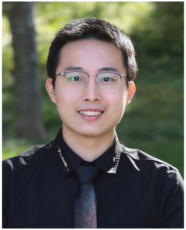
Dynamic Resource Allocation to Strategic Agents under Cost Constraints

Yan Dai

Operations Research Center, MIT



(project overseen by Prof. Negin Golrezaei and Prof. Patrick Jaillet)

The Modern Allocation Trilemma

Dynamic Allocation of Reusable Resources

- T rounds, 1 planner, K agents, 1 indivisible resource

The Modern Allocation Trilemma

Dynamic Allocation of Reusable Resources

- T rounds, 1 planner, K agents, 1 indivisible resource
- Agents have agent- & round- dependent value – **private**

The Modern Allocation Trilemma

Dynamic Allocation of Reusable Resources

- T rounds, 1 planner, K agents, 1 indivisible resource
- Agents have agent- & round- dependent value – **private**
- Agents craft arbitrary reports on their own – **self-interested**

The Modern Allocation Trilemma

Dynamic Allocation of Reusable Resources

- T rounds, 1 planner, K agents, 1 indivisible resource
- Agents have agent- & round- dependent value – **private**
- Agents craft arbitrary reports on their own – **self-interested**
- Allocation incurs d-dimensional cost – **facing constraints**

The Modern Allocation Trilemma

Dynamic Allocation of Reusable Resources

- T rounds, 1 planner, K agents, 1 indivisible resource
- Agents have agent- & round- dependent value – **private**
- Agents craft arbitrary reports on their own – **self-interested**
- Allocation incurs d-dimensional cost – **facing constraints**

Trilemma: Efficiency, Incentives, & Feasibility

- **Efficiency.** Max social welfare (allocate to whom in need)

The Modern Allocation Trilemma

Dynamic Allocation of Reusable Resources

- T rounds, 1 planner, K agents, 1 indivisible resource
- Agents have agent- & round- dependent value – **private**
- Agents craft arbitrary reports on their own – **self-interested**
- Allocation incurs d-dimensional cost – **facing constraints**

Trilemma: Efficiency, Incentives, & Feasibility

- **Efficiency.** Max social welfare (allocate to whom in need)
- **Incentives.** Handle strategic manipulations

The Modern Allocation Trilemma

Dynamic Allocation of Reusable Resources

- T rounds, 1 planner, K agents, 1 indivisible resource
- Agents have agent- & round- dependent value – **private**
- Agents craft arbitrary reports on their own – **self-interested**
- Allocation incurs d-dimensional cost – **facing constraints**

Trilemma: Efficiency, Incentives, & Feasibility

- **Efficiency.** Max social welfare (allocate to whom in need)
- **Incentives.** Handle strategic manipulations
- **Feasibility.** Obey long-term constraints (*e.g.*, cost, energy)

The Modern Allocation Trilemma

Dynamic Allocation of Reusable Resources

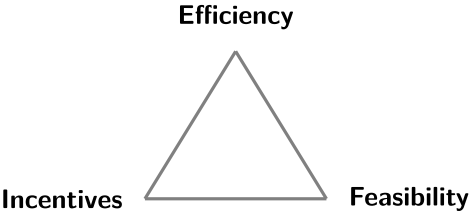
- T rounds, 1 planner, K agents, 1 indivisible resource
- Agents have agent- & round- dependent value – **private**
- Agents craft arbitrary reports on their own – **self-interested**
- Allocation incurs d-dimensional cost – **facing constraints**

Trilemma: Efficiency, Incentives, & Feasibility

- **Efficiency.** Max social welfare (allocate to whom in need)
- **Incentives.** Handle strategic manipulations
- **Feasibility.** Obey long-term constraints (*e.g.*, cost, energy)

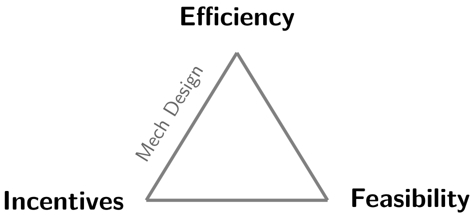
Question. Can all three be achieved simultaneously?

No “3-in-1” Approach in the Literature



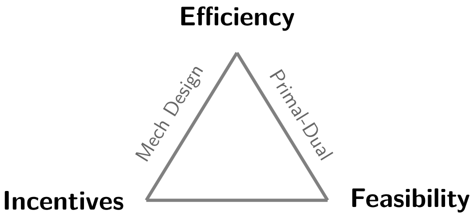
No "3-in-1" Approach in the Literature

- **Efficiency + Incentives.** Classical mechanism design (*e.g.*, VCG [Vickrey, 1961; Clarke, 1971; Groves, 1973] & many variants)



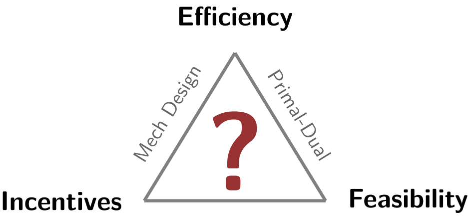
No "3-in-1" Approach in the Literature

- **Efficiency + Incentives.** Classical mechanism design (*e.g.*, VCG [Vickrey, 1961; Clarke, 1971; Groves, 1973] & many variants)
- **Efficiency + Feasibility.** Online linear programming (many primal-dual framework [Li et al., 2023; Balseiro et al., 2023])



No "3-in-1" Approach in the Literature

- **Efficiency + Incentives.** Classical mechanism design (*e.g.*, VCG [Vickrey, 1961; Clarke, 1971; Groves, 1973] & many variants)
- **Efficiency + Feasibility.** Online linear programming (many primal-dual framework [Li et al., 2023; Balseiro et al., 2023])
- **Efficiency + Incentives + Feasibility? No** unless super restrictive assumptions (*e.g.*, homogeneous agents [Yin et al., 2022] & "fair share"-like constraints & non-social-welfare objective [Gorokh et al., 2021]) **"Impossible triangle"?**



Standard Methods Fail: The Strategic Gap

Standard Primal-Dual Methods

Decide *dual* $\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_T$ ("shadow prices" for cost constraints)
Give dual-adjusted *primal allocation* ($\tilde{i}_t^* := \text{argmax}_i(v_{t,i} - \boldsymbol{\lambda}_t^\mathsf{T} \boldsymbol{c}_{t,i})$)

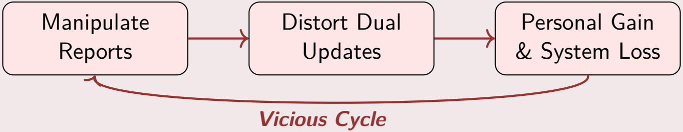
Standard Methods Fail: The Strategic Gap

Standard Primal-Dual Methods

Decide *dual* $\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_T$ ("shadow prices" for cost constraints)

Give dual-adjusted *primal allocation* ($\tilde{i}_t^* := \operatorname{argmax}_i(v_{t,i} - \boldsymbol{\lambda}_t^\top \boldsymbol{c}_{t,i})$)

Fragile to strategic manipulation due to **frequent dual updates**



Standard Methods Fail: The Strategic Gap

Standard Primal-Dual Methods

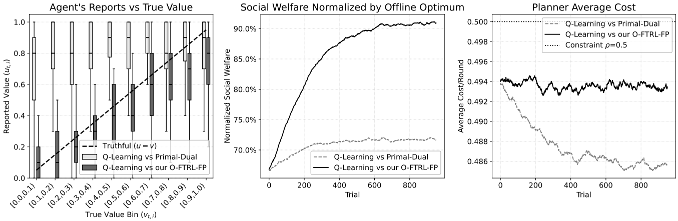
Decide *dual* $\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_T$ ("shadow prices" for cost constraints)

Give dual-adjusted *primal allocation* $(\tilde{i}_t^* := \operatorname{argmax}_i(v_{t,i} - \boldsymbol{\lambda}_t^\top \boldsymbol{c}_{t,i}))$

Fragile to strategic manipulation due to **frequent dual updates**

Manipulate Reports → Distort Dual Updates → Personal Gain & System Loss → Manipulate Reports

Vicious Cycle



Setup & Contribution

- T rounds, 1 planner, K agents, 1 indivisible resource

Setup & Contribution

- T rounds, 1 planner, K agents, 1 indivisible resource
- **Agent** i private value $v_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{V}_i$ – **dist. fixed but unknown**

Setup & Contribution

- T rounds, 1 planner, K agents, 1 indivisible resource
- **Agent** i private value $v_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{V}_i$ – **dist. fixed but unknown**
- **Agent** i public cost $\boldsymbol{c}_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{C}_i$ – **dist. fixed but unknown**

Setup & Contribution

- T rounds, 1 planner, K agents, 1 indivisible resource
- **Agent** i private value $v_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{V}_i$ – **dist. fixed but unknown**
- **Agent** i public cost $\boldsymbol{c}_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{C}_i$ – **dist. fixed but unknown**
- **Agent** i gives arbitrary report $u_{t,i}$ – **very strategic**

Setup & Contribution

- T rounds, 1 planner, K agents, 1 indivisible resource
- **Agent** i private value $v_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{V}_i$ – **dist. fixed but unknown**
- **Agent** i public cost $\boldsymbol{c}_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{C}_i$ – **dist. fixed but unknown**
- **Agent** i gives arbitrary report $u_{t,i}$ – **very strategic**
- **Planner** decides allocation i_t & payment p_{t,i_t}

Setup & Contribution

- T rounds, 1 planner, K agents, 1 indivisible resource
- **Agent** i private value $v_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{V}_i$ – **dist. fixed but unknown**
- **Agent** i public cost $\boldsymbol{c}_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{C}_i$ – **dist. fixed but unknown**
- **Agent** i gives arbitrary report $u_{t,i}$ – **very strategic**
- **Planner** decides allocation i_t & payment p_{t,i_t}

Agents. max $\mathbb{E}[\sum_t \gamma^t \mathbb{1}[i_t = i](v_{t,i} - p_{t,i})]$ (γ-discounted value-pay)

Setup & Contribution

- T rounds, 1 planner, K agents, 1 indivisible resource
- **Agent** i private value $v_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{V}_i$ – **dist. fixed but unknown**
- **Agent** i public cost $\boldsymbol{c}_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{C}_i$ – **dist. fixed but unknown**
- **Agent** i gives arbitrary report $u_{t,i}$ – **very strategic**
- **Planner** decides allocation i_t & payment p_{t,i_t}

Agents. max $\mathbb{E}[\sum_t \gamma^t \mathbb{1}[i_t = i](v_{t,i} - p_{t,i})]$ (γ-discounted value-pay)
Planner. max $\mathbb{E}[\sum_t v_{t,i_t}]$ (undiscounted total social welfare)
subject to $\frac{1}{T} \sum_t \boldsymbol{c}_{t,i_t} \le \boldsymbol{\rho}$ (d-dimensional cost constraint)

Setup & Contribution

- T rounds, 1 planner, K agents, 1 indivisible resource
- **Agent** i private value $v_{t,i} \overset{\textit{i.i.d.}}{\sim} \mathcal{V}_i$ – **dist. fixed but unknown**
- **Agent** i public cost $\boldsymbol{c}_{t,i} \overset{\textit{i.i.d.}}{\sim} \mathcal{C}_i$ – **dist. fixed but unknown**
- **Agent** i gives arbitrary report $u_{t,i}$ – **very strategic**
- **Planner** decides allocation i_t & payment p_{t,i_t}

Agents. max $\mathbb{E}[\sum_t \gamma^t \mathbb{1}[i_t = i](v_{t,i} - p_{t,i})]$ (γ-discounted value-pay)
Planner. max $\mathbb{E}[\sum_t v_{t,i_t}]$ (undiscounted total social welfare)
subject to $\frac{1}{T} \sum_t \boldsymbol{c}_{t,i_t} \le \boldsymbol{\rho}$ (d-dimensional cost constraint)

Main Result: $\widetilde{\mathcal{O}}(\sqrt{T})$ Social Welfare Regret & 0 Constr Violation

Regret. $\mathbb{E}[\sum_t (v_{t,i_t^*} - v_{t,i_t})] = \widetilde{\mathcal{O}}(\sqrt{T})$ ($\{i_t^*\} :=$ offline optimum)
Constr Violation. $\frac{1}{T} \sum_t \boldsymbol{c}_{t,i_t} \le \boldsymbol{\rho}$ *a.s.* (0 constraint violation)

Primal Allocation: Incentive-Aware Framework

Goal. Given dual $\boldsymbol{\lambda}_t$, make allocations – despite strategic reports

Primal Allocation: Incentive-Aware Framework

Goal. Given dual $\boldsymbol{\lambda}_t$, make allocations – despite strategic reports

Incentive-Aware Primal Allocation Framework: 3 Innovations

1. **Epoch-Based Lazy Updates.**
   - Fix dual variables / shadow prices $\boldsymbol{\lambda}_\ell$ within long "epochs"
   - $\Longrightarrow$ Reduce agents' manipulation incentive & ability

Primal Allocation: Incentive-Aware Framework

Goal. Given dual $\boldsymbol{\lambda}_t$, make allocations – despite strategic reports

Incentive-Aware Primal Allocation Framework: 3 Innovations

1. **Epoch-Based Lazy Updates.**
   - Fix dual variables / shadow prices $\boldsymbol{\lambda}_\ell$ within long "epochs"
   - $\Longrightarrow$ Reduce agents' manipulation incentive & ability
2. **Uniform Exploration.**
   - With low prob offer random prices to random agents
   - $\Longrightarrow$ Impose immediate utility loss for misreporting

Primal Allocation: Incentive-Aware Framework

Goal. Given dual $\boldsymbol{\lambda}_t$, make allocations – despite strategic reports

Incentive-Aware Primal Allocation Framework: 3 Innovations

1. **Epoch-Based Lazy Updates.**
   - Fix dual variables / shadow prices $\boldsymbol{\lambda}_\ell$ within long "epochs"
   - $\Longrightarrow$ Reduce agents' manipulation incentive & ability
2. **Uniform Exploration.**
   - With low prob offer random prices to random agents
   - $\Longrightarrow$ Impose immediate utility loss for misreporting
3. **Dual-Adjusted Payments.**
   - VCG-like rule aligning social welfare & dual-adjusted value
   - $\Longrightarrow$ Truth-telling is optimal in normal rounds

Primal Allocation: Incentive-Aware Framework

Goal. Given dual $\boldsymbol{\lambda}_t$, make allocations – despite strategic reports

Incentive-Aware Primal Allocation Framework: 3 Innovations

1. **Epoch-Based Lazy Updates.**
   - Fix dual variables / shadow prices $\boldsymbol{\lambda}_\ell$ within long "epochs"
   - $\Longrightarrow$ Reduce agents' manipulation incentive & ability
2. **Uniform Exploration.**
   - With low prob offer random prices to random agents
   - $\Longrightarrow$ Impose immediate utility loss for misreporting
3. **Dual-Adjusted Payments.**
   - VCG-like rule aligning social welfare & dual-adjusted value
   - $\Longrightarrow$ Truth-telling is optimal in normal rounds

Theorem. $\widetilde{\mathcal{O}}(1)$ misreports & $\widetilde{\mathcal{O}}(1)$ misallocations per epoch

Dual Updates: Breaking $\Omega(T^{2/3})$ via Predictability

Goal. Update duals to capture constraints – despite lazy updates

Dual Updates: Breaking $\Omega(T^{2/3})$ via Predictability

Goal. Update duals to capture constraints – despite lazy updates

Challenge: Lazy Updates ⇒ Slow Learning

- Standard online learning (*e.g.*, FTRL, FTPL) gives $\widetilde{\mathcal{O}}(T^{2/3})$
- **Learning Barrier.** Lazy updates $\Longrightarrow \Omega(T^{2/3})$ [Dekel et al., 2014]

Dual Updates: Breaking $\Omega(T^{2/3})$ via Predictability

Goal. Update duals to capture constraints – despite lazy updates

Challenge: Lazy Updates ⇒ Slow Learning

- Standard online learning (*e.g.*, FTRL, FTPL) gives $\widetilde{\mathcal{O}}(T^{2/3})$
- **Learning Barrier.** Lazy updates $\Longrightarrow \Omega(T^{2/3})$ [Dekel et al., 2014]

Key Insight: Utilize Near-Truthfulness of Agents

Incentive-aware primal allocation $\Longrightarrow$ near-truthful reports

Dual Updates: Breaking $\Omega(T^{2/3})$ via Predictability

Goal. Update duals to capture constraints – despite lazy updates

Challenge: Lazy Updates $\Rightarrow$ Slow Learning

- Standard online learning (*e.g.*, FTRL, FTPL) gives $\widetilde{\mathcal{O}}(T^{2/3})$
- **Learning Barrier.** Lazy updates $\Longrightarrow \Omega(T^{2/3})$ [Dekel et al., 2014]

Key Insight: Utilize Near-Truthfulness of Agents

Incentive-aware primal allocation $\Longrightarrow$ near-truthful reports

1. $\Longrightarrow$ **near-*i.i.d.*** future allocations & cost consumption

Dual Updates: Breaking $\Omega(T^{2/3})$ via Predictability

Goal. Update duals to capture constraints – despite lazy updates

Challenge: Lazy Updates $\Rightarrow$ Slow Learning

- Standard online learning (*e.g.*, FTRL, FTPL) gives $\widetilde{\mathcal{O}}(T^{2/3})$
- **Learning Barrier.** Lazy updates $\Longrightarrow \Omega(T^{2/3})$ [Dekel et al., 2014]

Key Insight: Utilize Near-Truthfulness of Agents

Incentive-aware primal allocation $\Longrightarrow$ near-truthful reports

1. $\Longrightarrow$ **near-*i.i.d.*** future allocations & cost consumption
2. $\Longrightarrow$ **near-truthful** historical reports for reliable predictions

Dual Updates: Breaking $\Omega(T^{2/3})$ via Predictability

Goal. Update duals to capture constraints – despite lazy updates

Challenge: Lazy Updates ⇒ Slow Learning

- Standard online learning (*e.g.*, FTRL, FTPL) gives $\widetilde{\mathcal{O}}(T^{2/3})$
- **Learning Barrier.** Lazy updates $\Longrightarrow \Omega(T^{2/3})$ [Dekel et al., 2014]

Key Insight: Utilize Near-Truthfulness of Agents

Incentive-aware primal allocation $\Longrightarrow$ near-truthful reports

1. $\Longrightarrow$ **near-*i.i.d.*** future allocations & cost consumption
2. $\Longrightarrow$ **near-truthful** historical reports for reliable predictions

Novel Online Learning Framework: O-FTRL-FP

Equip Optimistic FTRL [Rakhlin and Sridharan, 2013] with **Fixed Points**

Allow action-dependent predictions: If round-t loss func $f_t(x)$ depends on round-t action x_t, we allow $\hat{f}_t(x; x_t)$-style predictions instead of only $\tilde{f}_t(x)$

Main Results & Takeaway

Main Contribution

First dynamic mechanism achieving the **trilemma**:

- **Efficiency.** Optimal $\widetilde{\mathcal{O}}(\sqrt{T})$ regret (matching non-strategic LB)
- **Incentives.** Robust to strategic agents ($\exists$ near-truthful PBE)
- **Feasibility.** Zero constraint violation (with probability 1)

Key Techniques

- **Incentive-Aware Primal Allocations.** Novel mixture of lazy updates, uniform exploration, & dual-adjusted payments
- **Dual Learning via Predictions.** Truthful $\Rightarrow$ predictability (nearly) & novel O-FTRL-FP framework for online learning

Questions are more than welcomed!

References

Santiago R Balseiro, Haihao Lu, and Vahab Mirrokni. The best of many worlds:: Dual mirror descent for online allocation problems. *Operations Research*, 71(1):101–119, 2023.

Edward H Clarke. Multipart pricing of public goods. *Public choice*, pages 17–33, 1971.

Ofer Dekel, Jian Ding, Tomer Koren, and Yuval Peres. Bandits with switching costs: T 2/3 regret. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 459–467, 2014.

Artur Gorokh, Siddhartha Banerjee, and Krishnamurthy Iyer. The remarkable robustness of the repeated fisher market. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, pages 562–562, 2021.

Theodore Groves. Incentives in teams. *Econometrica: Journal of the Econometric Society*, pages 617–631, 1973.

Xiaocheng Li, Chunlin Sun, and Yinyu Ye. Simple and fast algorithm for binary integer and online linear programming. *Mathematical Programming*, 200(2):831–875, 2023.

Alexander Rakhlin and Karthik Sridharan. Online learning with predictable sequences. In *Conference on Learning Theory*, pages 993–1019. PMLR, 2013.

William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of finance*, 16(1):8–37, 1961.

Steven Yin, Shipra Agrawal, and Assaf Zeevi. Online allocation and learning in the presence of strategic agents. *Advances in Neural Information Processing Systems*, 35:6333–6344, 2022.