Adaptive Wage Setting

A Prior-Free Theory of Adverse Selection and Monopsony Markets

Carlos Gonzalez

University of Oxford, Department of Economics

June, 2023

Motivation I

- Characterisation of equilibrium dynamics under imperfect information and adverse selection mechanisms is of utmost interest in the fields of mechanism design and public policy evaluation
- Equilibrium existence is futile if agents are not endowed with **simple strategies** which gets them close to those equilibria [Hart and Mas-Colell, 2013]
- Especially relevant in environments with asymmetric information, limited feedback or inaccurate priors

Motivation II

- Increasing awareness of monopsony power in labour markets
- Growing policy and academic interest [Furman and Orszag, 2018] [Manning, 2003] [Manning, 2021]
- Impact of Minimum Wage Policies, Limited Information Processing Capacity and Productivity shocks on monopsony firms


The Economist explains

Are labour markets becoming less competitive?

The demise of collective bargaining has allowed firms to flex their "monopsony power" and squeeze wages


Monopsony, Rigidity, and the Wage Puzzle (Wonkish)

By Paul Krugman
Opinion Columnist

Key Results I

Economic Theory

- Develop a novel approach to the monopsony wage setting problem (GMP) **without priors** on the joint distribution of workers' productivity and reservation wage
- Gain further insights on **adverse selection** mechanisms in dynamic games
- Revisit an equivalence between **Hannan Consistency and equilibrium convergence** [Hart and Mas-Colell, 2001a]

Key Results II

Online Learning and Adaptive Policy Design

- Constructively, show the existence of Hannan Consistent policies in the GMP embedded in Algorithm 1
  - Without any priors
  - Under **limited feedback** structures
  - For any **arbitrary** (even adversarial!) sequence of outcomes
- Show **near optimality** of Algorithm 1 with bounds of $\tilde{\mathcal{O}}(K^{2/3})$
- Introduction of new **feedback structures** (asymmetric feedback) to the field of Adaptive Policy Design
- Empirical discussion of the risk associated to **greedy parameter selection**

Key Results III

Policy Analysis

- Structural Policy Analysis using our new toolkit
  - Minimum Wage
  - Limited Information Processing Capacity
  - Productivity Shocks

Literature Review

- **Adverse selection** in static sames. Competitive equilibrium [Akerlof, 1978], Monopoly setting [Mas-Colell et al., 1995]
- **Convergence to equilibrium**. Econ Theory [Hart and Mas-Colell, 2001a] [Hart and Mas-Colell, 2000] [Hart and Mas-Colell, 2001b] [Hart and Mas-Colell, 2013], Online Learning [Auer et al., 2002] [Cesa-Bianchi and Lugosi, 2006] [Bubeck et al., 2012]
- **Adaptive policy design**. Monopoly pricing [Kleinberg and Leighton, 2003], Bilateral trade [Cesa-Bianchi et al., 2021], Optimal tax [Cesa-Bianchi et al., 2022] more
- **Online Learning in Economics** Behavioral eqm with adverse selection [Esponda, 2008], Manipulation-proof ML [Björkegren et al., 2020], Imperfect competition in dynamic games [Ericson and Pakes, 1995] [Doraszelski and Pakes, 2007]

Set-Up

$$\mathcal{R}(\pi, \nu)_K = \mathbb{E}[\sup_{x \in [0,1]} \sum_i^K S_i(x; u_i, v_i) - \sum_i^K S_i(x_i; u_i, v_i)] \tag{1}$$

- where $S_i(\cdot)$ is the welfare function
- Sequence of K workers characterised by a pair $(u_i, v_i) \in [0,1]^2$ following $F_{U,V}$. Let u_i be the productivity and v_i the reservation wage of worker i
- $\pi : H_i \to [0,1]^K$, where H_i is the history of outcomes and actions up to period i
- **Policy Target:** Select policy π to minimize $\mathcal{R}(\pi, \nu)$, where ν is the *environment* or class of policies available to the adversary (**stochastic** vs oblivious **adversarial**) more

Offline Monopsony Problem

$$\$ = \int_{[0,1]\times[0,1]} \mathbb{1}(x \geq v) \cdot (u - x + \lambda(x - v)) \;\; dF_{U,V} \tag{2}$$

- where $(x - v)$ is the workers' surplus
- $\lambda < 1$ preferences towards workers' surplus (GMP)
- We can also write the one individual analogue of equation (2)

$$S_i^{\text{GMP}} = G_i^v(x_i) \cdot (u_i - x_i) + \lambda \int_0^x G_i^v(x') \;\; dx' \tag{3}$$

- where $G_i^v(x') = \mathbb{1}(x \geq v)$ is the *demand function* and $\mathbb{1}(x_i \geq v_i)(x_i - v_i) = \max(x_i - v_i, 0) = \int_0^x G_i^v(x') \;\; dx'$

Partial Information in the Offline GMP

- The belief-conditional best reply (BNE) is given by

$$x^{\mathrm{P}}_{\mathrm{GMP}} = \arg\max_{x} \mathbb{E}_{V,U}\left[\mathbb{1}(x \geq v) \cdot \left(u - x + \lambda \cdot (x - v)\right)\right] \tag{4}$$

- Adverse Selection mechanisms and market unraveling
- Elegant..., but unsatisfactory $\rightarrow$ Calls for a **theory of learning** which ideally
  - **Prior-free**
  - Realistic **limited feedback** $(x, \mathbb{1}(x \geq v), \psi^{\emptyset}((x \geq v), u))$
  - For any **arbitrary distribution** of outcomes (u_i, v_i) (even adversarial!)

Hannan Consistency

- Consider some notion of equilibrium (best reply) defined as $x^* := \sup_{x\in[0,1]} \sum_i S_i(x)$

Definition 1: Hannan Consistency

A policy is Hannan consistent if the sequence $\{x_i\}_i^K$ induced by policy π yields

$$\limsup_{K\to\infty} \frac{1}{K}\Big(\sup_{x\in[0,1]} \sum_i^K S_i(x) - \sum_i^K S_i(x_i)\Big) = 0 \text{ with probability 1} \tag{5}$$

- There is an intuitive connection between Hannan Consistency and convergence to x^*

Equilibrium Convergence

Proposition 2: Equilibrium Convergence

Fix a sequence $(u_i, v_i)_{i=1}^K$ and a discrete policy space with $B+1$ arms. Let policy π be Hannan Consistent for a K-period one-player game with bounded rewards $S_i(x)$, then

$$p_{x^*}^{\pi,K} = \frac{1}{K}\sum_{i}^{K} \mathbb{1}(x_i = x^*) \xrightarrow{p} 1 \text{ as } K \text{ goes to } \infty \tag{6}$$

Proof

Interpretation

- **Interpretation.** In a game which satisfies the conditions in Proposition 2, the implementation of a Hannan Consistent strategy guarantees that the induced actions converge to the optimal set of actions x^* in the Partial Information context **with correct beliefs**. [Hart and Mas-Colell, 2001a] [Cesa-Bianchi and Lugosi, 2006]

- **Do Hannan Consistent policies exist in a PF-LF-AD scenario?** Yes! We show this constructively

Observation 3: Sub-linear Regret

Under equation (1), let a policy π embedded in Algorithm A be sub-linear for some environment ν, then policy π is HC

Feedback Info Req OCO

Algorithm

Algorithm 1 Tempered Exp3 for the GMP

Input B, λ, η, γ
Set $x_b = (b-1)/B$ for $b \in \{1, 2, \ldots, B+1\}$, $\widehat{\mathbb{G}}_{1b} = 0$, $\widehat{\mathbb{U}}_{1b} = 0$
for $i = 1, 2, \ldots, K$
  for $b = 1, \ldots, B+1$
    Set $\widehat{\mathbb{S}}_{ib} = \widehat{\mathbb{U}}_{ib} - x_b \cdot \widehat{\mathbb{G}}_{ib} + \frac{\lambda}{B} \sum_{b'<b} \widehat{\mathbb{G}}_{ib'}$ $p_{ib} = (1-\gamma)\frac{\exp(\eta\ \widehat{\mathbb{S}}_{ib})}{\sum_{b'} \exp(\eta\ \widehat{\mathbb{S}}_{ib'})} + \frac{\gamma}{B+1}$
  end for
  Sample $b_i \sim p_{ib}$ and observe $\mathbb{1}(x_{b_i} \geq v_i)$
  If $\mathbb{1}(x_{b_i} \geq v_i) = 1$ observe u_i
  for $b = 1, \ldots, B+1$
    Update $\widehat{\mathbb{G}}_{i+1,b} = \widehat{\mathbb{G}}_{ib} + \mathbb{1}(x_{b_i} \geq v_i)\frac{\mathbb{1}(b_i = b)}{p_{ib}}$ $\widehat{\mathbb{U}}_{i+1,b} = \widehat{\mathbb{U}}_{ib} + u_i \cdot \mathbb{1}(x_{b_i} \geq v_i)\frac{\mathbb{1}(b_i = b)}{p_{ib}}$
  end for
end for SGD

Upper Bound

Theorem 3: Adversarial Upper Bound on Algorithm 1

Consider a sequence $\{x_i\}_{i=1}^K$ as given by Algorithm 1 with parameters $\gamma = c_1 \cdot \left(\frac{log(K)}{K}\right)^{\frac{1}{3}}$, $\eta = c_2 \cdot \gamma^2$ and $B = \frac{c_3}{\gamma}$ for some $c_1, c_2, c_3 \in \mathbb{R}$. It follows that for any arbitrary sequence $\{(u_i, v_i)\}_{i=1}^K$, there exist a constant $c_4 < \infty$ such that

$$\mathbb{E}[\sup_x \sum_i^K S_i(x) - \sum_i^K S_i(x_i)] \leq c_4 \cdot \log(K)^{\frac{1}{3}} \cdot K^{2/3} \tag{7}$$

- Upper bound on the adversarial case returns a bound on the *iid* case (c.f. Theorem 4)

Corollaries Upper Bound

Corollary 7: Hannan Consistency of Algorithm 1

Policy π embedded in Algorithm 1 is Hannan Consistent for any arbitrary distribution $\{u_i, v_i\}_{i=1}^{K}$ under limited feedback structures and without imposing any priors on the learner

Corollary 8: Convergence in Probability to Equilibrium

The empirical probability $p_{x^*}^{\pi}$ of playing the BNE x^* by an agent who implements policy π embedded in Algorithm 1 $\xrightarrow{p}$ 1 as $K \to \infty$. Equivalently, the distribution of actions induced by policy π converges in probability to the Monopsony Equilibrium (best reply) under Partial Information and Correct Beliefs

Lower Bound I

Theorem 5: Stochastic Lower Bound on the GMP

Consider the problem of sequentially choosing $\{x_i\}_{i=1}^{K}$ in the GMP set-up. There exists a constant $C > 0$ such that for any policy π, and horizon $K \in \mathbb{N}$ there exists a distribution $F_{U,V}$ such that

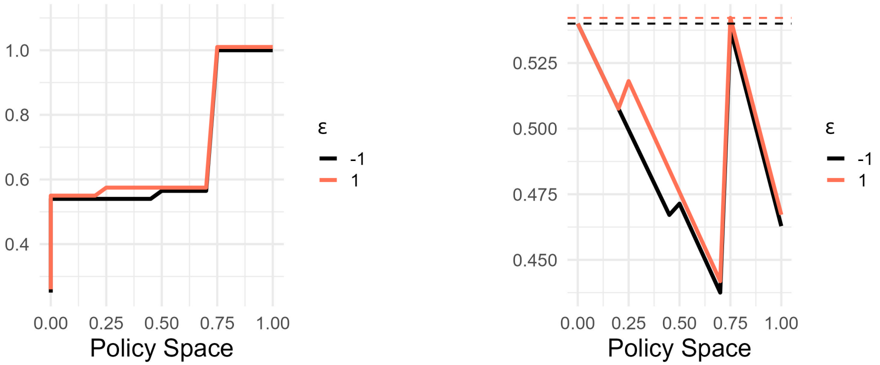
$$\mathbb{E}_{F_{U,V}}[\mathcal{R}(\pi)] \geq C \cdot K^{2/3} \tag{8}$$

- A lower bound on the stochastic *iid* case immediately returns a bound on the adversarial case (c.f. Corollary 6)
- The GMP exhibits an excess of regret compared to standard (adversarial) bandits $\mathcal{O}(\sqrt{K})$
- Global information requirement in the objective function (integral component)

Lower Bound II

- Adversary could create a distribution $F^{\epsilon}_{U,V}$ that necessitates the policymaker to infer the sign of ϵ to determine the optimality between policies x' and x''
- $F^{\epsilon}_{U,V}$ can be defined such that ϵ can only be inferred by sampling from a clearly sub-optimal region [Cesa-Bianchi et al., 2022]

Lower Bound Graphical Intuition



(a) $\mathbb{P}_{F^{\epsilon}_{U,V}}(x \geq v)$ (b) $\mathbb{E}_{F^{\epsilon}_{U,V}}[S_i(x)]$

Figure 2: Lower Bound on the GMP. There are only two candidates to optimal policy $x = 0$ and $x = 3/4$, but the policymaker needs to sample from the sub-optimal region $[1/4, 1/2]$ to infer the sign

Corollary Lower Bound

Corollary 9: Optimality of Algorithm 1

Algorithm 1 is essentially unimprovable (near-optimal) up to logarithmic factors provided we have presented matching upper and lower bounds under stochastic and adversarial specifications of $\mathcal{O}(K^{2/3})$

Uniform Linear Degenerate Case I

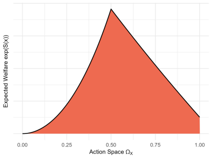
- $U \sim \mathcal{U}[0,1], \ V = 0.5 \cdot u$

$$\arg\max_x \mathbb{E}_{U,V}\left[\mathbb{1}(x \geq v) \cdot ((u - x) + \lambda(x - v))\right] = \int_0^{2x} u - x + \lambda\left(x - \frac{1}{2}\right) du \tag{9}$$

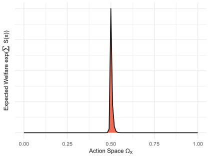
- Equation (9) is maximised at $x^* = 1/2$
- Bounds are not tight to ULDC (potentially faster), sub-optimal parameter selection (potentially slower) more

Uniform Linear Degenerate Case II

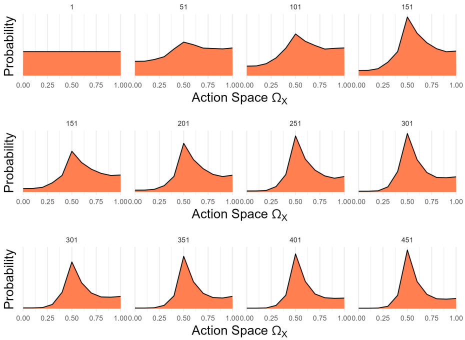
Figure 3: Algorithm 1 with $U \sim \mathcal{U}[0,1]$, $v = 0.5u$



(a) $\mathbb{E}_{F_{U,V}}[\exp(S_i(x))]$



(b) $\mathbb{E}_{F_{U,V}}[\exp(\mathbb{S}_i(x))]$

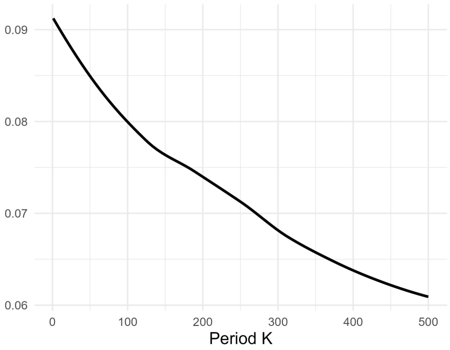


(c) Empirical Distribution of Probs across K

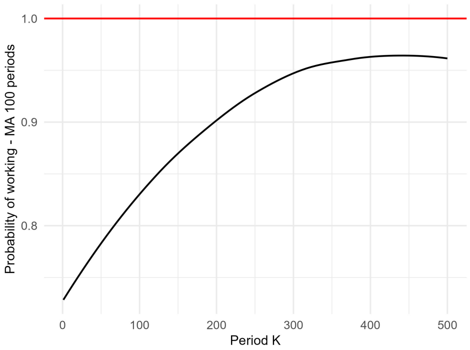
Average 1,000 simulations. $\lambda = 0.7$. $K = 500$. $\eta = 0.132$, $B = 10$, $\gamma = 0.029$.

Uniform Linear Degenerate Case III

Figure 4: Algorithm 1 given $U \sim \mathcal{U}[0,1]$, $V = 0.5 \cdot u$



(a) Avg Cum Regret

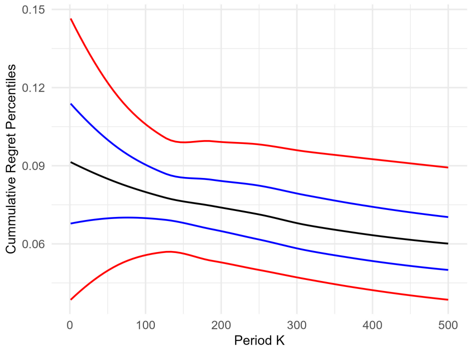


(b) Prob of Work $\mathbb{P}(v \leq x)$

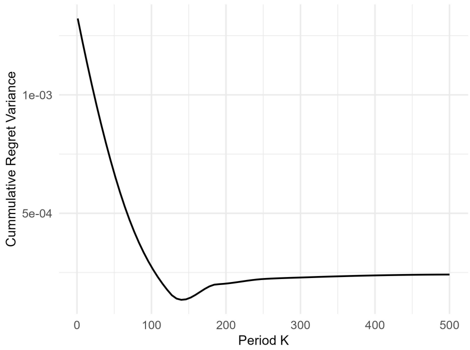
1,000 simulations $\lambda = 0.7$ $K = 500$ $\eta = 0.132$ $B = 10$ $\gamma = 0.029$ MA $= 100$

Uniform Linear Degenerate Case IV

Figure 5: Algorithm 1 given $U \sim \mathcal{U}[0,1]$, $V = 0.5 \cdot u$



(a) Avg Regret, p5-p95, p25-p75, p50

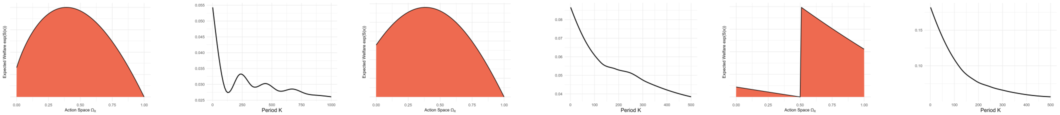


(b) Cross-Sim Regret Variance

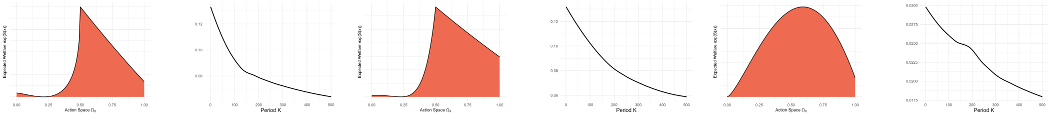
1,000 simulations $\lambda = 0.7$ $K = 500$ $\eta = 0.132$ $B = 10$ $\gamma = 0.029$

Further Simulation Evidence

Figure 6: Further Simulation Evidence of Algorithm 1



(a) $U \sim \mathcal{U}[0,1], V = u^2$ (b) $U, V \sim \mathcal{U}[0,1]$ (c) $U \sim B(0.7), V = 0.5u$



(d) $U \sim \mathcal{B}(.5,.5), V = 0.5u$ (e) $U \sim \mathcal{B}(2,1), V = 0.5u$ (f) $U \sim \mathcal{B}(2,1), V = u^2$

1,000 simulations $\lambda = 0.7$ $K = 500$ $\eta = 0.132$ $B = 10$ $\gamma = 0.029$[1]

[1] For the Uniform Degenerate Non-linear Case $K = 1,000, \eta = 0.025$

Small Recap

- So far, we have redefined the wage-setting problem of the firm in an online game in a PF-LF-AD framework, and we have shown convergence to partial information best reply outcomes
- Our model stands out as a reasonable modelling device in broader micro and macro theory including structural policy analysis
- Certainly, many limitations (so take it as an heuristic!)
  - Infinite pool of workers
  - Strict monopsony considerations
  - Linearity of the production function wrt productivity
  - No budget constraints [Gonzalez, 2023]

Minimum Wage

- Large literature on monopsony power in labour markets [Furman and Orszag, 2018] [Manning, 2003] [Manning, 2021] with an interest on the role of uncertainty on workers' welfare and inequality [Dube et al., 2016] [Card et al., 2012]

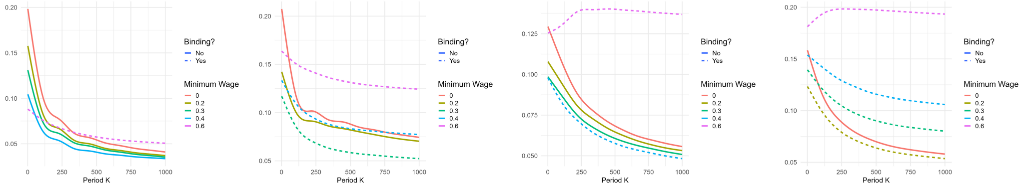
- Model minimum wage as a restriction to the policy space to $[m, 1]$ with $m > 0$

- We show that if $x_{ib^*} \notin [m, 1]$ AND $x_{ib^*} \geq v_i$ OR $x_{ib^*} \notin [m, 1]$, $x_{ib^*} \geq v_i$ AND $x_{ib^*,\mathrm{MW}} \geq v_i$, profit losses are potentially unbounded

- However, in a stochastic context one can analyse which features of the DGP are likely to increase welfare loss
  - High V reduces profit loss (reservation wages as entry barriers)
  - Ambigous effect of workers' productivity on profit loss
  - *Information gains* of small increase in MW

Simulation Evidence of Minimum Wage

Figure 7: Algorithm 1 under MW restrictions



(a) High U, High V (b) High U, Low V (c) Low U, High V (d) Low U, Low V

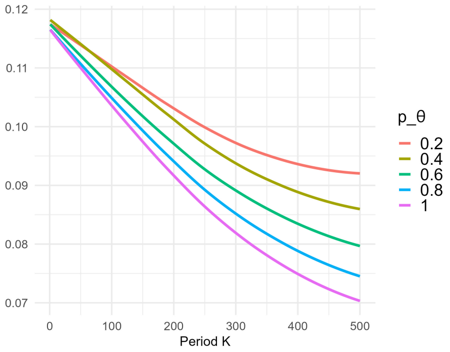
High $U = \mathcal{U}[1/2, 1]$, Low $U = \mathcal{U}[1/4, 3/4]$, High $V = 0.5u$, Low $V = 0.25u$
1,000 simulations $\lambda = 0.7$ $K = 500$ $B = 10$ $\gamma = 0.029 \eta = 0.132$

Limited Information Processing Capacity

- Increasing interest for the role of **costly information acquisition** in macro [Sims, 2003] [Maćkowiak et al., 2023], labour [Acharya and Wee, 2020], finance [Van Nieuwerburgh and Veldkamp, 2010]
- Many ways to model restricted information. Our approach: **label efficient**
- **Intuition:** Adversary independent variable $\Theta \sim \mathscr{Bern}(p_\theta)$, such that if $\Theta = 1$, feedback $= \psi_i$, else $= \emptyset$
- **Problem:** Algorithm 1 not optimal anymore, but a version of it does! conditional on $p_\theta \geq \mathcal{O}\left(\left(\frac{\log(K)^{\frac{1}{3}}}{K^{\frac{1}{3}}}\right)^{1/2}\right)$ (c.f. Proposition 10) Algorithm 2

Graphical Evidence LIPC

Figure 8: Algorithm 2 under LIPC



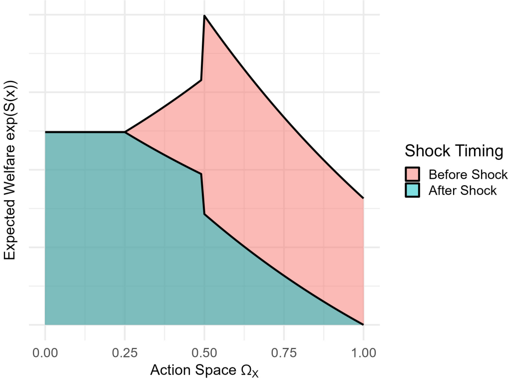
$$U = \mathcal{U}[1/4, 3/4], V = 0.5 \cdot u$$

Productivity Shocks I (VERY Preliminary)

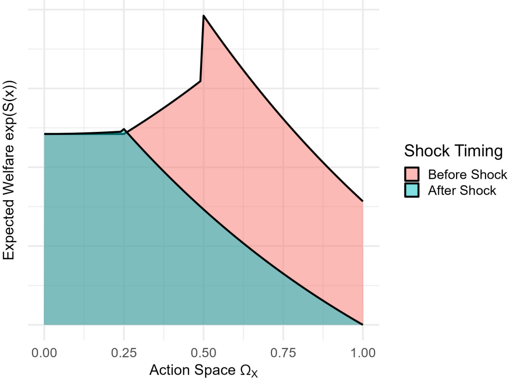
- Very little understanding on the role of Technology/Productivity shocks on RBC, labour inputs and outcome [Ramey, 2016]
- **Heuristics:** The firm learns for $K_0 < K$ periods. Shock hits but the firm remains ignorant
- Analyse regret
  - (i) Fair and Unfair competitor class
  - (ii) Decrease vs no-decrease of associated reservation wages
- Analyse evolution of labour inputs

Productivity Shocks II (VERY Preliminary)

Figure 9: $\mathbb{E}[\exp(S_i(x))]$ for $U \sim [1/2, 1]$ pre and $U \sim [0, 1/2]$ post. $V = 0.5 \cdot u^*$



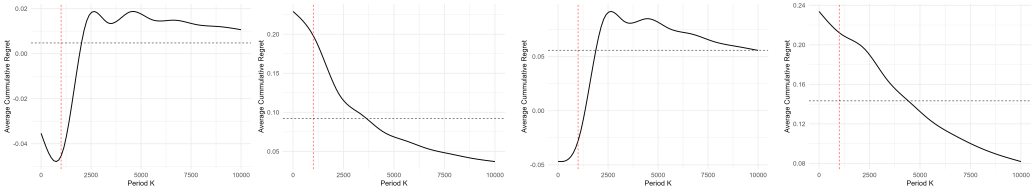
(a) $\mathbb{E}[\exp(S_i(x))]$ with unchanged V



(b) $\mathbb{E}[\exp(S_i(x))]$ with change in V

Productivity Shocks Graphical Evidence (VERY Preliminary)

Figure 10: Avg cum regret Algorithm 1 with productivity shocks



(a) Unch V, Fair (b) Unch V, Unfair (c) Changed V, Fair (d) Changed V, Unfair

1,000 simulations $\lambda = 0.1 K_0 = 1,000 K = 10,00 B = 10 \gamma = 0.029 \eta = 0.00267$

1. Introduction
2. Set-Up
3. Equilibrium Convergence
4. Algorithmic Bounds
5. Simulation Analysis
6. Structural Analysis
7. **Conclusion**

Conclusion

- Showed existence of strategies converging to best reply in **PF-LF-AD** environments
- Showed convergence at a **near-optimal** rate
- Auxiliary contributions to the literature: **Asymmetric Feedback and greedy parameter selection**
- **Economic modelling and structural policy analysis** potential of bandit learning

Thank you!

carlos.gonzalezperez@economics.ox.ac.uk

presidente-carlos.github.io

8 Additional Material

9 Bibliography

Introduction to Bandit Problems

- ν refers to the class of DGP available to the adversary
- Two limiting cases
  - **Stochastic:** Nature selects $F_{U,V}$. (u_i, v_i) are *iid* realisations of such distribution. Expectations are taken wrt $F_{U,V}$ (and possibly) any randomness in the algorithm
  - Oblivious **Adversary:** Any arbitrary distribution of outcomes (u_i, v_i), possibly depending on the algorithm of the learner. Any deterministic algorithm incurs in linear regret. Cannot depend on H_i. Sequence is considered to be fixed, so expectations are taken wrt to the randomness in the algorithm of the learner only.

back

Three Canonical Problems

	Monopoly Pricing	Bilateral Trade	Optimal Tax
Objective Function	$\mathbb{1}(x \leq v) \cdot x$	$\mathbb{1}(x \leq v^b) \cdot \max(x - v^s, 0) +$ $\mathbb{1}(x \geq v^s) \cdot \max(v^b - x, 0)$	$x \cdot \mathbb{1}(x \leq v) +$ $\lambda \cdot \max(v - x, 0)$
Welfare	Pointwise	Global	Global
Gradient	Local	Local	Global
Bounds	$\mathcal{O}(K^{1/2})$	$\mathcal{O}(K)$	$\mathcal{O}(K^{2/3})$

Table 1: Three Canonical Problems

back

Equilibrium Convergence Proof I

Proof: Assume $p_{x^*}^{\pi,K} \overset{p}{\nrightarrow} 1$, then $\lim_{K\to\infty} \mathbb{P}\big(1 - \frac{1}{K}\sum_i^K \mathbb{1}(x_i^\pi = x^*) > \epsilon\big) > 0$ for some $\epsilon > 0$.

Call this event E, with $\mathbb{P}(E) > 0$.

Define the set $K \supseteq I^* = \{i : x_i^{\pi,K} \neq x^*\}$. Rewrite $\mathbb{P}(E)$ as $\mathbb{P}\big(\frac{|I^*|}{K} > \epsilon\big) > 0 \implies \mathbb{P}(|I^*| > K \cdot \epsilon) > 0$

Let event E hold. Define $x' = \arg\max_{x \neq x^*} \limsup_{K\to\infty} \sum_{i\in I^*} S_i(x)$.
By construction of x^*,

$$\limsup_{K\to\infty} \frac{1}{|I^*|}\sum_{i\in I^*} S_i(x^*) - \frac{1}{|I^*|}\sum_{i\in I^*} S_i(x') = S_{\min} > 0 \tag{10}$$

Equilibrium Convergence Proof II

$$\limsup_{K\to\infty} \frac{1}{K}\Big(\sup_{x\in\Omega_X} \sum_i^K S_i(x) - \sum_i^K S_i(x_i)\Big) \geq \limsup_{K\to\infty} \frac{1}{K}\Big(\sum_{i\in I^*} S_i(x^*) - \sum_{i\in I^*} S_i(x')\Big) = \\ \limsup_{K\to\infty} \frac{1}{K} \cdot |I^*| \cdot S_{\min} \geq \limsup_{K\to\infty} \frac{1}{K} \cdot K \cdot \epsilon \cdot S_{\min} = \epsilon \cdot S_{\min} > 0 \tag{11}$$

$$\mathbb{P}\left(\limsup_{K\to\infty} \tfrac{1}{K}\big(\textstyle\sum_i^K S_i(x^*) - \sum_i^K S_i(x_i)\big) > \delta = \epsilon \cdot S_{\min}/2 \right) > 0 \quad \square$$

back

Feedback Structure I

- Literature focuses on two limiting cases
  - **Full feedback:** Variable Z is recovered at the end of the period
  - **Realistic feedback:** Only $\mathbb{1}(x \geq z)$ is recovered at the end of the period
- Feedback in the GMP is arguably different: $(x, \mathbb{1}(x \geq v), \psi^{\emptyset}((x \geq v), u))$
- Realistic on V, and $\mathbb{1}(x \geq v)$-asymmetric on U
- Does not really matter. Consider $y_i = \mathbb{1}(x_i \geq v_i) \cdot u_i$ with $\phi_y : (x, u, v) \mapsto (x, \mathbb{1}(x \geq v), y)$

Feedback Structure II

- Under full-feedback GMP is a standard bandit problem $\mathcal{O}(\sqrt{KB \ln B}) << \mathcal{O}(K^{2/3})$ [Cesa-Bianchi and Lugosi, 2006]
- Under realistic feedback we **conjecture** GMP is $\mathcal{O}(K)$ [Cesa-Bianchi et al., 2021] [Cesa-Bianchi et al., 2022] more
- Overall, not clear difficulty relation

$$S_i^{\mathrm{TMP}}(x_i) = \mathbb{1}(x_i \geq v_i)(u_i - x_i) + \lambda_2 \cdot u_i \cdot x \tag{12}$$

- Previous and future research can benefit from asymmetric feedback considerations (ubiquitous in Economics!). Example,
  - Gap in [Cesa-Bianchi et al., 2021] full $\mathcal{O}(\sqrt{K})$ vs realistic $\mathcal{O}(K)$[2]

back

[2] In the absence of strong assumptions on the DGP, namely independence across U, V and bounded densities

Information Requirements

- Discussion introduced in [Cesa-Bianchi et al., 2022]
- **Intuition:** Analyse the information requirements of the **objective function** and its **gradient**
- Use GMP in integral form $S_i^{\text{GMP}} = G_i^v(x_i) \cdot (u_i - x_i) + \lambda \int_0^x G_i^v(x') \; dx'$
- $\nabla S_i^{\text{GMP}} = G^{v'}(x) \cdot (u - x) - (1 - \lambda) \cdot G^v(x)$
- Thus the objective function depends globally on x, and its gradient locally. This makes the problem most similar to [Cesa-Bianchi et al., 2022] but more difficult!

back

Online Convex Optimisation I

- Interesting connection between OCO and (adversarial) bandits

$$\mathcal{R}_K^{\text{OCO}} = \sum_i^K f_i(x_i) - \min_{x \in \mathcal{K}} \sum_i^K f_i(x) \tag{13}$$

- where $\mathcal{K}$ is a convex decision set and f_i is a sequence of convex loss-function realisations
- Rewards and losses can be interchanged without loss
- Convexity of the policy space? Redefine the policy space as the B arm simplex Δ_B
- Rewrite expected losses as $\mathbb{E}[f_i(x)] = \sum_b p_{ib} \cdot f_{ib}(x)$, where p_{ib} is the probability of selecting arm b and f_{ib} is its associated one-period loss
- Highlights the importance of randomisation (once again)

Online Convex Optimisation II

- It enables the importation of OCO results like SGD
- Can we recover $\nabla f_i(x_i)$ using $f_i(x_i)$?[3] Yes! Key insight: Characterise the decision set as the simplex of the policy space

$$\tilde{\nabla}_{ib} = \frac{1}{p_{ib}} \cdot \nabla_p f_i(x_i) = \frac{1}{p_{ib}} \cdot \nabla_p (f_{ib} \cdot p_{ib}) = f_{ib}(x_i) \frac{\mathbb{1}(x_i = x_b)}{p_{ib}} \tag{14}$$

- We show that $\tilde{\nabla}_i$ is an unbiased estimate of the true gradient of $f_i(x_i)$

back

[3] In fact, the learner in the GMP does not even recover $f_i(x_i)$, but feedback ψ_i. In the paper we show that feedback ψ_i can be used to recover unbiased estimates of $\nabla f_i(x_i)$

Algorithm 1 as Penalised SGD

- Standard (constrained) SGD algorithms update by setting
$$z_{i+1} = x_i + \eta_i \cdot \tilde{\nabla}_i \tag{15}$$
- where η_i is the learning rate in i and $\tilde{\nabla}_i$ is the ith realisation of a rv $\tilde{\nabla}$ such that $\mathbb{E}[\tilde{\nabla}_i] = \nabla_i$
- Finally, it projects back by setting $x_{i+1} = \Pi_{\mathcal{K}}(z_{i+1})$
- In the GMP, the decision set is the probability simplex hence the softmax function $x_{i+1,b} = \frac{x_{ib} \cdot \exp(-\eta_i \tilde{\nabla}_{ib})}{\sum_b \exp(-\eta_i \tilde{\nabla}_{ib})}$ is a sensible projection
- The update step in Algorithm 1 follows these intuitions with $-\tilde{\nabla}_{ib} = \widehat{\mathbb{S}}_{ib} \frac{\mathbb{1}(b_i = b)}{p_{ib}}$ and $\eta_i = \eta$ (up to the regularisation term $\frac{\gamma}{B+1}$)

back

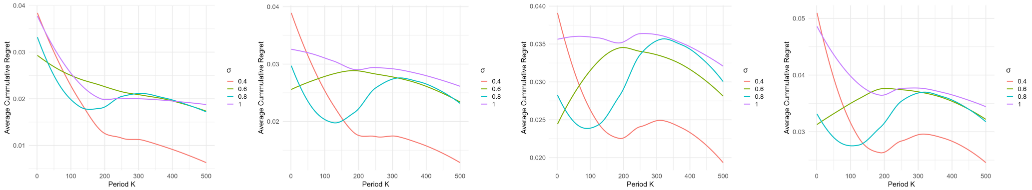
Greedy Parameter Selection I

- Optimality-ensuring parameters are in practice overly conservative
- Policymaker might be tempted to *speed-up* the process by selecting a more agressive learning rate η (i.e. ULDC: $0.132 >> 0.0027$)
- Things can go badly shall the policymaker ignore the variance of the DGP
- In very noisy DGP it is optimal to restrict the effect of single realisations[4]
- **Example:** Consider $U \sim \mathcal{U}[0,1]$ and $V = U + \phi_\sigma$ where $\phi_\sigma \sim N(0, \sigma^2)$ with

[4] From a theory standpoint, this intuition is correct provided that upper-bound derivation, mediated by the learning rate η, relies on bounding the second order moment of the gradient

Greedy Parameter Selection Graphical Evidence

Figure 11: Algorithm 1 given $U \sim \mathcal{U}[0,1]$, $V = u + \phi_\sigma$



(a) $\eta = 0.25$ (b) $\eta = 0.5$ (c) $\eta = 1$ (d) $\eta = 2$

1,000 simulations $\lambda = 0.7$ $K = 500$ $B = 10$ $\gamma = 0.029$

Greedy Parameter Selection II

- Increase in η leads to uniform decrease in performance
- More volatile DGP seem disproportionately affected
- Greedy parameter selection can lead to important reductions in welfare back

Algorithm II

Algorithm 2 Tempered Exp3 for the LE-GMP

Input B, λ, η, γ

Set $x_b = (b-1)/B$ for $b \in \{1, 2, \dots, B+1\}$, $\widehat{\mathbb{G}}_{1b} = 0$, $\widehat{\mathbb{U}}_{1b} = 0$

for $i = 1, 2, \dots, K$

  for $b = 1, \dots, B+1$

    Set $\widehat{\mathbb{S}}_{ib} = \widehat{\mathbb{U}}_{ib} - x_b \cdot \widehat{\mathbb{G}}_{ib} + \frac{\lambda}{B} \sum_{b'<b} \widehat{\mathbb{G}}_{ib'}$ $p_{ib} = (1-\gamma) \frac{\exp(\eta\ \widehat{\mathbb{S}}_{ib})}{\sum_{b'} \exp(\eta\ \widehat{\mathbb{S}}_{ib'})} + \frac{\gamma}{B+1}$

  end for

  Sample $b_i \sim p_{ib}$ and **observe** $\Theta_i = \theta$. **If** $\Theta_i = 1$

    Observe $\mathbb{1}(x_{b_i} \geq v_i)$, **If** $\mathbb{1}(x_{b_i} \geq v_i) = 1$ observe u_i

    for $b = 1, \dots, B+1$

      Update $\widehat{\mathbb{G}}_{i+1,b} = \widehat{\mathbb{G}}_{ib} + \mathbb{1}(x_{b_i} \geq v_i) \frac{\mathbb{1}(b_i = b)}{p_{ib} \cdot p_\theta}$ $\widehat{\mathbb{U}}_{i+1,b} = \widehat{\mathbb{U}}_{ib} + u_i \cdot \mathbb{1}(x_{b_i} \geq v_i) \frac{\mathbb{1}(b_i = b)}{p_{ib} \cdot p_\theta}$

    end for

  else $\widehat{\mathbb{G}}_{i+1,b} = \widehat{\mathbb{G}}_{ib}$, $\widehat{\mathbb{U}}_{i+1,b} = \widehat{\mathbb{U}}_{ib}$

end for back

Bibliography I

[Acharya and Wee, 2020] Acharya, S. and Wee, S. L. (2020).
Rational inattention in hiring decisions.
American Economic Journal: Macroeconomics, 12(1):1–40.

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

[Auer et al., 2002] Auer, P., Cesa-Bianchi, N., Freund, Y., and Schapire, R. E. (2002).
The nonstochastic multiarmed bandit problem.
SIAM journal on computing, 32(1):48–77.

[Björkegren et al., 2020] Björkegren, D., Blumenstock, J. E., and Knight, S. (2020).
Manipulation-proof machine learning.
arXiv preprint arXiv:2004.03865.

Bibliography II

[Bubeck et al., 2012] Bubeck, S., Cesa-Bianchi, N., et al. (2012).
Regret analysis of stochastic and nonstochastic multi-armed bandit problems.
Foundations and Trends® in Machine Learning, 5(1):1–122.

[Card et al., 2012] Card, D., Mas, A., Moretti, E., and Saez, E. (2012).
Inequality at work: The effect of peer salaries on job satisfaction.
American Economic Review, 102(6):2981–3003.

[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.

Bibliography III

[Cesa-Bianchi and Lugosi, 2006] Cesa-Bianchi, N. and Lugosi, G. (2006).
Prediction, learning, and games.
Cambridge university press.

[Doraszelski and Pakes, 2007] Doraszelski, U. and Pakes, A. (2007).
A framework for applied dynamic analysis in io.
Handbook of industrial organization, 3:1887–1966.

[Dube et al., 2016] Dube, A., Lester, T. W., and Reich, M. (2016).
Minimum wage shocks, employment flows, and labor market frictions.
Journal of Labor Economics, 34(3):663–704.

[Ericson and Pakes, 1995] Ericson, R. and Pakes, A. (1995).
Markov-perfect industry dynamics: A framework for empirical work.
The Review of economic studies, 62(1):53–82.

Bibliography IV

[Esponda, 2008] Esponda, I. (2008).
Behavioral equilibrium in economies with adverse selection.
American Economic Review, 98(4):1269–1291.

[Furman and Orszag, 2018] Furman, J. and Orszag, P. (2018).
1. a firm-level perspective on the role of rents in the rise in inequality.
In *Toward a Just Society*, pages 19–47. Columbia University Press.

[Gonzalez, 2023] Gonzalez, C. (2023).
Hiring decisions with knapsack bandits.

[Hart and Mas-Colell, 2000] Hart, S. and Mas-Colell, A. (2000).
A simple adaptive procedure leading to correlated equilibrium.
Econometrica, 68(5):1127–1150.

Bibliography V

[Hart and Mas-Colell, 2001a] Hart, S. and Mas-Colell, A. (2001a).
A general class of adaptive strategies.
Journal of Economic Theory, 98(1):26–54.

[Hart and Mas-Colell, 2001b] Hart, S. and Mas-Colell, A. (2001b).
A reinforcement procedure leading to correlated equilibrium.
Springer.

[Hart and Mas-Colell, 2013] Hart, S. and Mas-Colell, A. (2013).
Simple adaptive strategies: from regret-matching to uncoupled dynamics, volume 4.
World Scientific.

[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.

Bibliography VI

[Maćkowiak et al., 2023] Maćkowiak, B., Matějka, F., and Wiederholt, M. (2023).
Rational inattention: A review.
Journal of Economic Literature, 61(1):226–273.

[Manning, 2003] Manning, A. (2003).
The real thin theory: monopsony in modern labour markets.
Labour economics, 10(2):105–131.

[Manning, 2021] Manning, A. (2021).
Monopsony in labor markets: A review.
ILR Review, 74(1):3–26.

[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.

Bibliography VII

[Ramey, 2016] Ramey, V. A. (2016).
Macroeconomic shocks and their propagation.
Handbook of macroeconomics, 2:71–162.

[Sims, 2003] Sims, C. A. (2003).
Implications of rational inattention.
Journal of monetary Economics, 50(3):665–690.

[Van Nieuwerburgh and Veldkamp, 2010] Van Nieuwerburgh, S. and Veldkamp, L. (2010).
Information acquisition and under-diversification.
The Review of Economic Studies, 77(2):779–805.