Dynamic Capacity Management for Deferred Surgeries

Kartikey Sharma

kartikey.sharma@zib.de
Zuse Institute Berlin

Spotlight Talk

Joint work with Eojin Han, Kristian Singh, and Omid Nohadani

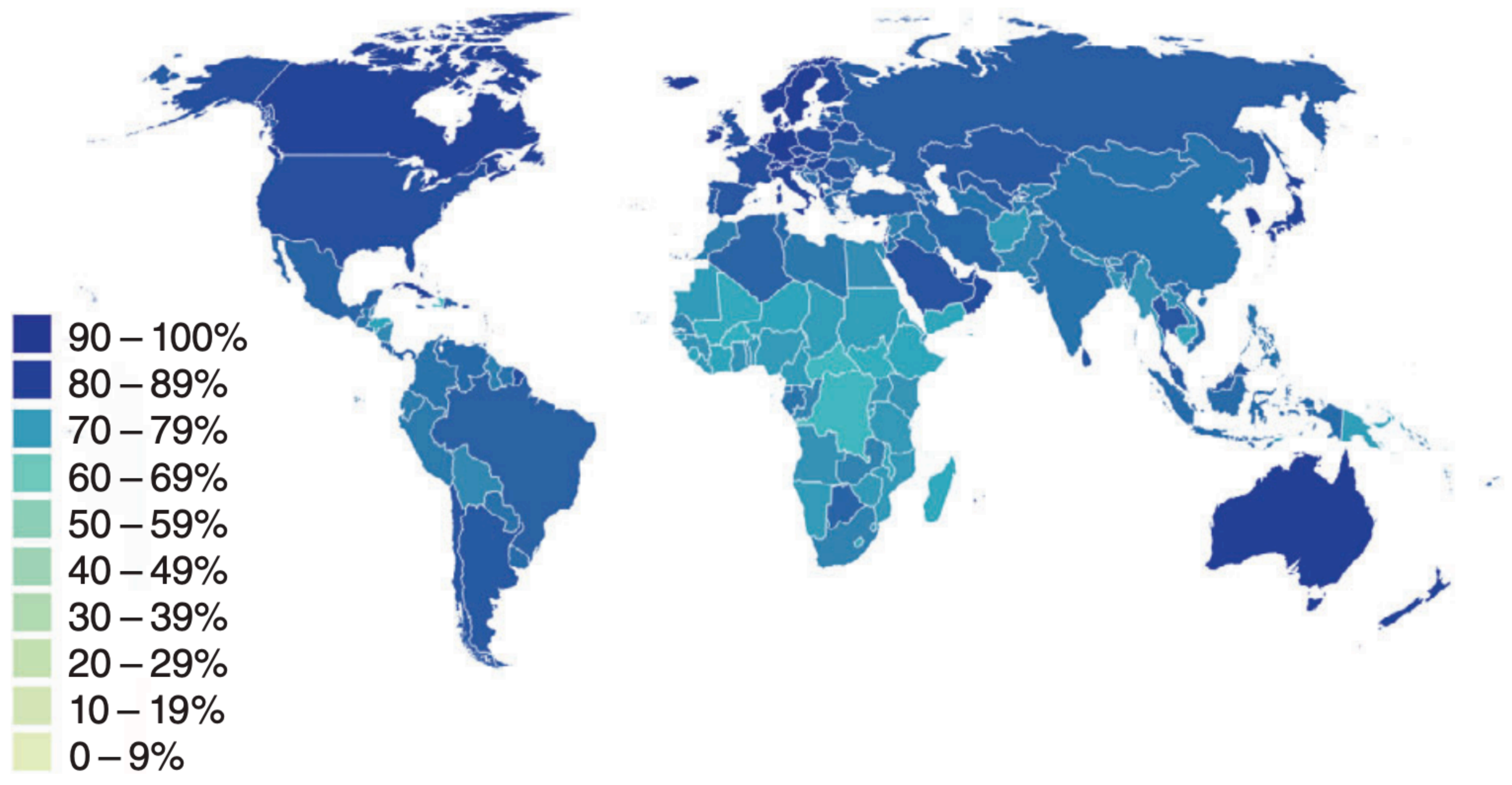
SMU LYLE SCHOOL OF ENGINEERING

ZIB ZUSE INSTITUTE BERLIN

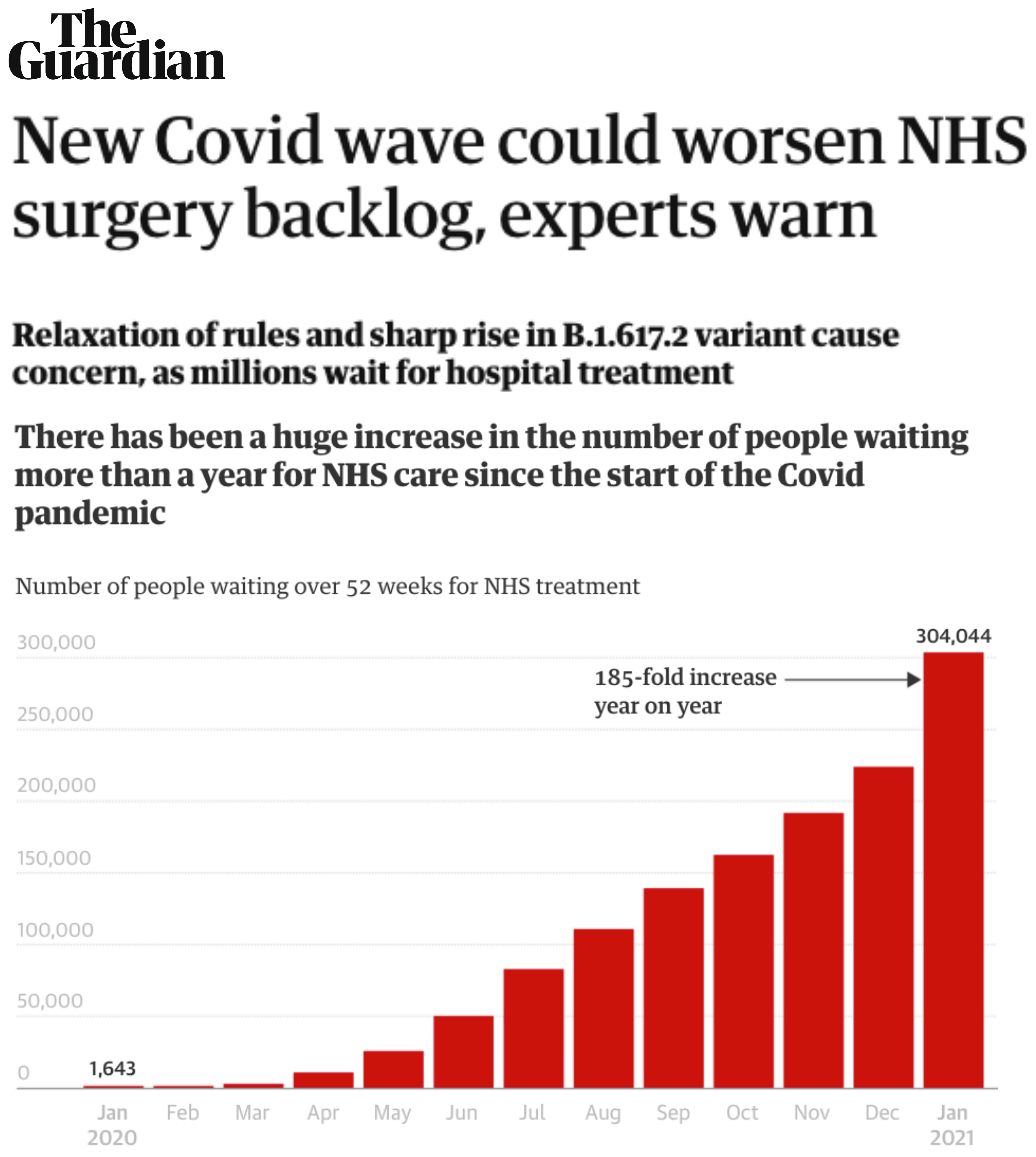
Benefits Science™ Technologies

Deferred Surgeries due to COVID-19

- 12-week cancellation rates of surgery for benign disease (March to May 2020)



Source: COVIDSurg Collaborative (2020) Elective surgery cancellations due to the COVID-19 pandemic: global predictive modeling to inform surgical recovery plans. British Journal of Surgery, 107(11): 1440-1449.


The Guardian

New Covid wave could worsen NHS surgery backlog, experts warn

Relaxation of rules and sharp rise in B.1.617.2 variant cause concern, as millions wait for hospital treatment

There has been a huge increase in the number of people waiting more than a year for NHS care since the start of the Covid pandemic

Number of people waiting over 52 weeks for NHS treatment

Source: D. Campbell. ‘A truly frightening backlog’: ex-NHS chief warns of delays in vital care. The Guardian, April 2, 2021 / N. Davis and D. Campbell. New Covid wave could worsen NHS surgery backlog, experts warn. The Guardian, May 20, 2021.

Cost of Deferred Elective Surgeries

Cost of Deferred Elective Surgeries

- Potentially, worst health care outcomes for patients due to delayed treatment

Cost of Deferred Elective Surgeries

- Potentially, worst health care outcomes for patients due to delayed treatment

- Increased financial costs for hospitals and insurers due to worsened diseases

Cost of Deferred Elective Surgeries

- Potentially, worst health care outcomes for patients due to delayed treatment
- Increased financial costs for hospitals and insurers due to worsened diseases
- Significant financial loss for hospitals
  - Average monthly loss of revenue of the U.S. hospitals is $50.7 billion for March-June 2020 (Meredith et al. 2020).
  - Elective surgeries account for 43% of gross revenue of the U.S. hospitals (Tonna et al. 2020).

Source: Meredith, High, and Freischlag (2020) Preserving elective surgeries in the COVID-19 pandemic and the future. JAMA 324(17):1725-1726. Tonna, Hanson, Cohan, McCrum, Horns, Brooke, Das, Kelly, Campbell, and Hotaling (2020) Balancing revenue generation with capacity generation: case distribution, financial impact and hospital capacity changes from cancelling or resuming elective surgeries in the US during COVID-19. BMC Health Services Research 20(1):1-7.

Capacity Management for Deferred Surgeries

Capacity Management for Deferred Surgeries

- Expanding surgical capacity of hospitals will expedite patient treatment
  - Improved health outcomes
  - Lower treatment costs

Capacity Management for Deferred Surgeries

- Expanding surgical capacity of hospitals will expedite patient treatment
  - Improved health outcomes
  - Lower treatment costs

- Current policies are rather adhoc
  - No expansion, or expanding capacities by pre-determined factors.

Capacity Management for Deferred Surgeries

- Expanding surgical capacity of hospitals will expedite patient treatment
  - Improved health outcomes
  - Lower treatment costs

- Current policies are rather adhoc
  - No expansion, or expanding capacities by pre-determined factors.

- The continuously changing patient demand requires the capacity has to adjust *dynamically*.

Capacity Management for Deferred Surgeries

- Expanding surgical capacity of hospitals will expedite patient treatment
  - Improved health outcomes
  - Lower treatment costs

- Current policies are rather adhoc
  - No expansion, or expanding capacities by pre-determined factors.

- The continuously changing patient demand requires the capacity has to adjust *dynamically*.

➡ **Silver Bullet**: An *optimization-based methodology* to dynamically manage surgical capacity for deferred surgeries, while balancing the profit with service requirements.

Problem Set-up

Problem Set-up

Expansion Decisions

- $\mathbf{C}_B = (C_{B,1}, \cdots, C_{B,t})$
- $\mathbf{C} = (C_1, \cdots, C_t)$

Problem Set-up

Expansion Decisions

- $\mathbf{C}_B = (C_{B,1}, \cdots, C_{B,t})$
- $\mathbf{C} = (C_1, \cdots, C_t)$

Surgeries and Deferrals

- $\mathbf{u}_t = (u_t^{(-L)}, \cdots, u_t^{(t)})$: deferred surgeries.
  - $u_t^{(\tau)}$: surgeries scheduled at τ and carried out at t.
- $\mathbf{x}_t = (x_t^{(-L)}, \cdots, x_t^{(t)})$: performed surgeries.

Problem Set-up

Expansion Decisions

- $\mathbf{C}_B = (C_{B,1}, \cdots, C_{B,t})$
- $\mathbf{C} = (C_1, \cdots, C_t)$

Surgeries and Deferrals

- $\mathbf{u}_t = (u_t^{(-L)}, \cdots, u_t^{(t)})$: deferred surgeries.
  - $u_t^{(\tau)}$: surgeries scheduled at τ and carried out at t.
- $\mathbf{x}_t = (x_t^{(-L)}, \cdots, x_t^{(t)})$: performed surgeries.

Uncertainty Sources

- d_t : demand
- $\mathbf{w}_t(\theta_t) = (w_t^{(-L)}, \cdots, w_t^{(t)})$: departing patients.
- $\xi_t = (\theta_t, d_t)$

Problem Set-up

Expansion Decisions

- $\mathbf{C}_B = (C_{B,1}, \cdots, C_{B,t})$
- $\mathbf{C} = (C_1, \cdots, C_t)$

Surgeries and Deferrals

- $\mathbf{u}_t = (u_t^{(-L)}, \cdots, u_t^{(t)})$: deferred surgeries.
  - $u_t^{(\tau)}$: surgeries scheduled at τ and carried out at t.
- $\mathbf{x}_t = (x_t^{(-L)}, \cdots, x_t^{(t)})$: performed surgeries.

Uncertainty Sources

- d_t : demand
- $\mathbf{w}_t(\theta_t) = (w_t^{(-L)}, \cdots, w_t^{(t)})$: departing patients.
- $\xi_t = (\theta_t, d_t)$

State dynamics and constraints:

Problem Set-up

Expansion Decisions

- $\mathbf{C}_B = (C_{B,1}, \cdots, C_{B,t})$
- $\mathbf{C} = (C_1, \cdots, C_t)$

Surgeries and Deferrals

- $\mathbf{u}_t = (u_t^{(-L)}, \cdots, u_t^{(t)})$: deferred surgeries.
  - $u_t^{(\tau)}$: surgeries scheduled at τ and carried out at t.
- $\mathbf{x}_t = (x_t^{(-L)}, \cdots, x_t^{(t)})$: performed surgeries.

Uncertainty Sources

- d_t : demand
- $\mathbf{w}_t(\theta_t) = (w_t^{(-L)}, \cdots, w_t^{(t)})$: departing patients.
- $\xi_t = (\theta_t, d_t)$

State dynamics and constraints:

State dynamics →

$$u_{t+1}^{(\tau)} = u_t^{(\tau)} - x_t^{(\tau)} - w_t^{(\tau)} \quad \forall \tau = -L, \ldots, t$$

Problem Set-up

Expansion Decisions

- $\mathbf{C}_B = (C_{B,1}, \cdots, C_{B,t})$
- $\mathbf{C} = (C_1, \cdots, C_t)$

Surgeries and Deferrals

- $\mathbf{u}_t = (u_t^{(-L)}, \cdots, u_t^{(t)})$: deferred surgeries.
  - $u_t^{(\tau)}$: surgeries scheduled at τ and carried out at t.
- $\mathbf{x}_t = (x_t^{(-L)}, \cdots, x_t^{(t)})$: performed surgeries.

Uncertainty Sources

- d_t : demand
- $\mathbf{w}_t(\theta_t) = (w_t^{(-L)}, \cdots, w_t^{(t)})$: departing patients.
- $\xi_t = (\theta_t, d_t)$

State dynamics and constraints:

State dynamics

$$u_{t+1}^{(\tau)} = u_t^{(\tau)} - x_t^{(\tau)} - w_t^{(\tau)} \quad \forall \tau = -L, \ldots, t$$

Demand & capacity constraints

$$x_t^{(\tau)} \leq u_t^{(\tau)} \qquad \sum_{\tau=-L}^{t} x_t^{(\tau)} \leq \widehat{C}_t + C_{B,t} + C_t$$

Problem Set-up

Expansion Decisions

- $\mathbf{C}_B = (C_{B,1}, \cdots, C_{B,t})$
- $\mathbf{C} = (C_1, \cdots, C_t)$

Surgeries and Deferrals

- $\mathbf{u}_t = (u_t^{(-L)}, \cdots, u_t^{(t)})$: deferred surgeries.
  - $u_t^{(\tau)}$: surgeries scheduled at τ and carried out at t.
- $\mathbf{x}_t = (x_t^{(-L)}, \cdots, x_t^{(t)})$: performed surgeries.

Uncertainty Sources

- d_t : demand
- $\mathbf{w}_t(\theta_t) = (w_t^{(-L)}, \cdots, w_t^{(t)})$: departing patients.
- $\xi_t = (\theta_t, d_t)$

State dynamics and constraints:

State dynamics → $$u_{t+1}^{(\tau)} = u_t^{(\tau)} - x_t^{(\tau)} - w_t^{(\tau)} \quad \forall \tau = -L, \ldots, t$$

Demand & capacity constraints

$$x_t^{(\tau)} \leq u_t^{(\tau)} \qquad \sum_{\tau=-L}^{t} x_t^{(\tau)} \leq \widehat{C}_t + C_{B,t} + C_t$$

$$(\mathbf{C}_B, \mathbf{C}) \in \mathcal{C}$$

Expansion constraints

Dynamic Programming Formulation

- Cost at time t :

$$H_t(C_{B,t}, C_t, \mathbf{u}_t, \mathbf{x}_t, \mathbf{w}_t) := b_{B,t}(\hat{C}_t + C_{B,t}) + b_t C_t + c_t \sum_{\tau=-L}^{t} x_t^{(\tau)} + \sum_{\tau=-L}^{t} p_{t-\tau}(u_t^{(\tau)} - x_t^{(\tau)} - w_t^{(\tau)}) + \sum_{\tau=-L}^{t} f_{t-\tau} w_t^{(\tau)}$$

Dynamic Programming Formulation

- Cost at time t :

$b_{B,t}$: Base expansion cost

$$H_t(C_{B,t}, C_t, \mathbf{u}_t, \mathbf{x}_t, \mathbf{w}_t) := b_{B,t}(\hat{C}_t + C_{B,t}) + b_t C_t + c_t \sum_{\tau=-L}^{t} x_t^{(\tau)} + \sum_{\tau=-L}^{t} p_{t-\tau}(u_t^{(\tau)} - x_t^{(\tau)} - w_t^{(\tau)}) + \sum_{\tau=-L}^{t} f_{t-\tau} w_t^{(\tau)}$$

Dynamic Programming Formulation

- Cost at time t :

$b_{B,t}$: Base expansion cost b_t : Expedited expansion cost

$$H_t(C_{B,t}, C_t, \mathbf{u}_t, \mathbf{x}_t, \mathbf{w}_t) := b_{B,t}(\hat{C}_t + C_{B,t}) + b_t C_t$$
$$+ c_t \sum_{\tau=-L}^{t} x_t^{(\tau)} + \sum_{\tau=-L}^{t} p_{t-\tau}(u_t^{(\tau)} - x_t^{(\tau)} - w_t^{(\tau)}) + \sum_{\tau=-L}^{t} f_{t-\tau} w_t^{(\tau)}$$

Dynamic Programming Formulation

- Cost at time t :

$b_{B,t}$: Base expansion cost

b_t : Expedited expansion cost

$$H_t(C_{B,t}, C_t, \mathbf{u}_t, \mathbf{x}_t, \mathbf{w}_t) := b_{B,t}(\hat{C}_t + C_{B,t}) + b_t C_t$$
$$+ c_t \sum_{\tau=-L}^{t} x_t^{(\tau)} + \sum_{\tau=-L}^{t} p_{t-\tau}(u_t^{(\tau)} - x_t^{(\tau)} - w_t^{(\tau)}) + \sum_{\tau=-L}^{t} f_{t-\tau} w_t^{(\tau)}$$

c_t : Surgery cost

Dynamic Programming Formulation

- Cost at time t :

$b_{B,t}$: Base expansion cost b_t : Expedited expansion cost

$$H_t(C_{B,t}, C_t, \mathbf{u}_t, \mathbf{x}_t, \mathbf{w}_t) := b_{B,t}(\hat{C}_t + C_{B,t}) + b_t C_t + c_t \sum_{\tau=-L}^{t} x_t^{(\tau)} + \sum_{\tau=-L}^{t} p_{t-\tau}(u_t^{(\tau)} - x_t^{(\tau)} - w_t^{(\tau)}) + \sum_{\tau=-L}^{t} f_{t-\tau} w_t^{(\tau)}$$

c_t : Surgery cost $p_{t-\tau}$: Deferral cost

Dynamic Programming Formulation

- Cost at time t :

$b_{B,t}$: Base expansion cost

b_t : Expedited expansion cost

$$H_t(C_{B,t}, C_t, \mathbf{u}_t, \mathbf{x}_t, \mathbf{w}_t) := b_{B,t}(\hat{C}_t + C_{B,t}) + b_t C_t + c_t \sum_{\tau=-L}^{t} x_t^{(\tau)} + \sum_{\tau=-L}^{t} p_{t-\tau}(u_t^{(\tau)} - x_t^{(\tau)} - w_t^{(\tau)}) + \sum_{\tau=-L}^{t} f_{t-\tau} w_t^{(\tau)}$$

c_t : Surgery cost

$p_{t-\tau}$: Deferral cost

$f_{t-\tau}$: Departure cost

Dynamic Programming Formulation

- Cost at time t :

$b_{B,t}$: Base expansion cost b_t : Expedited expansion cost

$$H_t(C_{B,t}, C_t, \mathbf{u}_t, \mathbf{x}_t, \mathbf{w}_t) := b_{B,t}(\hat{C}_t + C_{B,t}) + b_t C_t + c_t \sum_{\tau=-L}^{t} x_t^{(\tau)} + \sum_{\tau=-L}^{t} p_{t-\tau}(u_t^{(\tau)} - x_t^{(\tau)} - w_t^{(\tau)}) + \sum_{\tau=-L}^{t} f_{t-\tau} w_t^{(\tau)}$$

c_t : Surgery cost $p_{t-\tau}$: Deferral cost $f_{t-\tau}$: Departure cost

- Dynamic programming (DP) model:

$$\min_{\mathbf{C}_B, C_1} \mathbb{E}_{d_1}\left[\min_{\mathbf{x}_1} \mathbb{E}_{\mathbf{w}_1}\left[H_1(\cdot) + \min_{C_2} \mathbb{E}_{d_2}\left[\min_{\mathbf{x}_2} \mathbb{E}_{\mathbf{w}_2}\left[H_2(\cdot) + \cdots + \min_{C_T} \mathbb{E}_{d_T}\left[\min_{\mathbf{x}_T} \mathbb{E}_{\mathbf{w}_T}\left[H_T(\cdot)\right]\right] \cdots \right]\right]\right]\right]$$

Dynamic Programming Formulation

- Dynamic programming (DP) model:

$$\min_{\mathbf{C}_B, C_1} \mathbb{E}_{d_1}\left[\min_{\mathbf{x}_1} \mathbb{E}_{\mathbf{w}_1}\left[H_1(\cdot) + \min_{C_2} \mathbb{E}_{d_2}\left[\min_{\mathbf{x}_2} \mathbb{E}_{\mathbf{w}_2}\left[H_2(\cdot) + \cdots + \min_{C_T} \mathbb{E}_{d_T}\left[\min_{\mathbf{x}_T} \mathbb{E}_{\mathbf{w}_T}\left[H_T(\cdot)\right]\right] \cdots\right]$$

Dynamic Programming Formulation

- Dynamic programming (DP) model:

$$\min_{\mathbf{C}_B, C_1} \mathbb{E}_{d_1}\left[\min_{\mathbf{x}_1} \mathbb{E}_{\mathbf{w}_1}\left[H_1(\cdot) + \min_{C_2} \mathbb{E}_{d_2}\left[\min_{\mathbf{x}_2} \mathbb{E}_{\mathbf{w}_2}\left[H_2(\cdot) + \cdots + \min_{C_T} \mathbb{E}_{d_T}\left[\min_{\mathbf{x}_T} \mathbb{E}_{\mathbf{w}_T}\left[H_T(\cdot)\right]\right] \cdots\right]\right.\right.\right.$$

- Challenges

Dynamic Programming Formulation

- Dynamic programming (DP) model:

$$\min_{\mathbf{C}_B, C_1} \mathbb{E}_{d_1}\left[\min_{\mathbf{x}_1} \mathbb{E}_{\mathbf{w}_1}\left[H_1(\cdot) + \min_{C_2} \mathbb{E}_{d_2}\left[\min_{\mathbf{x}_2} \mathbb{E}_{\mathbf{w}_2}\left[H_2(\cdot) + \cdots + \min_{C_T} \mathbb{E}_{d_T}\left[\min_{\mathbf{x}_T} \mathbb{E}_{\mathbf{w}_T}\left[H_T(\cdot)\right]\right]\cdots\right]$$

- Challenges
  - Lack of distributional information

Dynamic Programming Formulation

- Dynamic programming (DP) model:

$$\min_{\mathbf{C}_B, C_1} \mathbb{E}_{d_1}\left[\min_{\mathbf{x}_1} \mathbb{E}_{\mathbf{w}_1}\left[H_1(\cdot) + \min_{C_2} \; \mathbb{E}_{d_2}\left[\min_{\mathbf{x}_2} \mathbb{E}_{\mathbf{w}_2}\left[H_2(\cdot) + \cdots + \min_{C_T} \mathbb{E}_{d_T}\left[\min_{\mathbf{x}_T} \mathbb{E}_{\mathbf{w}_T}\left[H_T(\cdot)\right]\right] \cdots \right]\right.\right.\right.\right.$$

- Challenges
  - Lack of distributional information
  - Model difficult to solve

Demand and Departure Uncertainty

- Departing patients $\mathbf{w}_t$ depends on $\mathbf{u}_t$ and $\mathbf{x}_t$:

$$\mathbf{w}_t \leq \mathbf{u}_t - \mathbf{x}_t \text{ almost surely.}$$

Demand and Departure Uncertainty

- Departing patients $\mathbf{w}_t$ depends on $\mathbf{u}_t$ and $\mathbf{x}_t$:

$$\mathbf{w}_t \leq \mathbf{u}_t - \mathbf{x}_t \text{ almost surely.}$$

- Introduce **departure uncertainty** $\theta_t \in [0,1]$ such that $\mathbf{w}_t = (1 - \theta_t)(\mathbf{u}_t - \mathbf{x}_t)$.

  ➡ θ_t is an uncertain proportion of non-departing patients at time t.

Demand and Departure Uncertainty

- Departing patients $\mathbf{w}_t$ depends on $\mathbf{u}_t$ and $\mathbf{x}_t$:

$$\mathbf{w}_t \leq \mathbf{u}_t - \mathbf{x}_t \text{ almost surely.}$$

- Introduce **departure uncertainty** $\theta_t \in [0,1]$ such that $\mathbf{w}_t = (1 - \theta_t)(\mathbf{u}_t - \mathbf{x}_t)$.

  ➡ θ_t is an uncertain proportion of non-departing patients at time t.

- Now $\mathbf{u}_t$ is described via **multilinear** functions of θ_t , d_t , and $\mathbf{x}_t$ as

$$u_t^{(\tau)} = \left(\prod_{k=\max(\tau,1)}^{t-1} \theta_k \right) d_\tau - \sum_{t'=\max(\tau,1)}^{t-1} \left(\prod_{k=t'}^{t-1} \right) x_{t'}^{(\tau)} \qquad \forall \tau = -L, \cdots, t \ \forall t \in [T] \, .$$

Demand and Departure Uncertainty

- Departing patients $\mathbf{w}_t$ depends on $\mathbf{u}_t$ and $\mathbf{x}_t$:

$$\mathbf{w}_t \le \mathbf{u}_t - \mathbf{x}_t \text{ almost surely.}$$

- Introduce **departure uncertainty** $\theta_t \in [0,1]$ such that $\mathbf{w}_t = (1-\theta_t)(\mathbf{u}_t - \mathbf{x}_t)$.

  ➡ θ_t is an uncertain proportion of non-departing patients at time t.

- Now $\mathbf{u}_t$ is described via **multilinear** functions of θ_t , d_t , and $\mathbf{x}_t$ as

$$u_t^{(\tau)} = \left(\prod_{k=\max(\tau,1)}^{t-1} \theta_k \right) d_\tau - \sum_{t'=\max(\tau,1)}^{t-1} \left(\prod_{k=t'}^{t-1} \right) x_{t'}^{(\tau)} \qquad \forall \tau = -L, \cdots, t \ \forall t \in [T] .$$

- We take a (distributionally) robust optimization approach to address this multilinearity.

Outline of Methods

Outline of Methods

- Robust Optimization (RO)
  - Uncertainties are described via polyhedral and box sets.
  - Decisions are made to minimize the worst-case cost.
  - Introduce the *tree of uncertainty products* and leverage McCormick relaxations to handle multilinear uncertainty.

$$\mathcal{U}_w(\mathbf{u}_T, \mathbf{x}_T) =$$

$$\{\mathbf{w}_t \mid \mathbf{w}_t = \rho_t(\mathbf{u}_t - x_t),\ \boldsymbol{\rho} \in \mathcal{U}_\rho\}$$

$$\mathbf{d} \in \mathcal{U}$$

Outline of Methods

- Robust Optimization (RO)
  - Uncertainties are described via polyhedral and box sets.
  - Decisions are made to minimize the worst-case cost.
  - Introduce the *tree of uncertainty products* and leverage McCormick relaxations to handle multilinear uncertainty.

$$\mathcal{U}_w(\mathbf{u}_T, \mathbf{x}_T) =$$
$$\{\mathbf{w}_t \mid \mathbf{w}_t = \rho_t(\mathbf{u}_t - x_t),\ \boldsymbol{\rho} \in \mathcal{U}_\rho\}$$
$$\mathbf{d} \in \mathcal{U}$$

- Distributionally Robust Optimization (DRO)
  - Uncertainties are described via unknown distributions, which are described via sets.
  - Decisions are made to minimize the worst-case expected cost.
  - Leverage the *mean-absolute deviation (MAD) based ambiguity sets*

$$F \in \mathcal{M}_+ \text{ s.t.}$$
$$\mathbf{P}_F\left(\xi_t \in [\underline{\xi}_t\ ,\ \overline{\xi}_t]\right) = 1$$
$$\mathbf{E}_F\left[|\xi_t - \widehat{\xi}_t|\right] \leq \lambda_{\xi_t}$$
$$\mathbf{E}_F[\xi_t] = \widehat{\xi}_t$$

Outline of Methods

- Robust Optimization (RO)
  - Uncertainties are described via polyhedral and box sets.
  - Decisions are made to minimize the worst-case cost.
  - Introduce the *tree of uncertainty products* and leverage McCormick relaxations to handle multilinear uncertainty.

$$\mathcal{U}_w(\mathbf{u}_T, \mathbf{x}_T) =$$
$$\{\mathbf{w}_t \mid \mathbf{w}_t = \rho_t(\mathbf{u}_t - x_t),\ \boldsymbol{\rho} \in \mathcal{U}_\rho\}$$
$$\mathbf{d} \in \mathcal{U}$$

- Distributionally Robust Optimization (DRO)
  - Uncertainties are described via unknown distributions, which are described via sets.
  - Decisions are made to minimize the worst-case expected cost.
  - Leverage the *mean-absolute deviation (MAD) based ambiguity sets*

$$F \in \mathcal{M}_+ \text{ s.t.}$$
$$\mathbf{P}_F\left(\xi_t \in [\underline{\xi}_t\ ,\ \overline{\xi}_t]\right) = 1$$
$$\mathbf{E}_F\left[|\xi_t - \widehat{\xi}_t|\right] \le \lambda_{\xi_t}$$
$$\mathbf{E}_F[\xi_t] = \widehat{\xi}_t$$

- Numerical Experiments

Outline of Methods

- Robust Optimization (RO)
  - Uncertainties are described via polyhedral and box sets.
  - Decisions are made to minimize the worst-case cost.
  - Introduce the *tree of uncertainty products* and leverage McCormick relaxations to handle multilinear uncertainty.

$$\mathcal{U}_w(\mathbf{u}_T, \mathbf{x}_T) =$$
$$\{\mathbf{w}_t \mid \mathbf{w}_t = \rho_t(\mathbf{u}_t - x_t),\ \boldsymbol{\rho} \in \mathcal{U}_\rho\}$$
$$\mathbf{d} \in \mathcal{U}$$

- Distributionally Robust Optimization (DRO)
  - Uncertainties are described via unknown distributions, which are described via sets.
  - Decisions are made to minimize the worst-case expected cost.
  - Leverage the *mean-absolute deviation (MAD) based ambiguity sets*

$$F \in \mathcal{M}_+ \text{ s.t.}$$
$$\mathbf{P}_F\left(\xi_t \in [\underline{\xi}_t, \overline{\xi}_t]\right) = 1$$
$$\mathbf{E}_F\left[|\xi_t - \widehat{\xi}_t|\right] \leq \lambda_{\xi_t}$$
$$\mathbf{E}_F[\xi_t] = \widehat{\xi}_t$$

- Numerical Experiments

Overall Problem

- Overall problem:

$$\min_{C_t(\cdot),\mathbf{x}_t(\cdot)} \max_{\theta_{[T]}\in\Theta, d_{[T]}\in\mathcal{U}} \sum_{t\in[T]} G_t\left(C_t(\theta_{[t-1]}, d_{[t-1]}), \mathbf{x}_{[t]}(\theta_{[t-1]}, d_{[t]}), \theta_{[t]}, d_{[t]}\right)$$

$$\text{s.t.} \sum_{t'=\max(\tau,1)}^{t}\left(\prod_{k=t'}^{t-1}\theta_k\right) x_{t'}^{(\tau)}(\theta_{[t'-1]}, d_{[t']}) \leq \left(\prod_{k=\max(\tau,1)}^{t-1}\theta_k\right) d_\tau \quad \forall\theta_{[T]}\in\Theta, d_{[T]}\in\mathcal{U}, \tau\in[-L:t], t\in[T]$$

$$\sum_{\tau\in[-L:t]} x_t^{(\tau)}(\theta_{[t-1]}, d_{[t]}) \leq \hat{C}_t + C_{B,t} + C_t(\theta_{[t-1]}, d_{[t-1]}) \quad \forall\theta_{[T]}\in\Theta, d_{[T]}\in\mathcal{U}, t\in[T]$$

$$\mathbf{x}_t(\theta_{[t-1]}, d_{[t]}) \in \mathbb{R}_+^{t+L} \quad \forall\theta_{[T]}\in\Theta, d_{[T]}\in\mathcal{U}, t\in[T]$$

$$(\mathbf{C}_B, C_1, C_2(\theta_1, d_1), \cdots, C_T(\theta_{[T-1]}, d_{[T-1]})) \in \mathcal{C} \quad \forall\theta_{[T]}\in\Theta, d_{[T]}\in\mathcal{U},$$

where $G_t(C_t, \mathbf{x}_{[t]}, \theta_{[t]}, d_{[t]}) :=$

$$b_{B,t}(\hat{C}_t + C_{B,t}) + b_t C_t + \sum_{\tau=-L}^{t} c_t x_t^{(\tau)} + \sum_{\tau=-L}^{t} f_{t-\tau}\left[\left(\prod_{k=\max(\tau,1)}^{t-1}\theta_k\right) d_\tau - \sum_{t'=\max(\tau,1)}^{t}\left(\prod_{k=t'}^{t-1}\theta_k\right) x_{t'}^{(\tau)}\right]$$

$$+ \sum_{\tau=-L}^{t}(p_{t-\tau} - f_{t-\tau})\left[\left(\prod_{k=\max(\tau,1)}^{t}\theta_k\right) d_\tau - \sum_{t'=\max(\tau,1)}^{t}\left(\prod_{k=t'}^{t}\theta_k\right) x_{t'}^{(\tau)}\right].$$

Overall Problem

- Overall problem:

Can be $\mathbb{E}$ (stochastic), or $\sup \mathbb{E}$ (distributionally robust)

$$\min_{C_t(\cdot),\mathbf{x}_t(\cdot)} \max_{\theta_{[T]}\in\Theta, d_{[T]}\in\mathcal{U}} \sum_{t\in[T]} G_t\left(C_t(\theta_{[t-1]}, d_{[t-1]}), \mathbf{x}_{[t]}(\theta_{[t-1]}, d_{[t]}), \theta_{[t]}, d_{[t]}\right)$$

$$\begin{aligned}
\text{s.t.} \quad & \sum_{t'=\max(\tau,1)}^{t} \left(\prod_{k=t'}^{t-1} \theta_k\right) x_{t'}^{(\tau)}(\theta_{[t'-1]}, d_{[t']}) \le \left(\prod_{k=\max(\tau,1)}^{t-1} \theta_k\right) d_\tau && \forall \theta_{[T]} \in \Theta, d_{[T]} \in \mathcal{U}, \tau \in [-L:t], t \in [T] \\
& \sum_{\tau\in[-L:t]} x_t^{(\tau)}(\theta_{[t-1]}, d_{[t]}) \le \hat{C}_t + C_{B,t} + C_t(\theta_{[t-1]}, d_{[t-1]}) && \forall \theta_{[T]} \in \Theta, d_{[T]} \in \mathcal{U}, t \in [T] \\
& \mathbf{x}_t(\theta_{[t-1]}, d_{[t]}) \in \mathbb{R}_+^{t+L} && \forall \theta_{[T]} \in \Theta, d_{[T]} \in \mathcal{U}, t \in [T] \\
& (\mathbf{C}_B, C_1, C_2(\theta_1, d_1), \cdots, C_T(\theta_{[T-1]}, d_{[T-1]})) \in \mathcal{C} && \forall \theta_{[T]} \in \Theta, d_{[T]} \in \mathcal{U},
\end{aligned}$$

where $G_t(C_t, \mathbf{x}_{[t]}, \theta_{[t]}, d_{[t]}) :=$

$$b_{B,t}(\hat{C}_t + C_{B,t}) + b_t C_t + \sum_{\tau=-L}^{t} c_t x_t^{(\tau)} + \sum_{\tau=-L}^{t} f_{t-\tau} \left[\left(\prod_{k=\max(\tau,1)}^{t-1} \theta_k\right) d_\tau - \sum_{t'=\max(\tau,1)}^{t} \left(\prod_{k=t'}^{t-1} \theta_k\right) x_{t'}^{(\tau)}\right]$$

$$+ \sum_{\tau=-L}^{t} (p_{t-\tau} - f_{t-\tau}) \left[\left(\prod_{k=\max(\tau,1)}^{t} \theta_k\right) d_\tau - \sum_{t'=\max(\tau,1)}^{t} \left(\prod_{k=t'}^{t} \theta_k\right) x_{t'}^{(\tau)}\right].$$

Overall Problem

- Overall problem:

Can be $\mathbb{E}$ (stochastic), or $\sup \mathbb{E}$ (distributionally robust)

$$\min_{C_t(\cdot),\mathbf{x}_t(\cdot)} \max_{\theta_{[T]}\in\Theta, d_{[T]}\in\mathcal{U}} \sum_{t\in[T]} G_t\left(C_t(\theta_{[t-1]}, d_{[t-1]}), \mathbf{x}_{[t]}(\theta_{[t-1]}, d_{[t]}), \theta_{[t]}, d_{[t]}\right)$$

$$\text{s.t.} \sum_{t'=\max(\tau,1)}^{t} \left(\prod_{k=t'}^{t-1} \theta_k\right) x_{t'}^{(\tau)}(\theta_{[t'-1]}, d_{[t']}) \le \left(\prod_{k=\max(\tau,1)}^{t-1} \theta_k\right) d_\tau \quad \forall \theta_{[T]} \in \Theta, d_{[T]} \in \mathcal{U}, \tau \in [-L:t], t \in [T]$$

$$\sum_{\tau\in[-L:t]} x_t^{(\tau)}(\theta_{[t-1]}, d_{[t]}) \le \hat{C}_t + C_{B,t} + C_t(\theta_{[t-1]}, d_{[t-1]}) \quad \forall \theta_{[T]} \in \Theta, d_{[T]} \in \mathcal{U}, t \in [T]$$

$$\mathbf{x}_t(\theta_{[t-1]}, d_{[t]}) \in \mathbb{R}_+^{t+L} \quad \forall \theta_{[T]} \in \Theta, d_{[T]} \in \mathcal{U}, t \in [T]$$

$$(\mathbf{C}_B, C_1, C_2(\theta_1, d_1), \cdots, C_T(\theta_{[T-1]}, d_{[T-1]})) \in \mathcal{C} \quad \forall \theta_{[T]} \in \Theta, d_{[T]} \in \mathcal{U},$$

where $G_t(C_t, \mathbf{x}_{[t]}, \theta_{[t]}, d_{[t]}) :=$

$$b_{B,t}(\hat{C}_t + C_{B,t}) + b_t C_t + \sum_{\tau=-L}^{t} c_t x_t^{(\tau)} + \sum_{\tau=-L}^{t} f_{t-\tau}\left[\left(\prod_{k=\max(\tau,1)}^{t-1} \theta_k\right) d_\tau - \sum_{t'=\max(\tau,1)}^{t} \left(\prod_{k=t'}^{t-1} \theta_k\right) x_{t'}^{(\tau)}\right]$$

$$+ \sum_{\tau=-L}^{t} (p_{t-\tau} - f_{t-\tau})\left[\left(\prod_{k=\max(\tau,1)}^{t} \theta_k\right) d_\tau - \sum_{t'=\max(\tau,1)}^{t} \left(\prod_{k=t'}^{t} \theta_k\right) x_{t'}^{(\tau)}\right].$$

Multilinear uncertainty

Tree of Uncertainty Products

Tree of Uncertainty Products

- The problem consists of uncertain terms of the form

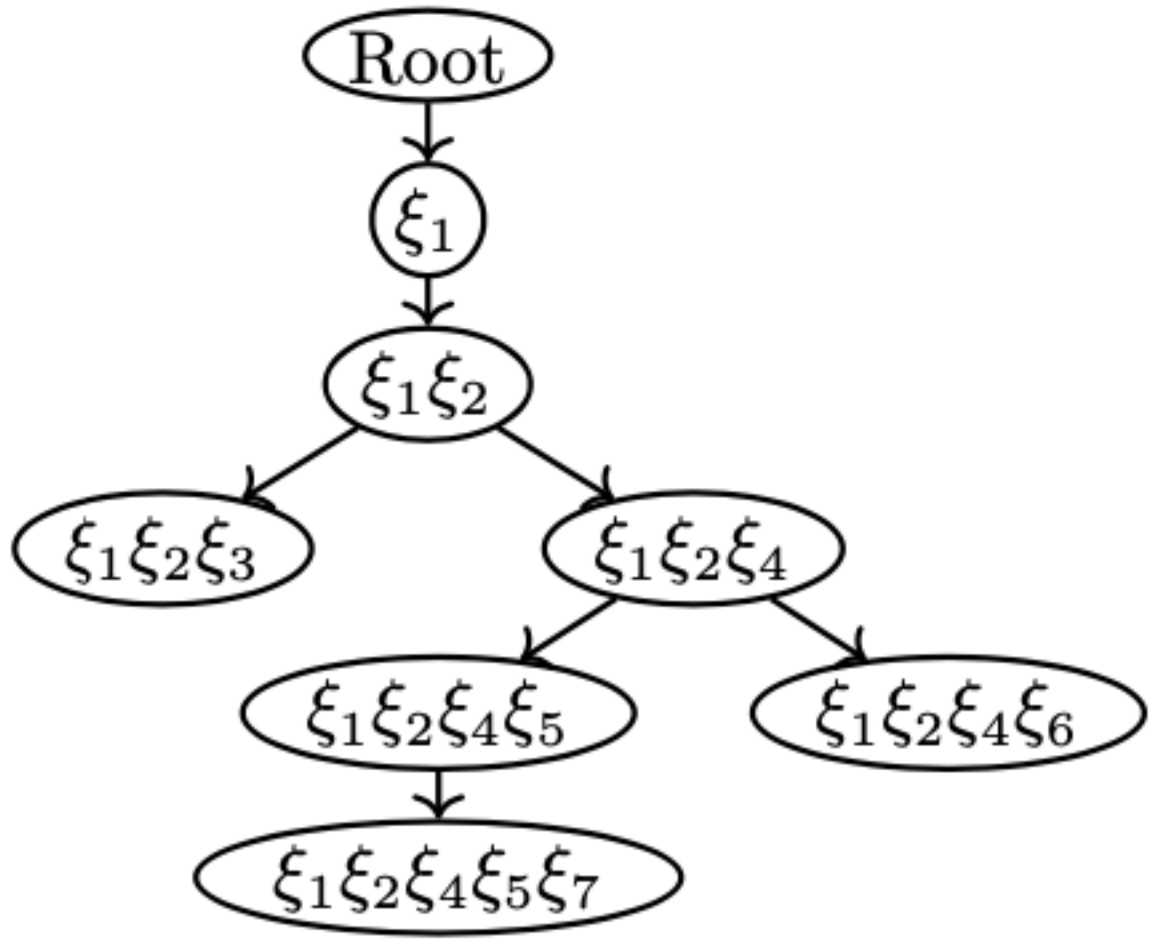
$$\sum_{k\in[K]} q^{(k)}\xi_k + \sum_{n\in[N]} q_g^{(n)} \prod_{i\in S_n} \xi_i \geq q_0 \quad \forall \boldsymbol{\xi} \in \mathcal{U}$$

- This constraint involves sums of multilinear terms.

Tree of Uncertainty Products

- The problem consists of uncertain terms of the form

$$\sum_{k\in[K]} q^{(k)}\xi_k + \sum_{n\in[N]} q_g^{(n)} \prod_{i\in S_n} \xi_i \geq q_0 \quad \forall \boldsymbol{\xi} \in \mathcal{U}$$



- This constraint involves sums of multilinear terms.

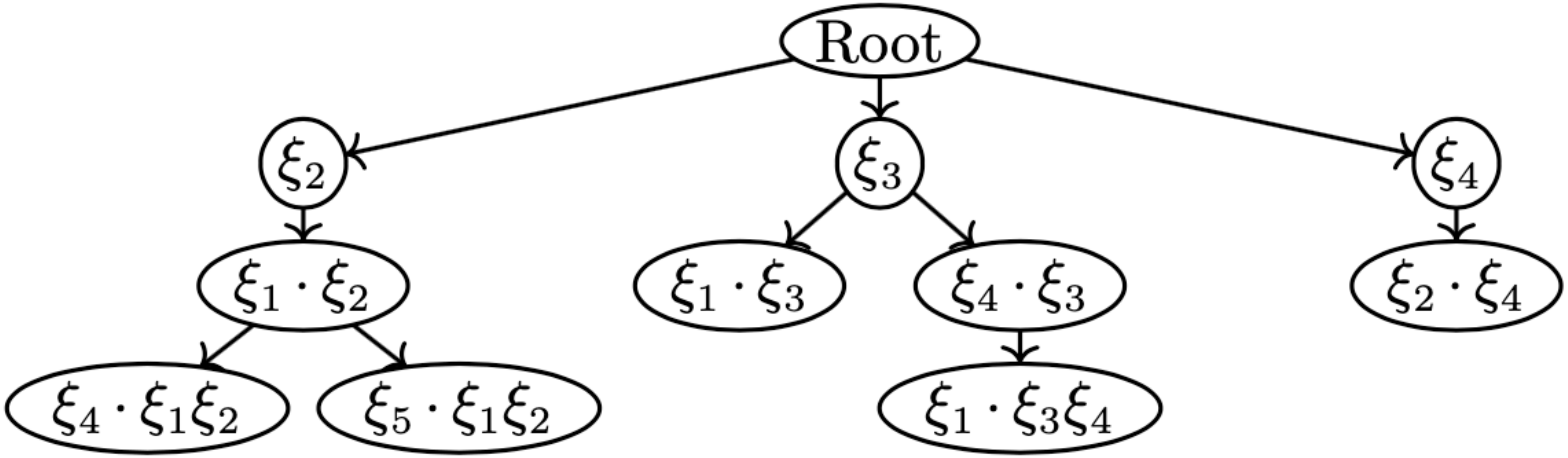
- We show that if
  - the multilinear products are in the form of leaves of tree, and
  - No two leaves without common ancestor share that uncertain component

then the constraint is equivalent to its McCormick relaxation

Tree of Uncertainty Products: Example

$\mathbf{z}^{(0)} = (0,0,0,0,0)$

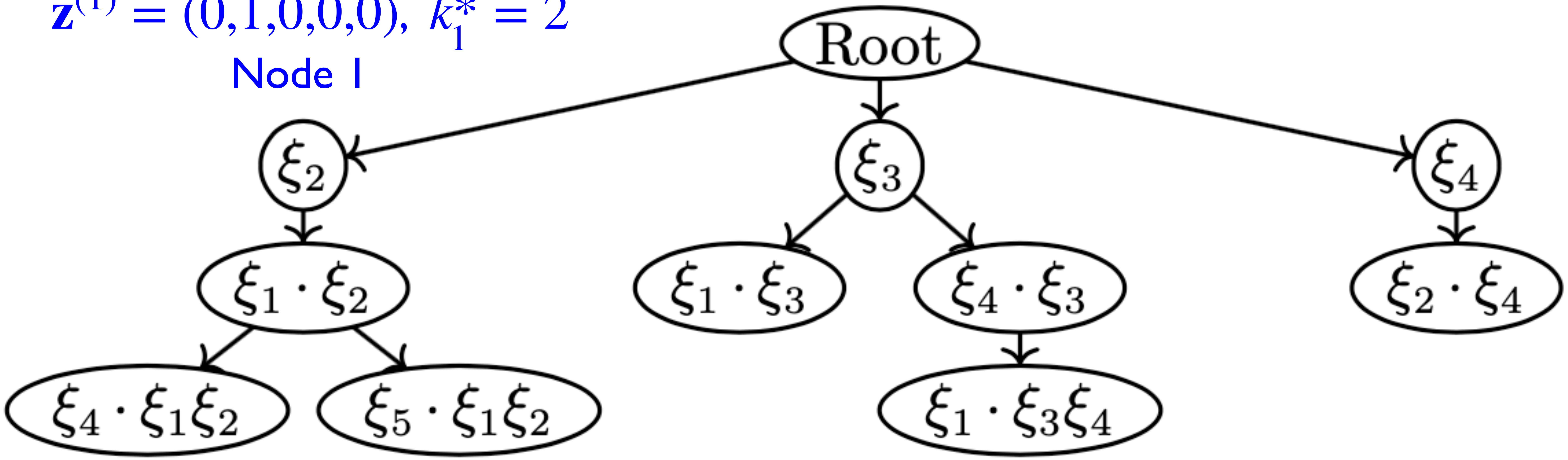
Root node (Node 0)



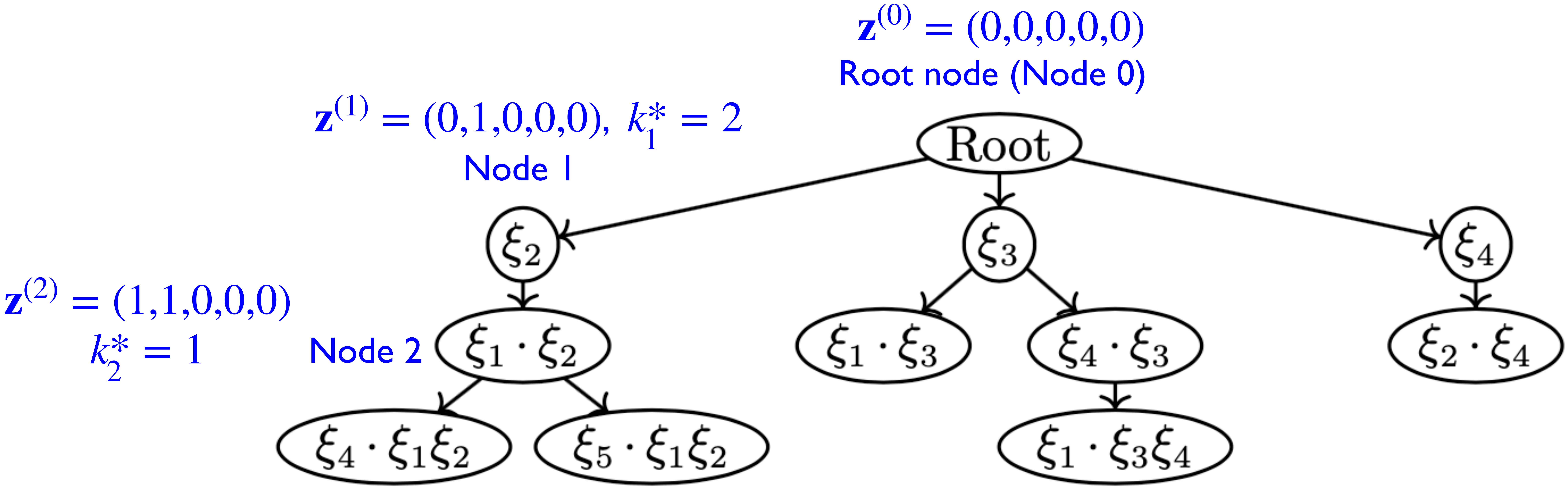
Tree of Uncertainty Products: Example

$\mathbf{z}^{(0)} = (0,0,0,0,0)$
Root node (Node 0)

$\mathbf{z}^{(1)} = (0,1,0,0,0),\ k_1^* = 2$
Node 1



Tree of Uncertainty Products: Example



Tree of Uncertainty Products: Example

$\mathbf{z}^{(0)} = (0,0,0,0,0)$

Root node (Node 0)

$\mathbf{z}^{(1)} = (0,1,0,0,0),\ k_1^* = 2$

Node 1

$\mathbf{z}^{(2)} = (1,1,0,0,0)$

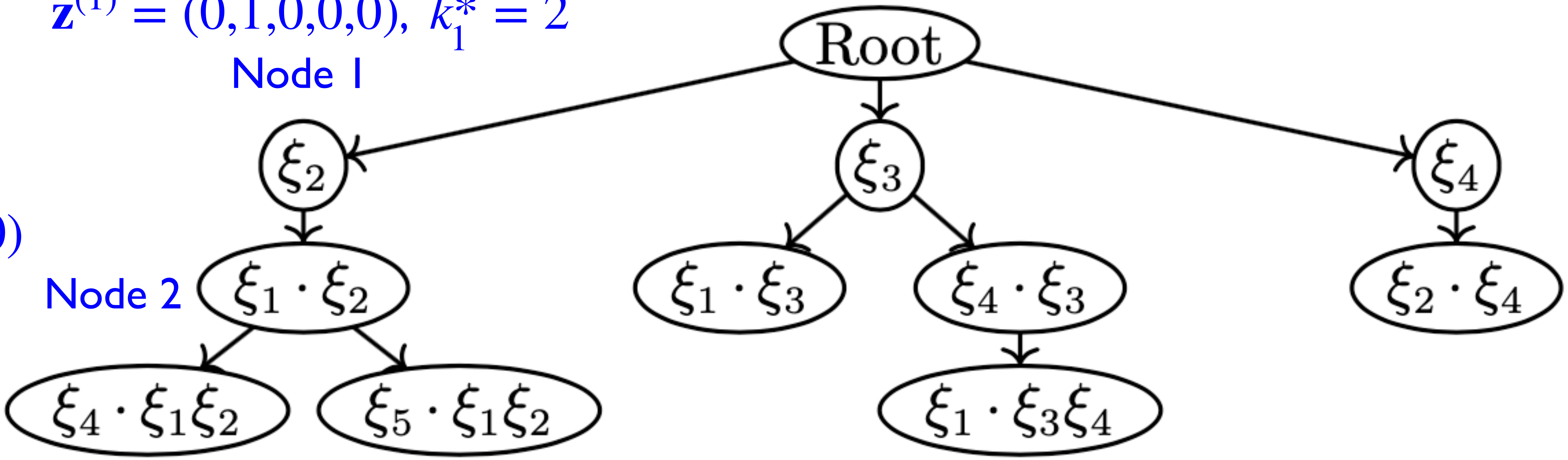
$k_2^* = 1$

Node 2

Node 3

$\mathbf{z}^{(3)} = (1,1,0,1,0)$

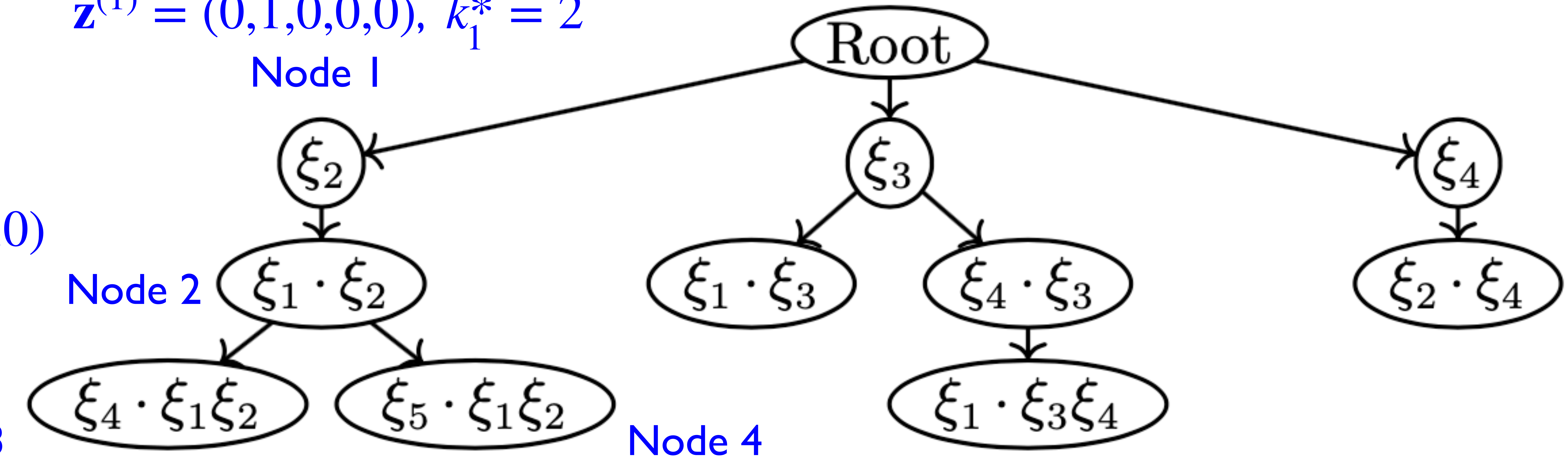
$k_3^* = 4$



Tree of Uncertainty Products: Example

$\mathbf{z}^{(0)} = (0,0,0,0,0)$
Root node (Node 0)

$\mathbf{z}^{(1)} = (0,1,0,0,0),\ k_1^* = 2$
Node 1

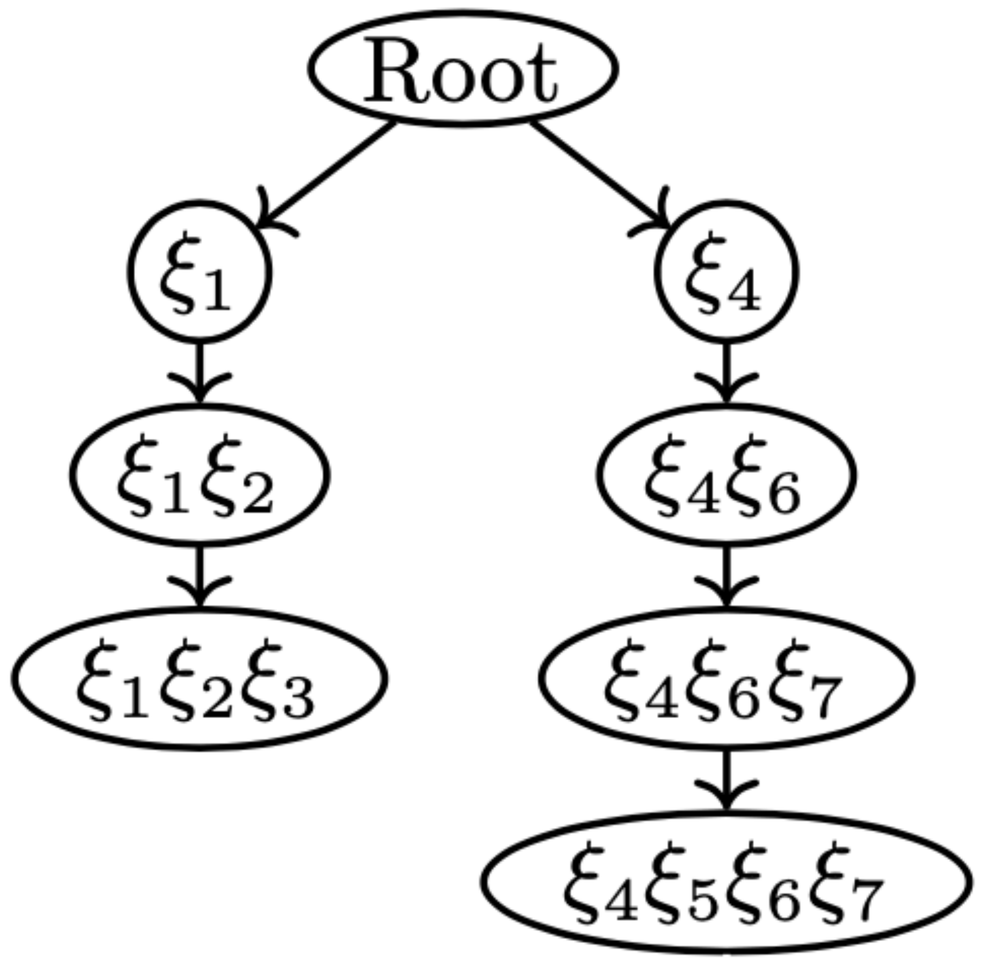


$\mathbf{z}^{(2)} = (1,1,0,0,0)$
$k_2^* = 1$
Node 2

Node 3
$\mathbf{z}^{(3)} = (1,1,0,1,0)$
$k_3^* = 4$

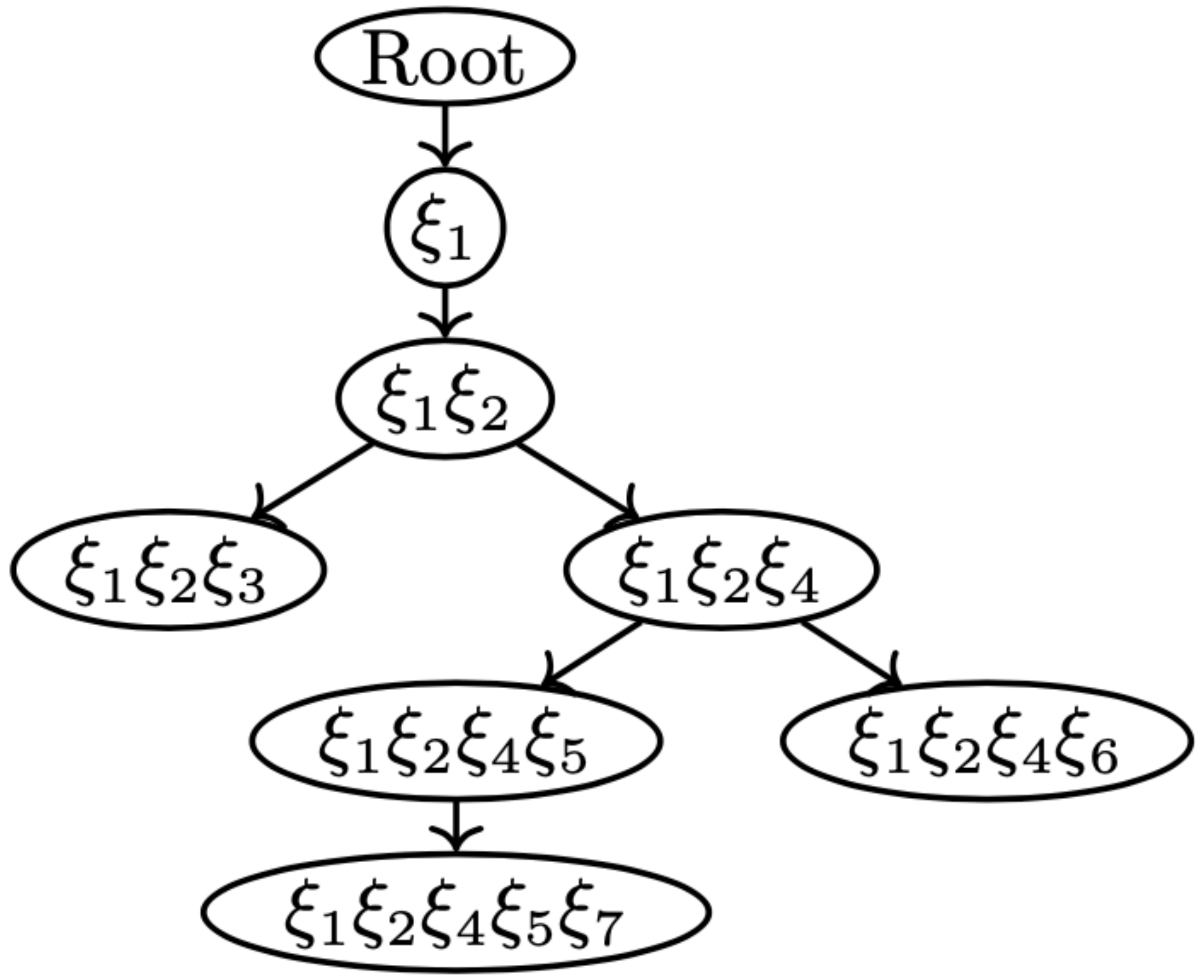
Node 4
$\mathbf{z}^{(4)} = (1,1,0,0,1)$
$k_4^* = 5$

Tree of Uncertainty Products



Exact

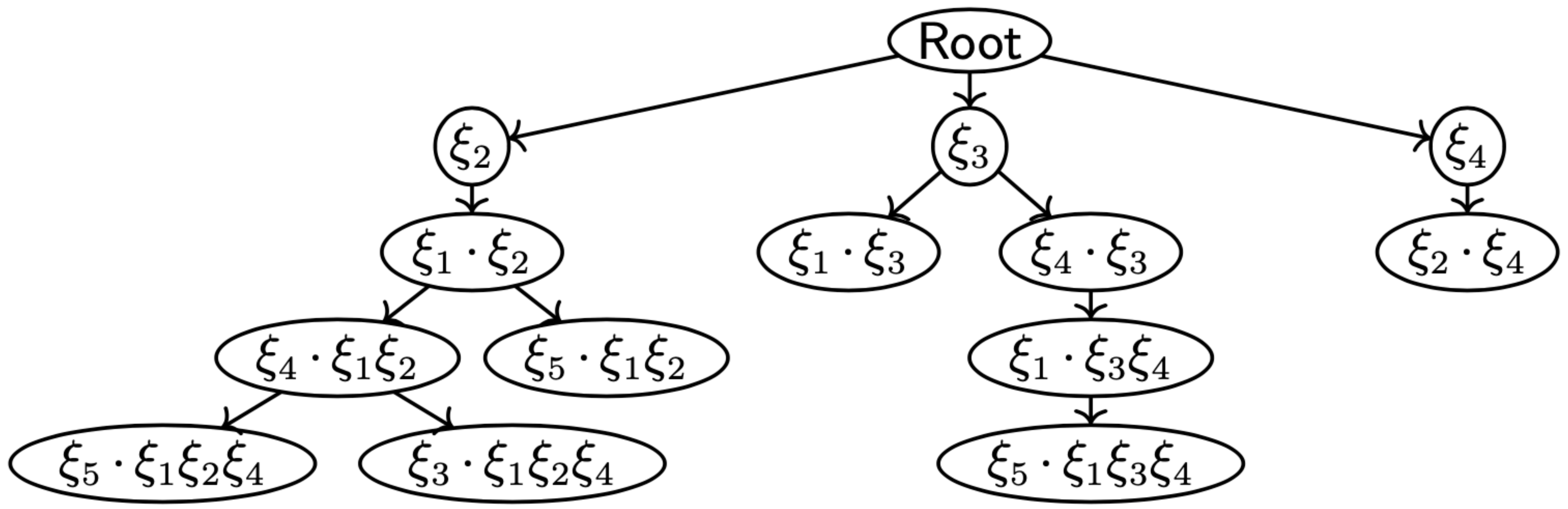
No shared ξ



Exact

No shared ξ

except from ancestor



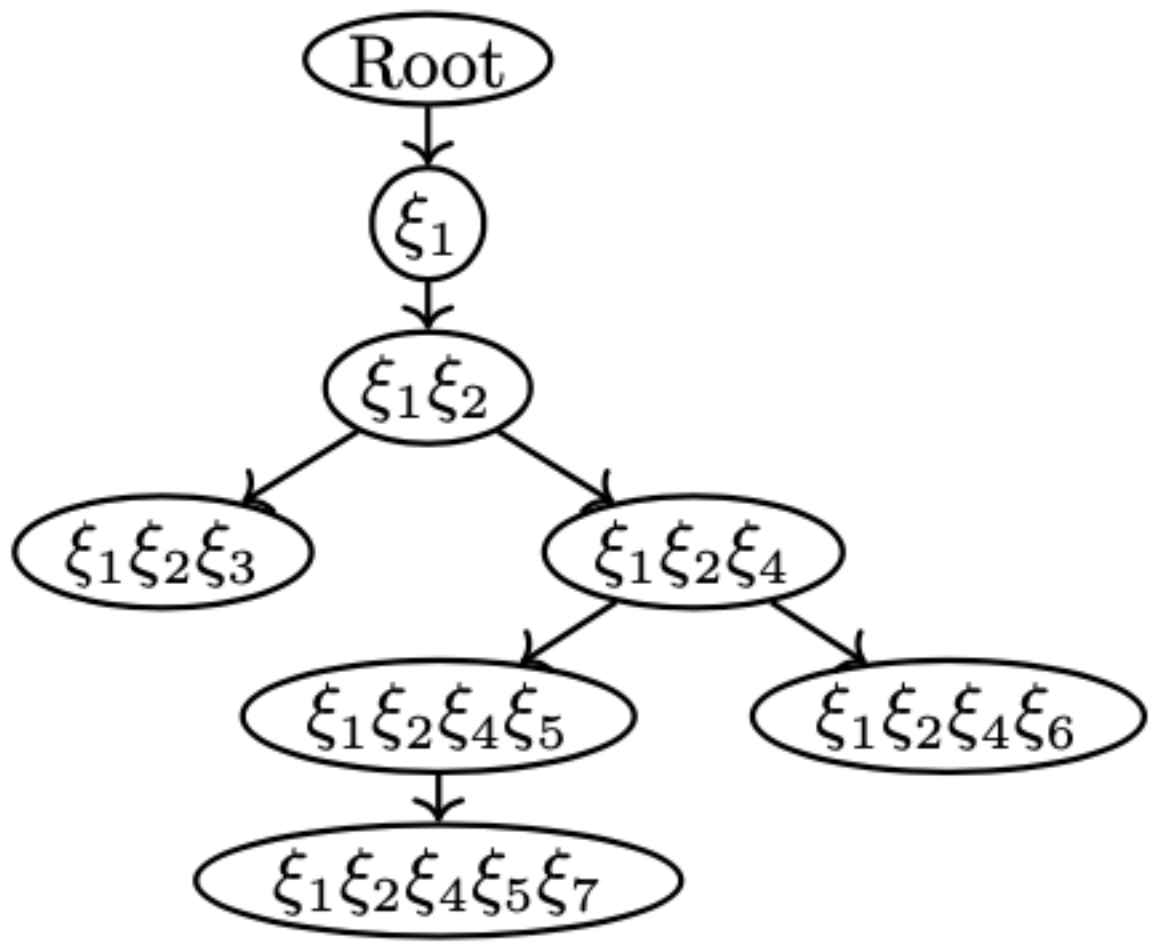
Not Exact

Shared ξ

Tree of Uncertainty Products

Tree of Uncertainty Products

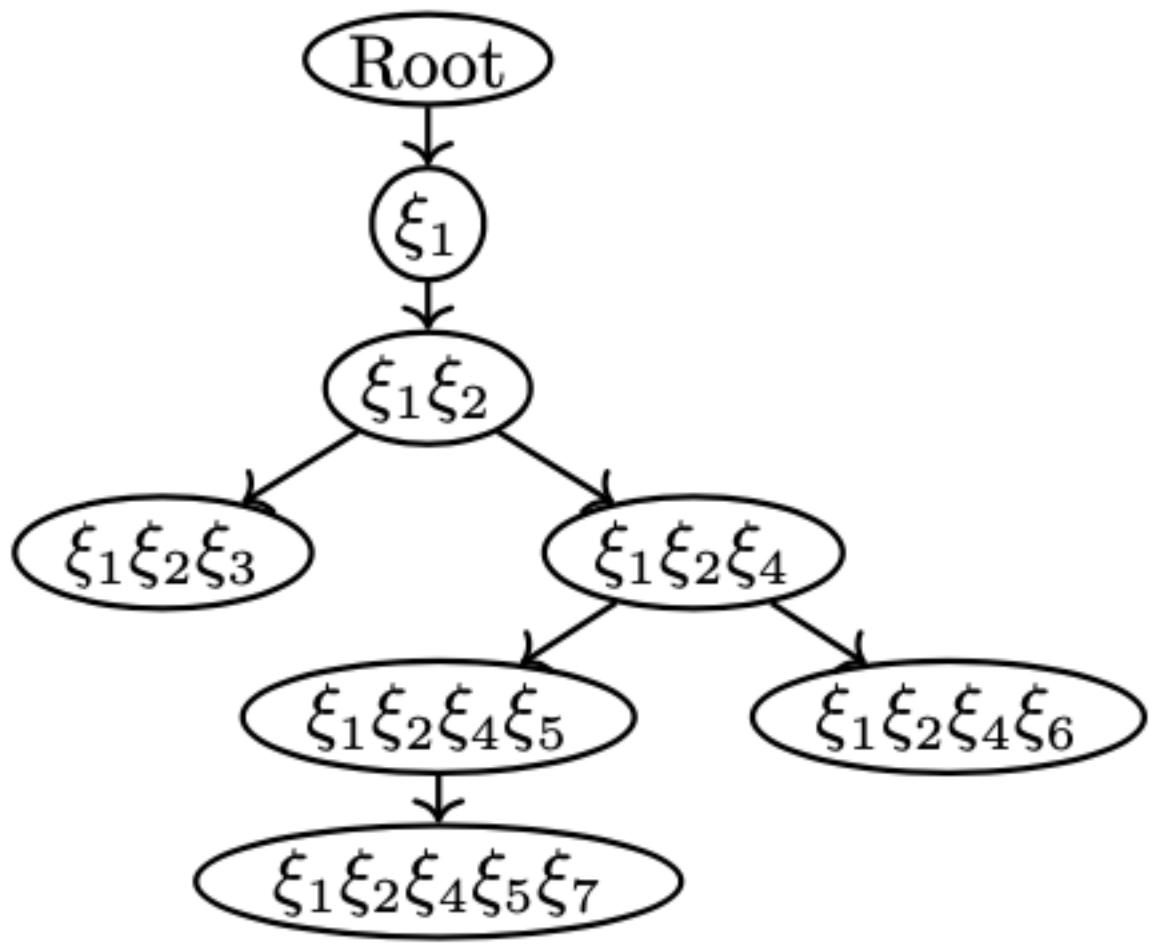
Key Result



Tree of Uncertainty Products

Key Result

$$\sum_{k\in[K]} q^{(k)}\xi_k + \sum_{n\in[N]} q_g^{(n)} \prod_{i\in S_n} \xi_i \geq q_0 \quad \forall \boldsymbol{\xi} \in \mathcal{U}$$



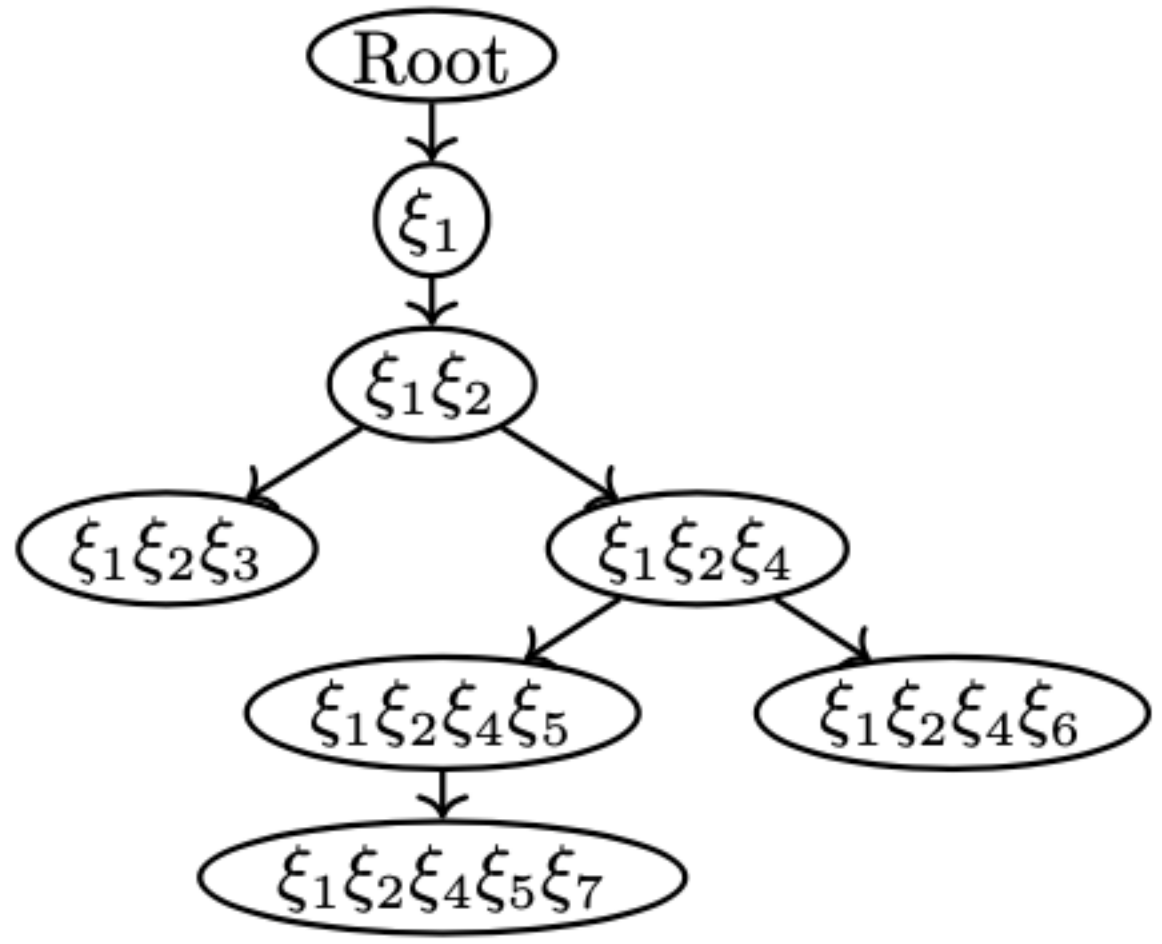
Tree of Uncertainty Products

Key Result

$$\sum_{k\in[K]} q^{(k)}\xi_k + \sum_{n\in[N]} q_g^{(n)} \prod_{i\in S_n} \xi_i \geq q_0 \quad \forall \boldsymbol{\xi} \in \mathcal{U}$$

- Conservative Approximation

$$\sum_{k\in[K]} q^{(k)}\xi_k + \sum_{n\in[N]} q_g^{(n)} \eta_n \geq q_0 \quad \forall (\boldsymbol{\xi}, \boldsymbol{\eta}) \in \overline{\Xi}$$



$$\overline{\Xi} := \left\{ (\boldsymbol{\xi}, \boldsymbol{\eta}) \in \mathbb{R}^{K+N} \;\middle|\; \begin{array}{ll} \boldsymbol{\xi} \in \Xi & \\ \eta_i = \xi_{k_i^*} & \forall i : \ell(i) = 0 \\ \eta_i \geq \overline{\eta}_{\ell(i)} \xi_{k_i^*} + \overline{\xi}_{k_i^*} \eta_{\ell(i)} - \overline{\eta}_{\ell(i)} \overline{\xi}_{k_i^*} & \forall i : \ell(i) \neq 0 \\ \eta_i \geq \underline{\eta}_{\ell(i)} \xi_{k_i^*} + \underline{\xi}_{k_i^*} \eta_{\ell(i)} - \underline{\eta}_{\ell(i)} \underline{\xi}_{k_i^*} & \forall i : \ell(i) \neq 0 \\ \eta_i \leq \overline{\eta}_{\ell(i)} \xi_{k_i^*} + \underline{\xi}_{k_i^*} \eta_{\ell(i)} - \overline{\eta}_{\ell(i)} \underline{\xi}_{k_i^*} & \forall i : \ell(i) \neq 0 \\ \eta_i \leq \underline{\eta}_{\ell(i)} \xi_{k_i^*} + \overline{\xi}_{k_i^*} \eta_{\ell(i)} - \underline{\eta}_{\ell(i)} \overline{\xi}_{k_i^*} & \forall i : \ell(i) \neq 0 \end{array} \right\}.$$

Tree of Uncertainty Products

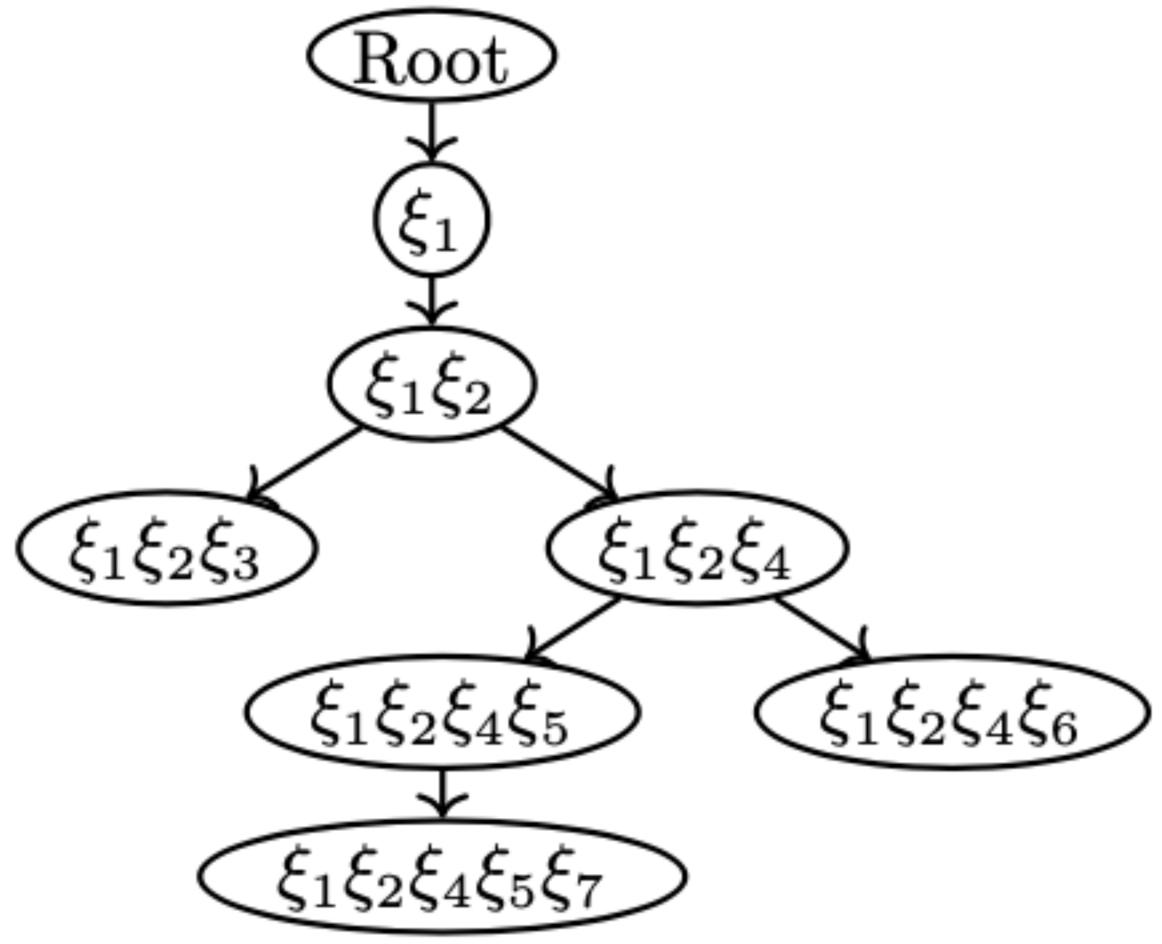
Key Result

$$\sum_{k\in[K]} q^{(k)}\xi_k + \sum_{n\in[N]} q_g^{(n)} \prod_{i\in S_n} \xi_i \geq q_0 \quad \forall \boldsymbol{\xi} \in \mathcal{U}$$

- Conservative Approximation

$$\sum_{k\in[K]} q^{(k)}\xi_k + \sum_{n\in[N]} q_g^{(n)} \eta_n \geq q_0 \quad \forall (\boldsymbol{\xi}, \boldsymbol{\eta}) \in \overline{\Xi}$$

- Exact Reformulation
  - If $\mathscr{U}$ is a box and
  - Tree of Uncertainty Products has no overlap



$$\overline{\Xi} := \left\{ (\boldsymbol{\xi}, \boldsymbol{\eta}) \in \mathbb{R}^{K+N} \;\middle|\; \begin{array}{ll} \boldsymbol{\xi} \in \Xi & \\ \eta_i = \xi_{k_i^*} & \forall i : \ell(i) = 0 \\ \eta_i \geq \overline{\eta}_{\ell(i)} \xi_{k_i^*} + \overline{\xi}_{k_i^*} \eta_{\ell(i)} - \overline{\eta}_{\ell(i)} \overline{\xi}_{k_i^*} & \forall i : \ell(i) \neq 0 \\ \eta_i \geq \underline{\eta}_{\ell(i)} \xi_{k_i^*} + \underline{\xi}_{k_i^*} \eta_{\ell(i)} - \underline{\eta}_{\ell(i)} \underline{\xi}_{k_i^*} & \forall i : \ell(i) \neq 0 \\ \eta_i \leq \overline{\eta}_{\ell(i)} \xi_{k_i^*} + \underline{\xi}_{k_i^*} \eta_{\ell(i)} - \overline{\eta}_{\ell(i)} \underline{\xi}_{k_i^*} & \forall i : \ell(i) \neq 0 \\ \eta_i \leq \underline{\eta}_{\ell(i)} \xi_{k_i^*} + \overline{\xi}_{k_i^*} \eta_{\ell(i)} - \underline{\eta}_{\ell(i)} \overline{\xi}_{k_i^*} & \forall i : \ell(i) \neq 0 \end{array} \right\}.$$

Tractable Approximations

Step 1. Employ **decision rules**, e.g., Linear decision rules

$$x_t^{(\tau)} = w_t^{(\tau)} + \sum_{t'=1}^{t-1} W_{t,t'}^{(\tau)} \theta_{t'} + \sum_{t'=1}^{t} \hat{W}_{t,t'}^{(\tau)} d_{t'}\,, C_t = v_t + \sum_{t'=1}^{t-1} V_{t,t'} \theta_{t'} + \sum_{t'=1}^{t-1} \hat{V}_{t,t'} d_{t'}.$$

Step 2. Substitute to the overall problem and identify all multilinear functions of uncertainties.

Step 3. Form a tree of uncertainty products and approximate with lifted uncertainty sets.

Tractable Approximations

Step 1. Employ **decision rules**, e.g., Linear decision rules

$$x_t^{(\tau)} = w_t^{(\tau)} + \sum_{t'=1}^{t-1} W_{t,t'}^{(\tau)} \theta_{t'} + \sum_{t'=1}^{t} \hat{W}_{t,t'}^{(\tau)} d_{t'}\,,\, C_t = v_t + \sum_{t'=1}^{t-1} V_{t,t'} \theta_{t'} + \sum_{t'=1}^{t-1} \hat{V}_{t,t'} d_{t'}.$$

Step 2. Substitute to the overall problem and identify all multilinear functions of uncertainties.

Step 3. Form a tree of uncertainty products and approximate with lifted uncertainty sets.

Proposition

Under linear decision rules, the multistage problem is approximated as a static robust optimization problem with $\mathcal{O}(T^3)$ uncertain parameters and decision variables.

Tractable Approximations

Step 1. Employ **decision rules**, e.g., Linear decision rules

$$x_t^{(\tau)} = w_t^{(\tau)} + \sum_{t'=1}^{t-1} W_{t,t'}^{(\tau)} \theta_{t'} + \sum_{t'=1}^{t} \hat{W}_{t,t'}^{(\tau)} d_{t'}\ , C_t = v_t + \sum_{t'=1}^{t-1} V_{t,t'} \theta_{t'} + \sum_{t'=1}^{t-1} \hat{V}_{t,t'} d_{t'}.$$

Step 2. Substitute to the overall problem and identify all multilinear functions of uncertainties.

Step 3. Form a tree of uncertainty products and approximate with lifted uncertainty sets.

Proposition

Under linear decision rules, the multistage problem is approximated as a static robust optimization problem with $\mathcal{O}(T^3)$ uncertain parameters and decision variables.

Generalizable to *multilinear* decision rules!

Outline of Methods

- Robust Optimization (RO)
  - Uncertainties are described via polyhedral and box sets.
  - Decisions are made to minimize the worst-case cost.
  - Introduce the *tree of uncertainty products* and leverage McCormick relaxations to handle multilinear uncertainty.

$$\mathcal{U}_w(\mathbf{u}_T, \mathbf{x}_T) =$$
$$\{\mathbf{w}_t \mid \mathbf{w}_t = \rho_t(\mathbf{u}_t - x_t),\ \boldsymbol{\rho} \in \mathcal{U}_\rho\}$$
$$\mathbf{d} \in \mathcal{U}$$

- Distributionally Robust Optimization (DRO)
  - Uncertainties are described via unknown distributions, which are described via sets.
  - Decisions are made to minimize the worst-case expected cost.
  - Leverage the *mean-absolute deviation (MAD) based ambiguity sets*

$$F \in \mathcal{M}_+ \text{ s.t.}$$
$$\mathbf{P}_F\left(\xi_t \in [\underline{\xi}_t\ ,\ \overline{\xi}_t]\right) = 1$$
$$\mathbf{E}_F\left[|\xi_t - \widehat{\xi}_t|\right] \leq \lambda_{\xi_t}$$
$$\mathbf{E}_F[\xi_t] = \widehat{\xi}_t$$

- Numerical Experiments

Outline of DRO Approach

- Uncertainties are described via unknown distributions, which are described via ***ambiguity sets***.
- Decisions are made to minimize the worst-case expected cost.
- Overall formulation:

$$\min_{\mathbf{C}_B, C_1} \sup_{F_{d_1}} \mathbb{E}_{d_1}\left[\min_{\mathbf{x}_1} \sup_{F_{\mathbf{w}_1}} \mathbb{E}_{\mathbf{w}_1}\left[H_1(\cdot) + \min_{C_2} \sup_{F_{d_2}} \mathbb{E}_{d_2}\left[\min_{\mathbf{x}_2} \sup_{F_{\mathbf{w}_2}} \mathbb{E}_{\mathbf{w}_2}\left[H_2(\cdot) + \cdots + \min_{C_T} \sup_{F_{d_T}} \mathbb{E}_{d_T}\left[\min_{\mathbf{x}_T} \sup_{F_{\mathbf{w}_T}} \mathbb{E}_{\mathbf{w}_T}\left[H_T(\cdot)\right]\right] \cdots\right]\right]\right]\right]$$

Mean-MAD Ambiguity Sets

Definition

For the set of non-negative Borel measurable functions $\mathscr{M}_+(\mathbb{R}^{2T}), \lambda_{\theta_t}, \lambda_{d_t} \geq 0$, $0 \leq \underline{\theta}_t < \hat{\theta}_t < \overline{\theta}_t \leq 1$, and $0 \leq \underline{d}_t < \hat{d}_t < \overline{d}_t$, mean-MAD ambiguity set $\mathscr{F}$ is defined as

$$\mathcal{F} = \left\{ F \in \mathcal{M}_+(\mathbb{R}^{2T}) \;\middle|\; \begin{array}{ll} \mathbb{P}_F\left(\theta_t \in [\underline{\theta}_t, \overline{\theta}_t]\right) = 1,\ \mathbb{E}_F[\theta_t] = \widehat{\theta}_t,\ \mathbb{E}_F\left[|\theta_t - \widehat{\theta}_t|\right] \leq \lambda_{\theta_t} & \forall t \in [T] \\ \mathbb{P}_F\left(d_t \in [\underline{d}_t, \overline{d}_t]\right) = 1,\ \mathbb{E}_F[d_t] = \widehat{d}_t,\ \mathbb{E}_F\left[|d_t - \widehat{d}_t|\right] \leq \lambda_{d_t} & \forall t \in [T] \\ \{\theta_{[T]}, d_{[T]}\} \text{ are mutually independent} & \end{array} \right\}.$$

- $\underline{\theta}_t$, $\overline{\theta}_t$, $\underline{d}_t$, $\overline{d}_t$: lower and upper support of θ_t and d_t.
- $\hat{\theta}_t$, $\hat{d}_t$: expectation of θ_t and d_t.
- λ_{θ_t}, λ_{d_t} : mean-absolute deviation bound of θ_t and d_t.

All of them can be easily estimated from (small) data!

Mean-MAD Ambiguity Sets

$$\mathcal{F}_Y = \left\{F \in \mathcal{M}_+(\mathbb{R}) \;\middle|\; \mathbb{P}(\overline{Y} \in [\underline{y}, \overline{y}]) = 1,\; \mathbb{E}[\overline{Y}] = y_0,\; \mathbb{E}[|\overline{Y} - y_0|] \leq \overline{\lambda}\right\} \,,$$

Mean-MAD Ambiguity Sets

$$\mathcal{F}_Y = \left\{F \in \mathcal{M}_+(\mathbb{R}) \;\middle|\; \mathbb{P}(\overline{Y} \in [\underline{y}, \overline{y}]) = 1,\ \mathbb{E}[\overline{Y}] = y_0,\ \mathbb{E}[|\overline{Y} - y_0|] \leq \overline{\lambda}\right\} ,$$

- The supremum of a convex function over this set is a 3 point distribution (Ben-Tal & Hochman 1972).

Mean-MAD Ambiguity Sets

$$\mathcal{F}_Y = \left\{ F \in \mathcal{M}_+(\mathbb{R}) \;\middle|\; \mathbb{P}(\overline{Y} \in [\underline{y}, \overline{y}]) = 1,\; \mathbb{E}[\overline{Y}] = y_0,\; \mathbb{E}[|\overline{Y} - y_0|] \leq \overline{\lambda} \right\},$$

- The supremum of a convex function over this set is a 3 point distribution (Ben-Tal & Hochman 1972).
- We extend existing results to the MAD set with inequality.

Mean-MAD Ambiguity Sets

$$\mathcal{F}_Y = \left\{F \in \mathcal{M}_+(\mathbb{R}) \;\middle|\; \mathbb{P}(\overline{Y} \in [\underline{y}, \overline{y}]) = 1,\ \mathbb{E}[\overline{Y}] = y_0,\ \mathbb{E}[|\overline{Y} - y_0|] \leq \overline{\lambda}\right\},$$

- The supremum of a convex function over this set is a 3 point distribution (Ben-Tal & Hochman 1972).
- We extend existing results to the MAD set with inequality.
- This allows us to reformulate the DRO problem as a Stochastic Optimization problem.
  - We solve this problem using Sample Average Approximation

Mean-MAD Ambiguity Sets

$$\mathcal{F}_Y = \left\{ F \in \mathcal{M}_+(\mathbb{R}) \;\middle|\; \mathbb{P}(\overline{Y} \in [\underline{y}, \overline{y}]) = 1,\ \mathbb{E}[\overline{Y}] = y_0,\ \mathbb{E}[|\overline{Y} - y_0|] \leq \overline{\lambda} \right\} ,$$

Mean-MAD Ambiguity Sets

$$\mathcal{F}_Y = \left\{ F \in \mathcal{M}_+(\mathbb{R}) \;\middle|\; \mathbb{P}(\overline{Y} \in [\underline{y}, \overline{y}]) = 1,\ \mathbb{E}[\overline{Y}] = y_0,\ \mathbb{E}[|\overline{Y} - y_0|] \leq \overline{\lambda} \right\} ,$$

Key Result

- The supremum of a convex function over this set is a 3 point distribution on
  - $\underline{y}, \overline{y}$ and y_0 with probabilities
  - $\dfrac{\overline{\lambda}}{2(y_0 - \underline{y})}, \dfrac{\overline{\lambda}}{2(\overline{y} - y_0)}$ and $1 - \dfrac{\overline{\lambda}}{2(y_0 - \underline{y})} - \dfrac{\overline{\lambda}}{2(\overline{y} - y_0)}$

Mean-MAD Ambiguity Sets

$$\mathcal{F}_Y = \left\{ F \in \mathcal{M}_+(\mathbb{R}) \;\middle|\; \mathbb{P}(\overline{Y} \in [\underline{y}, \overline{y}]) = 1,\; \mathbb{E}[\overline{Y}] = y_0,\; \mathbb{E}[|\overline{Y} - y_0|] \le \overline{\lambda} \right\} ,$$

Key Result

- The supremum of a convex function over this set is a 3 point distribution on
  - $\underline{y}, \overline{y}$ and y_0 with probabilities
  - $\dfrac{\bar{\lambda}}{2(y_0 - \underline{y})}, \dfrac{\bar{\lambda}}{2(\bar{y} - y_0)}$ and $1 - \dfrac{\bar{\lambda}}{2(y_0 - \underline{y})} - \dfrac{\bar{\lambda}}{2(\bar{y} - y_0)}$

Multistage DRO → **Multistage SO**

Reformulation

Theorem

With the mean-MAD ambiguity set $\mathscr{F}$, the multistage DRO problem is reformulated as a stochastic optimization problem with *three-points discrete distributions* for each uncertain parameter.

Reformulation

Theorem

With the mean-MAD ambiguity set $\mathscr{F}$, the multistage DRO problem is reformulated as a stochastic optimization problem with *three-points discrete distributions* for each uncertain parameter.

$$\min_{\mathbf{C}_B, C_1} \sup_{F_{d_1}} \mathbb{E}_{d_1}\left[\min_{\mathbf{x}_1} \sup_{F_{\mathbf{w}_1}} \mathbb{E}_{\mathbf{w}_1}\left[H_1(\cdot) + \min_{C_2} \sup_{F_{d_2}} \mathbb{E}_{d_2}\left[\min_{\mathbf{x}_2} \sup_{F_{\mathbf{w}_2}} \mathbb{E}_{\mathbf{w}_2}\left[H_2(\cdot) + \cdots + \min_{C_T} \sup_{F_{d_T}} \mathbb{E}_{d_T}\left[\min_{\mathbf{x}_T} \sup_{F_{\mathbf{w}_T}} \mathbb{E}_{\mathbf{w}_T}\left[H_T(\cdot)\right]\right] \cdots\right]\right]\right]\right]$$

Reformulation

Theorem

With the mean-MAD ambiguity set $\mathscr{F}$, the multistage DRO problem is reformulated as a stochastic optimization problem with *three-points discrete distributions* for each uncertain parameter.

$$\min_{\mathbf{c}_B, C_1} \mathbb{E}_{d_1 \sim F^*_{d_1}} \left[\min_{\mathbf{x}_1} \mathbb{E}_{\mathbf{w}_1 \sim F^*_{\mathbf{w}_1}} \left[H_1(\cdot) + \min_{C_2} \mathbb{E}_{d_2 \sim F^*_{d_2}} \left[\min_{\mathbf{x}_2} \mathbb{E}_{\mathbf{w}_2 \sim F^*_{\mathbf{w}_2}} \left[H_2(\cdot) + \cdots + \min_{C_T} \mathbb{E}_{d_T \sim F^*_{d_T}} \left[\min_{\mathbf{x}_T} \mathbb{E}_{\mathbf{w}_T \sim F^*_{\mathbf{w}_T}} \left[H_T(\cdot) \right] \right] \cdots \right] \right] \right] \right]$$

Reformulation

Theorem

With the mean-MAD ambiguity set $\mathscr{F}$, the multistage DRO problem is reformulated as a stochastic optimization problem with *three-points discrete distributions* for each uncertain parameter.

$$\min_{\mathbf{c}_B, C_1} \mathbb{E}_{d_1 \sim F^*_{d_1}} \left[\min_{\mathbf{x}_1} \mathbb{E}_{\mathbf{w}_1 \sim F^*_{\mathbf{w}_1}} \left[H_1(\cdot) + \min_{C_2} \mathbb{E}_{d_2 \sim F^*_{d_2}} \left[\min_{\mathbf{x}_2} \mathbb{E}_{\mathbf{w}_2 \sim F^*_{\mathbf{w}_2}} \left[H_2(\cdot) + \cdots + \min_{C_T} \mathbb{E}_{d_T \sim F^*_{d_T}} \left[\min_{\mathbf{x}_T} \mathbb{E}_{\mathbf{w}_T \sim F^*_{\mathbf{w}_T}} \left[H_T(\cdot) \right] \right] \cdots \right. \right. \right. \right.$$

- Under mean and MAD constraints, the worst-case probability distribution is *always* fixed, supported over lower and upper bounds, and their means.
- **Insight**: There exists a class of stochastic optimization problems *whose solutions are distributionally robust*!

Outline of Methods

- Robust Optimization (RO)
  - Uncertainties are described via polyhedral and box sets.
  - Decisions are made to minimize the worst-case cost.
  - Introduce the *tree of uncertainty products* and leverage McCormick relaxations to handle multilinear uncertainty.

$$\mathcal{U}_w(\mathbf{u}_T, \mathbf{x}_T) = \{\mathbf{w}_t \mid \mathbf{w}_t = \rho_t(\mathbf{u}_t - x_t),\ \boldsymbol{\rho} \in \mathcal{U}_\rho\}$$

$$\mathbf{d} \in \mathcal{U}$$

- Distributionally Robust Optimization (DRO)
  - Uncertainties are described via unknown distributions, which are described via sets.
  - Decisions are made to minimize the worst-case expected cost.
  - Leverage the *mean-absolute deviation (MAD) based ambiguity sets*

$$F \in \mathcal{M}_+ \text{ s.t.}$$

$$\mathbf{P}_F\left(\xi_t \in [\underline{\xi}_t\ ,\ \overline{\xi}_t]\right) = 1$$

$$\mathbf{E}_F\left[|\xi_t - \widehat{\xi}_t|\right] \le \lambda_{\xi_t}$$

$$\mathbf{E}_F[\xi_t] = \widehat{\xi}_t$$

- Numerical Experiments

Case Study of Hernia Surgery

Case Study of Hernia Surgery

- Hernia dataset contains all claim records of patients in network from 2017 to 2020.
  - Dates of office visit, surgery (if performed)
  - All payment information with dates for all medical procedures and drug transaction history

Case Study of Hernia Surgery

- Hernia dataset contains all claim records of patients in network from 2017 to 2020.
  - Dates of office visit, surgery (if performed)
  - All payment information with dates for all medical procedures and drug transaction history
- Cost parameters and demand/departure uncertainty information is estimated from the hernia dataset.

Case Study of Hernia Surgery

- Hernia dataset contains all claim records of patients in network from 2017 to 2020.
  - Dates of office visit, surgery (if performed)
  - All payment information with dates for all medical procedures and drug transaction history
- Cost parameters and demand/departure uncertainty information is estimated from the hernia dataset.
- Our analysis estimates current backlog as 4 months of average (pre-pandemic) monthly demand.

Case Study of Hernia Surgery

- Hernia dataset contains all claim records of patients in network from 2017 to 2020.
  - Dates of office visit, surgery (if performed)
  - All payment information with dates for all medical procedures and drug transaction history
- Cost parameters and demand/departure uncertainty information is estimated from the hernia dataset.
- Our analysis estimates current backlog as 4 months of average (pre-pandemic) monthly demand.
- Four methods are implemented and compared:
  - RO: robust optimization-based method
  - DRO: distributionally robust optimization-based method
  - Det60: temporally increase capacity by at most 60% (for ~7 months)
  - Det100: temporally increase capacity by at most 100% (for ~5 months)

Performance Improvement

Departure Level	DRO		RO		Det60		Det100	
	Mean	CVaR90	Mean	CVaR90	Mean	CVaR90	Mean	CVaR90
More Departure	-3969 (10.0)	-2882 (6.31)	-3833 (6.25)	-2927 (7.97)	-2740 (-24.1)	-1871 (-31.0)	-3608 (0.0)	-2711 (0.0)

Performance Improvement

Departure Level	DRO		RO		Det60		Det100	
	Mean	CVaR90	Mean	CVaR90	Mean	CVaR90	Mean	CVaR90
More Departure	-3969 (10.0)	-2882 (6.31)	-3833 (6.25)	-2927 (7.97)	-2740 (-24.1)	-1871 (-31.0)	-3608 (0.0)	-2711 (0.0)

- Both RO and DRO policies achieve better objective values (costs) than deterministic policies.

Performance Improvement

Departure Level	DRO		RO		Det60		Det100	
	Mean	CVaR90	Mean	CVaR90	Mean	CVaR90	Mean	CVaR90
More Departure	-3969 (10.0)	-2882 (6.31)	-3833 (6.25)	-2927 (7.97)	-2740 (-24.1)	-1871 (-31.0)	-3608 (0.0)	-2711 (0.0)

- Both RO and DRO policies achieve better objective values (costs) than deterministic policies.
- DRO performs better in expectation (mean), but RO performs better at higher risk (CVaR90).

Performance Improvement

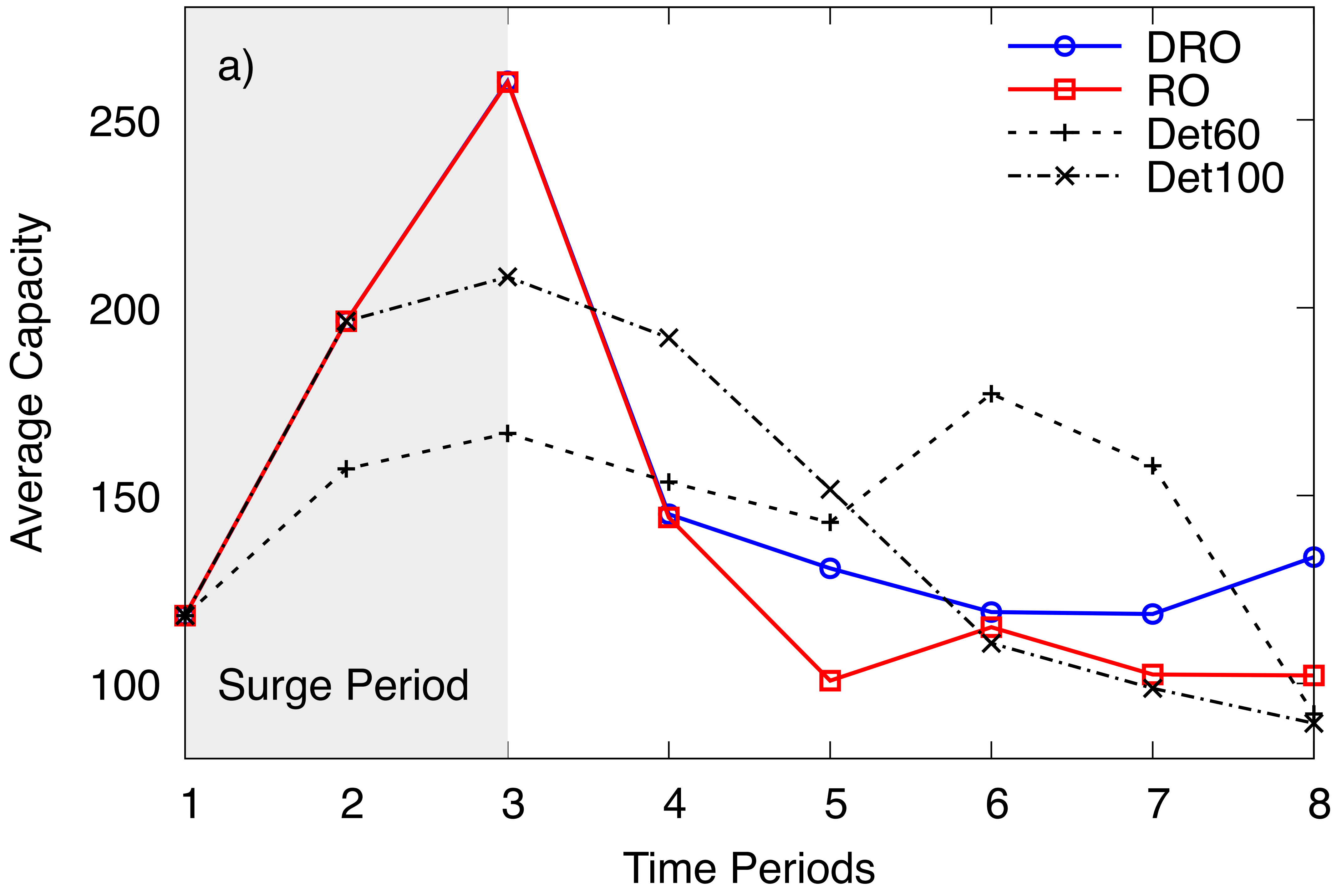
Departure Level	DRO		RO		Det60		Det100	
	Mean	CVaR90	Mean	CVaR90	Mean	CVaR90	Mean	CVaR90
More Departure	-3969 (10.0)	-2882 (6.31)	-3833 (6.25)	-2927 (7.97)	-2740 (-24.1)	-1871 (-31.0)	-3608 (0.0)	-2711 (0.0)
Less Departure	-5078 (7.02)	-4284 (5.49)	-4906 (3.39)	-4306 (6.03)	-4078 (-14.1)	-3446 (-15.1)	-4745 (0.0)	-4061 (0.0)

- Both RO and DRO policies achieve better objective values (costs) than deterministic policies.
- DRO performs better in expectation (mean), but RO performs better at higher risk (CVaR90).

Structure of Expansion Policies

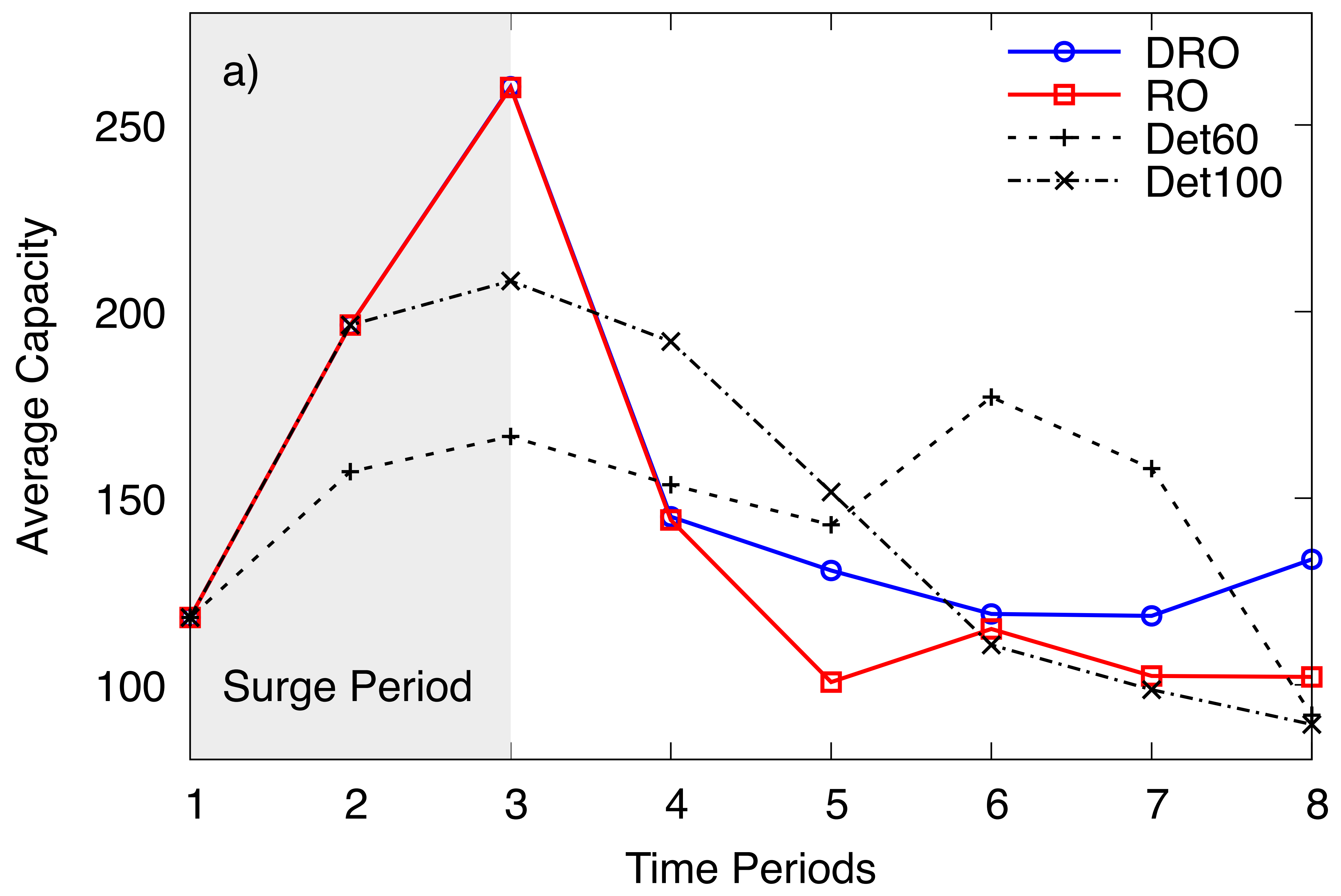
Structure of Expansion Policies

a) More Departure



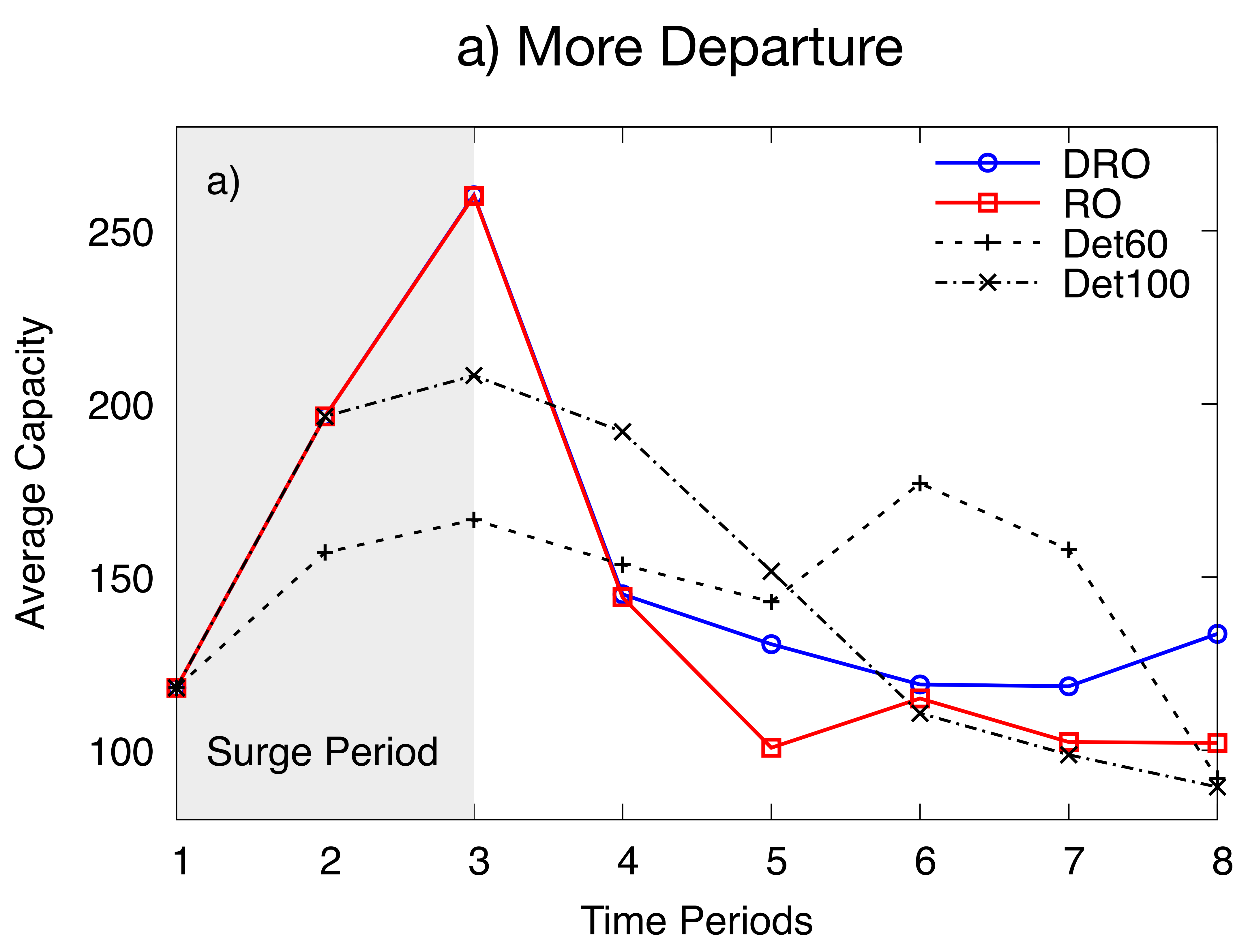
Structure of Expansion Policies

a) More Departure



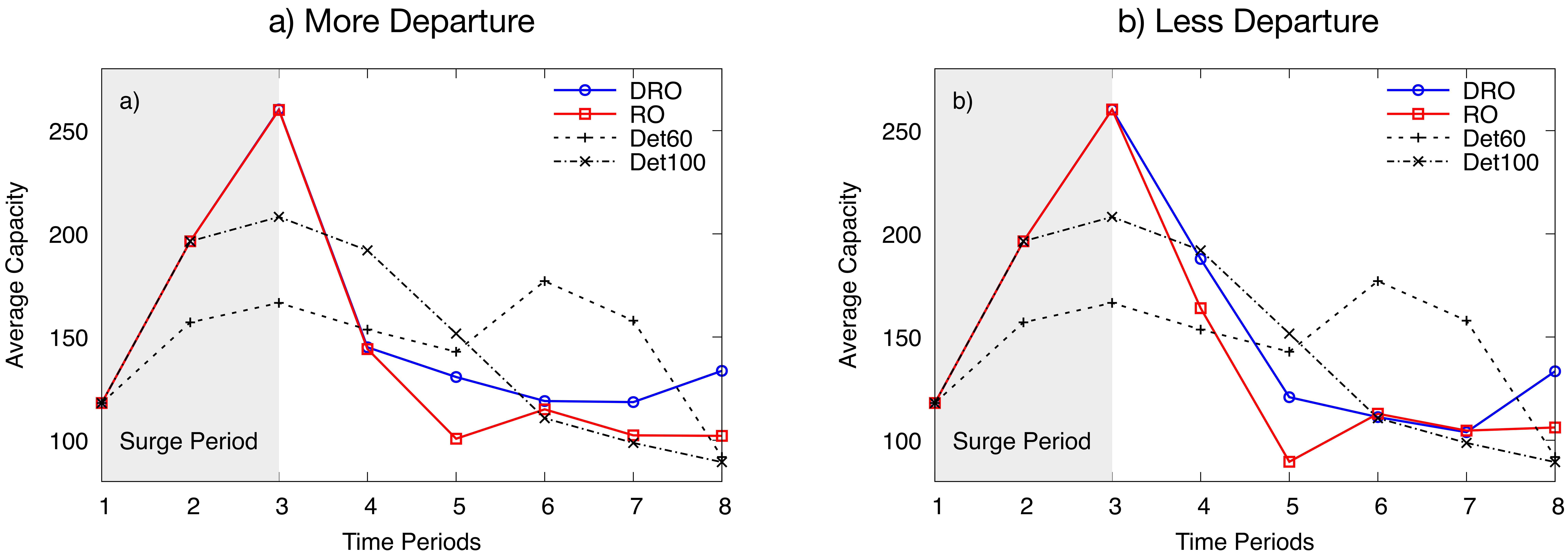
- Both RO and DRO keep maximum capacity for the first three months (surge period).

Structure of Expansion Policies



- Both RO and DRO keep maximum capacity for the first three months (surge period).
- DRO keeps higher capacity than RO after the surge period.

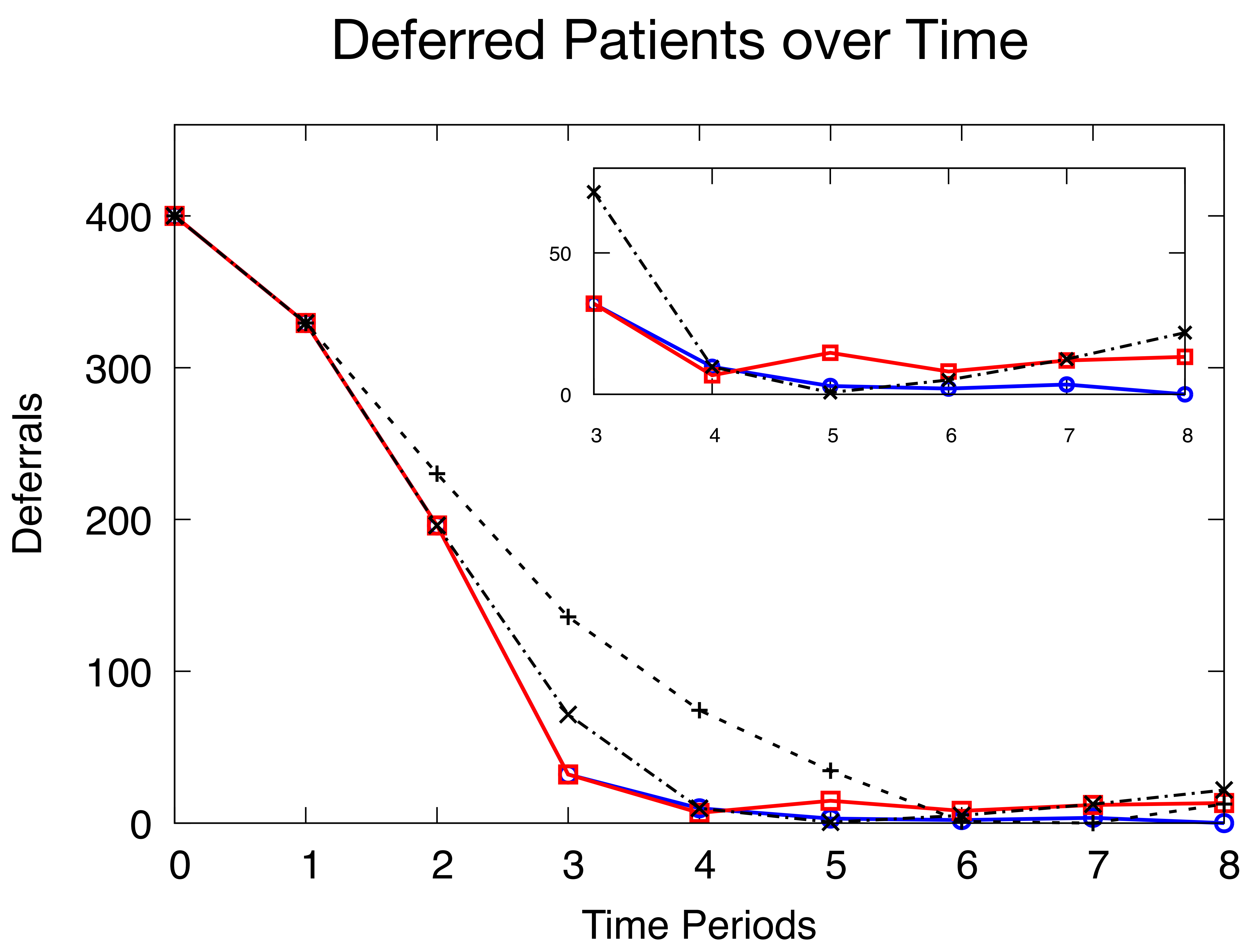
Structure of Expansion Policies



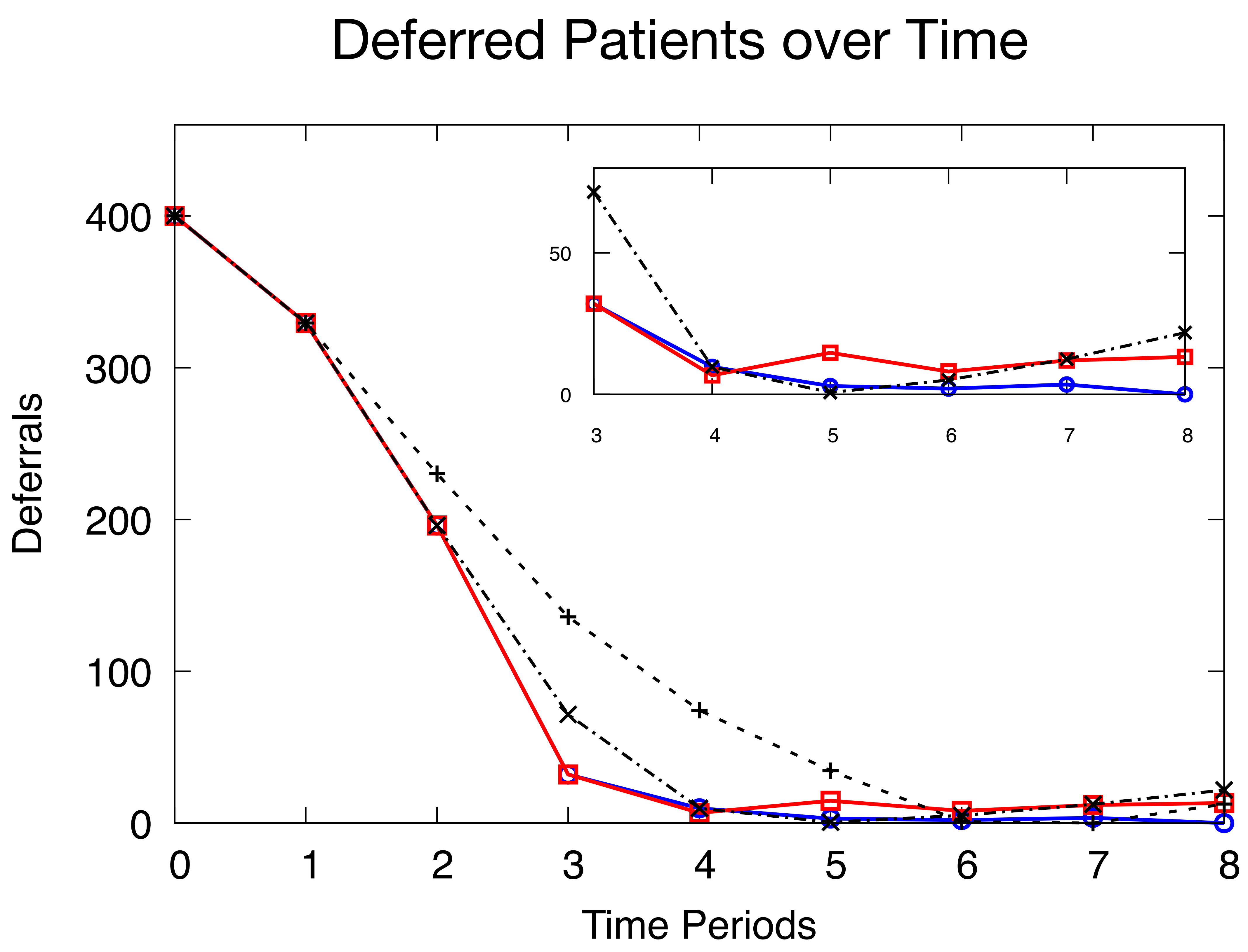
- Both RO and DRO keep maximum capacity for the first three months (surge period).
- DRO keeps higher capacity than RO after the surge period.

Deferred and Departed Patients

Deferred and Departed Patients

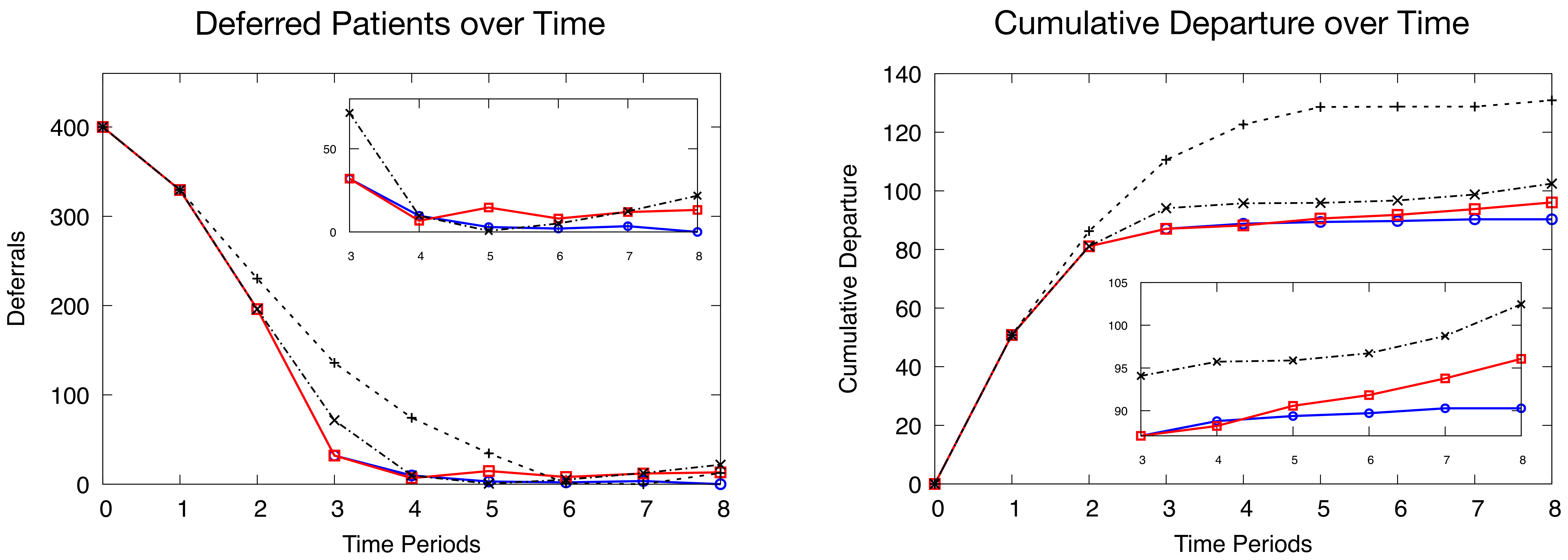


Deferred and Departed Patients



- Both RO and DRO policies achieve less numbers of deferrals and departures than deterministic policies.

Deferred and Departed Patients



- Both RO and DRO policies achieve less numbers of deferrals and departures than deterministic policies.

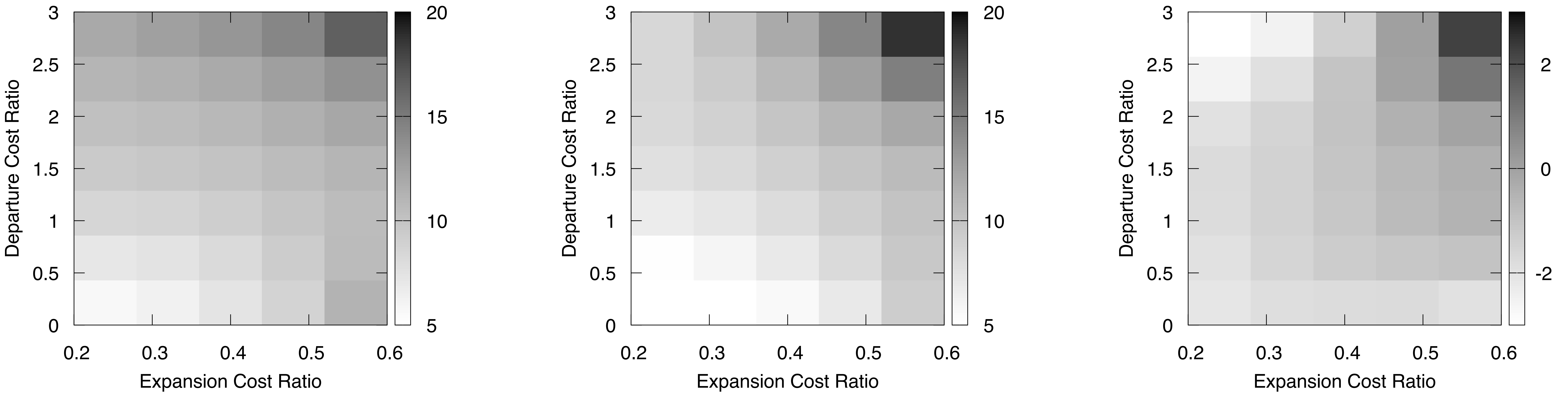
Comparison of Policies

	Lower demand (Mean 94)			Nominal demand (Mean 100)			Higher demand (Mean 106)		
	Static	Hybrid	Dynamic	Static	Hybrid	Dynamic	Static	Hybrid	Dynamic
RO	7.17%	8.91%	11.15%	5.61%	6.25%	9.90%	6.31%	7.89%	10.29%
DRO	2.61%	4.11%	4.36%	10.29%	10.00%	13.31%	11.12%	10.65%	13.10%
Det60	-23.8%	-24.2%	-24.9%	-25.6%	-24.1%	-23.3%	-27.4%	-26.6%	-26.2%

- Both RO and DRO policies improve over the deterministic policies
- RO is robust to higher and lower demand scenarios, but DRO is only guaranteed to protect against high demand settings.

Analysis of Outcomes

- Objective improvement (in percentage) over deterministic policies for different costs

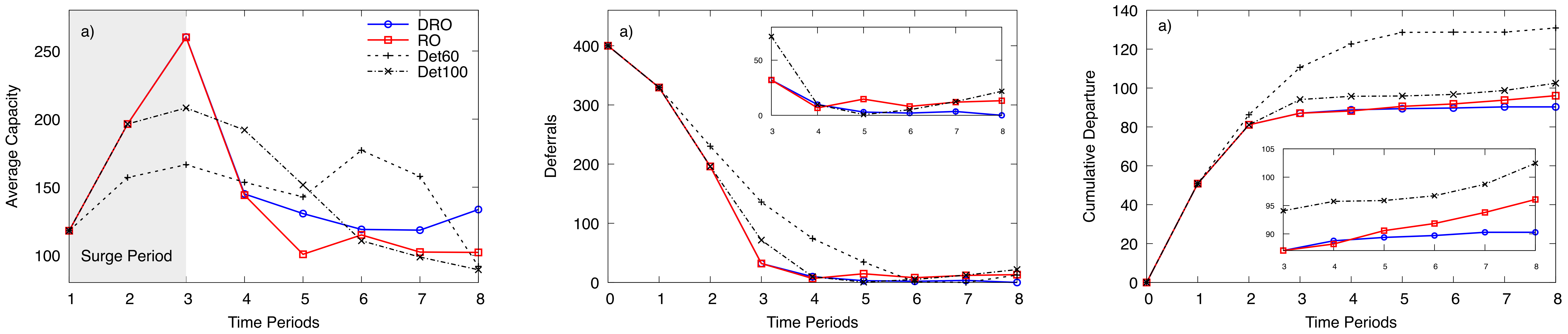


(a) Heatmap of DRO (b) Heatmap of RO (c) Heatmap of RO-DRO

- RO becomes more preferable than DRO when a decision-maker faces both higher expansion and departure costs.

Conclusions

- Dynamic expansion of surgical capacity is necessary to manage a large number of deferred surgeries.
- We develop two optimization methods, based on RO and DRO.
- We introduce the notion of tree of uncertainty products to make RO models tractable.
- Proposed methods significantly improve objectives (5~10%) over deterministic policies in the hernia case study.



Email: kartikey.sharma@zib.de