

# MS&E 233

# Game Theory, Data Science and AI

## Lecture 10

Vasilis Syrgkanis

Assistant Professor

Management Science and Engineering

(by courtesy) Computer Science and Electrical Engineering

Institute for Computational and Mathematical Engineering

# Computational Game Theory for Complex Games

- 1
  - Basics of game theory and zero-sum games (T)
  - Basics of online learning theory (T)
  - Solving zero-sum games via online learning (T)
  - *HW1: implement simple algorithms to solve zero-sum games*
  - Applications to ML and AI (T+A)
  - *HW2: implement boosting as solving a zero-sum game*

- 2
  - Basics of extensive-form games
  - Solving extensive-form games via online learning (T)
  - *HW3: implement agents to solve very simple variants of poker*

- 3
  - General games, equilibria and online learning (T)
  - Online learning in general games
  - *HW4: implement no-regret algorithms that converge to correlated equilibria in general games*

## Data Science for Auctions and Mechanisms

- 4
  - Basics and applications of auction theory (T+A)
  - **Basic Auctions and Learning to bid in auctions (T)**
  - *HW5: implement bandit algorithms to bid in ad auctions*

- 5
  - **Optimal auctions and mechanisms (T)**
  - **Simple vs optimal mechanisms (T)**
  - *HW6: calculate equilibria in simple auctions, implement simple and optimal auctions, analyze revenue empirically*

- 6
  - **Optimizing mechanisms from samples (T)**
  - **Online optimization of auctions and mechanisms (T)**
  - *HW7: implement procedures to learn approximately optimal auctions from historical samples and in an online manner*

## Further Topics

- 7
  - **Econometrics in games and auctions (T+A)**
  - **A/B testing in markets (T+A)**
  - *HW8: implement procedure to estimate values from bids in an auction, empirically analyze inaccuracy of A/B tests in markets*

## Guest Lectures

- Mechanism Design for LLMs, Renato Paes Leme, Google Research
- Auto-bidding in Sponsored Search Auctions, Kshipra Bhawalkar, Google Research

# ***Sum: Auction Applications***

- Traditionally, selling of luxury goods, art
- Digital auction markets for goods (eBay)
- Energy markets
- Digital ad markets (sponsored search, display ads, amazon ads)
- Spectrum auctions
- Government procurement auctions
- Web3.0 transaction protocols

# ***Sum: First Price***

- First Price is arguably the simplest auction rule
- It can be hard to strategize in such an auction
- The auction can lead to inefficient allocations
- Though approximately efficient
- Still used in practice in many settings (e.g. online advertising, government procurement)
- Primarily because it has very transparent rules

# ***Sum: Second Price***

- Second Price is arguably the simplest truthful auction rule
  - It is very easy to strategize in such an auction (be truthful)
  - Auction always leads to efficient allocations (highest value wins)
  - Auction can be run very quickly (computationally efficient)
- 
- Still not always the auction used in many places
  - Primarily because it has not very transparent rules
  - Susceptible to collusion and manipulations by the auctioneer

# Sponsored Search Auctions

# Sponsored Search Auctions

- Now we have many items to sell
- Slots on a web impression
- Higher slots get more clicks!
- Each slot has some probability of click
$$a_1 > a_2 > \dots > a_m$$
- Bidders have a value-per-click  $v_i$

Google

digital advertising

All Images News Videos Shopping More

About 6,620,000,000 results (0.44 seconds)

**Sponsored**

Reddit  
https://www.redditforbusiness.com

[Advertise on Reddit](#)

Reach over 100K communities — Connect with passionate communities that deliver results for brands across all industries. Create impact & own top communities in your target category for 24 hours. Try Reddit ads.

**Sponsored**

Microsoft  
https://about.ads.microsoft.com › advertising › start-now

[Microsoft Advertising® | Get a \\$500 Advertising Credit](#)

We'll Help You Find Your Customers and Reach Searchers Across The Microsoft Network. Plus, Receive a \$500 Microsoft **Advertising** Credit When You Spend Just \$250! Free Sign Up.

**Sponsored**

coseom  
https://www.coseom.com

[Pay Per Click Company](#)

COSEOM™ — Generate Leads For Your Business Using Advanced PPC Strategies. Request A Proposal Today!

**Sponsored**

Simpli.fi  
https://www.simpli.fi

[Simpli.fi | Advertising Success Platform](#)

Established in 2010 — Enjoy the perks of multi-channel targeting, measurement, & reporting with our interface. CTV.

# Generalized First Price (GFP) Auction

- Bidders submit a bid-per-click  $b_i$
- Slots allocated in decreasing order of bids
- Bidder  $i$  is allocated slot  $j_i(b)$
- Bidder pays their bid when clicked

$$u_i(b; v_i) = a_{j_i(b)} \cdot (v_i - b_i)$$

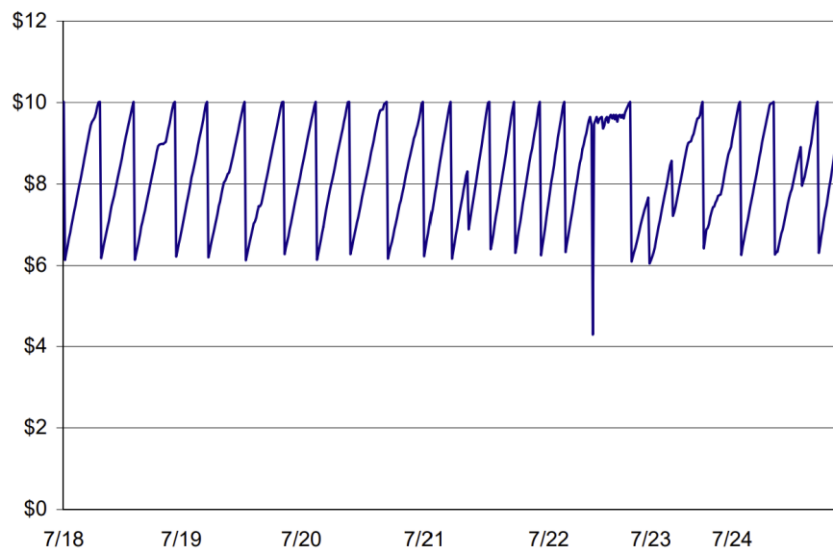
Google search results for "digital advertising". The results show four sponsored ads, each with a bid  $b_i$  and a value  $a_i$ .

Bid	Advertiser	Value
$b_{(1)}$	Reddit	$a_1$
$b_{(2)}$	Microsoft	$a_2$
$b_{(3)}$	COSEOM	$a_3$
$b_{(4)}$	Simpli.fi	$a_4$



# Generalized First Price (GFP) Auction

- The first auction that was used by Overture in late 90s
- Lead to weird bidding patterns



(b) 1 week

Google digital advertising

About 6,620,000,000 results (0.44 seconds)

**Sponsored**

**Reddit**  
https://www.redditforbusiness.com  
**Advertise on Reddit**  
Reach over 100K communities — Connect with passionate communities that deliver results for brands across all industries. Create impact & own top communities in your target category for 24 hours. Try Reddit ads.

**Sponsored**

**Microsoft**  
https://about.ads.microsoft.com › advertising › start-now  
**Microsoft Advertising® | Get a \$500 Advertising Credit**  
We'll Help You Find Your Customers and Reach Searchers Across The Microsoft Network. Plus, Receive a \$500 Microsoft **Advertising** Credit When You Spend Just \$250! Free Sign Up.

**Sponsored**

**coseom**  
https://www.coseom.com  
**Pay Per Click Company**  
COSEOM™ — Generate Leads For Your Business Using Advanced PPC Strategies. Request A Proposal Today!

**Sponsored**

**Simpli.fi**  
https://www.simpli.fi  
**Simpli.fi | Advertising Success Platform**  
Established in 2010 — Enjoy the perks of multi-channel targeting, measurement, & reporting with our interface. CTV.

$a_1$

$a_2$

$a_3$

$a_4$

$b_{(1)}$

$b_{(2)}$

$b_{(3)}$

$b_{(4)}$

IV

IV

IV

# Generalized Second Price (GSP) Auction

- Bidders submit a bid-per-click  $b_i$
- Slots allocated in decreasing order of bids
- Bidder  $i$  is allocated slot  $j_i(b)$
- Bidder pays the next highest bid when clicked

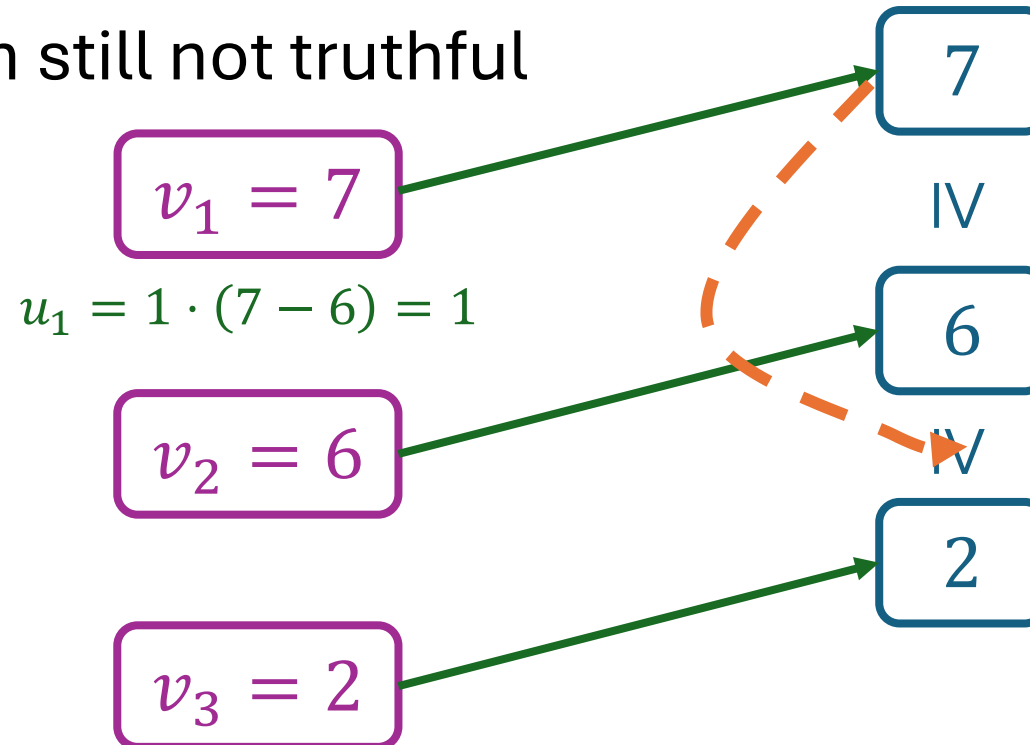
$$u_i(b; v_i) = a_{j_i(b)} \cdot (v_i - b_{(j_i(b)+1)})$$

The screenshot shows Google search results for the query "digital advertising". The results are sorted by relevance, but the top four results are marked as "Sponsored" and are highlighted with red boxes, indicating they are the top four highest bids in the auction. To the left of the search results, four blue boxes labeled  $b_{(1)}$ ,  $b_{(2)}$ ,  $b_{(3)}$ , and  $b_{(4)}$  are stacked vertically, with a Roman numeral "IV" between each box, representing the bids of the top four bidders. To the right of the search results, four green boxes labeled  $a_1$ ,  $a_2$ ,  $a_3$ , and  $a_4$  are stacked vertically, representing the ad prices (the next highest bid) for each slot. The sponsored results are:

- Slot 1:** Sponsored by Reddit. Ad title: "Advertise on Reddit". Description: "Reach over 100K communities — Connect with passionate communities that deliver results for brands across all industries. Create impact & own top communities in your target category for 24 hours. Try Reddit ads." Price:  $a_1$ .
- Slot 2:** Sponsored by Microsoft. Ad title: "Microsoft Advertising® | Get a \$500 Advertising Credit". Description: "We'll Help You Find Your Customers and Reach Searchers Across The Microsoft Network. Plus, Receive a \$500 Microsoft Advertising Credit When You Spend Just \$250! Free Sign Up." Price:  $a_2$ .
- Slot 3:** Sponsored by coseom. Ad title: "Pay Per Click Company". Description: "COSEOM™ — Generate Leads For Your Business Using Advanced PPC Strategies. Request A Proposal Today!" Price:  $a_3$ .
- Slot 4:** Sponsored by Simpli.fi. Ad title: "Simpli.fi | Advertising Success Platform". Description: "Established in 2010 — Enjoy the perks of multi-channel targeting, measurement, & reporting with our interface. CTV." Price:  $a_4$ .

# Generalized Second Price (GSP) Auction

- The auction of choice in current sponsored search systems
- Even though still not truthful



Google digital advertising

All Images News Videos Shopping More

About 6,620,000,000 results (0.44 seconds)

**Sponsored**

Reddit  
https://www.redditforbusiness.com

**Advertise on Reddit**

Reach over 100K communities — Connect with passionate communities that deliver results for brands across all industries. Create impact & own top communities in your target category for 24 hours. Try Reddit ads.

**Sponsored**

Microsoft  
https://about.ads.microsoft.com › advertising › start-now

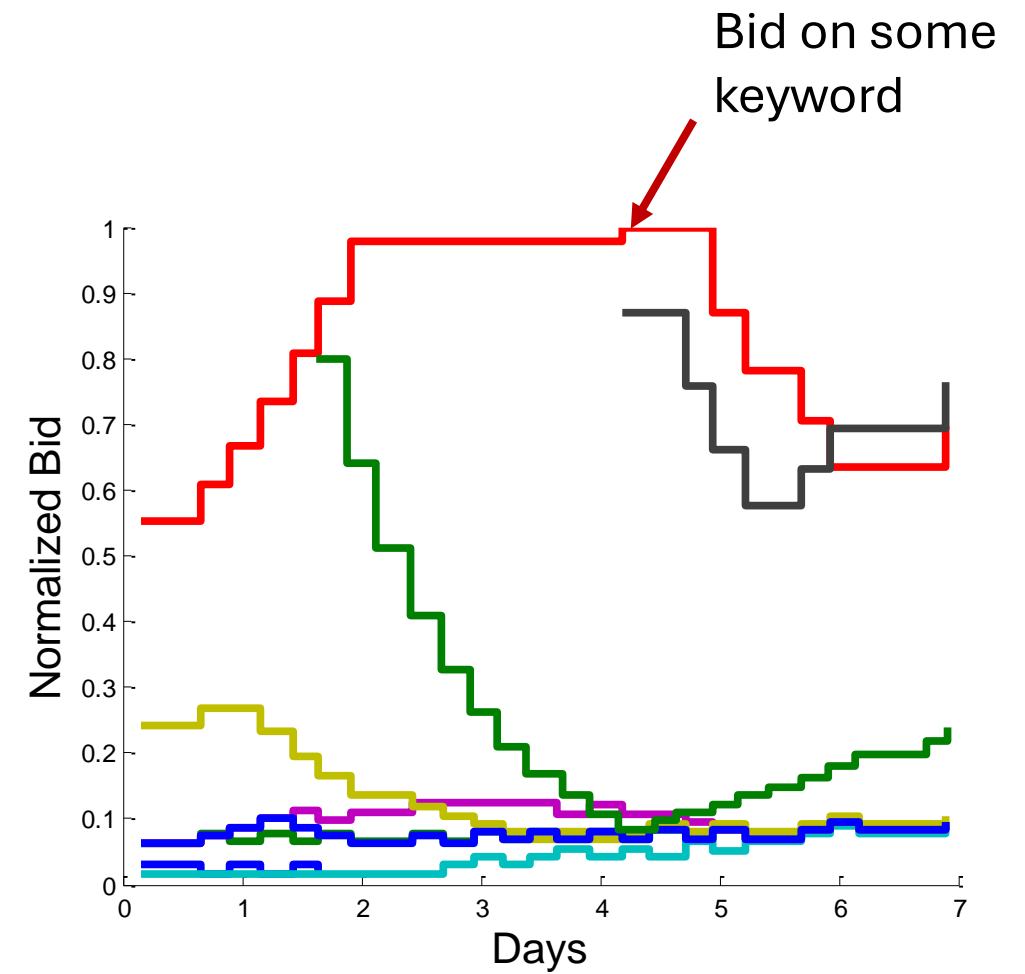
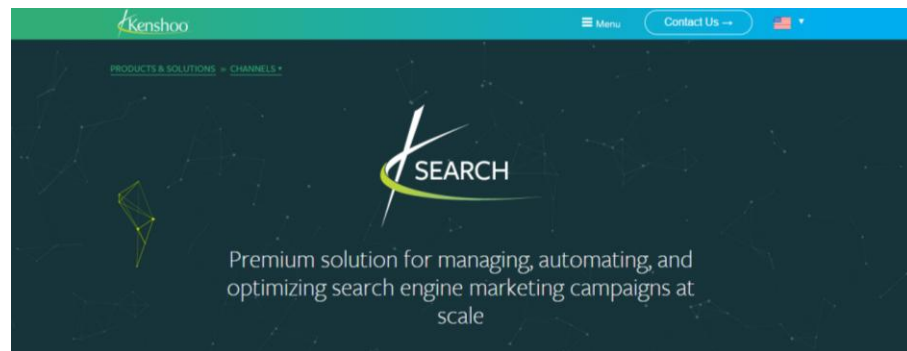
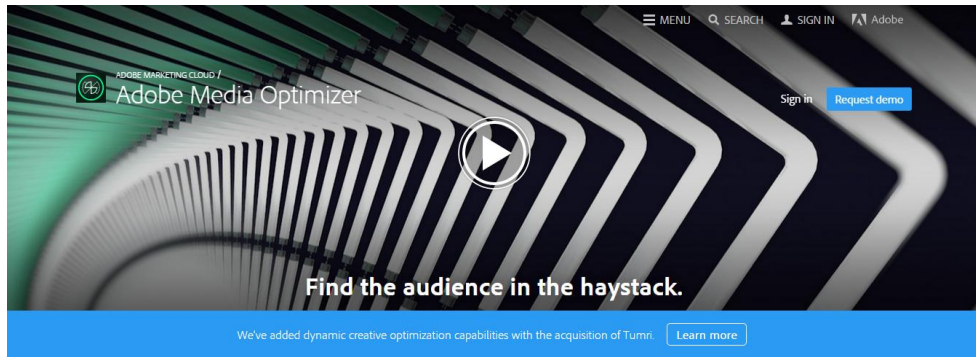
**Microsoft Advertising® | Get a \$500 Advertising Credit**

We'll Help You Find Your Customers and Reach Searchers Across The Microsoft Network. Plus, Receive a \$500 Microsoft **Advertising** Credit When You Spend Just \$250! Free Sign Up.

1

0.5

$u'_1 = .5 \cdot (7 - 2) = 2.5$



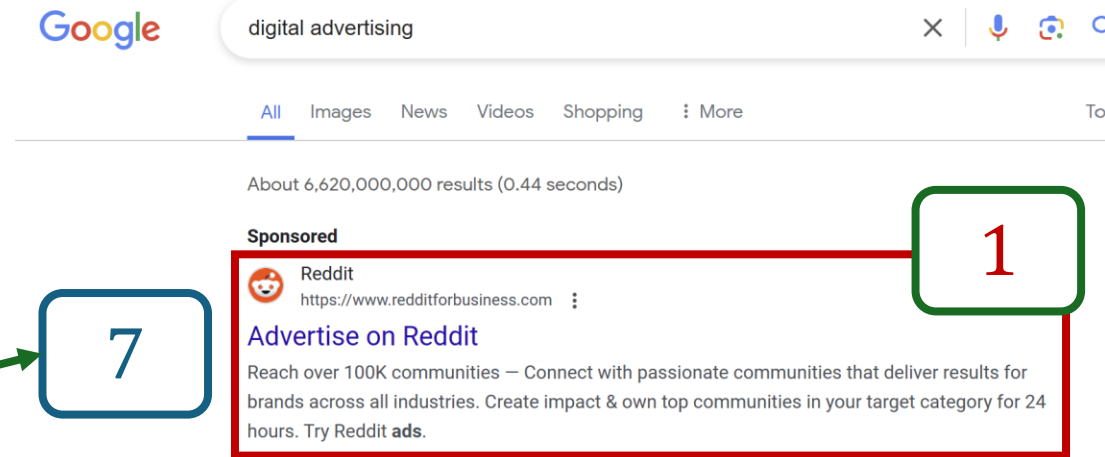
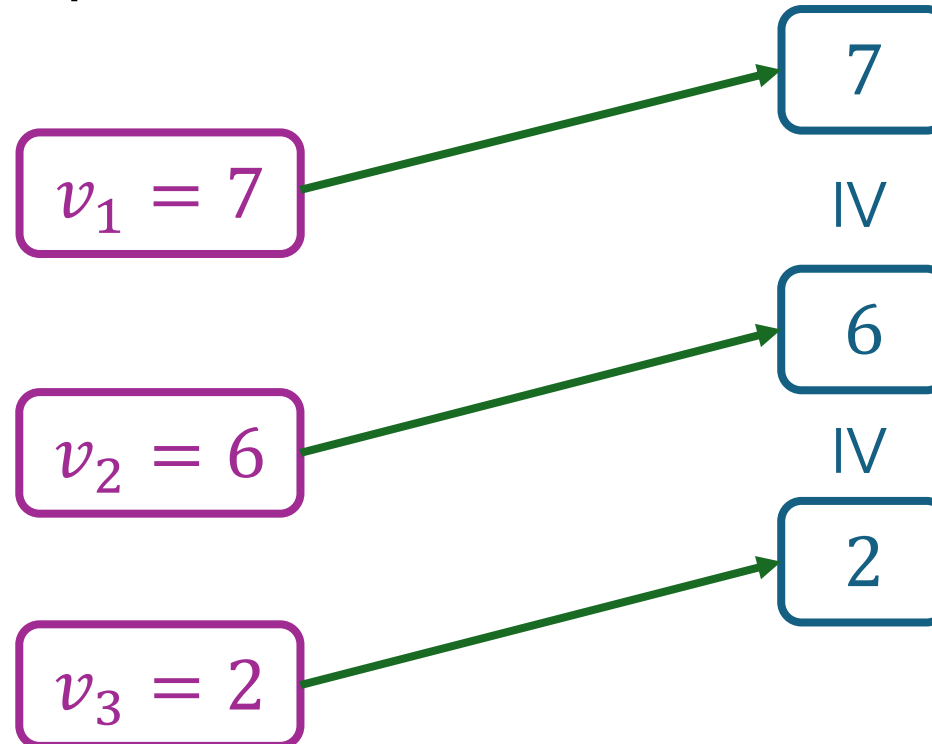
How would you turn GSP  
truthful?

# Right intuition, why Second-Price is truthful

- Second price is truthful **not because** we charge next highest bid
- Second price is truthful **not because** we charge smallest bid to maintain the same allocation
- Second price is truthful **because** we charged the winner their “externalities to the rest of society”

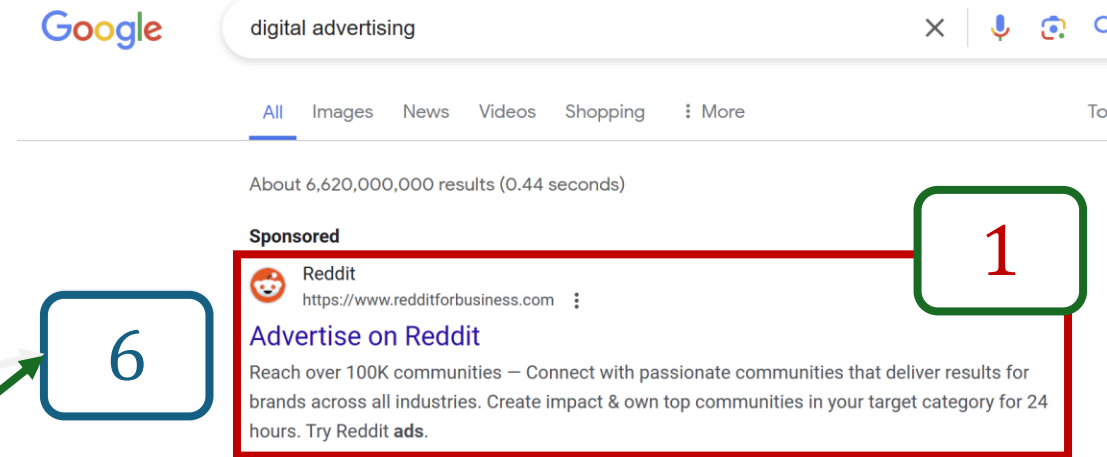
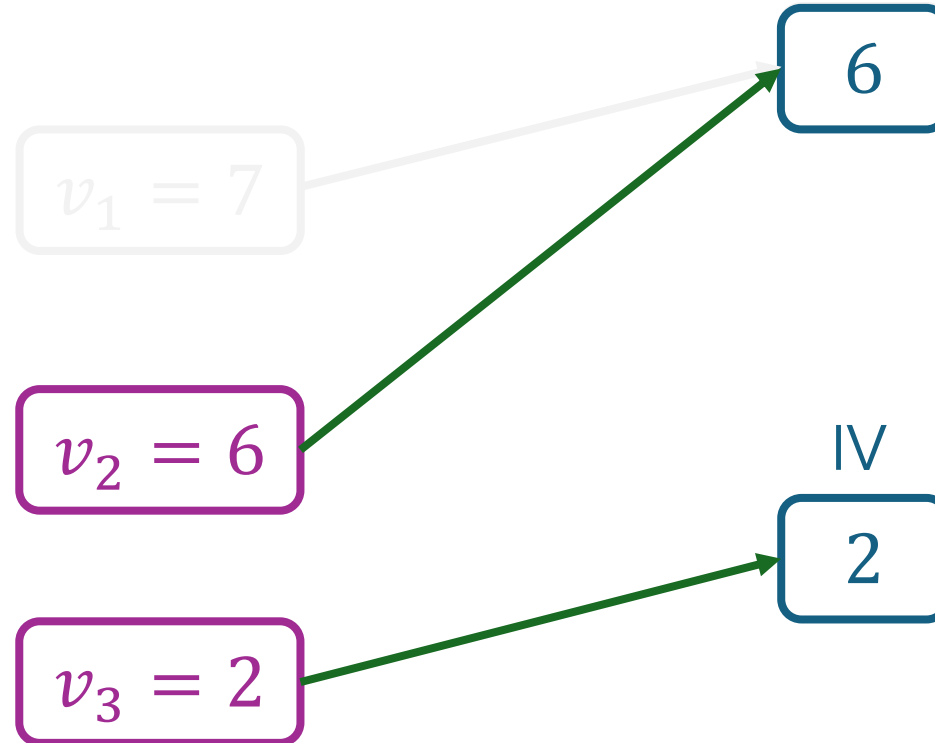
# The Deep Reason why SP is Truthful

- When highest bidder exists, rest of players achieve reported welfare of 0



# The Deep Reason why SP is Truthful

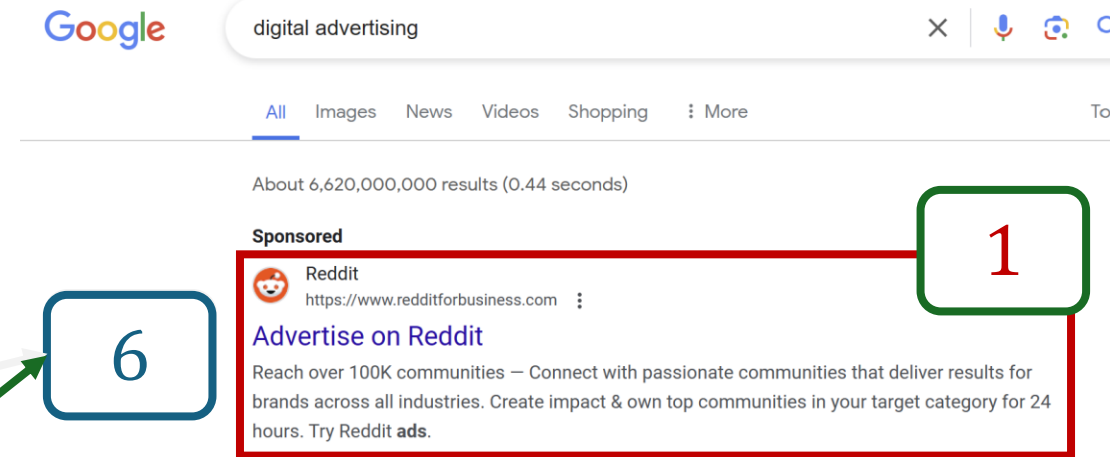
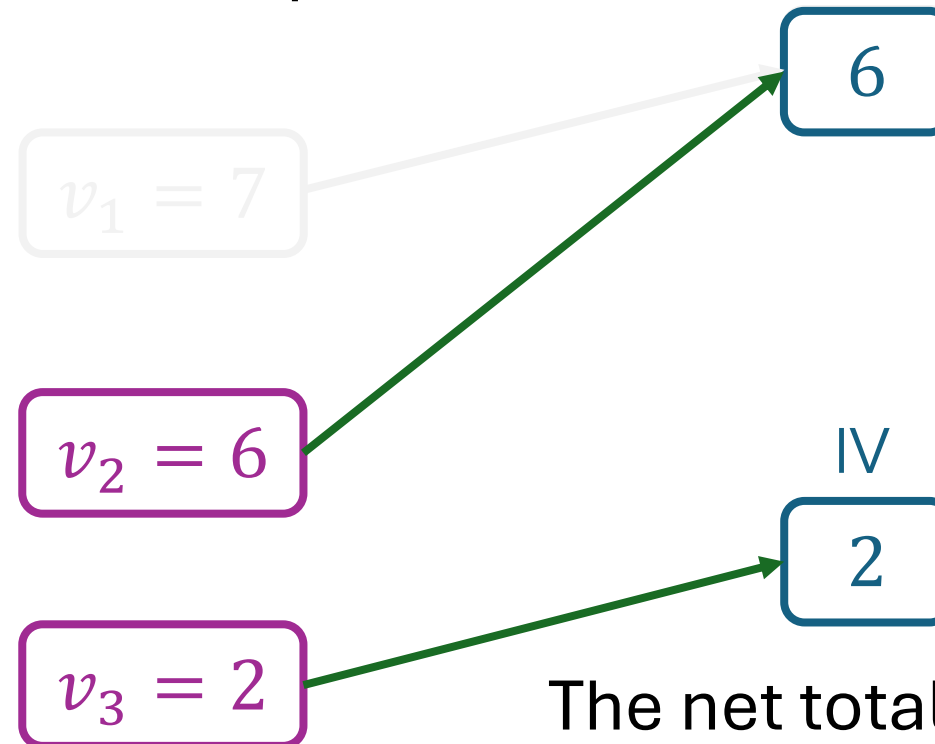
- When highest bidder does not exist, rest of players achieve reported welfare of 6





# The Deep Reason why SP is Truthful

- When highest bidder does not exist, rest of players achieve reported welfare of 6



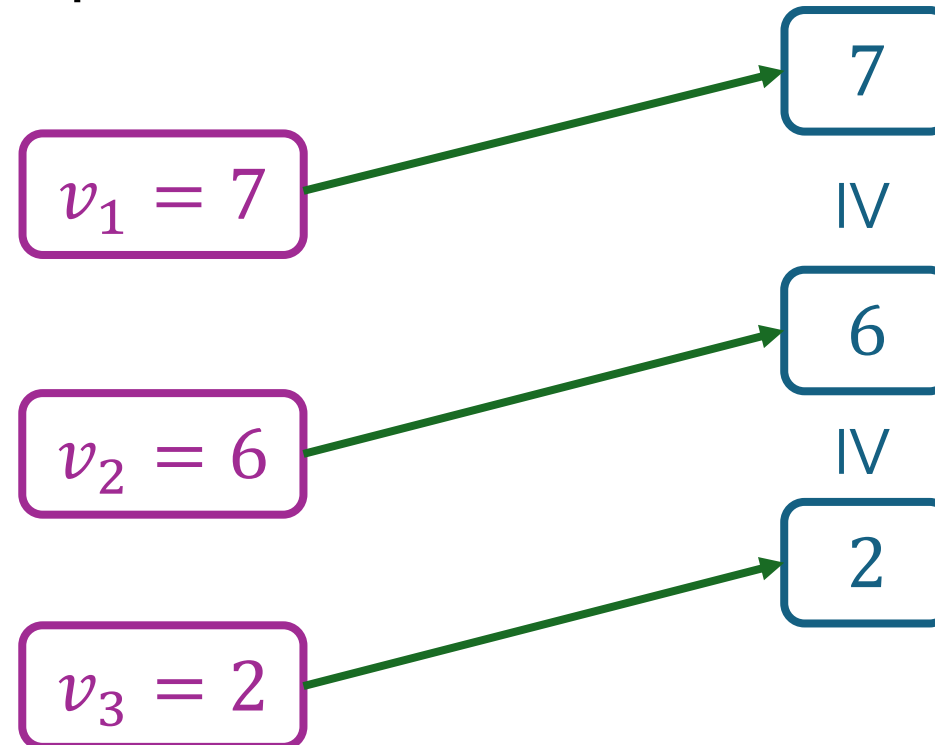
The net total gain to the rest of the bidders, from bidder 1 vanishing is 6

# Right intuition, why Second-Price is truthful

- Second price is truthful **because** we charged the winner their “externalities to the rest of society”
- When highest bidder exists, rest of players achieve reported welfare 0
- When highest bidder vanishes, rest of players achieve reported welfare  
 $b_{(2)}$  = second highest bid
- The net total gain to the rest of the bidders, from bidder 1 vanishing is  
 $b_{(2)}$  = second highest bid
- That’s what we should charge the winner!

# Let's repeat this exercise with two slots

- When highest bidder exists, rest of players achieve reported welfare of ...?



Google search results for "digital advertising". The top two sponsored results are highlighted with red boxes, corresponding to the highest and second-highest bidders from the diagram.

Player	Bid	Ad Content
1	1	<b>Sponsored</b> Reddit https://www.redditforbusiness.com <b>Advertise on Reddit</b> Reach over 100K communities — Connect with passionate communities that deliver results for brands across all industries. Create impact & own top communities in your target category for 24 hours. Try Reddit ads.
0.5	0.5	<b>Sponsored</b> Microsoft https://about.ads.microsoft.com › advertising › start-now <b>Microsoft Advertising®   Get a \$500 Advertising Credit</b> We'll Help You Find Your Customers and Reach Searchers Across The Microsoft Network. Plus, Receive a \$500 Microsoft <b>Advertising</b> Credit When You Spend Just \$250! Free Sign Up.
2	2	(Partially visible)

When the highest value bidder exists the rest of the players get a reported welfare of

1

2

3

4

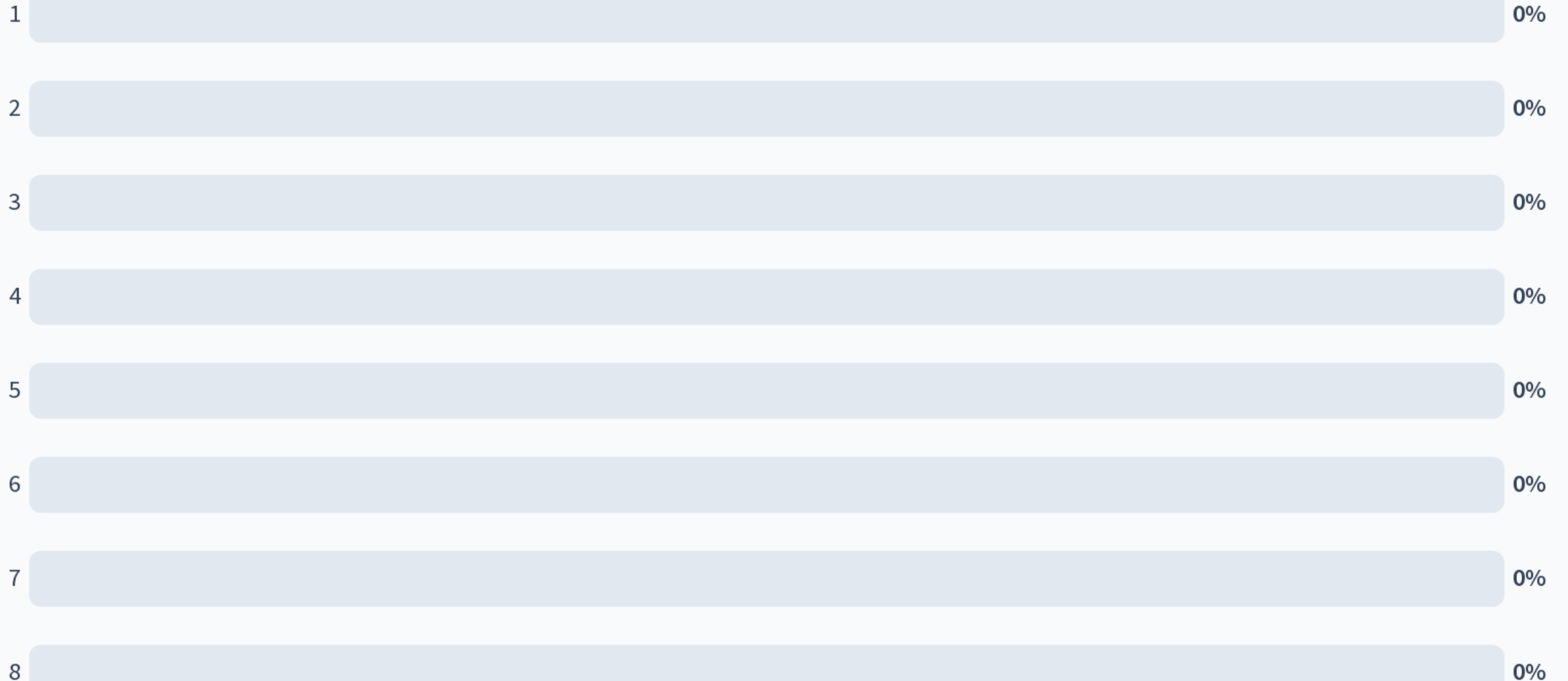
5

6

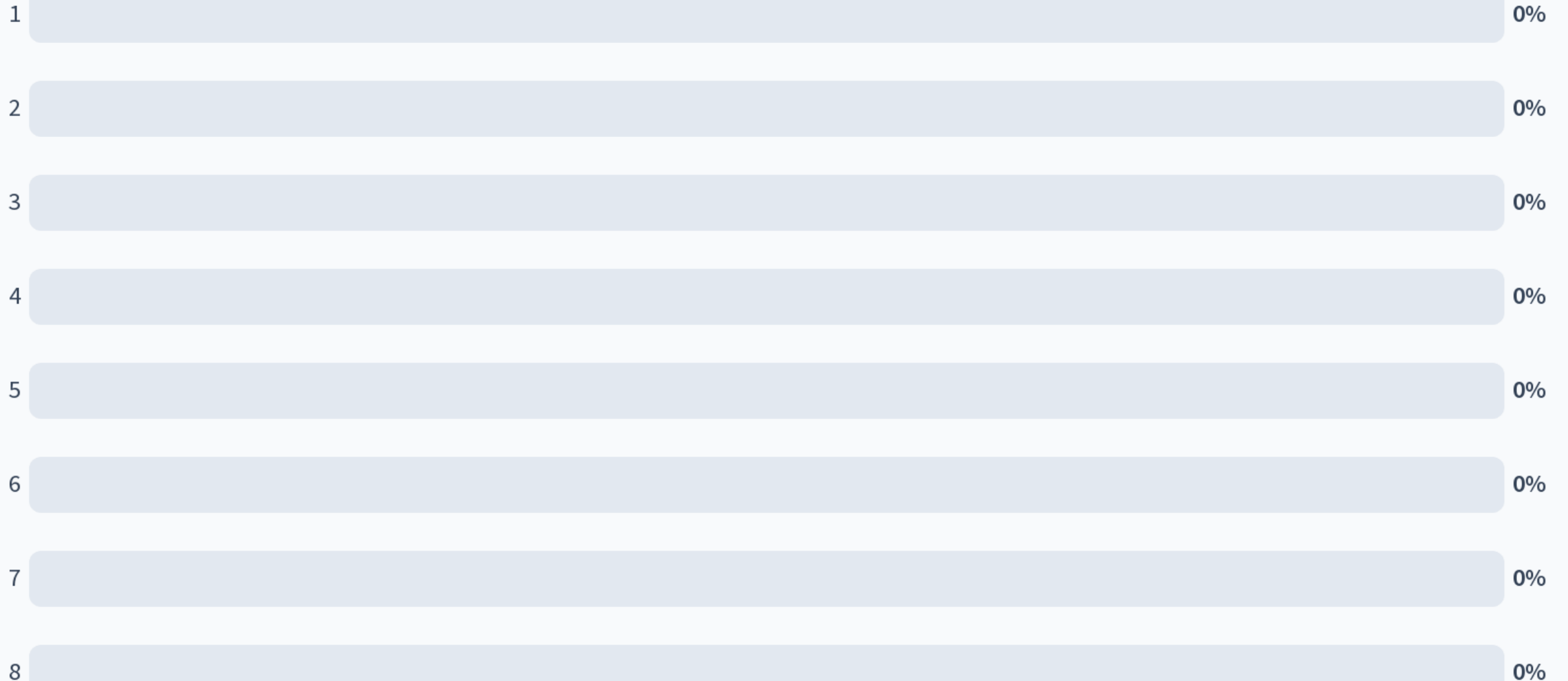
7

8

When the highest value bidder exists the rest of the players get a reported welfare of

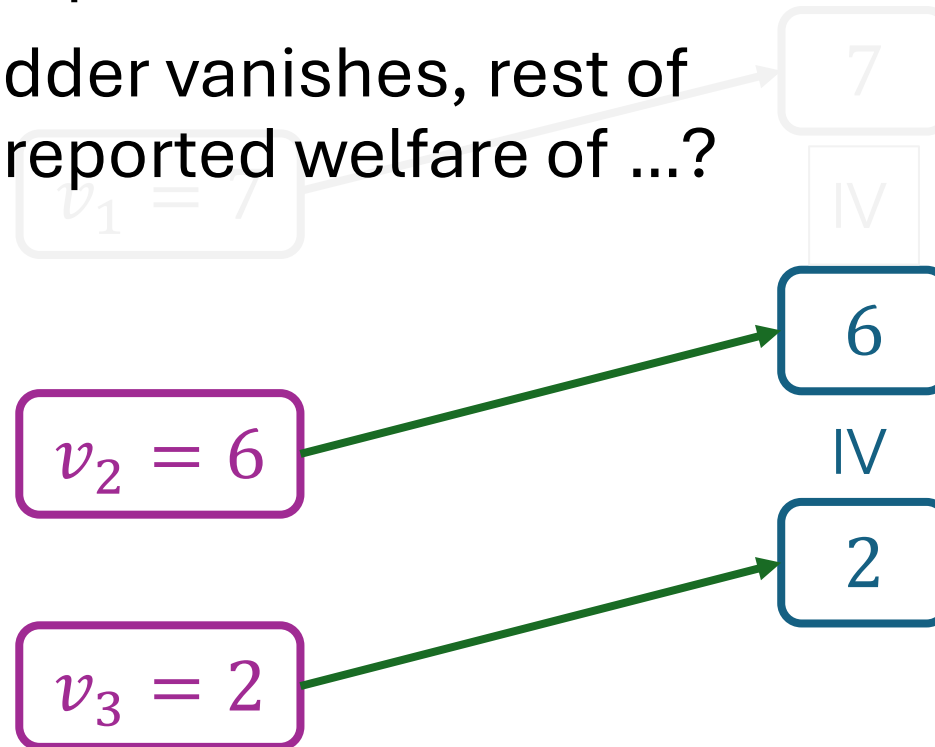


When the highest value bidder exists the rest of the players get a reported welfare of



# Let's repeat this exercise with two slots

- When highest bidder exists, rest of players achieve reported welfare of ...?
- When highest bidder vanishes, rest of players achieve reported welfare of ...?



The screenshot shows a Google search results page for the query "digital advertising". The search bar at the top shows the query and the Google logo. Below the search bar, there are tabs for "All", "Images", "News", "Videos", "Shopping", and "More". The search results show "About 6,620,000,000 results (0.44 seconds)". Two sponsored ads are visible:

- Sponsored** (Reddit): <https://www.redditforbusiness.com>. **Advertise on Reddit**. Reach over 100K communities — Connect with passionate communities that deliver results for brands across all industries. Create impact & own top communities in your target category for 24 hours. Try Reddit ads. (Welfare: 1)
- Sponsored** (Microsoft): <https://about.ads.microsoft.com> > advertising > start-now. **Microsoft Advertising® | Get a \$500 Advertising Credit**. We'll Help You Find Your Customers and Reach Searchers Across The Microsoft Network. Plus, Receive a \$500 Microsoft **Advertising** Credit When You Spend Just \$250! Free Sign Up. (Welfare: 0.5)

When the highest value bidder vanishes the rest of the players get a reported welfare of

1

2

3

4

5

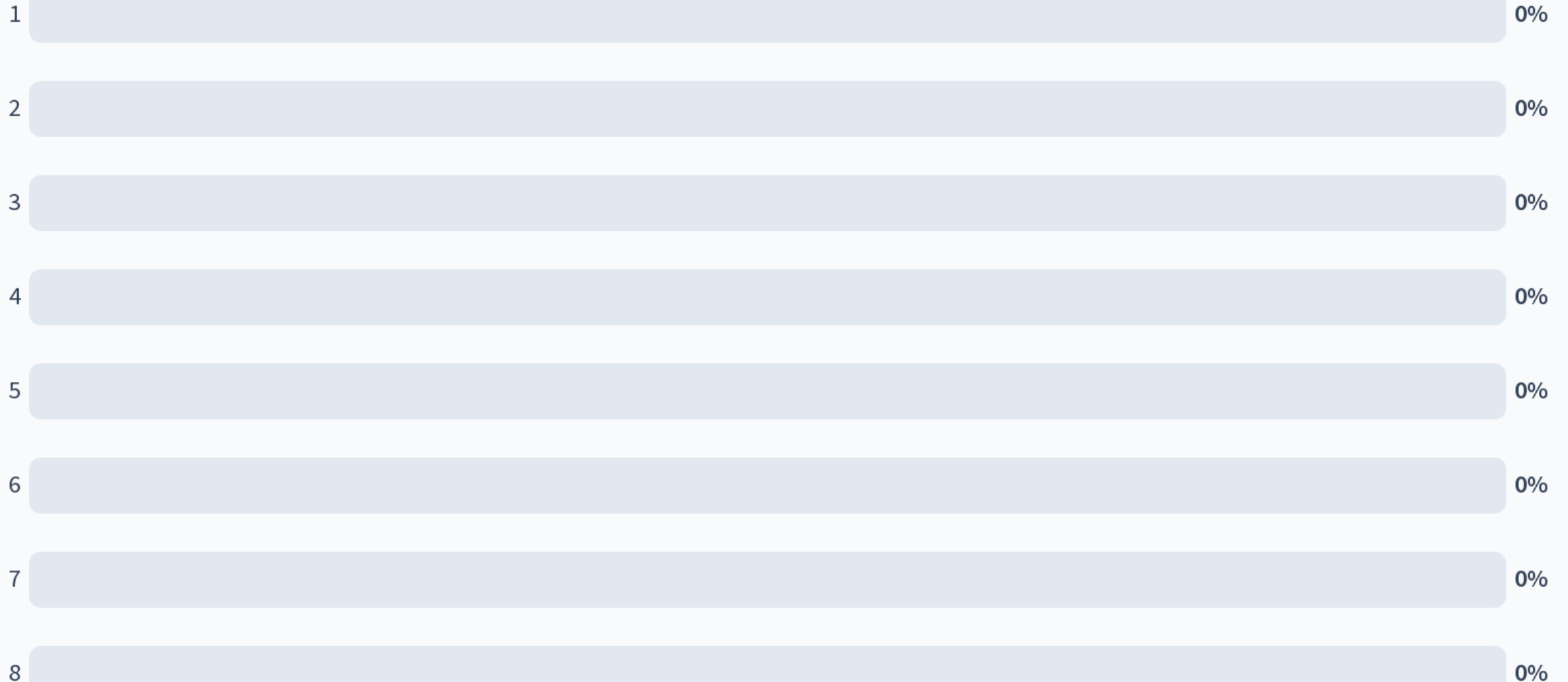
6

7

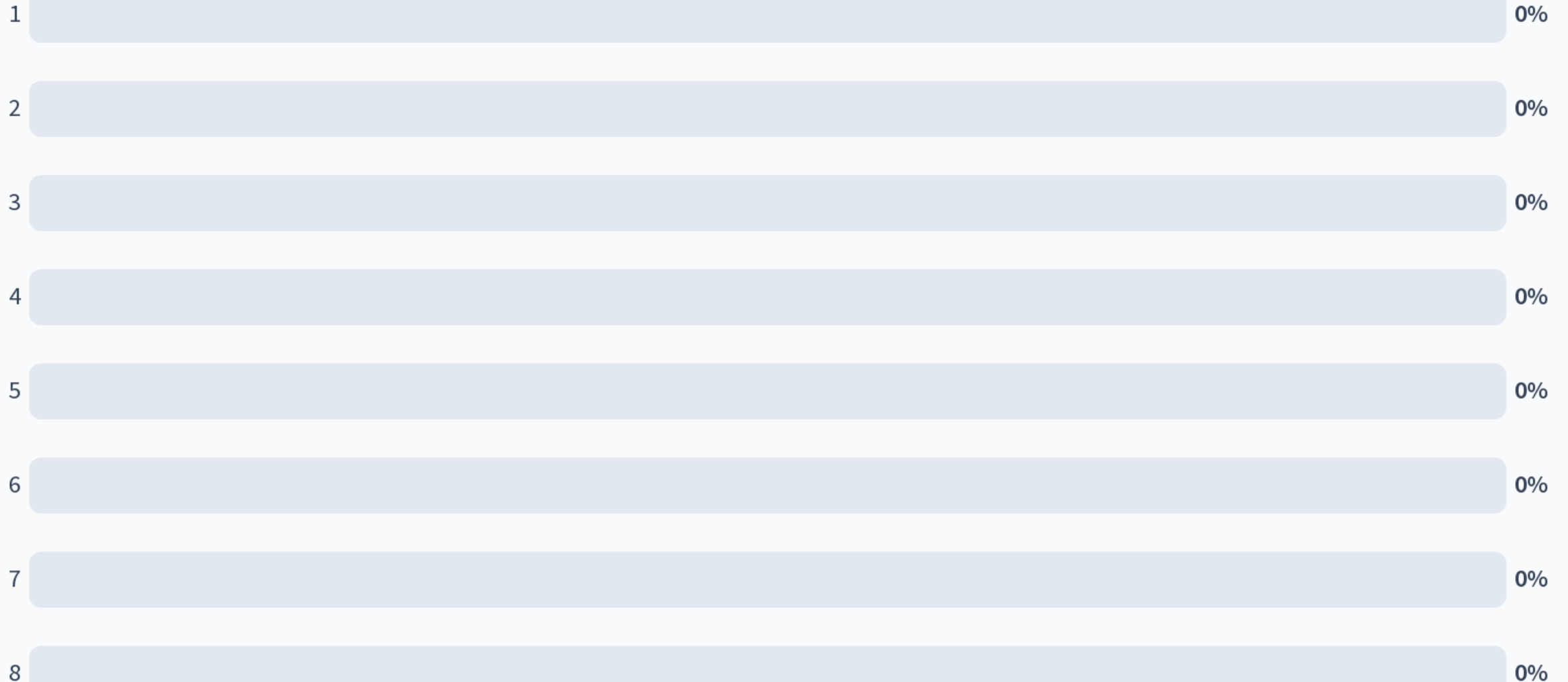
8



When the highest value bidder vanishes the rest of the players get a reported welfare of

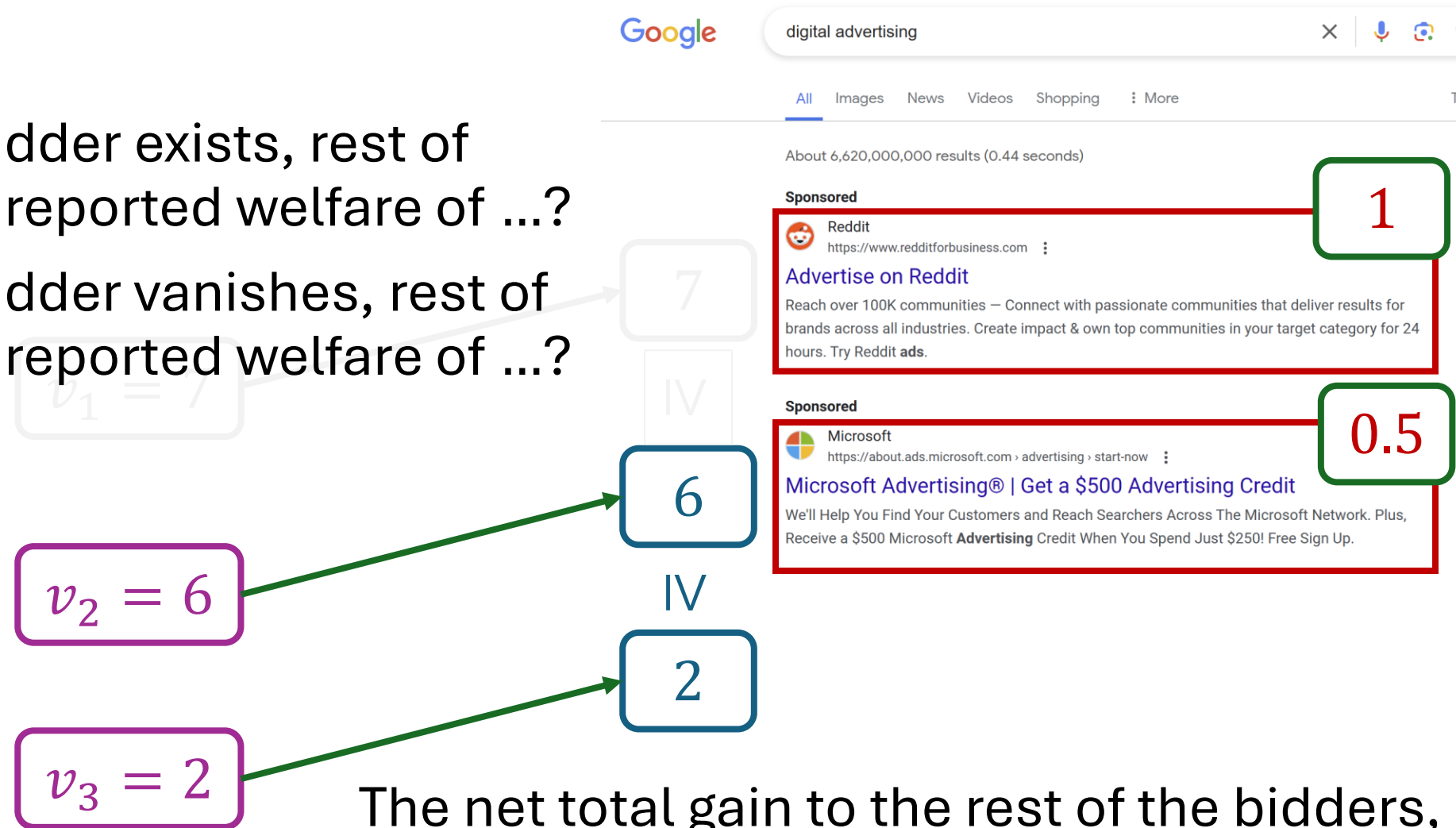


When the highest value bidder vanishes the rest of the players get a reported welfare of



# Let's repeat this exercise with two slots

- When highest bidder exists, rest of players achieve reported welfare of ...?
- When highest bidder vanishes, rest of players achieve reported welfare of ...?



The net total gain to the rest of the bidders, from bidder 1 vanishing is 4

# What about the second highest bidder?

- When second highest bidder exists, rest of players achieve reported welfare of 7

$$v_1 = 7$$

7

IV

- When second highest bidder vanishes, rest of players achieve reported welfare of  $7 + 1$

$$v_3 = 2$$

6

IV

2

Google digital advertising

All Images News Videos Shopping More

About 6,620,000,000 results (0.44 seconds)

**Sponsored**

Reddit  
https://www.redditforbusiness.com

**Advertise on Reddit**

Reach over 100K communities — Connect with passionate communities that deliver results for brands across all industries. Create impact & own top communities in your target category for 24 hours. Try Reddit ads.

**Sponsored**

Microsoft  
https://about.ads.microsoft.com › advertising › start-now

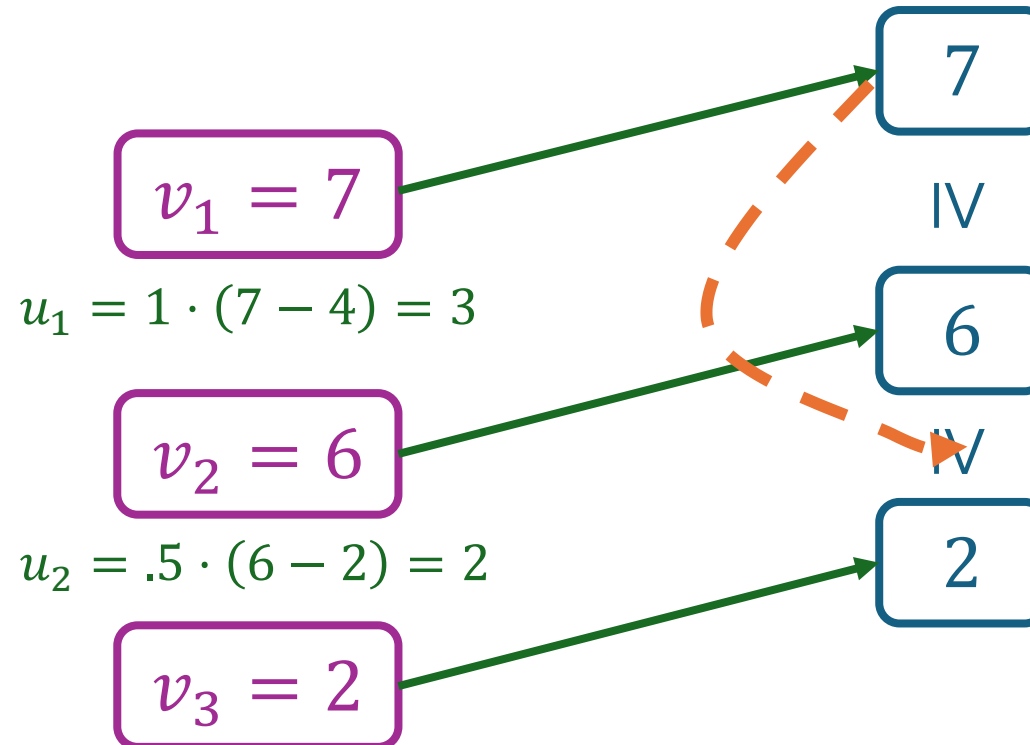
**Microsoft Advertising® | Get a \$500 Advertising Credit**

We'll Help You Find Your Customers and Reach Searchers Across The Microsoft Network. Plus, Receive a \$500 Microsoft **Advertising** Credit When You Spend Just \$250! Free Sign Up.

I should charge a total price of 1  
(equivalently a price-per-click of 2)

# Bidders now don't have incentive to deviate

- Unlike GSP, highest bidder doesn't prefer reducing the bid to get the second slot



Google digital advertising

All Images News Videos Shopping More

About 6,620,000,000 results (0.44 seconds)

Sponsored

Reddit  
https://www.redditforbusiness.com

Advertise on Reddit

Reach over 100K communities — Connect with passionate communities that deliver results for brands across all industries. Create impact & own top communities in your target category for 24 hours. Try Reddit ads.

Sponsored

Microsoft  
https://about.ads.microsoft.com › advertising › start-now

Microsoft Advertising® | Get a \$500 Advertising Credit

We'll Help You Find Your Customers and Reach Searchers Across The Microsoft Network. Plus, Receive a \$500 Microsoft Advertising Credit When You Spend Just \$250! Free Sign Up.

$u'_1 = .5 \cdot (7 - 2) = 2.5$

# How much utility do bidders receive?

$$\begin{aligned}\text{Externality} &= \text{RWelfare of Others without me} - \overbrace{\text{RWelfare of Others with me}}^{\text{Reported Welfare}} \\ \text{Utility} &= \text{Value of my Allocation} - \text{Payment}\end{aligned}$$

# How much utility do bidders receive?

$$\begin{aligned}\text{Externality} &= \text{RWelfare of Others without me} - \overset{\text{Reported Welfare}}{\text{RWelfare of Others with me}} \\ \text{Utility} &= \text{Value of my Allocation} - \text{Payment}\end{aligned}$$

If we set payment = externality

$$\text{Value of my Allocation} - \text{RWelfare of Others without me} + \text{RWelfare of Others with me}$$

# How much utility do bidders receive?

$$\text{Externality} = \text{RWelfare of Others without me} - \overset{\text{Reported Welfare}}{\boxed{\text{RWelfare of Others with me}}}$$

$$\text{Utility} = \text{Value of my Allocation} - \text{Payment}$$

If we set payment = externality

$$\text{Value of my Allocation} - \text{RWelfare of Others without me} + \text{RWelfare of Others with me}$$

When I'm truthful:

$$\text{Value of my Allocation} + \text{RWelfare of Others with me} = \text{Total RWelfare with me}$$



# How much utility do bidders receive?

$$\text{Externality} = \text{RWelfare of Others without me} - \overset{\text{Reported Welfare}}{\boxed{\text{RWelfare of Others with me}}}$$

$$\text{Utility} = \text{Value of my Allocation} - \text{Payment}$$

If we set payment = externality

$$\text{Value of my Allocation} - \text{RWelfare of Others without me} + \text{RWelfare of Others with me}$$

When I'm truthful:

$$\text{Value of my Allocation} + \text{RWelfare of Others with me} = \text{Total RWelfare with me}$$

When I'm truthful my utility is as simple as:

$$\text{Utility} = \text{Total RWelfare with me} - \text{Total RWelfare without me}$$

# Can we ever charge bidders more than value?

- If we set payment = externality, and bidder is truthful

$$\text{Utility} = \text{Total RWelfare with me} - \text{Total RWelfare without me}$$

- If the auction always chooses the outcome that maximizes the reported welfare, then

$$\text{Total RWelfare with me} \geq \text{Total RWelfare without me}$$

# Why is the mechanism truthful?

- If we set payment = externality, and bidder is truthful

$$\mathbf{Utility} = \text{Total RWelfare with me} - \text{Total RWelfare without me}$$

- My bid does not affect the Total RWelfare without me!
- RWelfare only depends on the chosen allocation, not payments
- Trying to choose a bid  $b_i$  that leads to allocation  $x$  that maximizes

$$\text{Total RWelfare with me}(x)$$

# Intuition: Why is the mechanism truthful?

- If we set payment = externality, and bidder is truthful

$$\text{Utility} = \text{Total RWelfare with me} - \text{Total RWelfare without me}$$

- My bid does not affect the Total RWelfare without me!
- RWelfare only depends on the chosen allocation, not payments
- If I'm truthful the auctioneer chooses the allocation that maximizes exactly this quantity and hence that maximizes my utility.

# The Vickrey-Clarke-Groves (VCG) Mechanism

# General Auction (Mechanism Design) Setting

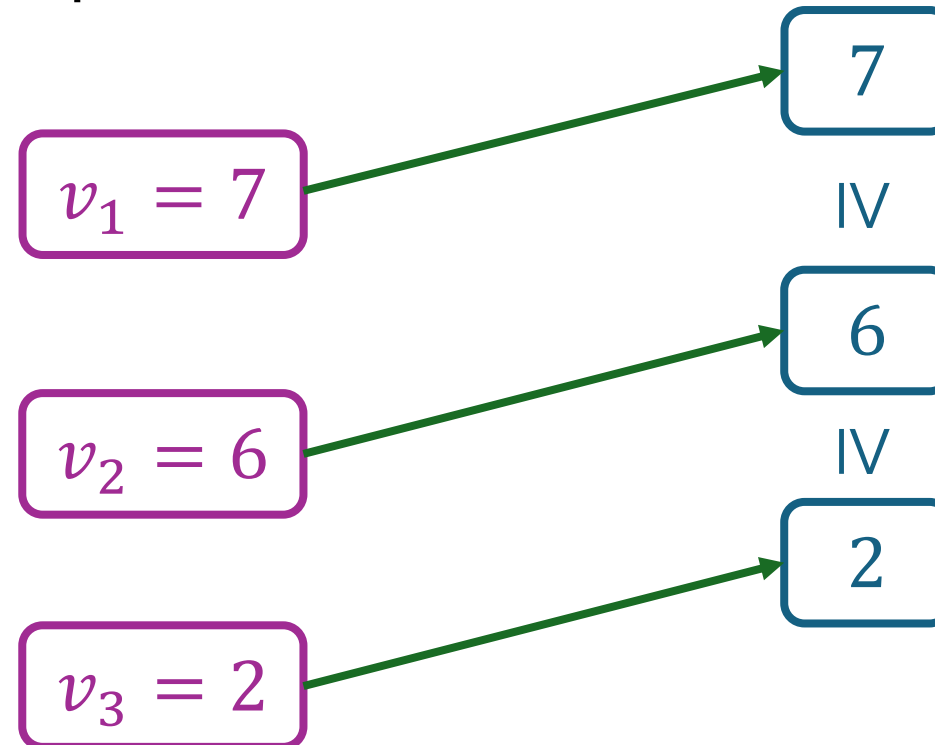
- Auctioneer (Designer) wants to choose among set of outcomes  $O$
- Each bidder  $i$  has some value for each outcome  $v_i(o) \in R$
- The value function  $v_i$  is called the **type** of player  $i$
- Designer elicits **types/bids** from players  $b = (b_1, \dots, b_n)$
- Designer chooses allocation that maximizes the reported welfare

$$x(b) = \operatorname{argmax}_{o \in O} RW(o; b) := \sum_{i=1}^n b_i(o)$$

Total Reported  
Welfare

# Let's repeat this exercise with two slots

- When highest bidder exists, rest of players achieve reported welfare of ...?



Google search results for "digital advertising". The top two sponsored results are highlighted with red boxes. The first result is from Reddit with a bid of 1. The second result is from Microsoft with a bid of 0.5. Green boxes highlight the bid values 1 and 0.5.

**Sponsored**

Reddit  
https://www.redditforbusiness.com

**Advertise on Reddit**

Reach over 100K communities — Connect with passionate communities that deliver results for brands across all industries. Create impact & own top communities in your target category for 24 hours. Try Reddit ads.

**Sponsored**

Microsoft  
https://about.ads.microsoft.com › advertising › start-now

**Microsoft Advertising® | Get a \$500 Advertising Credit**

We'll Help You Find Your Customers and Reach Searchers Across The Microsoft Network. Plus, Receive a \$500 Microsoft **Advertising** Credit When You Spend Just \$250! Free Sign Up.

# General Auction (Mechanism Design) Setting

- Designer chooses allocation that maximizes the reported welfare

$$x(b) = \operatorname{argmax}_{o \in O} RW(o; b) := \sum_{i=1}^n b_i(o)$$

- Charges to each player their externalities as payment

$$p_i(b) = \max_{o \in O} \sum_{j \neq i} b_j(o) - \sum_{j \neq i} b_j(x(b)) \geq 0$$

RWelfare of others  
without me      RWelfare of others  
with me

Why?



# How much utility do bidders receive?

- The utility of bidder  $i$  for reporting  $b_i$  when others report  $b_{-i}$

$$U_i(b) = \underbrace{v_i(x(b))}_{\text{My value}} - \underbrace{p(b)}_{\text{My payment}}$$

- If payment=externality

$$U_i(b) = \underbrace{v_i(x(b))}_{\text{My value}} - \underbrace{\max_{o \in O} \sum_{j \neq i} b_j(o)}_{\text{RWelfare of others without me}} + \underbrace{\sum_{j \neq i} b_j(x(b))}_{\text{RWelfare of others with me}}$$

# What is the optimal bid?

- If payment=externality

$$U_i(b) = \underbrace{v_i(x(b))}_{\text{My value}} + \underbrace{\sum_{j \neq i} b_j(x(b))}_{\text{RWelfare of others with me}} - \underbrace{\max_{o \in O} \sum_{j \neq i} b_j(o)}_{\text{RWelfare of others without me}}$$

- I want to choose a bid  $b_i$  that optimizes my utility

$$\max_{b_i} v_i(x(b)) + \sum_{j \neq i} b_j(x(b)) - \max_{o \in O} \sum_{j \neq i} b_j(o)$$

Does not depend on my bid

# What is the optimal bid?

- I want to choose a bid  $b_i$  that optimizes my utility

$$\max_{b_i} v_i(x(b)) + \sum_{j \neq i} b_j(x(b))$$

My value      RWelfare of others  
with me

- This only depends on the chosen allocation  $x(b)$
- Want to choose a bid that leads to an allocation  $x$  that maximizes

$$v_i(x) + \sum_{j \neq i} b_j(x)$$

# What is the optimal bid?

- Want to choose a bid that leads to an allocation  $x$  that maximizes

$$v_i(x) + \sum_{j \neq i} b_j(x)$$

My value      RWelfare of others with me

- Designer chooses allocation that maximizes reported welfare

$$b_i(x) + \sum_{j \neq i} b_j(x)$$

My bid      RWelfare of others with me

# What is the optimal bid? My true value

- Want to choose a bid that leads to an allocation  $x$  that maximizes

$$v_i(x) + \sum_{j \neq i} b_j(x)$$

My value      RWelfare of others with me

- Designer chooses allocation that maximizes reported welfare

$$b_i(x) + \sum_{j \neq i} b_j(x)$$

My bid      RWelfare of others with me

- If I'm **truthful** then auctioneer chooses the allocation that I want

# What is my utility under truthful reporting

- If payment=externality

$$U_i(b) = \underbrace{v_i(x(b)) + \sum_{j \neq i} b_j(x(b))}_{\text{Total RWelfare with me}} - \underbrace{\max_{o \in O} \sum_{j \neq i} b_j(o)}_{\text{RWelfare of others without me}}$$

- Since auctioneer optimizes reported welfare:

$$U_i(b) = \underbrace{\max_{o \in O} v_i(o) + \sum_{j \neq i} b_j(o)}_{\text{Total RWelfare with me}} - \underbrace{\max_{o \in O} \sum_{j \neq i} b_j(o)}_{\text{RWelfare of others without me}} \geq 0 \quad \text{Why?}$$



# Learning in Non-Truthful Auctions



# Non-Truthful Auctions

- Despite the universality of VCG, non-truthful auctions are frequently used
- More transparent and credible\* rules
- The mechanism used in government procurement and display ads

# Learning how to bid in auctions

- Given the complexity of digital auction markets
  - Given the hardness of strategizing in non-truthful auctions
  - Many of these auctions are repeated!
- 
- It makes sense to study learning over time, to decide how to bid
- 
- How do we learn over time when we repeatedly participate in an auction? Can we compete with the best fixed bid in hindsight?

# Generalized Second Price (GSP) Auction with Many Bells and Whistles

- Bidders submit a bid-per-click  $b_i$
- Each bidder assigned a **quality score**  $s_i$
- Slots allocated in decreasing order of **quality weighted bids**  $s_i \cdot b_i$
- Bidder  $i$  is allocated slot  $j_i(b)$
- Slots have **bidder-specific probability of click**  $a_{i,j_i(b)}$
- Each bidder pays, per-click, the **highest bid that still gives them the same slot**

$$p_i(b) = \frac{s_{(j_i(b)+1)} \cdot b_{(j_i(b)+1)}}{s_i}$$



digital advertising



All Images News Videos Shopping More

About 6,620,000,000 results (0.44 seconds)

Sponsored



Reddit

<https://www.redditforbusiness.com>

[Advertise on Reddit](#)

Reach over 100K communities — Connect with passionate communities that deliver results for brands across all industries. Create impact & own top communities in your target category for 24 hours. Try Reddit ads.

Sponsored



Microsoft

<https://about.ads.microsoft.com> > advertising > start-now

[Microsoft Advertising® | Get a \\$500 Advertising Credit](#)

We'll Help You Find Your Customers and Reach Searchers Across The Microsoft Network. Plus, Receive a \$500 Microsoft Advertising Credit When You Spend Just \$250! Free Sign Up.

Sponsored



coseom

<https://www.coseom.com>

[Pay Per Click Company](#)

COSEOM™ — Generate Leads For Your Business Using Advanced PPC Strategies. Request A Proposal Today!

Sponsored



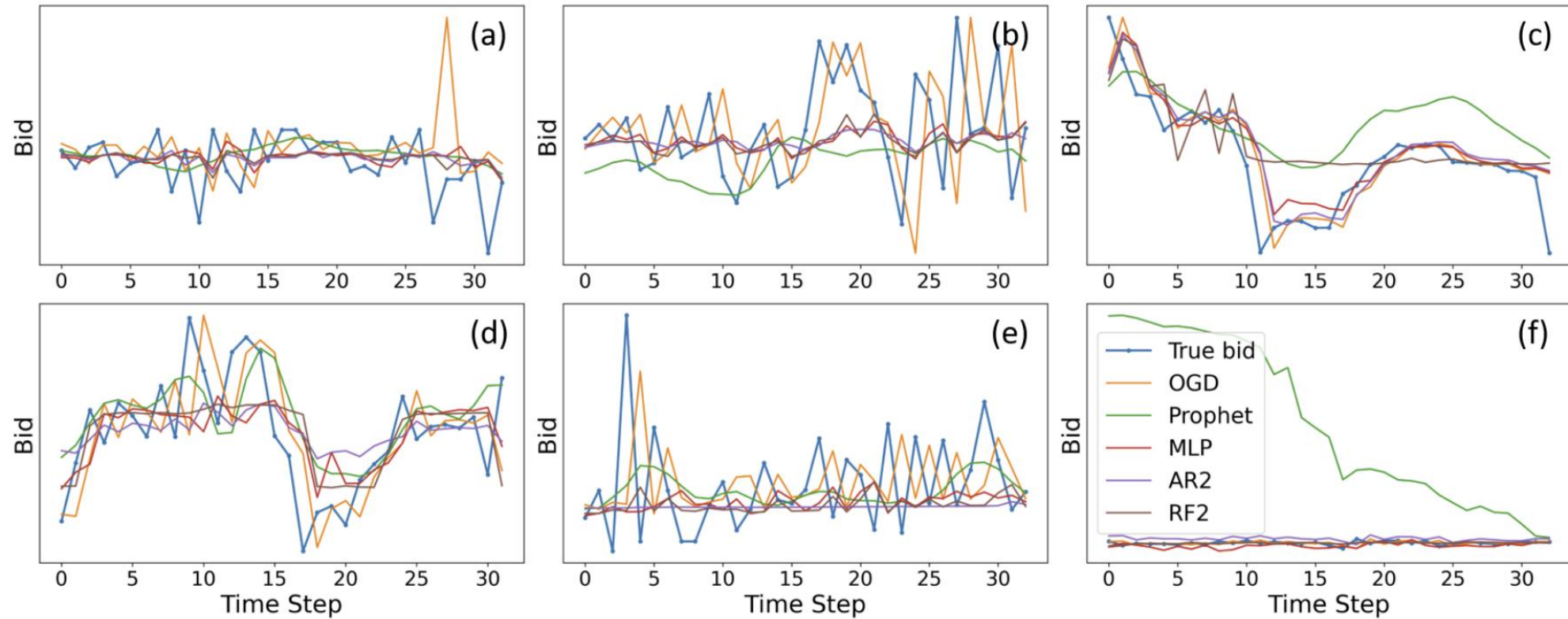
Simpli.fi

<https://www.simpli.fi>

[Simpli.fi | Advertising Success Platform](#)

Established in 2010 — Enjoy the perks of multi-channel targeting, measurement, & reporting with our interface. CTV.

# Simple learning dynamics are good predictors



(b) Stepahead Predictions

# No-Regret Learning in Auctions

At each period  $t \in \{1, \dots, T\}$

- An auction among  $n$  bidders takes place (GFP, GSP, FP)
- Each bidder  $i$  submits bid  $b_i$  from discrete set of  $N$  bids  $\{\epsilon, 2\epsilon \dots, 1\}$
- Each bidder learns their allocation and payment

$$x_i^t, p_i^t = x_i(b^t), p_i(b^t)$$

- e.g. in a first price auction, learn whether I won
- e.g. in a second price auction, learn whether I won and when I win, I learn the next highest bid.

# No-Regret Learning

- Want to choose my bids  $b_i^t$ , based on algorithm that guarantees

$$\frac{1}{T} \sum_{t=1}^T u_i(b^t) \geq \max_{b_i \in [N]} \frac{1}{T} \sum_{t=1}^T u_i(b_i, b^t) - \epsilon(T)$$

- for some  $\epsilon(T) \rightarrow 0$

## What algorithm should I use?

EXP

Optimistic EXP

Online Gradient Descent

None of the above

## What algorithm should I use?

EXP

0%

Optimistic EXP

0%

Online Gradient Descent

0%

None of the above

0%



## What algorithm should I use?

EXP

0%

Optimistic EXP

0%

Online Gradient Descent

0%

None of the above

0%

# No-Regret Learning with Limited Feedback

- Want to choose my bids  $b_i^t$ , based on algorithm that guarantees

$$\frac{1}{T} \sum_{t=1}^T u_i(b^t) \geq \max_{b_i \in [N]} \frac{1}{T} \sum_{t=1}^T u_i(b_i, b^t) - \epsilon(T)$$

- Seems like a standard  $N$  action no-regret problem
- **What's the catch!** I don't receive after each period the utility for all my actions. Only the utility for action I took!
- **Limited Feedback.** I cannot calculate how much I would have gotten with any other bid (e.g. in an FP, solely knowing whether I won or not).

# No-Regret Learning with Bandit Feedback

At each period  $t$

- Adversary chooses a loss vector  $\ell_t \in [0, 1]^N$
  - I choose an action  $i_t$  (not knowing  $\ell_t$ )
  - I observe loss of my chosen action  $\ell_t^{i_t}$
- 
- I want to guarantee small expected regret with any fixed action:

$$\max_{i \in N} E \left[ \frac{1}{T} \sum_{t=1}^T \ell_t^{i_t} - \ell_t^i \right] \leq \epsilon(T)$$

# Constructing Un-biased Estimates of Vector

- There is a hidden loss vector  $\ell_t = (\ell_t^1, \dots, \ell_t^N)$  (potential outcomes)
- At each period I choose action (treatment)  $j$  with probability  $p_t^j$
- I learn the loss  $\ell_t^j$  with probability  $p_t^j$
- **Remember:** no-regret algorithms work well, even if we have unbiased proxies of the true losses (e.g. Monte Carlo CFR)

**Question.** Can I construct a random variable that guarantees that in expectation over the choice of actions?

$$E[\tilde{\ell}_t] = \ell_t \Leftrightarrow \forall j: E[\tilde{\ell}_t^j] = \ell_t^j$$

# Constructing Un-biased Estimates of Vector

**Question.** Can I construct a random variable that guarantees that in expectation over the choice of actions?

$$E[\tilde{\ell}_t] = \ell_t \Leftrightarrow \forall j: E[\tilde{\ell}_t^j] = \ell_t^j$$

- Random variable can always depend on identity of chosen action  $j_t$ .  
When I choose  $j$  random variable can also depend on  $\ell_t^j$

$$\tilde{\ell}_t^j = 1\{j_t = j\}f_j(\ell_t^j) + 1\{j_t \neq j\}g_j(j_t)$$

- Let's make  $g_j$  zero, and  $f_j$  linear in  $\ell_t^j$

$$\tilde{\ell}_t^j = 1\{j_t = j\}a_j\ell_t^j \Rightarrow E[\tilde{\ell}_t^j] = p_t^j a_j \ell_t^j = \ell_t^j \Rightarrow a_j = \frac{1}{p_t^j}$$

# Inverse Propensity Estimates

At each period  $t$

- Consider the random variables

$$\tilde{\ell}_t^j = \frac{1\{j_t = j\}}{p_t^j} \ell_t^j$$

- The vector  $\tilde{\ell}_t$  can always be calculated  $\left(0, \dots, 0, \frac{\ell_t^{j_t}}{p_t^{j_t}}, 0, \dots, 0\right)$
- The vector  $\tilde{\ell}_t$  is an unbiased proxy of the true loss vector:

$$E[\tilde{\ell}_t] = \ell_t$$

# The EXP Algorithm with Bandit Feedback

Initialize  $\mathbf{p}_t$  to the uniform distribution

**For**  $t$  **in**  $1..T$

Draw action  $j_t$  based on distribution  $\mathbf{p}_t$

Observe loss of chosen action  $\mathbf{l}_t[j_t]$

Construct un-biased proxy loss vector

$$\mathbf{l}_{t\text{proxy}}[j] = \mathbf{1}(j_t=j) * \mathbf{l}_t[j_t] / \mathbf{p}_t[j_t]$$

Update probabilities based on EXP update

$$\mathbf{p}_t = \mathbf{p}_t * \exp(-\eta * \mathbf{l}_{t\text{proxy}})$$

$$\mathbf{p}_t = \mathbf{p}_t / \text{sum}(\mathbf{p}_t)$$

# Recap: Regret of FTRL

$$\text{(FTRL)} \quad x_t = \operatorname{argmin}_{x \in X} \underbrace{\sum_{\tau < t} \langle x, \ell_\tau \rangle}_{\substack{\text{Historical performance} \\ \text{of always choosing} \\ \text{strategy } x}} + \underbrace{\frac{1}{\eta} \mathcal{R}(x)}_{\substack{\text{1-strongly convex} \\ \text{function of } x \text{ that} \\ \text{stabilizes the maximizer}}}$$

**Theorem.** Assuming the utility function at each period  
 $f_t(x) = \langle x, \ell_t \rangle$

is  $L$ -Lipschitz with respect to some norm  $\|\cdot\|$  and the regularizer is 1-strongly convex with respect to the same norm then

$$\text{Regret} - \text{FTRL}(T) \leq \underbrace{\eta L}_{\substack{\text{Average stability} \\ \text{induced by regularizer}}} + \underbrace{\frac{1}{\eta T} \left( \max_{x \in X} \mathcal{R}(x) - \min_{x \in X} \mathcal{R}(x) \right)}_{\substack{\text{Average loss distortion} \\ \text{caused by regularizer}}}$$



**Problem!** The loss vector  $\tilde{\ell}_t$  is not in  $[0,1]$ .

It can take huge values, as probability of an action goes to 0!

**Intuition:** if probability goes to 0, then this action is chosen very infrequently. The loss vector very rarely takes this large value, i.e., the *variance* of the loss should be small.

# Variance of Loss Vector

- Variance is

$$E \left[ \left( \tilde{\ell}_t^j \right)^2 \right] - E \left[ \tilde{\ell}_t^j \right]^2 = E \left[ \left( \tilde{\ell}_t^j \right)^2 \right] - E \left[ \ell_t^j \right]^2$$

- Second term is in  $[0, 1]$ . We will focus on first term (call it “variance”)

$$E \left[ \left( \tilde{\ell}_t^j \right)^2 \right] = p_t^j \left( \frac{\ell_t^j}{p_t^j} \right)^2 = \frac{\left( \ell_t^j \right)^2}{p_t^j}$$

- And we collect this “variance” term only when end up choosing  $j$

$$\text{Average "Variance"} = \sum_j p_t^j \cdot E \left[ \left( \tilde{\ell}_t^j \right)^2 \right] = \sum_j \left( \ell_t^j \right)^2 \leq N$$

# Recap: Regret of FTRL

(FTRL) 
$$x_t = \operatorname{argmin}_{x \in X} \underbrace{\sum_{\tau < t} \langle x, \ell_\tau \rangle}_{\substack{\text{Historical performance} \\ \text{of always choosing} \\ \text{strategy } x}} + \underbrace{\frac{1}{\eta} \mathcal{R}(x)}_{\substack{\text{1-strongly convex} \\ \text{function of } x \text{ that} \\ \text{stabilizes the maximizer}}}$$

Can we replace  $L$  with the Average “Variance”?

**Theorem.** Assuming the utility function at each period  $f_t(x) = \langle x, \ell_t \rangle$

~~is  $L$  Lipschitz with respect to some norm  $\|\cdot\|$~~  and the regularizer is 1-strongly convex with respect to the same norm then

$$\text{Regret} - \text{FTRL}(T) \leq \underbrace{\eta L}_{\substack{\text{Average stability} \\ \text{induced by regularizer}}} + \underbrace{\frac{1}{\eta T} \left( \max_{x \in X} \mathcal{R}(x) - \min_{x \in X} \mathcal{R}(x) \right)}_{\substack{\text{Average loss distortion} \\ \text{caused by regularizer}}}$$

# ***Update:*** Regret of EXP

$$\begin{aligned} \text{(EXP)} \quad p_t &= \operatorname{argmin}_{p \in \Delta} \sum_{\tau < t} \langle p, \tilde{\ell}_\tau \rangle + \boxed{\frac{1}{\eta} \mathcal{R}(p)} \left( \begin{array}{c} \text{Negative} \\ \text{Entropy} \end{array} \right) \mathcal{R}(p) = \sum_{i=1}^n p_i \log(p_i) \\ p_t &\propto p_{t-1} \exp(-\eta \tilde{\ell}_{t-1}) \end{aligned}$$

**Theorem.** Assuming  $\tilde{\ell}_t$  are random proxies that, conditional on history, have expected value equal to true loss vector  $\ell_t$  and  $\tilde{\ell}_t \geq 0$ , then regret of **EXP** is bounded as:

$$\text{Regret} - \text{EXP}(T) \leq \frac{\eta}{T} \sum_t E \left[ \sum_j p_t^j \left( \tilde{\ell}_t^j \right)^2 \right] + \frac{\log(N)}{\eta T}$$

# ***Update:*** Regret of EXP

$$\begin{aligned} \text{(EXP)} \quad p_t &= \operatorname{argmin}_{p \in \Delta} \sum_{\tau < t} \langle p, \tilde{\ell}_\tau \rangle + \boxed{\frac{1}{\eta} \mathcal{R}(p)} \left( \begin{array}{c} \text{Negative} \\ \text{Entropy} \end{array} \right) \mathcal{R}(p) = \sum_{i=1}^n p_i \log(p_i) \\ p_t &\propto p_{t-1} \exp(-\eta \tilde{\ell}_{t-1}) \end{aligned}$$

**Theorem.** Assuming  $\tilde{\ell}_t$  are random proxies that, conditional on history, have expected value equal to true loss vector  $\ell_t$  and  $\tilde{\ell}_t \geq 0$ , then regret of **EXP** is bounded as:

$$\text{Regret} - \text{EXP}(T) \leq \frac{\eta}{T} \sum_t E \left[ \underbrace{\sum_j p_t^j E \left[ \left( \tilde{\ell}_t^j \right)^2 \right]}_{\text{Expected Average "Variance" ?}} \right] + \frac{\log(N)}{\eta T}$$

# ***Update:*** Regret of EXP

$$\begin{aligned} \text{(EXP)} \quad p_t &= \operatorname{argmin}_{p \in \Delta} \sum_{\tau < t} \langle p, \tilde{\ell}_\tau \rangle + \boxed{\frac{1}{\eta} \mathcal{R}(p)} \left( \begin{array}{c} \text{Negative} \\ \text{Entropy} \end{array} \right) \mathcal{R}(p) = \sum_{i=1}^n p_i \log(p_i) \\ p_t &\propto p_{t-1} \exp(-\eta \tilde{\ell}_{t-1}) \end{aligned}$$

**Theorem.** Assuming  $\tilde{\ell}_t$  are random proxies that, conditional on history, have expected value equal to true loss vector  $\ell_t$  and  $\tilde{\ell}_t \geq 0$ , then regret of **EXP** is bounded as:

$$\text{Regret} - \text{EXP}(T) \leq \frac{\eta}{T} \sum_t N + \frac{\log(N)}{\eta T}$$

For the inverse  
propensity proxies

# ***Update:*** Regret of EXP

$$\begin{aligned} \text{(EXP)} \quad p_t &= \operatorname{argmin}_{p \in \Delta} \sum_{\tau < t} \langle p, \tilde{\ell}_\tau \rangle + \boxed{\frac{1}{\eta} \mathcal{R}(p)} \left( \begin{array}{c} \text{Negative} \\ \text{Entropy} \end{array} \right) \mathcal{R}(p) = \sum_{i=1}^n p_i \log(p_i) \\ p_t &\propto p_{t-1} \exp(-\eta \tilde{\ell}_{t-1}) \end{aligned}$$

**Theorem.** Assuming  $\tilde{\ell}_t$  are random proxies that, conditional on history, have expected value equal to true loss vector  $\ell_t$  and  $\tilde{\ell}_t \geq 0$ , then regret of **EXP** is bounded as:

$$\text{Regret} - \text{EXP}(T) \leq \eta N + \frac{\log(N)}{\eta T}$$

For the inverse  
propensity proxies

# ***Update:*** Regret of EXP

$$\begin{aligned} \text{(EXP)} \quad p_t &= \operatorname{argmin}_{p \in \Delta} \sum_{\tau < t} \langle p, \tilde{\ell}_\tau \rangle + \boxed{\frac{1}{\eta} \mathcal{R}(p)} \left( \begin{array}{c} \text{Negative} \\ \text{Entropy} \end{array} \right) \mathcal{R}(p) = \sum_{i=1}^n p_i \log(p_i) \\ p_t &\propto p_{t-1} \exp(-\eta \tilde{\ell}_{t-1}) \end{aligned}$$

**Theorem.** Assuming  $\tilde{\ell}_t$  are random proxies that, conditional on history, have expected value equal to true loss vector  $\ell_t$  and  $\tilde{\ell}_t \geq 0$ , then regret of **EXP** is bounded as:

$$\text{Regret} - \text{EXP}(T) \leq \eta N + \frac{\log(N)}{\eta T} \Rightarrow \text{Regret} - \text{EXP}(T) \lesssim \sqrt{\frac{N \log(N)}{T}}$$

For  $\eta \sim \sqrt{\frac{\log(N)}{NT}}$



# Back to Bandit Learning in Auctions

# Bandit Learning in Auctions

- Want to choose my bids  $b_i^t$ , based on algorithm that guarantees

$$\frac{1}{T} \sum_{t=1}^T u_i(b^t) \geq \max_{b_i \in [N]} \frac{1}{T} \sum_{t=1}^T u_i(b_i, b^t) - \epsilon(T)$$

- We can apply EXP3 algorithm for each bidder
- We now have utilities, but EXP3 expects non-negative losses  
Maximizing utility = Minimizing (negative utility)
- However, to ensure losses are non-negative, add a large enough offset  
loss =  $H - \text{utility}$
- If for instance we know that utility  $\leq H$ , we can choose this  $H$  above

# ***Update:*** Regret of EXP

$$\begin{aligned} \text{(EXP)} \quad p_t &= \operatorname{argmin}_{p \in \Delta} \sum_{\tau < t} \langle p, \tilde{\ell}_\tau \rangle + \boxed{\frac{1}{\eta} \mathcal{R}(p)} \left( \begin{array}{c} \text{Negative} \\ \text{Entropy} \end{array} \right) \mathcal{R}(p) = \sum_{i=1}^n p_i \log(p_i) \\ p_t &\propto p_{t-1} \exp(-\eta \tilde{\ell}_{t-1}) \end{aligned}$$

**Theorem.** Assuming  $\tilde{\ell}_t$  are random proxies that, conditional on history, have expected value equal to true loss vector  $\ell_t$  and  $\tilde{\ell}_t \geq 0$ , then regret of **EXP** is bounded as:

$$\text{Regret} - \text{EXP}(T) \leq \eta N + \frac{\log(N)}{\eta T} \Rightarrow \text{Regret} - \text{EXP}(T) \lesssim \sqrt{\frac{N \log(N)}{T}}$$

For  $\eta \sim \sqrt{\frac{\log(N)}{NT}}$