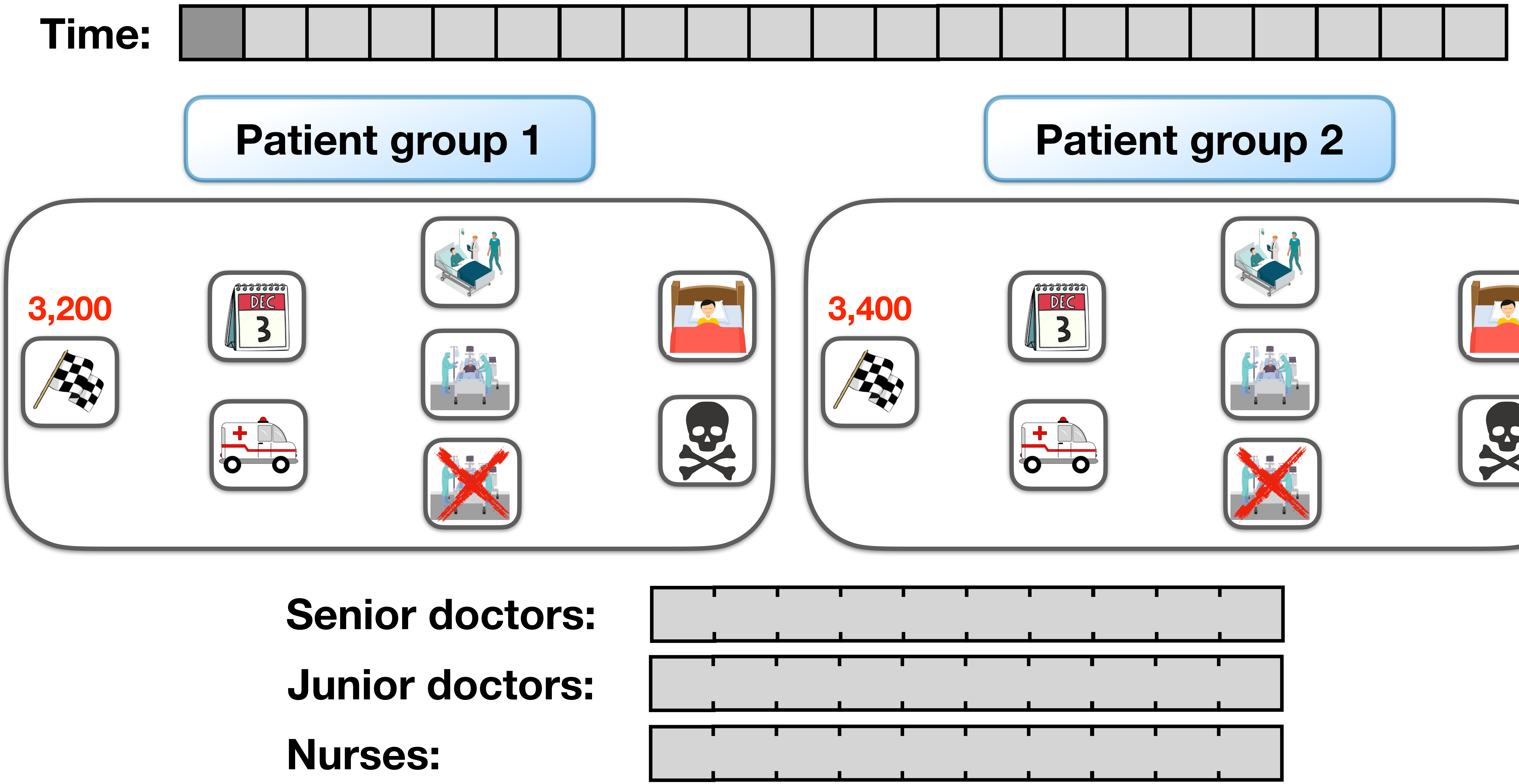
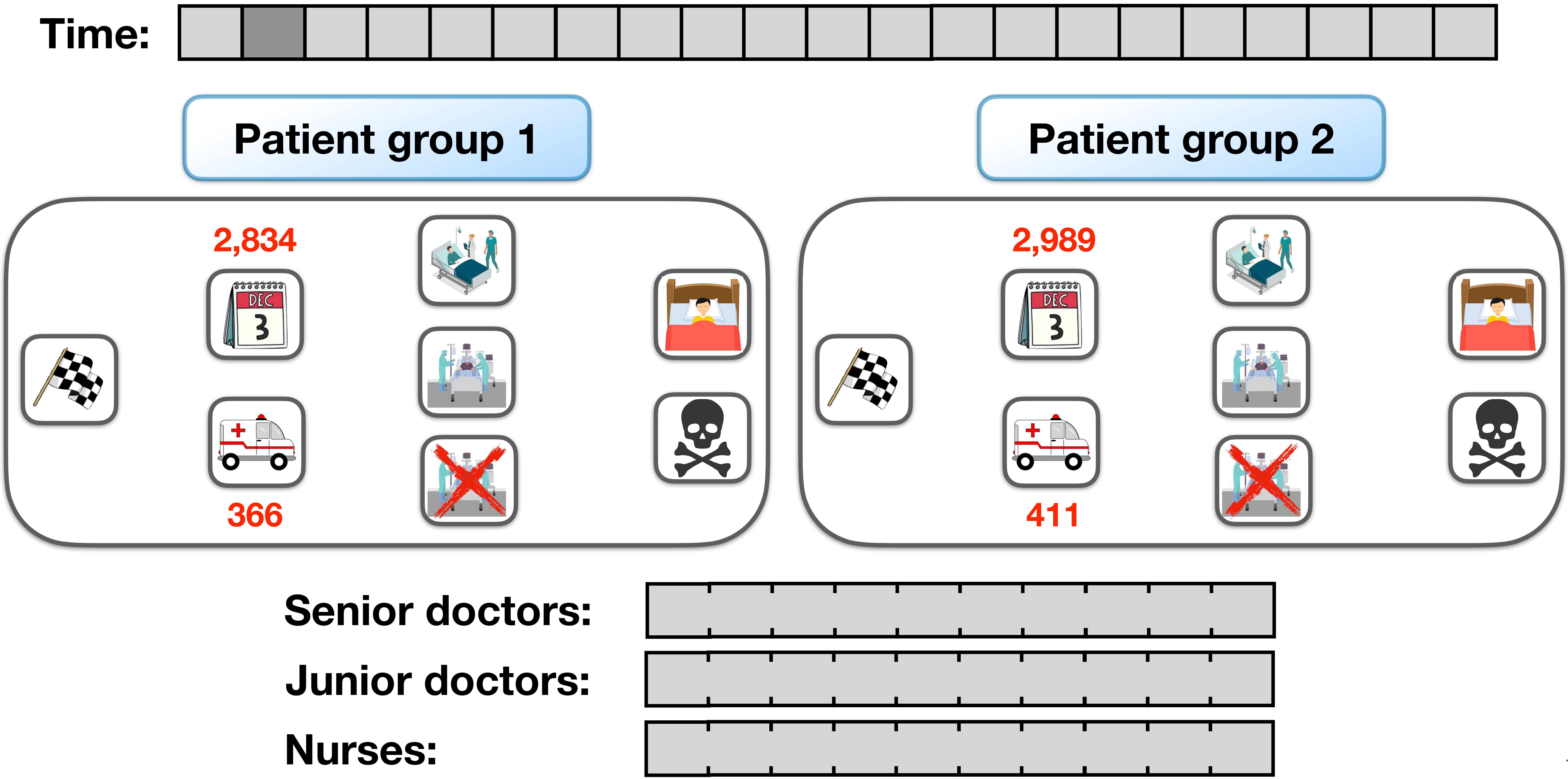
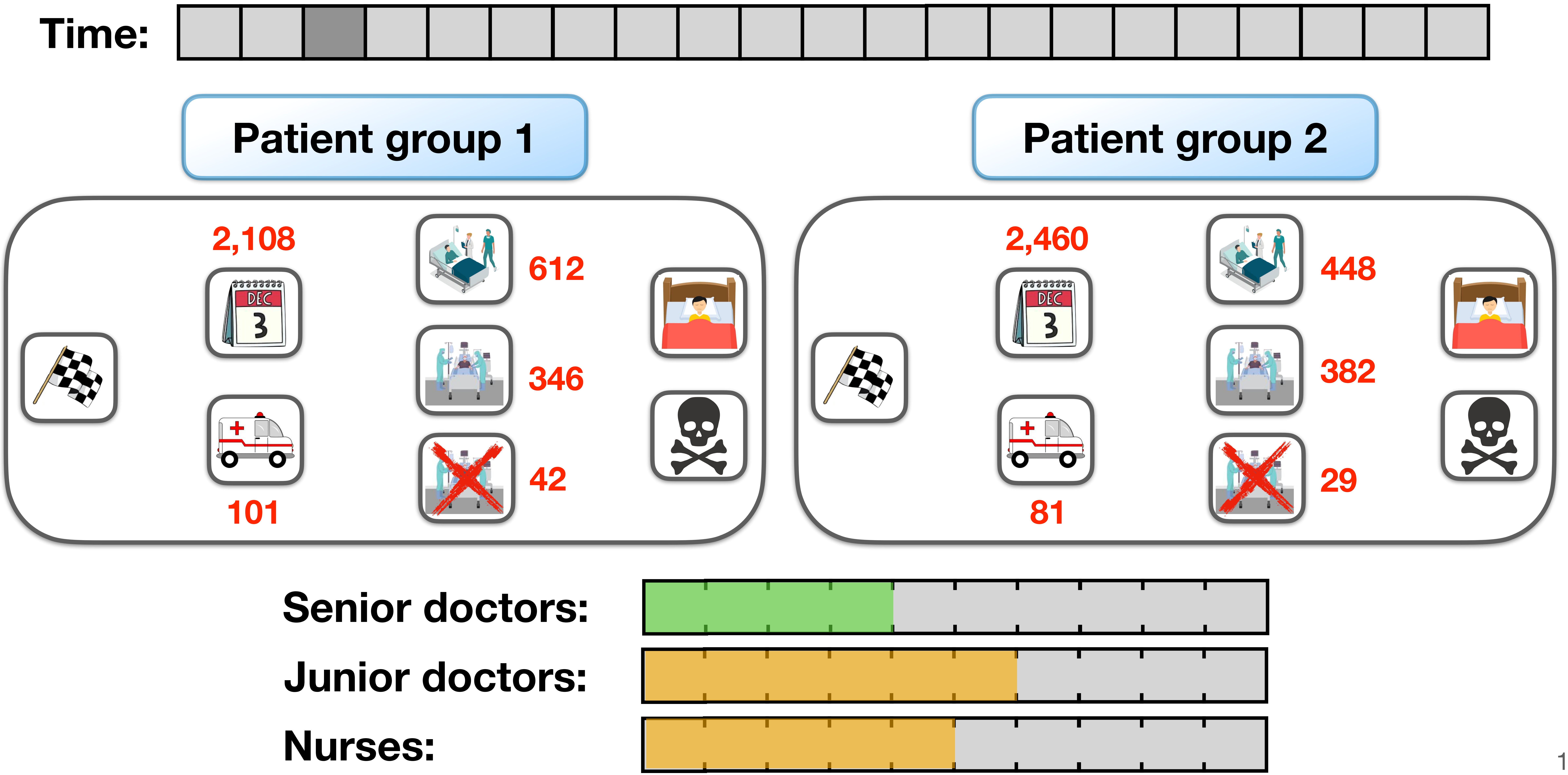
Weakly Coupled Counting DPs



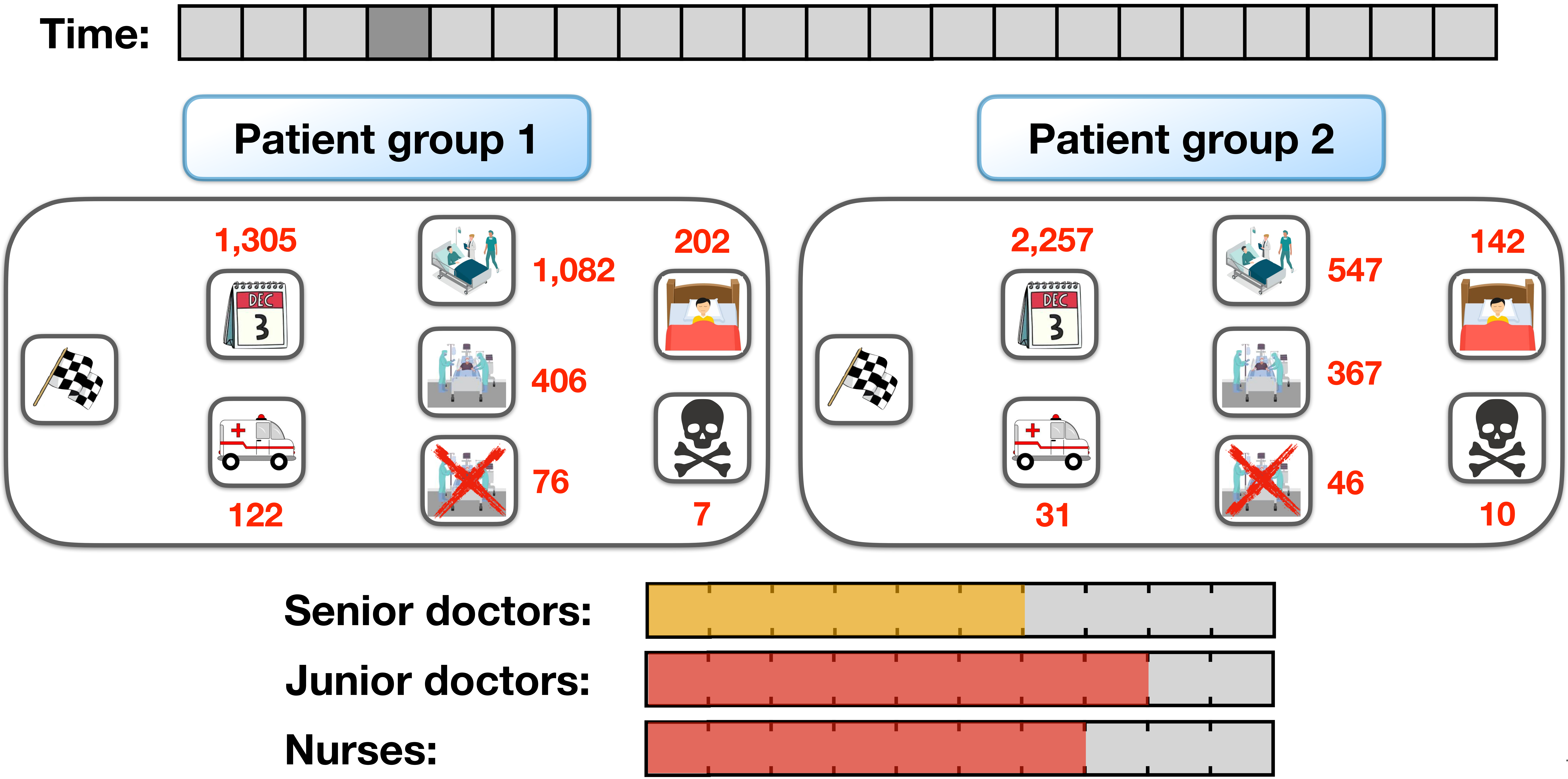
Weakly Coupled Counting DPs



Weakly Coupled Counting DPs



Weakly Coupled Counting DPs



Agenda

1. ~~Weakly Coupled Counting Dynamic Programs~~
2. **The Fluid Approximation**
3. Performance Guarantees
4. Case Study

An optimal policy for the fluid DP can be found via the fluid LP:

Fluid LP

$$\underset{\sigma,\pi \geq 0}{\text{maximize}} \quad \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} \sum_{s\in\mathcal{S}_j} \sum_{a\in\mathcal{A}_j} r_{jt}(s,a) \cdot \pi_{tj}(s,a)$$

$$\text{subject to} \quad \sigma_{1j}(s) = n_j \cdot q_j(s) \qquad \forall j, \forall s \in \mathcal{S}_j$$

$$\sigma_{t+1,j}(s') = \sum_{s\in\mathcal{S}_j} \sum_{a\in\mathcal{A}_j} p_{jt}(s' \mid s,a) \cdot \pi_{tj}(s,a) \qquad \forall j, \forall s' \in \mathcal{S}_j, \forall t$$

$$\sum_{j\in\mathcal{J}} \sum_{s\in\mathcal{S}_j} \sum_{a\in\mathcal{A}_j} c_{tlj}(s,a) \cdot \pi_{tj}(s,a) \leq b_{tl} \qquad \forall l, \forall t$$

$$\sum_{a\in\mathcal{A}_j} \pi_{tj}(s,a) = \sigma_{tj}(s) \qquad \forall j, \forall s \in \mathcal{S}_j, \forall t$$

An optimal policy for the fluid DP can be found via the fluid LP:

Fluid LP

maximize $\sigma, \pi \geq 0$ $\quad \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} r_{jt}(s,a) \cdot \pi_{tj}(s,a)$

$\sigma_{tj}(s)$: # of patients in patient group *j* that are in state *s* in week *t*

subject to $\quad$ $j, \forall s \in \mathcal{S}_j$

$$\sigma_{t+1,j}(s') = \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} p_{jt}(s' \mid s,a) \cdot \pi_{tj}(s,a) \qquad \forall j, \forall s' \in \mathcal{S}_j, \forall t$$

$$\sum_{j \in \mathcal{J}} \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} c_{tlj}(s,a) \cdot \pi_{tj}(s,a) \leq b_{tl} \qquad \forall l, \forall t$$

$$\sum_{a \in \mathcal{A}_j} \pi_{tj}(s,a) = \sigma_{tj}(s) \qquad \forall j, \forall s \in \mathcal{S}_j, \forall t$$

An optimal policy for the fluid DP can be found via the fluid LP:

Fluid LP

maximize $\sigma, \pi \geq 0$ $\quad \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} r_{jt}(s,a) \cdot \pi_{tj}(s,a)$

subject to $\quad \ldots \quad \forall j, \forall s \in \mathcal{S}_j$

$\ldots \quad \forall j, \forall s' \in \mathcal{S}_j, \forall t$

$\sum_{j \in \mathcal{J}} \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} \ldots \quad \forall l, \forall t$

$\sum_{a \in \mathcal{A}_j} \pi_{tj}(s,a) = \sigma_{tj}(s) \quad \forall j, \forall s \in \mathcal{S}_j, \forall t$

$\sigma_{tj}(s)$: # of patients in patient group *j* that are in state *s* in week *t*

$\pi_{tj}(s,a)$: # of patients from $\sigma_{tj}(s)$ that receive action *a*

An optimal policy for the fluid DP can be found via the fluid LP:

Fluid LP

objective function: maximize rewards

$$\max_{\sigma, \pi \geq 0} \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} r_{jt}(s,a) \cdot \pi_{tj}(s,a)$$

subject to

$$\sigma_{1j}(s) = n_j \cdot q_j(s) \qquad \forall j, \forall s \in \mathcal{S}_j$$

$$\sigma_{t+1,j}(s') = \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} p_{jt}(s' \mid s, a) \cdot \pi_{tj}(s,a) \qquad \forall j, \forall s' \in \mathcal{S}_j, \forall t$$

$$\sum_{j \in \mathcal{J}} \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} c_{tlj}(s,a) \cdot \pi_{tj}(s,a) \leq b_{tl} \qquad \forall l, \forall t$$

$$\sum_{a \in \mathcal{A}_j} \pi_{tj}(s,a) = \sigma_{tj}(s) \qquad \forall j, \forall s \in \mathcal{S}_j, \forall t$$

An optimal policy for the fluid DP can be found via the fluid LP:

Fluid LP

initial states: must follow q

$$\underset{\sigma,\pi\geq 0}{\text{maximize}} \quad \sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}}\sum_{s\in\mathcal{S}_j}\sum_{a\in\mathcal{A}_j} r_{jt}(s,a)\cdot\pi_{tj}(s,a)$$

$$\text{subject to} \quad \boxed{\sigma_{1j}(s) = n_j\cdot q_j(s)} \qquad \forall j, \forall s\in\mathcal{S}_j$$

$$\sigma_{t+1,j}(s') = \sum_{s\in\mathcal{S}_j}\sum_{a\in\mathcal{A}_j} p_{jt}(s'\,|\,s,a)\cdot\pi_{tj}(s,a) \qquad \forall j, \forall s'\in\mathcal{S}_j, \forall t$$

$$\sum_{j\in\mathcal{J}}\sum_{s\in\mathcal{S}_j}\sum_{a\in\mathcal{A}_j} c_{tlj}(s,a)\cdot\pi_{tj}(s,a)\leq b_{tl} \qquad \forall l, \forall t$$

$$\sum_{a\in\mathcal{A}_j}\pi_{tj}(s,a) = \sigma_{tj}(s) \qquad \forall j, \forall s\in\mathcal{S}_j, \forall t$$

An optimal policy for the fluid DP can be found via the fluid LP:

Fluid LP

transitions:
must follow p

maximize$_{\sigma, \pi \geq 0}$ $\sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} r_{jt}(s,a) \cdot \pi_{tj}(s,a)$

subject to

$$\sigma_{1j}(s) = n_j \cdot q_j(s) \qquad \forall j, \forall s \in \mathcal{S}_j$$

$$\sigma_{t+1,j}(s') = \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} p_{jt}(s' \mid s,a) \cdot \pi_{tj}(s,a) \qquad \forall j, \forall s' \in \mathcal{S}_j, \forall t$$

$$\sum_{j \in \mathcal{J}} \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} c_{tlj}(s,a) \cdot \pi_{tj}(s,a) \leq b_{tl} \qquad \forall l, \forall t$$

$$\sum_{a \in \mathcal{A}_j} \pi_{tj}(s,a) = \sigma_{tj}(s) \qquad \forall j, \forall s \in \mathcal{S}_j, \forall t$$

An optimal policy for the fluid DP can be found via the fluid LP:

Fluid LP

maximize $\sigma, \pi \geq 0$ $\quad \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} r_{jt}(s,a) \cdot \pi_{tj}(s,a)$

resources:
budgets must be kept

subject to

$$\sigma_{1j}(s) = n_j \cdot q_j(s) \qquad \forall j, \forall s \in \mathcal{S}_j$$

$$\sigma_{t+1,j}(s') = \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} p_{jt}(s' \mid s, a) \cdot \pi_{tj}(s,a) \qquad \forall j, \forall s' \in \mathcal{S}_j, \forall t$$

$$\sum_{j \in \mathcal{J}} \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} c_{tlj}(s,a) \cdot \pi_{tj}(s,a) \leq b_{tl} \qquad \forall l, \forall t$$

$$\sum_{a \in \mathcal{A}_j} \pi_{tj}(s,a) = \sigma_{tj}(s) \qquad \forall j, \forall s \in \mathcal{S}_j, \forall t$$

Fluid Approximation: Fluid LP

An optimal policy for the fluid DP can be found via the fluid LP:

Fluid LP

maximize $\sigma, \pi \geq 0$ $\quad \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} r_{jt}(s,a) \cdot \pi_{tj}(s, a$

“flow preservation”:
we cannot “drop” patients

subject to

$$\sigma_{1j}(s) = n_j \cdot q_j(s) \qquad \forall j, \forall s \in \mathcal{S}_j$$

$$\sigma_{t+1,j}(s') = \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} p_{jt}(s' \mid s, a) \cdot \pi_{tj}(s,a) \qquad \forall j, \forall s' \in \mathcal{S}_j, \forall t$$

$$\sum_{j \in \mathcal{J}} \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} c_{tlj}(s,a) \cdot \pi_{tj}(s,a) \leq b_{tl} \qquad \forall l, \forall t$$

$$\sum_{a \in \mathcal{A}_j} \pi_{tj}(s,a) = \sigma_{tj}(s) \qquad \forall j, \forall s \in \mathcal{S}_j, \forall t$$

An optimal policy for the fluid DP can be found via the fluid LP:

Fluid LP

$$\underset{\sigma, \pi \geq 0}{\text{maximize}} \quad \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} r_{jt}(s,a) \cdot \pi_{tj}(s,a)$$

subject to … $\forall s \in \mathcal{S}_j$

… $s' \in \mathcal{S}_j, \forall t$

$$\sum_{j \in \mathcal{J}} \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} c_{tlj}(s,a) \cdot \pi_{tj}(s,a) \leq b_{tl} \qquad \forall l, \forall t$$

$$\sum_{a \in \mathcal{A}_j} \pi_{tj}(s,a) = \sigma_{tj}(s) \qquad \forall j, \forall s \in \mathcal{S}_j, \forall t$$

Proposition

The fluid LP constitutes a relaxation of the weakly coupled counting DP.

Agenda

1. ~~Weakly Coupled Counting Dynamic Programs~~
2. ~~The Fluid Approximation~~
3. **Performance Guarantees**
4. Case Study

Recovery of Individual DP Policies: Randomized Approach

Algorithm

1. Solve Fluid LP to obtain the solution $(\sigma^\star, \pi^\star)$.

2. At each time stage $t \in \mathcal{T}$, implement for each patient of group $j \in \mathcal{J}$ the action $a \in \mathcal{A}_j$ with probability

$$\frac{\pi^\star_{tj}(s, a)}{\sigma^\star_{tj}(s)}$$

if the patient is in state $s \in \mathcal{S}_j$.

Recovery of Individual DP Policies: Randomized Approach

Theorem (Expected Case)

Let $\theta^\star$ be *expected* total reward of our randomized policy and θ^{DP} the expected total reward of the optimal DP policy. Then

$$\theta^\star \;\geq\; \theta^{\mathrm{DP}}$$

as well as, with high probability,

$$\sum_{j \in \mathcal{J}} \sum_{i=1}^{n_j} c_{tlj}(\tilde{s}_{t,(j,i)}, \tilde{a}_{t,(j,i)}) \;\leq\; b_{tl} + \epsilon \cdot \sum_{j \in \mathcal{J}} n_j \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} c_{tlj}(s, a) \qquad \forall t, \forall l\,.$$

Recovery of Individual DP Policies: Randomized Approach

Theorem (Worst Case)

Let $\tilde{\theta}^\star$ be *random* total reward of our randomized policy and θ^{DP} the expected total reward of the optimal DP policy. Then, with high probability,

$$\tilde{\theta}^\star \;\geq\; \theta^{\text{DP}} \;-\; \epsilon \cdot \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} n_j \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} r_{jt}(s,a)$$

as well as

$$\sum_{j \in \mathcal{J}} \sum_{i=1}^{n_j} c_{tlj}(\tilde{s}_{t,(j,i)}, \tilde{a}_{t,(j,i)}) \;\leq\; b_{tl} + \epsilon \cdot \sum_{j \in \mathcal{J}} n_j \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} c_{tlj}(s,a) \qquad \forall t, \forall l\,.$$

Recovery of Individual DP Policies: Deterministic Approach

Algorithm

1. Solve Fluid LP to obtain the solution $(\sigma^\star, \pi^\star)$.

2. At each time stage $t \in \mathcal{T}$, implement for group $j \in \mathcal{J}$ each action a in each state s

$$\left\{ \Pr \left[\frac{\pi^\star_{tj}(s,a)}{\sigma^\star_{tj}(s)} \cdot \sigma_{tj}(s) \right] \right\}_{s,a} \quad \text{many times}$$

if the group is in state $\sigma_j \in \mathfrak{S}_j$, where $\Pr$ is a projection.

Recovery of Individual DP Policies: Deterministic Approach

Theorem (Expected Case)

Let $\theta^\star$ be *expected* total reward of our deterministic policy and θ^{DP} the expected total reward of the optimal DP policy. Then

$$\theta^\star \;\geq\; \theta^{\text{DP}} \;-\; \sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}}\sum_{s\in\mathcal{S}_j}\sum_{a\in\mathcal{A}_j} r_{jt}(s,a)\cdot\epsilon_1$$

as well as, with high probability, for all t and l,

$$\sum_{j\in\mathcal{J}}\sum_{s\in\mathcal{S}_j}\sum_{a\in\mathcal{A}_j} c_{tlj}(s,a)\cdot[\pi_t^\star(\tilde{\sigma}_t^\star)]_j(s,a) \;\leq\; b_{tl} + \sum_{j\in\mathcal{J}}(1+\epsilon_2 n_j)\sum_{s\in\mathcal{S}_j}\sum_{a\in\mathcal{A}_j} c_{tlj}(s,a)\cdot\epsilon_1\,.$$

Recovery of Individual DP Policies: Deterministic Approach

Theorem (Worst Case)

Let $\tilde{\theta}^\star$ be *random* total reward of our deterministic policy and θ^{DP} the expected total reward of the optimal DP policy. Then, with high probability,

$$\tilde{\theta}^\star \geq \theta^{\text{DP}} - \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} \sum_{s\in\mathcal{S}_j} \sum_{a\in\mathcal{A}_j} r_{jt}(s,a)\cdot(1+\epsilon_2 n_j)\cdot\epsilon_1$$

as well as, for all t and l,

$$\sum_{j\in\mathcal{J}} \sum_{s\in\mathcal{S}_j} \sum_{a\in\mathcal{A}_j} c_{tlj}(s,a)\cdot[\pi_t^\star(\tilde{\sigma}_t^\star)]_j(s,a) \leq b_{tl} + \sum_{j\in\mathcal{J}}(1+\epsilon_2 n_j)\sum_{s\in\mathcal{S}_j} \sum_{a\in\mathcal{A}_j} c_{tlj}(s,a)\cdot\epsilon_1\,.$$

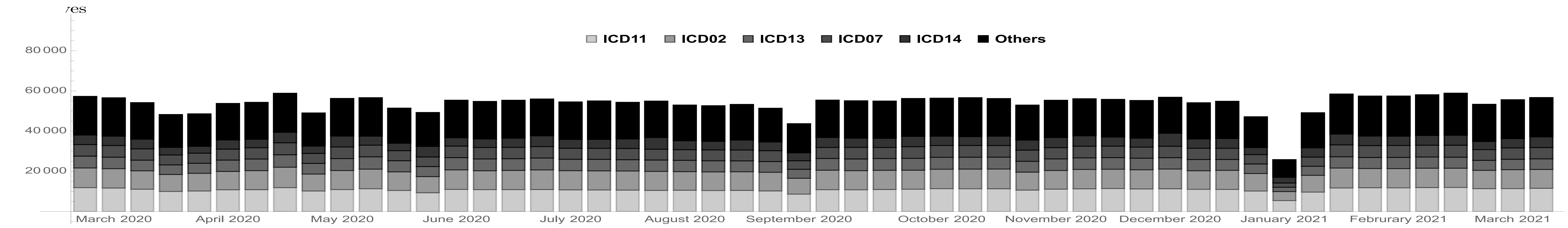
Agenda

1. ~~Weakly Coupled Counting Dynamic Programs~~
2. ~~The Fluid Approximation~~
3. ~~Performance Guarantees~~
4. **Case Study**

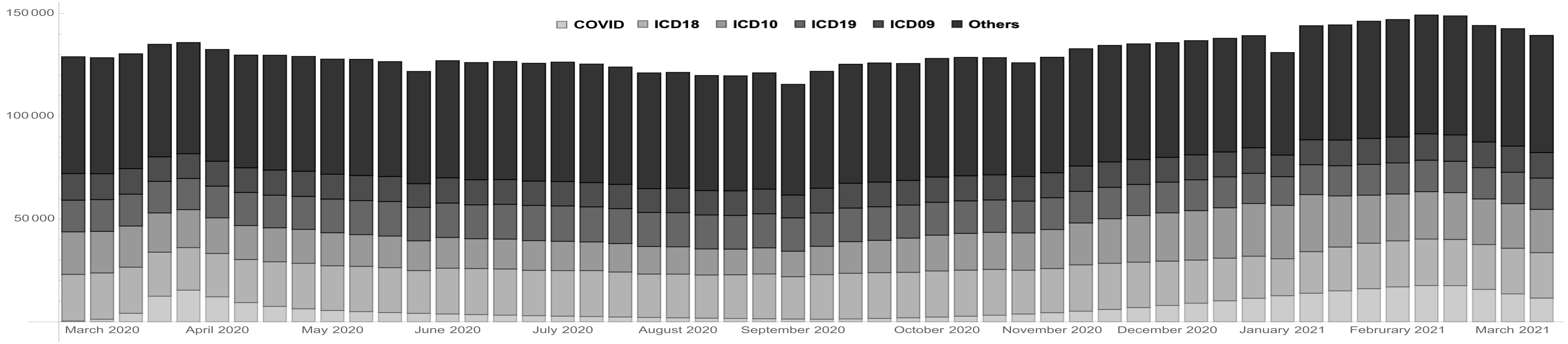
Input Data

	Capacity				Staff-to-bed Ratio		
	Beds	Senior Doctors	Junior Doctors	Nurses	Senior Doctors	Junior Doctors	Nurses
G&A	102,186	10,764	8,539	43,214	15	15	5
CC	4,122	1,013	963	18,856	15	8	1

Electives

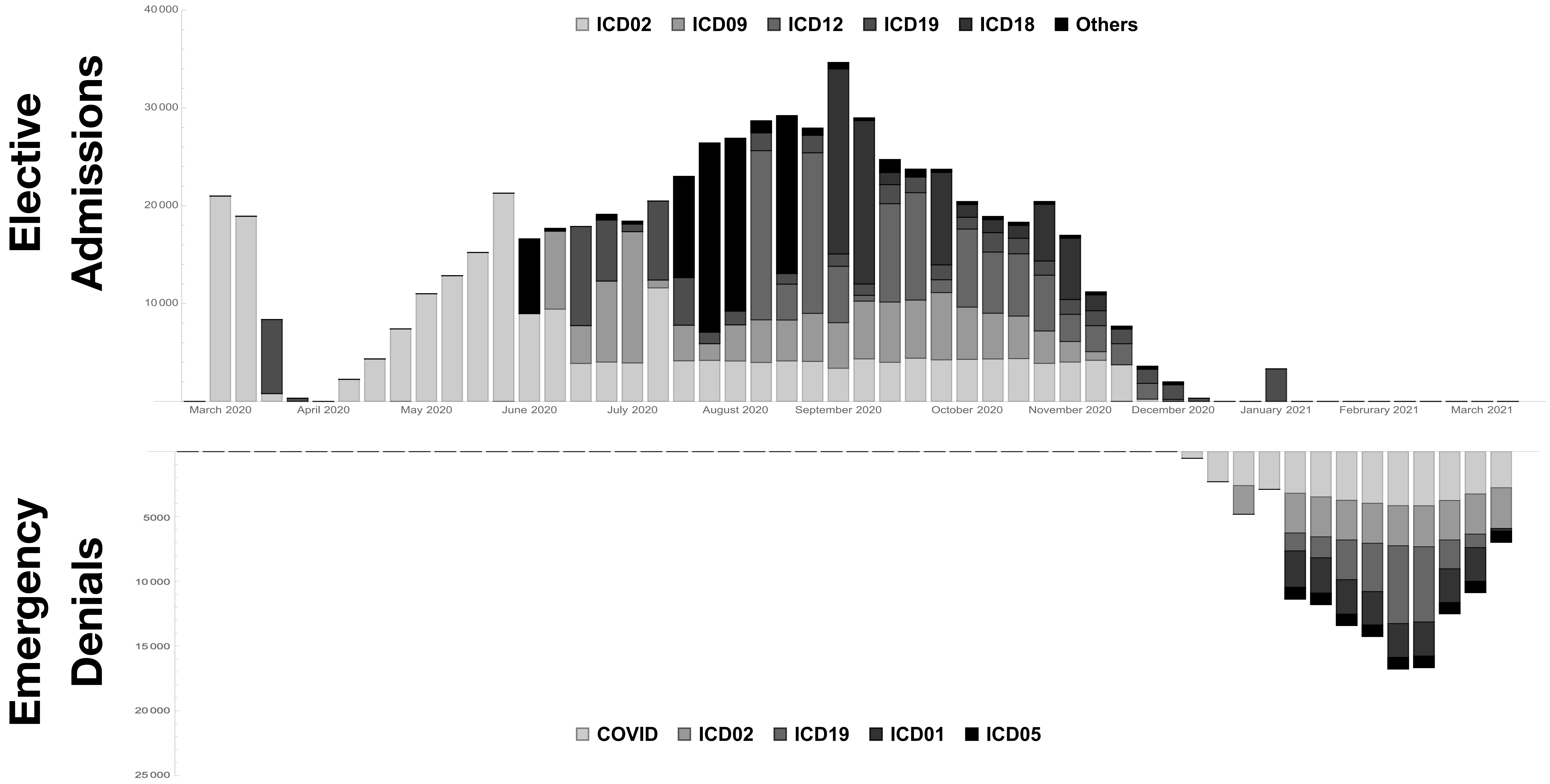


Emergencies



Admissions/Denials: Optimization vs Gov. Policy

Optimized Schedule

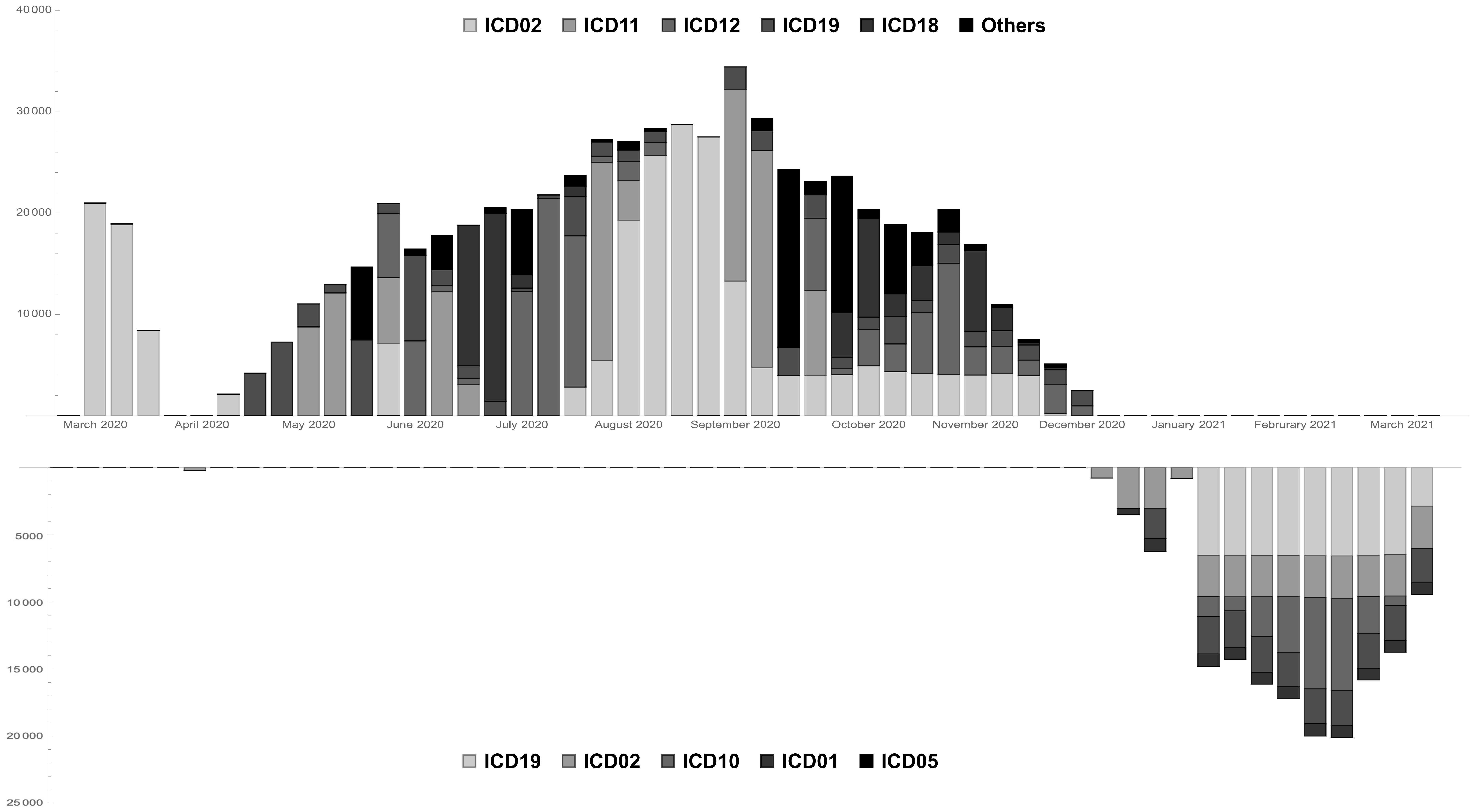


Admissions/Denials: Optimization vs Gov. Policy

Simulation of Government Policy

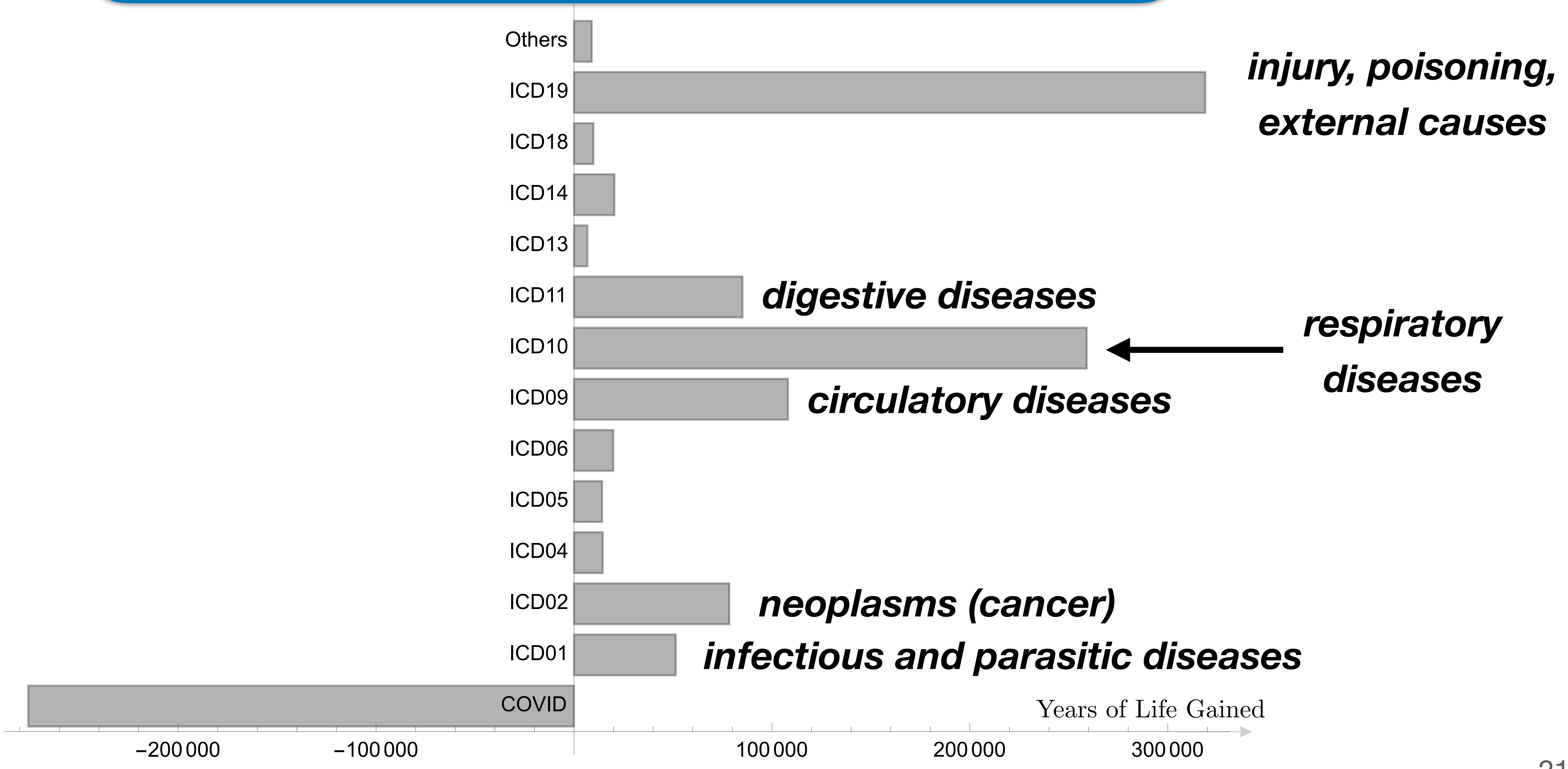
Elective Admissions

Emergency Denials

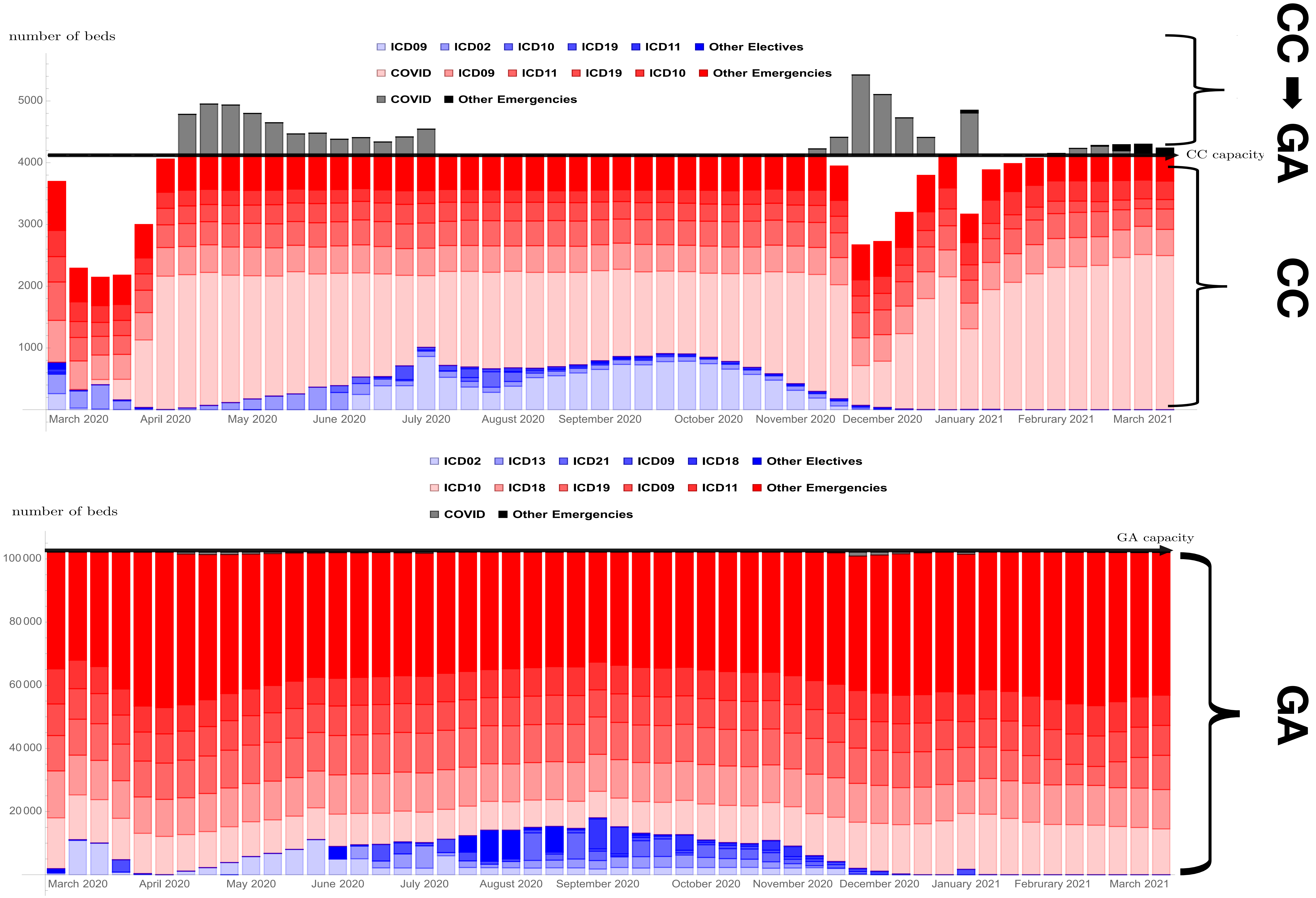


Comparison: Years of Life Gained

Years of Life Gained by Optimized Schedule



Optimized Schedule: Hospital Occupancy



Deterministic Rounding Approach:

- YLL +0.02%, G&A +0.04%, CC +0.31%

Randomization Approach:

- YLL -0.01%, G&A +0.05%, CC +1.56%

Hospital Occupancy:

		Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar
LP	G&A	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
	CC	63%	93%	100%	100%	100%	100%	100%	100%	100%	76%	91%	99%	100%
DR	G&A	100%	101%	100%	100%	100%	100%	100%	100%	100%	100%	99%	101%	100%
	CC	63%	94%	99%	100%	100%	100%	100%	98%	99%	78%	93%	101%	101%
R	G&A	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	99%	100%	100%
	CC	66%	95%	107%	108%	101%	104%	97%	96%	100%	76%	90%	100%	101%

Bibliography

D’Aeth et al. *Optimal Scheduling Rules for Elective Care to Minimize years of Life Lost During the SARS-CoV-2 Pandemic: An Application to England.* Nature Computational Science, 2021.

D’Aeth et al. *Optimal Hospital Care Scheduling During the SARS-CoV-2 Pandemic.* http://www.optimization-online.org/DB_FILE/2021/02/8261.pdf

Esma

Huikang

Shubhy

Stefano

Wolfram

BACKUP SLIDES

Recovery of Individual DP Policies: Randomized Approach

Proof (Expected Total Reward): $\theta^\star \geq \theta^{\text{DP}}$

$$\theta^\star = \mathbb{E}\left[\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}}\sum_{i=1}^{n_j} r_{jt}(\tilde{s}_{t,(j,i)}, \tilde{a}_{t,(j,i)})\right]$$

random state and action of DP i in group j at time t

Recovery of Individual DP Policies: Randomized Approach

Proof (Expected Total Reward): $\theta^\star \geq \theta^{\text{DP}}$

$$\theta^\star = \mathbb{E}\left[\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}}\sum_{i=1}^{n_j} r_{jt}(\tilde{s}_{t,(j,i)}, \tilde{a}_{t,(j,i)})\right]$$

$$= \mathbb{E}\left[\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}}\sum_{i=1}^{n_j}\sum_{s\in\mathcal{S}_j}\sum_{a\in\mathcal{A}_j} r_{jt}(s,a)\cdot \mathbf{1}\left[\tilde{s}_{t,(j,i)} = s \wedge \tilde{a}_{t,(j,i)} = a\right]\right]$$

summation over all possible state-action pairs

Recovery of Individual DP Policies: Randomized Approach

Proof (Expected Total Reward): $\theta^\star \geq \theta^{\mathrm{DP}}$

$$\theta^\star = \mathbb{E}\left[\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}}\sum_{i=1}^{n_j}\sum_{s\in\mathcal{S}_j}\sum_{a\in\mathcal{A}_j} r_{jt}(s,a)\cdot\mathbf{1}\left[\tilde{s}_{t,(j,i)}=s \wedge \tilde{a}_{t,(j,i)}=a\right]\right]$$

$$= \sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}}\sum_{s\in\mathcal{S}_j}\sum_{a\in\mathcal{A}_j} r_{jt}(s,a)\cdot\mathbb{E}\left[\sum_{i=1}^{n_j}\mathbf{1}\left[\tilde{s}_{t,(j,i)}=s \wedge \tilde{a}_{t,(j,i)}=a\right]\right]$$

move expectation inside

Recovery of Individual DP Policies: Randomized Approach

Proof (Expected Total Reward): $\theta^\star \geq \theta^{\text{DP}}$

$$\theta^\star = \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} \sum_{s\in\mathcal{S}_j} \sum_{a\in\mathcal{A}_j} r_{jt}(s,a) \cdot \mathbb{E}\left[\sum_{i=1}^{n_j} \mathbf{1}\left[\tilde{s}_{t,(j,i)} = s \wedge \tilde{a}_{t,(j,i)} = a\right]\right]$$

$$= \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} \sum_{s\in\mathcal{S}_j} \sum_{a\in\mathcal{A}_j} r_{jt}(s,a) \cdot \pi_{tj}^{\text{LP}}(s,a)$$

use definition of expectation

Recovery of Individual DP Policies: Randomized Approach

Proof (Expected Total Reward): $\theta^\star \;\geq\; \theta^{\mathrm{DP}}$

$$\theta^\star = \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} \sum_{s\in\mathcal{S}_j} \sum_{a\in\mathcal{A}_j} r_{jt}(s,a) \cdot \mathbb{E}\left[\sum_{i=1}^{n_j} \mathbf{1}\left[\tilde{s}_{t,(j,i)} = s \,\wedge\, \tilde{a}_{t,(j,i)} = a\right]\right]$$

$$= \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} \sum_{s\in\mathcal{S}_j} \sum_{a\in\mathcal{A}_j} r_{jt}(s,a) \cdot \pi^{\mathrm{LP}}_{tj}(s,a) \;=\; \theta^{\mathrm{LP}}$$

use definition of objective function

Recovery of Individual DP Policies: Randomized Approach

Proof (Expected Total Reward): $\theta^\star \geq \theta^{\text{DP}}$

$$\theta^\star = \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} \sum_{s\in\mathcal{S}_j} \sum_{a\in\mathcal{A}_j} r_{jt}(s,a) \cdot \mathbb{E}\left[\sum_{i=1}^{n_j} \mathbf{1}\left[\tilde{s}_{t,(j,i)} = s \wedge \tilde{a}_{t,(j,i)} = a\right]\right]$$

$$= \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} \sum_{s\in\mathcal{S}_j} \sum_{a\in\mathcal{A}_j} r_{jt}(s,a) \cdot \pi_{tj}^{\text{LP}}(s,a) = \theta^{\text{LP}} \geq \theta^{\text{DP}}$$

fluid LP is relaxation of fluid DP

Recovery of Individual DP Policies: Randomized Approach

Proof (Resource Violation): $\sum_{j \in \mathcal{J}} \sum_{i=1}^{n_j} c_{tlj}(\tilde{s}_{t,(j,i)}, \tilde{a}_{t,(j,i)}) \leq b_{tl} + \epsilon \cdot \sum_{j \in \mathcal{J}} n_j \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} c_{tlj}(s, a)$

$$\sum_{j \in \mathcal{J}} \sum_{i=1}^{n_j} c_{tlj}(\tilde{s}_{t,(j,i)}, \tilde{a}_{t,(j,i)})$$

Recovery of Individual DP Policies: Randomized Approach

Proof (Resource Violation): $\sum_{j \in \mathcal{J}} \sum_{i=1}^{n_j} c_{tlj}(\tilde{s}_{t,(j,i)}, \tilde{a}_{t,(j,i)}) \;\le\; b_{tl} + \epsilon \cdot \sum_{j \in \mathcal{J}} n_j \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} c_{tlj}(s, a)$

$$\sum_{j \in \mathcal{J}} \sum_{i=1}^{n_j} c_{tlj}(\tilde{s}_{t,(j,i)}, \tilde{a}_{t,(j,i)})$$

$$= \sum_{j \in \mathcal{J}} \sum_{i=1}^{n_j} \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} c_{tlj}(s, a) \cdot \mathbf{1}\left[\tilde{s}_{t,(j,i)} = s \,\wedge\, \tilde{a}_{t,(j,i)} = a\right]$$

summation over all
possible state-action pairs

Recovery of Individual DP Policies: Randomized Approach

Proof (Resource Violation): $\sum_{j \in \mathcal{J}} \sum_{i=1}^{n_j} c_{tlj}(\tilde{s}_{t,(j,i)}, \tilde{a}_{t,(j,i)}) \leq b_{tl} + \epsilon \cdot \sum_{j \in \mathcal{J}} n_j \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} c_{tlj}(s,a)$

$$\sum_{j \in \mathcal{J}} \sum_{i=1}^{n_j} \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} c_{tlj}(s,a) \cdot \mathbf{1}\left[\tilde{s}_{t,(j,i)} = s \wedge \tilde{a}_{t,(j,i)} = a\right]$$

$$= \sum_{j \in \mathcal{J}} \sum_{s \in \mathcal{S}_j} \sum_{a \in \mathcal{A}_j} c_{tlj}(s,a) \cdot \sum_{i=1}^{n_j} \mathbf{1}\left[\tilde{s}_{t,(j,i)} = s \wedge \tilde{a}_{t,(j,i)} = a\right]$$

move sum inside

Recovery of Individual DP Policies: Randomized Approach

Proof (Resource Violation): $\sum_{j\in\mathcal{J}}\sum_{i=1}^{n_j} c_{tlj}(\tilde{s}_{t,(j,i)}, \tilde{a}_{t,(j,i)}) \ \leq \ b_{tl} + \epsilon \cdot \sum_{j\in\mathcal{J}} n_j \sum_{s\in\mathcal{S}_j}\sum_{a\in\mathcal{A}_j} c_{tlj}(s,a)$

$$\sum_{j\in\mathcal{J}}\sum_{i=1}^{n_j}\sum_{s\in\mathcal{S}_j}\sum_{a\in\mathcal{A}_j} c_{tlj}(s,a)\cdot \mathbf{1}\left[\tilde{s}_{t,(j,i)} = s \ \wedge \ \tilde{a}_{t,(j,i)} = a\right]$$

$$= \sum_{j\in\mathcal{J}}\sum_{s\in\mathcal{S}_j}\sum_{a\in\mathcal{A}_j} c_{tlj}(s,a)\cdot \sum_{i=1}^{n_j}\mathbf{1}\left[\tilde{s}_{t,(j,i)} = s \ \wedge \ \tilde{a}_{t,(j,i)} = a\right]$$

$$\leq \ b_{tl} + \epsilon \sum_{j\in\mathcal{J}} n_j \cdot \sum_{s\in\mathcal{S}_j}\sum_{a\in\mathcal{A}_j} c_{tlj}(s,a)$$

Hoeffding's inequality
(fluid LP gives expected value)