Lack of Incentive Compatibility constraints pose new challenges in adaptive frameworks.
- **Asymmetric feedback**: Feedback is dependent to agent's actions. There are extra returns to exploration in a particular exploitation instance.
- **Target Distribution Structure**: I model concrete bounds for structurally dependent u, v. Of particular interest is the dependence structure $v_i \leq u_i$ for all i.

Adaptive Monopolistic Competition

- I start by characterizing a version of Adaptive Monopolistic Competition
- **Key Idea:** Create a model for equation (4) where $F_{U,V}$ remains unknown in $i = 0$
- This model remains novel, as introduces feedback asymmetries, which remain unexplored in the literature.
- Consider the following

$$S_i^{\mathrm{MC}} = \mathbb{1}(x_i > v_i)\big((u_i - x_i) + \lambda(x_i - v_i)\big) \tag{6}$$

Adaptive Monopolistic Competition

- **Timeline:** Agent i arrives, firm offers wage x_i based on H_i. Worker observes x_i and plays $J_i = \mathbb{1}(x_i > v_i)$.
- If $J_i = 1$, agent i works. Firm observes productivity u_i and welfare gains are realized.
- Crucially, productivity u_i (and consequently S_i) is only observed if $J_i = 1$. This introduces feedback asymmetry into the problem
- Optimal policy in this context is given by its known distribution analog. Regret is defined accordingly

Comparison with [Cesa-Bianchi et al., 2021]

- We may rewrite equation (6) following [Cesa-Bianchi et al., 2022] as

$$G_i^v(x_i)\int_x^{\infty} G_i^u(x')\ \ dx' + \lambda \int_0^x G_i^v(x')\ \ dx' \tag{7}$$

- Where we used that there is no loss in replacing $(u_i - x_i)$ by $\max(u_i - x_i, 0)$
- And we have defined where $G_i^v(x_i) = \mathbb{1}(x_i \geq v_i)$ and $G_i^u(x_i) = \mathbb{1}(x_i \leq u_i)$. Moreover, we use the fact that $\mathbb{1}(x_i > v_i)(x_i - v_i) = \max(x_i - v_i, 0) = \int_0^x G_i^v(x')\ \ dx'$ and $\max(u_i - x_i, 0) = \int_x^{\infty} G_i^u(x')\ \ dx'$

Comparison with [Cesa-Bianchi et al., 2021]

- This expression is rather similar to the one in [Cesa-Bianchi et al., 2022]

$$x_i G_i(x_i) + \lambda \int_x^1 G_i(x) \; dx \tag{8}$$

- And [Cesa-Bianchi et al., 2021]

$$G_i^b(x_i) \int_0^x G_i^s(x) \; dx + G_i^s(x_i) \int_x^1 G_i^b(x) \; dx \tag{9}$$

Comparison with [Cesa-Bianchi et al., 2021]

- In terms of Information requirements, our problem is more similar to the one by [Cesa-Bianchi et al., 2021]
- In particular, it requires global information for both the welfare and the gradient
- [Cesa-Bianchi et al., 2021] establishes optimal upper bounds for algorithms of $\mathcal{O}(N^{\frac{1}{2}})$ in the stochastic case when full feedback is recovered
- And of $\mathcal{O}(N)$ when only partial information G_i is revealed after each iteration. They also get $\mathcal{O}(N^{\frac{2}{3}})$ bounds in the partial information setting but under strong additional assumptions
- In the adversarial case, they get bounds $\mathcal{O}(N)$ in all cases

Comparison with [Cesa-Bianchi et al., 2021]

- **Conjecture:** The non-zero measure of the event "full-information" gives some hope for sublinear regret in the stochastic case
- **Conjecture:** I have little hope for sublinear regret in the adversarial case

Adaptive Competitive Equilibrium

- **Key Idea**: Create a model for equation (1) where $F_{U,V}$ remains unknown in $i = 0$
- **Challenge 1:** Reproduce competition in an adaptive setting is very difficult. Firm should have an idea of the wage setting mechanism of the other firm.
- **Challenge 2:** Cannot introduce constraints in expectation, given that the probability distribution is not defined in first place
- **Solution?** Introduce a penalization mechanism for firm profits and losses
- **Key idea:** This penalization **CANNOT** be linear, otherwise there will exist incentives to subsidize workers via firm losses

Naive Model Goes Wrong...

$$S_i = \max(x_i, v_i) + \lambda \mathbb{1}(x_i > v_i)(u_i - x_i) \tag{10}$$

- The policymaker finds profitable to subsidize the worker via losses for $\lambda < 1$
- Setting $\lambda > 1$ is not helping us neither $\implies \mathbb{E}[\Pi] > 0$
- We need to "disproportionately" penalize loses, while fostering worker's welfare. This is rather tricky

Adaptive Competitive Equilibrium

$$S_i = \max(x_i, v_i) + \mathbb{1}(x_i > v_i)[\lambda_1 \mathbb{1}(x_i \leq u_i)(u_i - x_i) + \lambda_2 \mathbb{1}(x_i > u_i)(x_i - u_i)] \quad (11)$$

$$S_i = \max(x_i, v_i) + \mathbb{1}(x_i > v_i)[\lambda_1 \mathbb{1} \max(u_i - x_i, 0) + \lambda_2 \max(x_i - u_i, 0)] \quad (12)$$

$$S_i \sim \max(x_i - v_i, 0) + \mathbb{1}(x_i > v_i)[\lambda_1 \mathbb{1} \max(u_i - x_i, 0) + \lambda_2 \max(x_i - u_i, 0)] \quad (13)$$

- Weights $\lambda_1 < 1$ and $\lambda_2 < -1$ ensure dislike for profits and loses

Did We Get It Right?

- Under full information equation (13) is maximized by setting $x_i = u_i$ with induced $J^1 = \{i : x_i = u_i \geq v_i\}$. Just like in equation (1) (classic result)
- However, under partial information our results will be in general different from Akerlof's $x_i = \mathbb{E}[u_i | i : x_i \geq v_i]$. Why? We broke asymmetry!

Does It Really Matter?

Consider two increasing sequences $\{\lambda_{1n}\}_1^N, \{\lambda_{2n}\}_1^N$ such that $\{\lambda_{1n}\}_1^N \to 1$, $\{\lambda_{2n}\}_1^N \to -1$. x_i is not well defined as the limit of the optimization problem BUT

Claim: For any $\epsilon > 0$ $\exists$ an $n \in \mathbb{N}$ such that $x_i - \mathbb{E}[u_i|x_i < v_i] < \epsilon$ where $x_i = \arg\max_x S_i(x, \lambda_{1n}, \lambda_{2n})$

Corollary: In general our problem characterizes a different equilibrium (a slightly more complicated object) than the one in Akerlof's static unknown distribution **BUT** we can get our solution as close as we want to his result.

Adaptive Competitive Equilibrium

- We may write equation (13) in integral form such that

$$S_i = \int_x^\infty G_i^v(x')\,dx' + (1 - G_i^v(x_i))\left(\lambda_1 \int_0^x G_i^u(x')\,dx' + \lambda_2 \int_x^\infty (1 - G_i^v(x'))\,dx'\right) \quad (14)$$

Adaptive Competitive Equilibrium

- Comments wrt [Cesa-Bianchi et al., 2021] remain valid
- **Conjecture:** Similar? I guess?

Conclusion

- Bandits are a very powerful tool for public policy design!
- This paper introduces analogs for Monopolistic and Competitive Equilibrium in adaptive settings which can be of relevance in many settings
- This paper introduces the concept of feedback asymmetry within the adaptive public policy literature
- This paper introduces competitive mechanisms within adaptive public policy literature. Results are not perfect, but not too bad!
- Previous results give me hope for sublinear regret bounds in the problems above

Thanks!

[Akerlof, 1978] Akerlof, G. A. (1978).
The market for "lemons" : Quality uncertainty and the market mechanism.
In *Uncertainty in economics*, pages 235–251. Elsevier.

[Cesa-Bianchi et al., 2021] Cesa-Bianchi, N., Cesari, T. R., Colomboni, R., Fusco, F., and Leonardi, S. (2021).
A regret analysis of bilateral trade.
In *Proceedings of the 22nd ACM Conference on Economics and Computation*, pages 289–309.

[Cesa-Bianchi et al., 2022] Cesa-Bianchi, N., Colomboni, R., and Kasy, M. (2022).
Adaptive maximization of social welfare.

[Cowell, 2018] Cowell, F. (2018).
Microeconomics: principles and analysis.
Oxford University Press.

[Kleinberg and Leighton, 2003] Kleinberg, R. and Leighton, T. (2003).
The value of knowing a demand curve: Bounds on regret for online posted-price auctions.
In *44th Annual IEEE Symposium on Foundations of Computer Science, 2003. Proceedings.*, pages 594–605. IEEE.

[Mas-Colell et al., 1995] Mas-Colell, A., Whinston, M. D., Green, J. R., et al. (1995).
Microeconomic theory, volume 1.
Oxford university press New York.