Nurse Staffing under Absenteeism:
A Distributionally Robust Optimization Approach

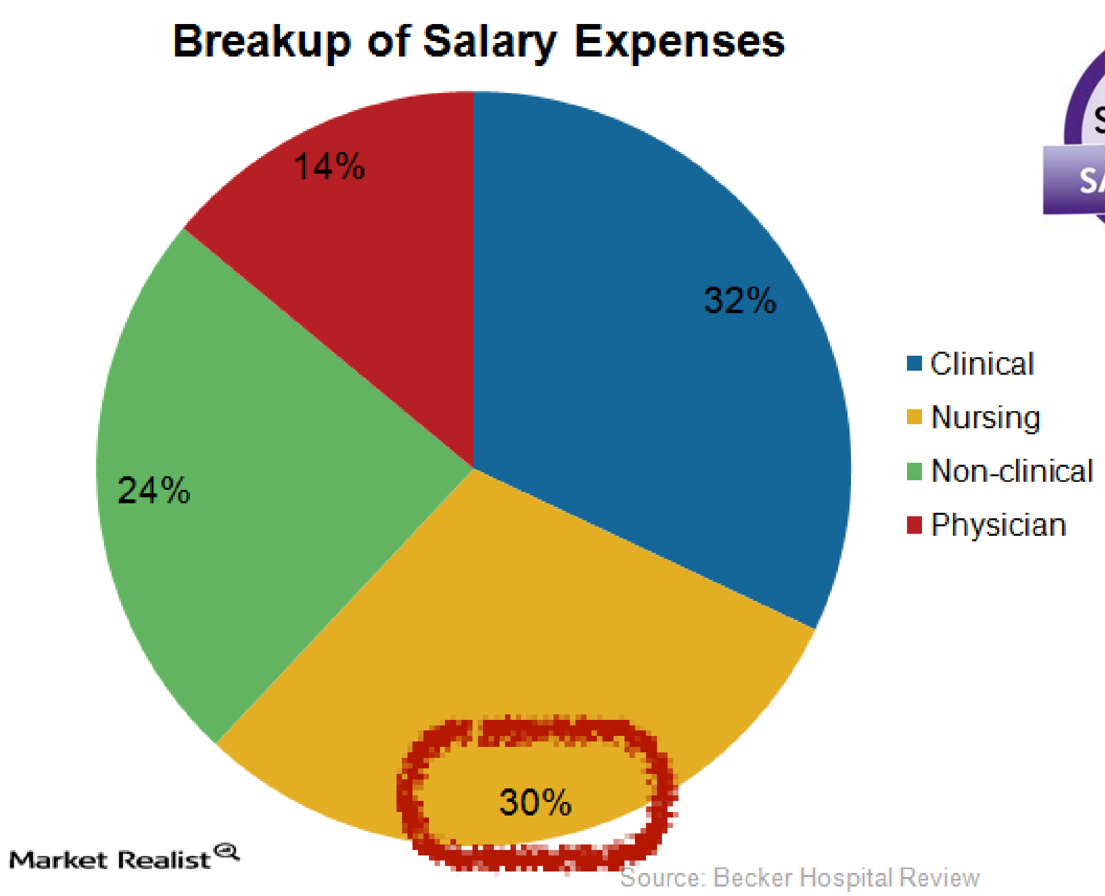
Ruiwei Jiang
IOE, University of Michigan

CHEPS Seminar Series
"Providing Better Healthcare through Systems Engineering"
University of Michigan

Joint work with Minseok Ryu (IOE, UMich)

Introduction

- Nurse staffing is significant.



Nurse staffing cost [1]

Safe nurse-to-patient ratio (NPR) [2]

[1] https://marketrealist.com/2014/11/analyzing-hospital-expenses
[2] https://www.nationalnursesunited.org/ratios

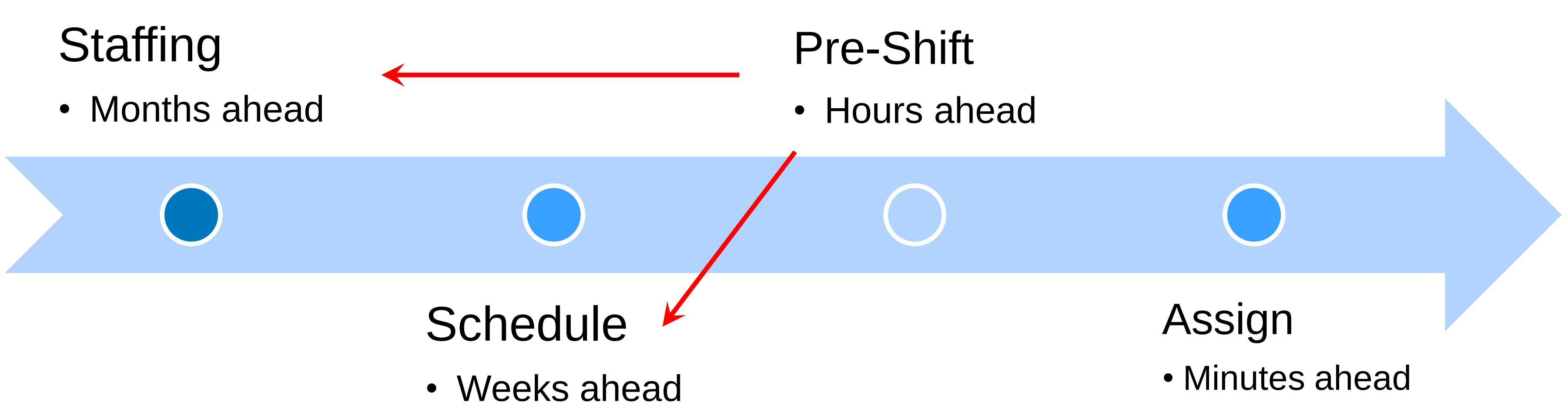
Introduction

Four Phases of nurse planning [3]:

1. Demand forecast and staffing.
2. Nurse shift scheduling.
3. Pre-shift staffing and re-scheduling.
4. Nurse-patient assignment.

Our research:

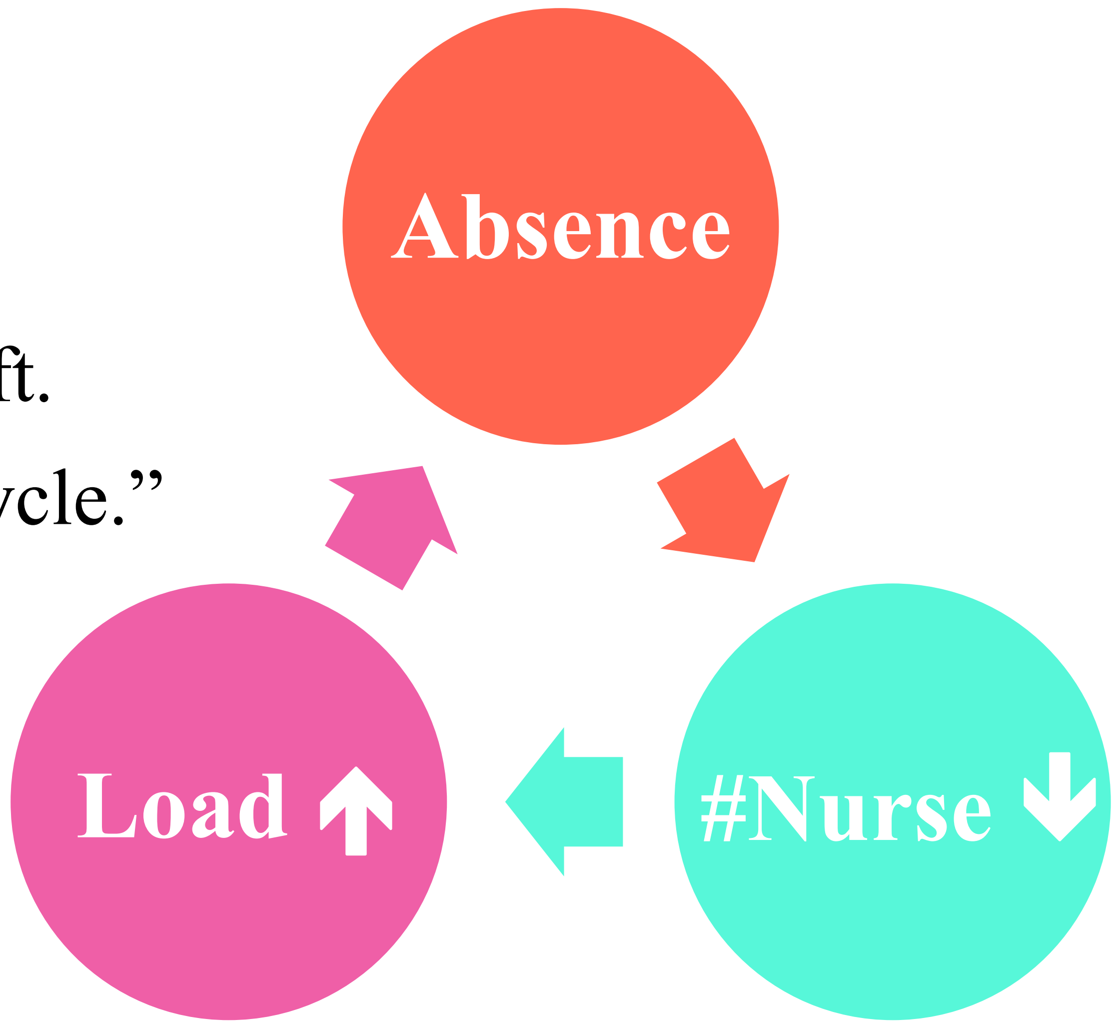
- Phases 1 and 3.
- Outputs can be used in Phases 2 and 4.



[3] Jonathan F Bard. Nurse scheduling models. *Wiley Encyclopedia of Operations Research and Management Science,* 2010.

Introduction

- Nurse staffing is <u>challenging</u>.
  - Uncertainties:
    - Demand side (nurse demand).
      - ✓ Random patient census.
      - ✓ Exogenous.
    - Supply side (nurse absenteeism).
      - ✓ # nurses showing up for the shift.
      - ✓ Work-related stress; “vicious cycle.”



Introduction

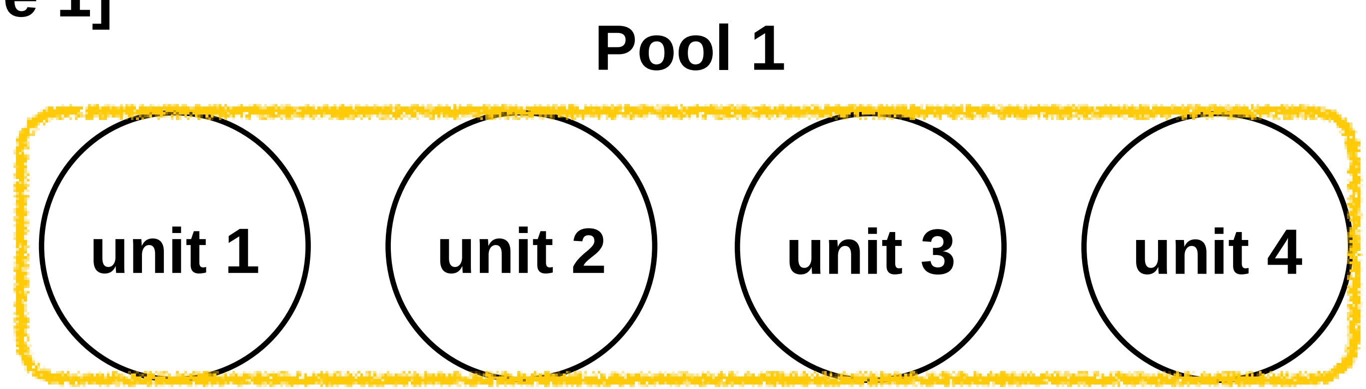
- Nurse staffing is challenging.
  - Uncertainties:
    - Demand side (nurse demand).
      - ✓ Random patient census.
      - ✓ Exogenous.
    - Supply side (nurse absenteeism).
      - ✓ # nurses showing up for the shift.
      - ✓ Work-related stress; “vicious cycle.”
      - ✓ Depends on the staffing level. [4]
      - ✓ Endogenous.
  - Temporary nurses:
    - Called in to maintain the safe NPR.

[4] Green et al., “Nursevendor Problem”: Personnel Staffing in the Presence of Endogenous Absenteeism. *Management Science,* 2013.

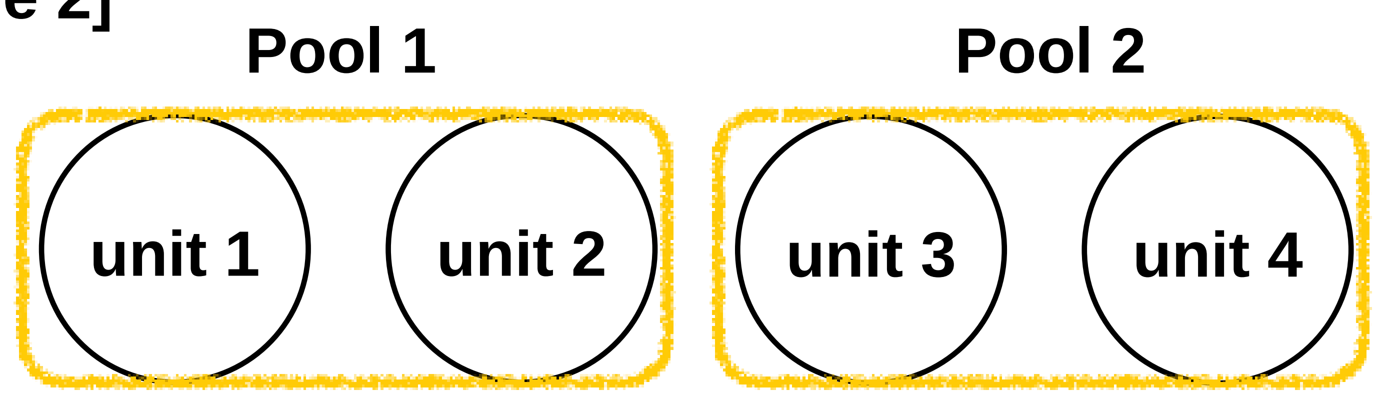
Introduction

- Potential solution:
  - Float pools.
    - Pool : a group of units.
    - Pool structure : which units belong to which pools.

[Example 1]

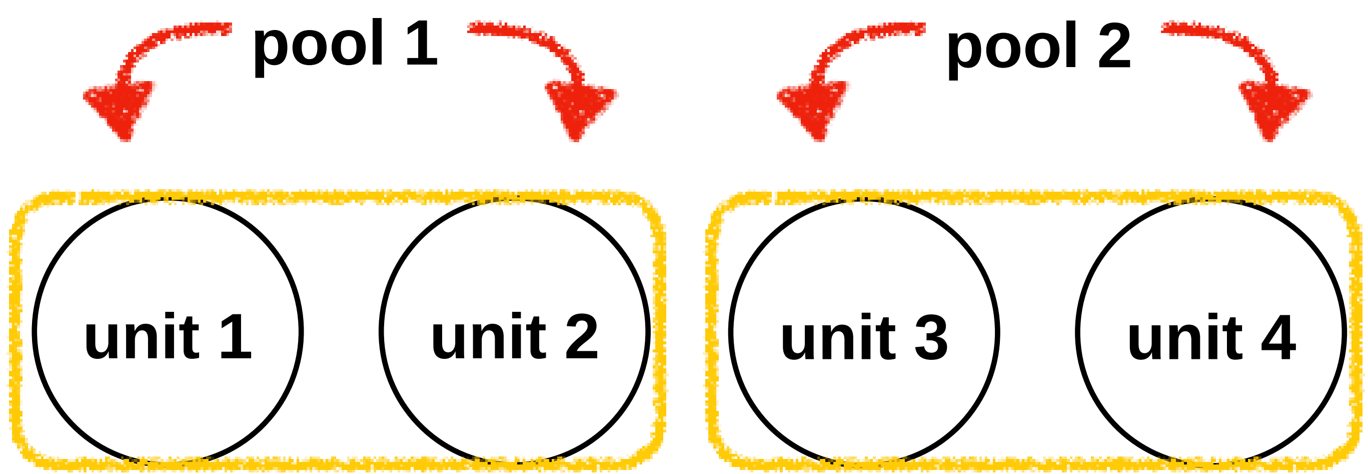


[Example 2]



Introduction

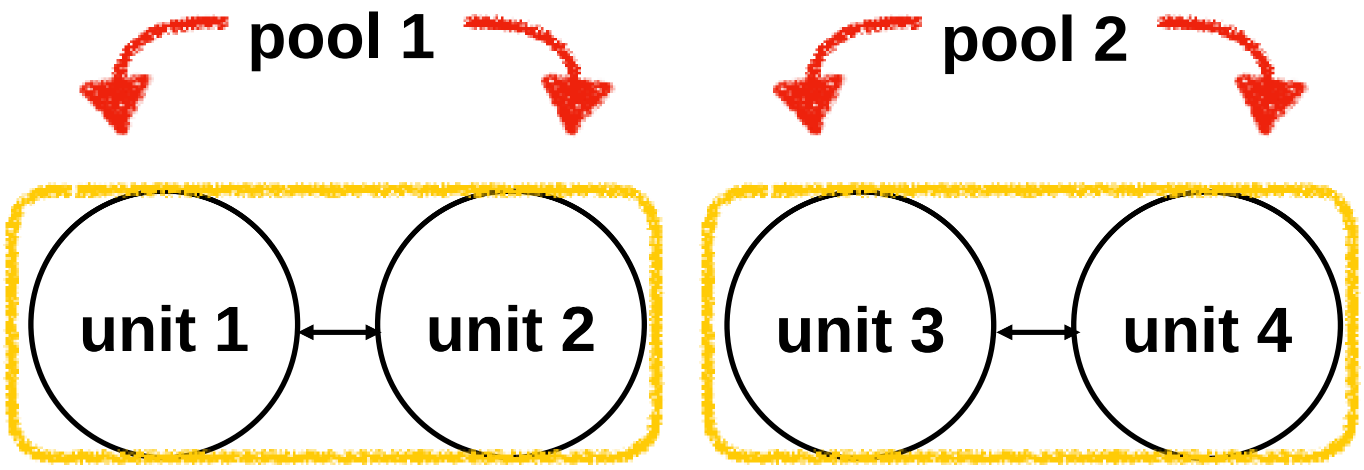
- Potential solution:
  - Pool nurses.
    - Assigned to a unit within the pool *after uncertainties are realized.*



    - Increases operational flexibility.
    - Reduce # temporary nurses.
    - Higher quality of care.
    - Lower nurse staffing cost.

Introduction

- Potential solution:
  - Cross-training.
    - Pool nurses trained for skills required by units within the pool.
    - Cross-training nurses causes an extra cost.



    - Minimize # cross-training by optimizing the pool structure.

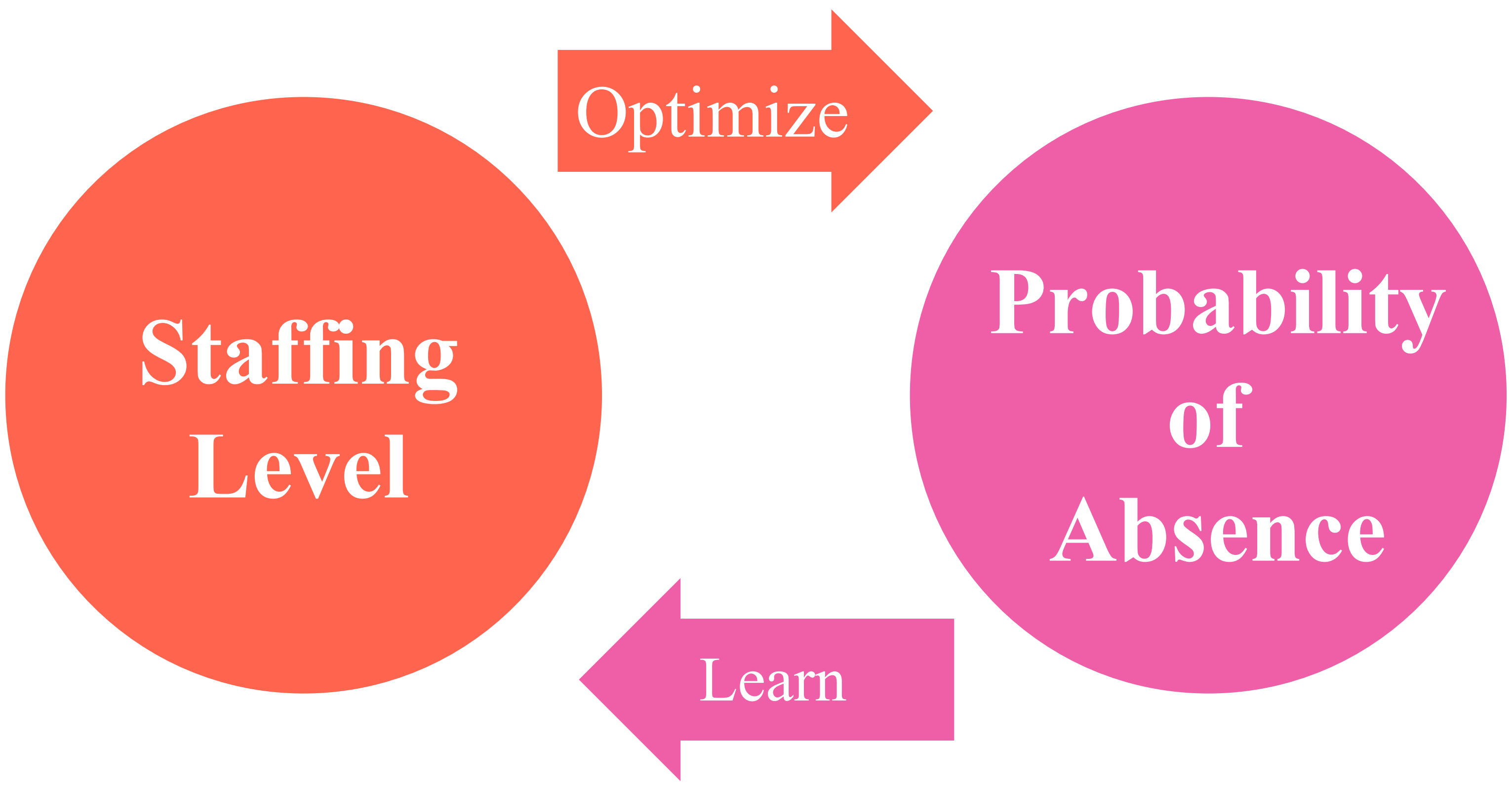
Very Brief Literature Review

- Majority: deterministic models:
  - Ignore demand and absenteeism uncertainties.
  - Underestimate the staffing costs. [5]
- Stochastic models:
  - Majority: demand uncertainty.
    - Ignore absenteeism; “results in under-staffing.” [4]
  - Absenteeism leads to endogenous uncertainty.
    - Computationally prohibitive. [6]
    - Heuristics typically used. [7], [8]

[5] Kao and Queyranne, Budgeting Costs of Nursing in a Hospital. *Management Science*, 1985.
[6] Easton, Service Completion Estimates for Cross-trained Workforce Schedules under Uncertain Attendance and Demand. *Production and Operations Management*, 2014.
[7] Kayse L. Maass et al., Incorporating nurse absenteeism into staffing with demand uncertainty. *Health care management science,* 20(1), 2017.
[8] Wang and Gupta, Nurse Absenteeism and Staffing Strategies for Hospital Inpatient Units. Manufacturing & Operations Service Management, 2014.

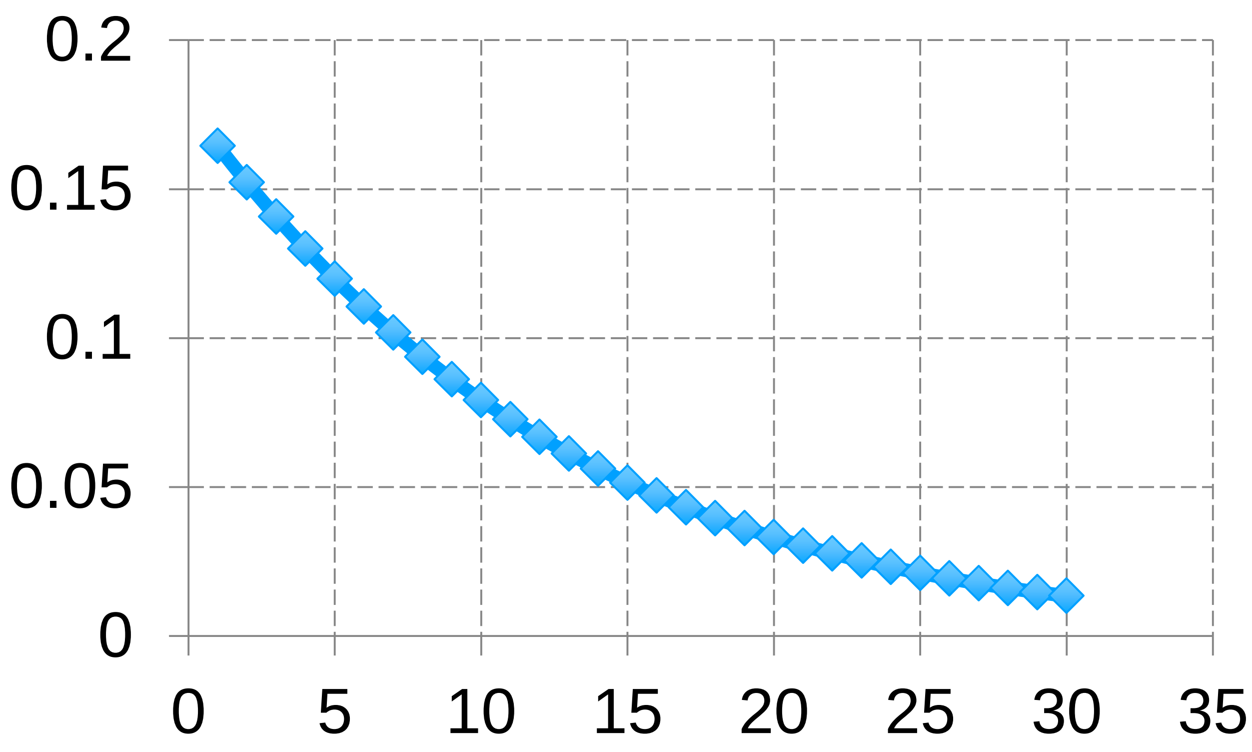
Very Brief Literature Review

- Challenge of modeling endogenous absenteeism.
  - Learn-and-optimize turns to under-staff. [4]



Very Brief Literature Review

- Potential walk-around:
  - Parametric models. [4]
    - Independence, homogeneous absence probability.
    - Data-driven; closed-form solution for 1-unit staffing.
    - Extension to multiple-unit & multiple-pool?
    - What if the chosen model is biased?



$$\text{Prob} = \frac{1}{1+\exp(1.5+0.1\times \#Nurse)}$$

Our Proposal

- Non-Parametric Model:
  - A family of probability distributions.
  - Moment-based ambiguity set:
    - Moments of nurse demands (support, mean, variance, etc.).
    - # nurses who show up $\leq$ staffing level.
    - Mean of # nurses who show up = f(staffing level).
    - Data-driven.
- Distributionally robust optimization:
  - Multiple-unit, multiple-pool.
  - MILP representable; global optimality.
  - Extension to optimal pool structure design.

Outline

- Distributionally Robust Nurse Staffing (DRNS) Model:
  - A two-stage model with a moment-based ambiguity set.
  - Recast it as a deterministic two-stage min-max formulation.
  - Separation approach.
- More Tractable Reformulation:
  - Special pool structures.

    (1) One pool, (2) Disjoint pools, (3) Chained pools.
  - A monolithic MILP reformulation.
- Optimal Nurse Pool Design (ONPD).
- Numerical Case Studies.

DRNS Model

- Variables (1st-stage, here-and-now):
  - w_j = # nurses assigned to unit j (regular nurses).
  - y_i = # nurses assigned to pool i (pool nurses).
- Objective:
  - Minimize the worst-case expected staffing cost.

$$\min_{w,y} \quad \overbrace{\sum_{j=1}^{J} c^{\mathrm{w}} w_j}^{\text{Regular Nurses}} + \overbrace{\sum_{i=1}^{I} c^{\mathrm{y}} y_i}^{\text{Pool Nurses}} + \overbrace{\sup_{\mathbb{P}\in\mathcal{D}} \mathbb{E}_{\mathbb{P}}\left[V(\tilde{w},\tilde{y},\tilde{\eta})\right]}^{\text{Temp Nurses}}$$

$$\begin{aligned}
\text{s.t.} \quad & \underline{N}_j \le w_j \le \overline{N}_j, \quad \forall j = 1,\ldots,J, \\
& \underline{M}_i \le y_i \le \overline{M}_i, \quad \forall i = 1,\ldots,I, \\
& y_i, w_j \in \mathbb{Z}_+, \quad \forall i = 1,\ldots,I, \ \forall j = 1,\ldots,J,
\end{aligned}$$

DRNS Model

- Ambiguity set:

$$\mathcal{D} = \left\{ \mathbb{P} : \begin{array}{l} \boxed{\mathbb{E}_{\mathbb{P}}[\tilde{\eta}_j^q] = \mu_{jq}, \quad \forall j \in [J], \ \forall q \in [Q],} \\ \boxed{\begin{array}{l} \mathbb{E}_{\mathbb{P}}[\tilde{w}_j] = f_j(w_j), \quad \forall j \in [J] : w_j \geq 1, \\ \mathbb{E}_{\mathbb{P}}[\tilde{y}_i] = g_i(y_i), \quad \forall i = [I] : y_i \geq 1 \end{array}} \end{array} \right\}$$

  - Nurse demand (*Exogenous*):
    - Marginal moment information.
  - Nurse absenteeism (*Endogenous*):
    - Mean of # nurses who show up = f (staffing level).
    - f, g: Piecewise linear function in general.

DRNS Model

- Ambiguity set:

$$\mathcal{D} = \left\{ \mathbb{P} : \boxed{\mathbb{E}_{\mathbb{P}}[\tilde{\eta}_j^q] = \mu_{jq}, \quad \forall j \in [J], \ \forall q \in [Q],} \right.$$

$$\boxed{\begin{array}{l} \mathbb{E}_{\mathbb{P}}[\tilde{w}_j] = a_j w_j + b_j, \quad \forall j \in [J] : w_j \geq 1, \\ \mathbb{E}_{\mathbb{P}}[\tilde{y}_i] = c_i y_i + d_i, \quad \forall i \in [I] : y_i \geq 1 \Big\} \end{array}}$$

  - Nurse demand (*Exogenous*):
    - Marginal moment information.
  - Nurse absenteeism (*Endogenous*):
    - Mean of # nurses who show up = f(staffing level).
    - f, g: Piecewise linear function in general.
    - This talk: linear f, g for simplicity.

DRNS Model

- Recourse problem (after uncertainties are realized):

$$
\begin{aligned}
V(\tilde{w}, \tilde{y}, \tilde{\eta}) = \min_{x,z,e} \quad & \sum_{j=1}^{J} (c^{\mathrm{x}} x_j - c^{\mathrm{e}} e_j) \\
\text{s.t.} \quad & \tilde{w}_j + \sum_{i:j \in P_i} z_{ij} + x_j - e_j = \tilde{\eta}_j, \quad \forall j = 1, \ldots, J, \\
& \sum_{j \in P_i} z_{ij} = \tilde{y}_i, \quad \forall i = 1, \ldots, I, \\
& x_j, e_j \in \mathbb{Z}_+, \quad \forall j = 1, \ldots, J, \\
& z_{ij} \in \mathbb{Z}_+, \quad \forall i = 1, \ldots, I, \ \forall j \in P_i.
\end{aligned}
$$

DRNS Model

- Recourse problem (after uncertainties are realized):

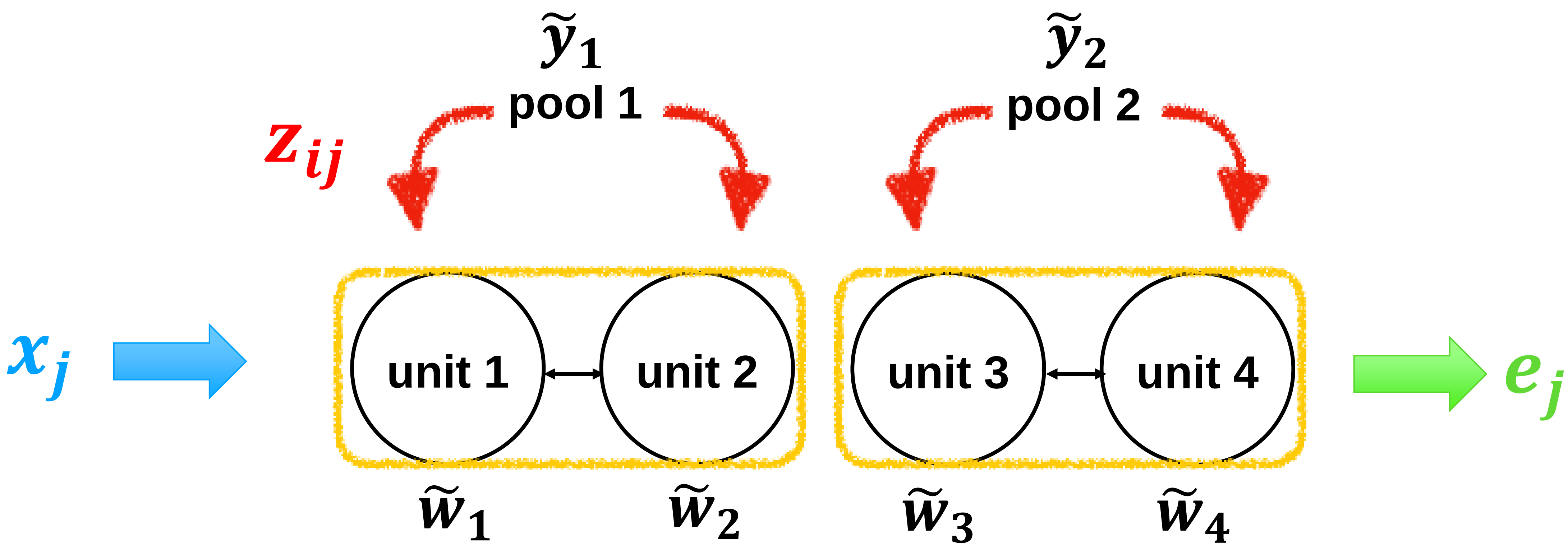
$$V(\tilde{w}, \tilde{y}, \tilde{\eta}) = \min_{x,z,e} \sum_{j=1}^{J} (c^{x} x_j - c^{e} e_j)$$

$$\text{s.t.} \quad \tilde{w}_j + \sum_{i:j \in P_i} z_{ij} + x_j - e_j = \tilde{\eta}_j, \quad \forall j = 1, \dots, J,$$

$$\sum_{j \in P_i} z_{ij} = \tilde{y}_i, \quad \forall i = 1, \dots, I,$$

$$x_j, e_j \in \mathbb{Z}_+, \quad \forall j = 1, \dots, J,$$

$$z_{ij} \in \mathbb{Z}_+, \quad \forall i = 1, \dots, I, \ \forall j \in P_i.$$



DRNS Model

- Recourse problem (after uncertainties are realized):

$$
\begin{aligned}
V(\tilde{w}, \tilde{y}, \tilde{\eta}) = \min_{x,z,e} \quad & \sum_{j=1}^{J} (c^{x} x_j - c^{e} e_j) \\
\text{s.t.} \quad & \tilde{w}_j + \sum_{i: j \in P_i} z_{ij} + x_j - e_j = \tilde{\eta}_j, \quad \forall j = 1, \ldots, J, \\
& \sum_{j \in P_i} z_{ij} = \tilde{y}_i, \quad \forall i = 1, \ldots, I, \\
& x_j, e_j \in \mathbb{Z}_+, \quad \forall j = 1, \ldots, J, \\
& z_{ij} \in \mathbb{Z}_+, \quad \forall i = 1, \ldots, I, \ \forall j \in P_i.
\end{aligned}
$$

Proposition.
The IP recourse problem is equivalent to the following linear program:

$$
V(\tilde{w}, \tilde{y}, \tilde{\eta}) \ = \ \max_{(\alpha,\beta)\in\Lambda} \sum_{j=1}^{J} (\tilde{\eta}_j - \tilde{w}_j)\alpha_j + \sum_{i=1}^{I} \tilde{y}_i \beta_i
$$

where

$$
\Lambda := \{\alpha \in \mathbb{R}^J, \beta \in \mathbb{R}^I \ : \beta_i + \alpha_j \leq 0, \ \forall i \in [I], \ \forall j \in P_i, \quad \alpha_j \in [c^{e}, c^{x}], \ \forall j \in [J]\}
$$

DRNS Model

Proposition.

For given w and y, the worst-case expectation $\sup_{\mathbb{P}\in\mathcal{D}} \mathbb{E}_{\mathbb{P}}[V(\tilde{w},\tilde{y},\tilde{\eta})]$ is equal to the optimal objective value of the following min-max optimization problem:

$$\min_{\gamma\geq 0,\ \lambda\geq 0,\ \rho} \max_{(\alpha,\beta)\in\Lambda} F(\alpha,\beta) - \sum_{i=1}^{I} (c_i y_i + d_i)\lambda_i + \sum_{j=1}^{J}\left[\sum_{q=1}^{Q} \mu_{jq}\rho_{jq} - (a_j w_j + b_j)\gamma_j\right]$$

$$s.t.\ \gamma_j \leq M w_j, \quad \forall j = 1,\dots,J,$$

$$\lambda_i \leq M y_i, \quad \forall i = 1,\dots,I,$$

where $F(\alpha,\beta) := \sum_{j=1}^{J}\Big[(-\alpha_j+\gamma_j)w_j\Big]^{+} + \sum_{i=1}^{I}\Big[(\beta_i+\lambda_i)y_i\Big]^{+} + \sum_{j=1}^{J} \sup_{\tilde{\eta}_j\in[\underline{\eta}_j,\overline{\eta}_j]_{\mathbb{Z}}} \left\{\alpha_j\tilde{\eta}_j - \sum_{q=1}^{Q}\rho_{jq}\tilde{\eta}_j^{q}\right\}$

- How do we solve the min-max problem?
  - *Separation approach* **:** $V \geq F(\alpha,\beta), \forall(\alpha,\beta)\in\Lambda$
  - *Reformulate* the inner maximization problem **:** $\max_{(\alpha,\beta)\in\Lambda} F(\alpha,\beta)$
  - Search the extreme point and direction of the polyhedron $\boldsymbol{\Lambda}$

DRNS Model

Proposition.

For given w, y, γ_j, λ_i, and ρ, problem $\max_{(\alpha,\beta)\in\Lambda} F(\alpha,\beta)$ has the same optimal objective value as the following integer (linear) program:

IP Sub-Problem

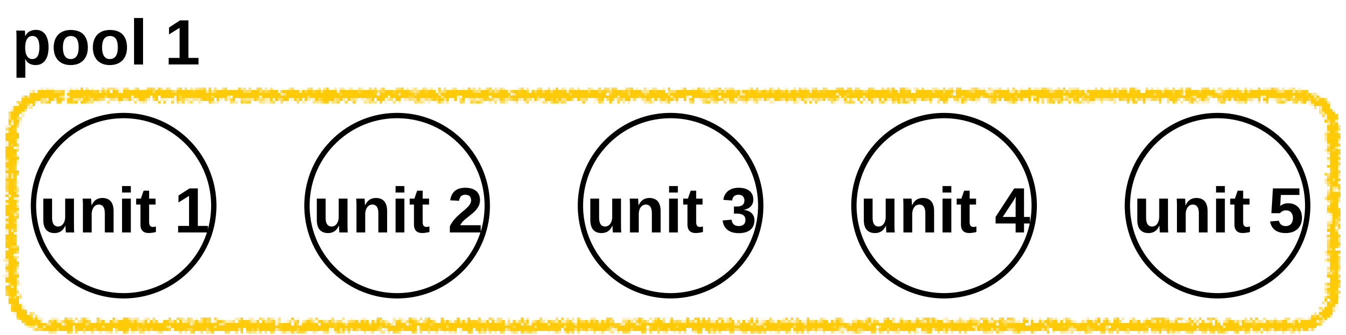
$$\begin{aligned} \max_{t,s,r,p} \quad & \sum_{j=1}^{J}(c_j^{\mathrm{t}} t_j + c_j^{\mathrm{r}} r_j) + \sum_{i=1}^{I}\Big(c_i^{\mathrm{p}} p_i + \sum_{j\in P_i} c_i^{\mathrm{s}} s_{ij}\Big) \\ \text{s.t.} \quad & (t,s)\in\mathcal{H} \\ & t_j + r_j = 1, \quad \forall j = 1,\ldots,J, \\ & \sum_{j\in P_i} s_{ij} + p_i = 1, \quad \forall i = 1,\ldots,I. \end{aligned}$$

★ The DRNS model = Deterministic two-stage min-max problem.

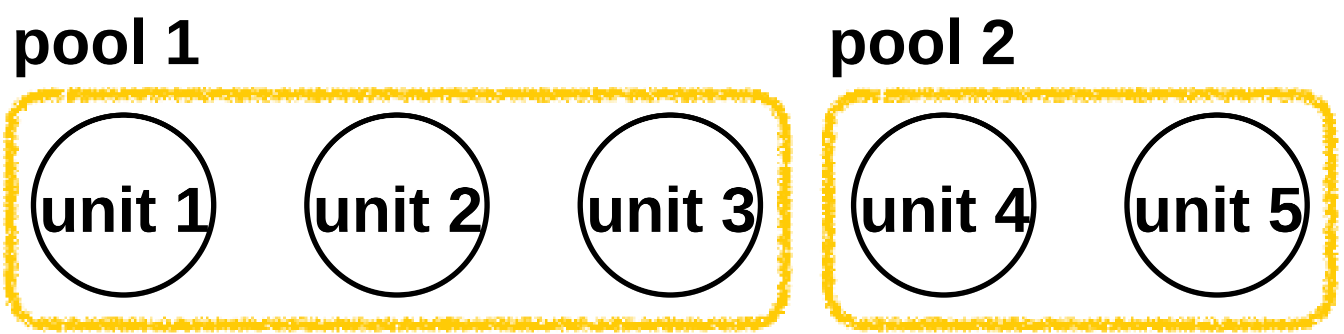
★ Solved by the separation approach.

More Tractable Reformulation

- Structure 1 : One pool



- Structure 2 : Disjoint pools



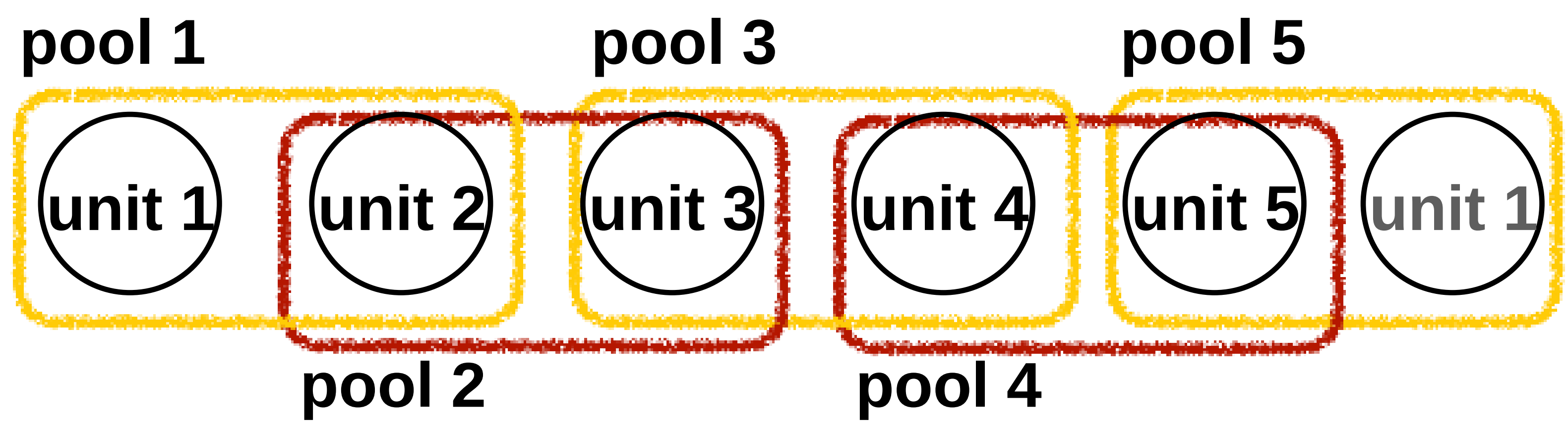
Theorem.

In structure 1 and 2,

LP relaxation of *H* = Convex hull of *H*

More Tractable Reformulation

- Structure 3 : Chained pools



  - In structure 3,

LP relaxation of H ≠ convex hull of H

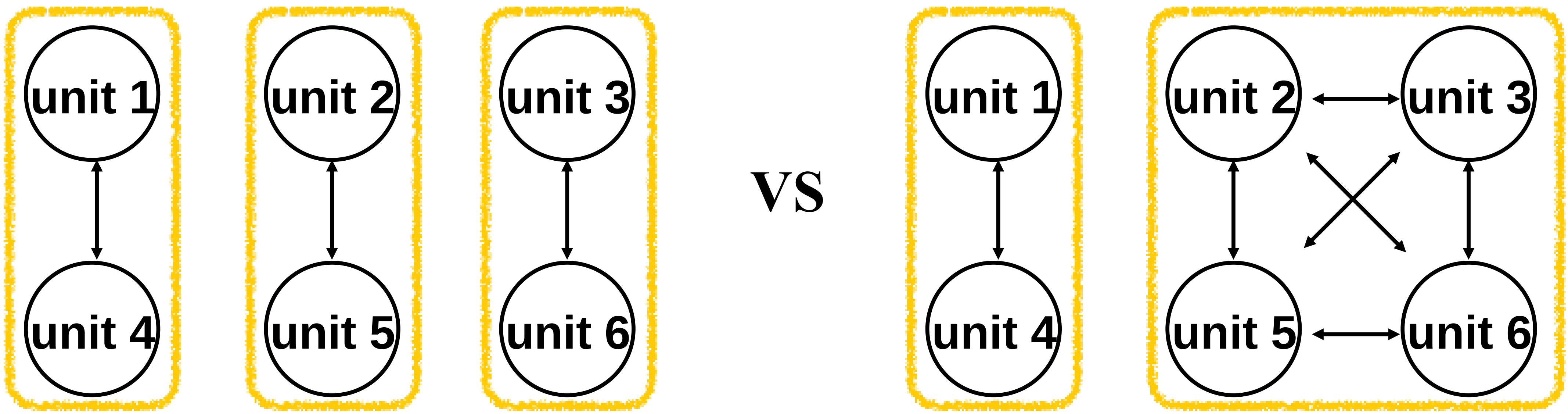
Proposition.

In structure 3, the IP sub-problem = two longest path problems.

★ The DRNS model under the one pool, disjoint pools, and chained pool structures can be expressed as *a monolithic MILP* which can be solved by the *branch-and-bound algorithm*.

ONPD Model

- Motivation
  - Two pool structures, same cost:



Staffing cost = \$15000
#cross-training = 3

Staffing cost = \$15000
#cross-training = 7

- A hospital hiring pool nurses should have a pool structure that minimizes the number of cross-training and the staffing cost.

ONPD Model

- The ONPD model is built upon the DRNS model for disjoint pools
  - Additional binary variables:

$$y_{ij} = \begin{cases} 1 \text{ if unit } j \text{ is in pool } i \\ 0 \text{ otherwise} \end{cases}$$

  - Objective:
    - minimize the #cross-training.
  - Additional constraints:
    - a target staffing cost (budget).

★ The ONPD model provides *an optimal disjoint pool structure* that minimizes the #cross-training while achieving a target staffing cost

Computational Results

1. Computational performance

Two-stage min-max problem (separation approach)

VS

A monolithic MILP (branch-and-bound)

- I = number of pools, J = number of units
- For each [I, J], we generate 10 instances to calculate the *average* computational time

One pool structure			Disjoint pools structure		
$[I, J]$	Two-stage	MILP	$[I, J]$	Two-stage	MILP
$[1, 5]$	2.28	0.09	$[3, 5]$	1.93	0.09
$[1, 7]$	8.24	0.10	$[3, 7]$	9.82	0.10
$[1, 10]$	44.42	0.13	$[3, 10]$	68.33	0.16
$[1, 20]$	> 3600	0.37	$[3, 20]$	> 3600	0.37
$[1, 50]$	> 3600	1.51	$[3, 50]$	> 3600	1.13

Average CPU Seconds

Computational Results

2. A small example

- 4 units with *one pool structure*

Input parameters (DRNS model)		
Nurse staffing costs	Bounds on number of nurses	Nurse demands
A regular nurse : \$400 A pool nurse : \$425 A temporary nurse : \$460	Unit 1: $w_1 \in [0, 10]$ Unit 2: $w_2 \in [0, 9]$ Unit 3: $w_3 \in [0, 11]$ Unit 4: $w_4 \in [0, 11]$ Pool: $y_1 \in [0, 10]$	Unit 1: $\tilde{\eta}_1 \in [1, 30]$, mean $= 22.48$ Unit 2: $\tilde{\eta}_2 \in [1, 28]$, mean $= 21.70$ Unit 3: $\tilde{\eta}_3 \in [1, 32]$, mean $= 25.18$ Unit 4: $\tilde{\eta}_4 \in [1, 31]$, mean $= 25.39$

Output results (DRNS model)	
Worst-case expected cost	$\text{WE}[\text{cost}] = 13468.2$
Nurse staffing decision	$w_1 = 6, w_2 = 6, w_3 = 7, w_4 = 7, y_1 = 0$

Computational performances			
	Time [s]	Node [#]	Gap [%]
Two-stage	1.49	2321	0
MILP	0.02	0	0

- GUROBI 7.0.1 (with Python 2.7)
- Processor: Intel® Core™i7-4850 HQ CPU@2.30GHz / RAM: 16GB / OS: 64bit

Computational Results

2. When should we construct float pools?

- Experiment
  - Inputs:

$$TCR = \frac{\text{cost of hiring a temporary nurse}}{\text{cost of hiring a regular nurse}}$$

$$PCR = \frac{\text{cost of hiring a pool nurse}}{\text{cost of hiring a regular nurse}}$$

$$ARR = \frac{\text{absenteeism rate of pool nurses}}{\text{absenteeism rate of regular nurses}}$$

  - Output:

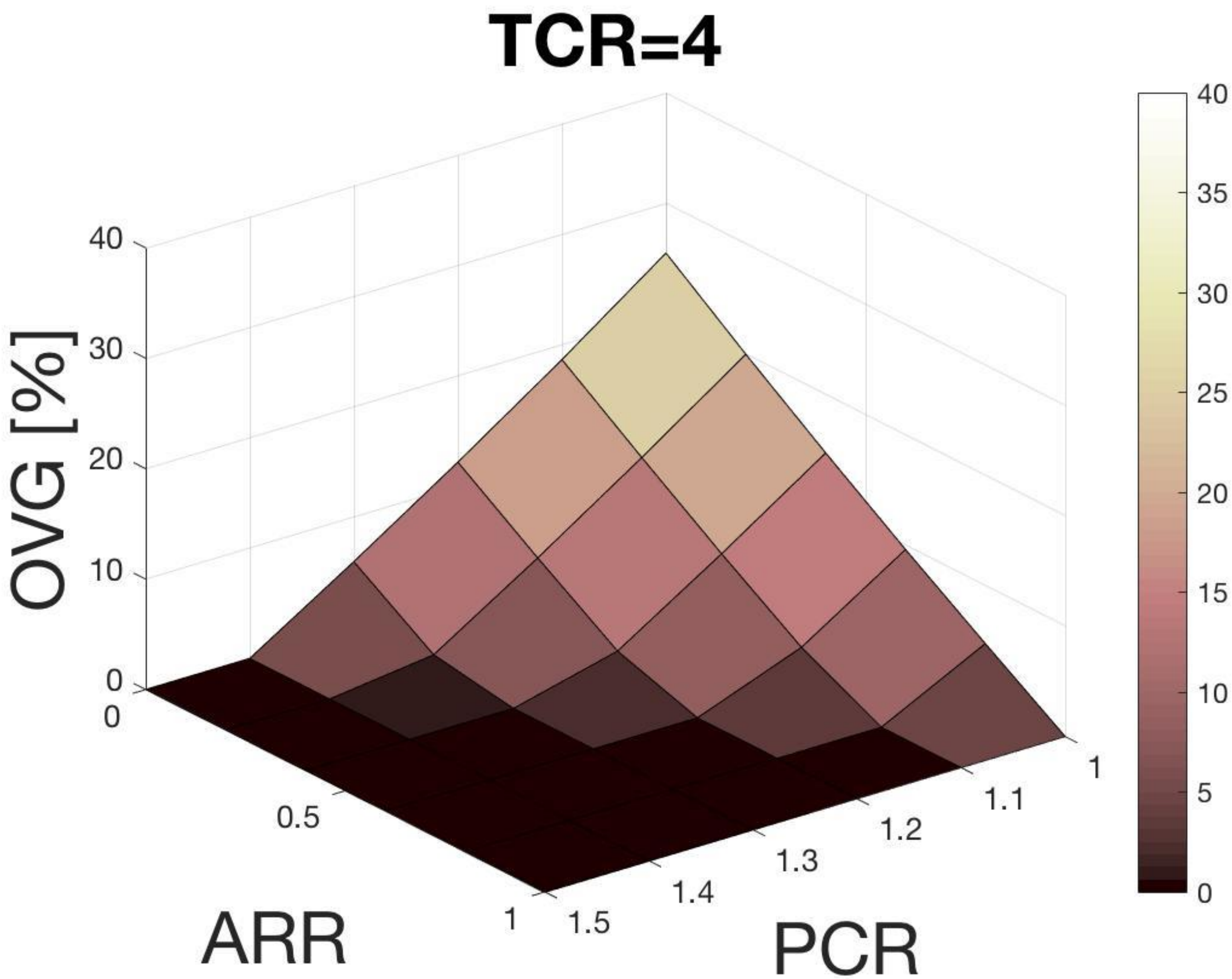
$$\mathrm{OVG}[\%] = \frac{\mathrm{WE}[\mathrm{cost}]_0 - \mathrm{WE}[\mathrm{cost}]_1}{\mathrm{WE}[\mathrm{cost}]_0} \times 100$$

$=$ Value of Float Pools

Computational Results

2. When should we construct float pools?

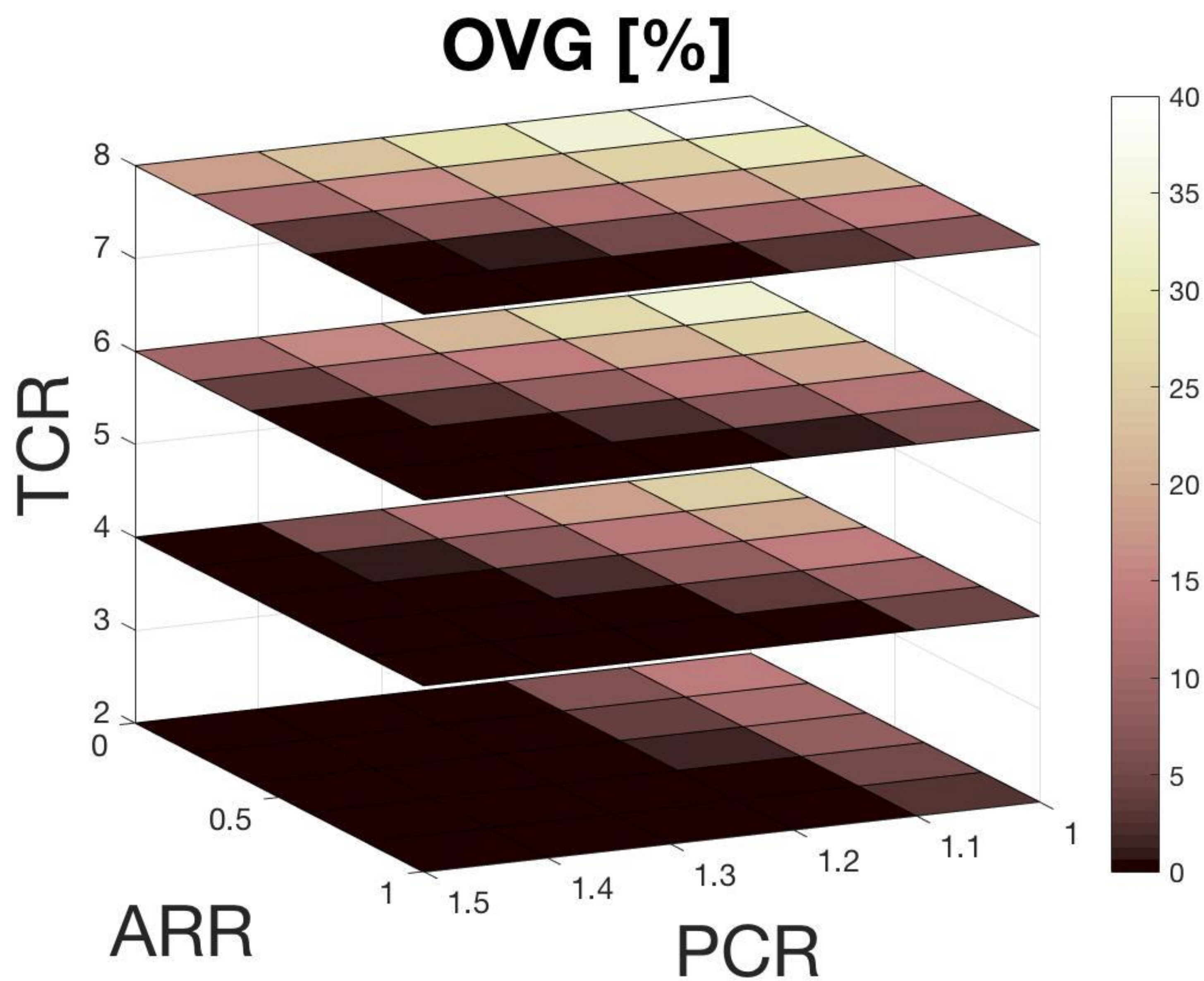
- Results:



Computational Results

2. When should we construct float pools?

- Results:



Computational Results

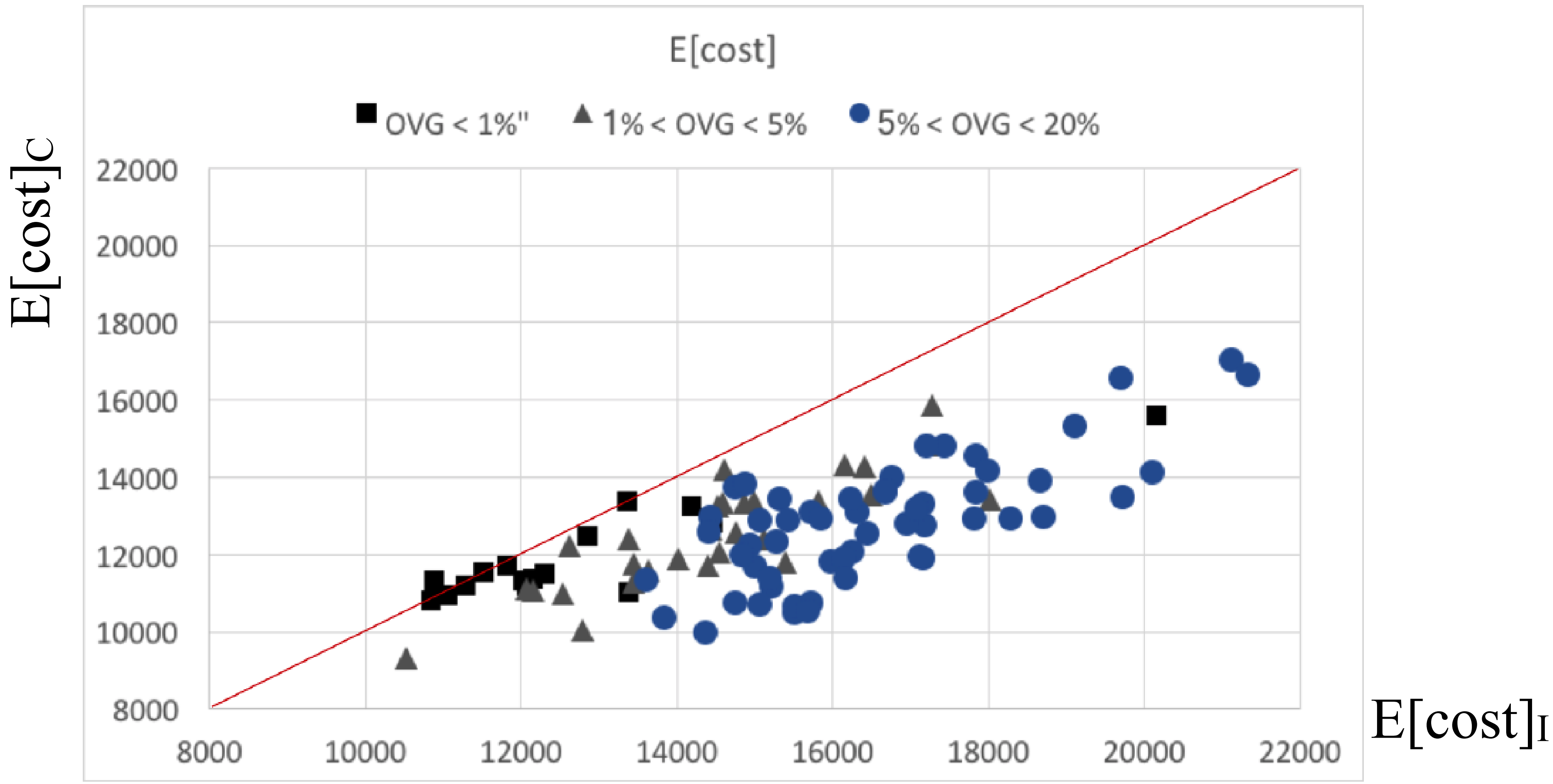
2. When should we construct float pools?

- Intuition: hiring pool nurses benefits when
  1. Cost of temporary nurses are high,
  2. Cost of pool nurses are low, and
  3. Pool nurses have low absenteeism.

Computational Results

3. Why should we model absenteeism?

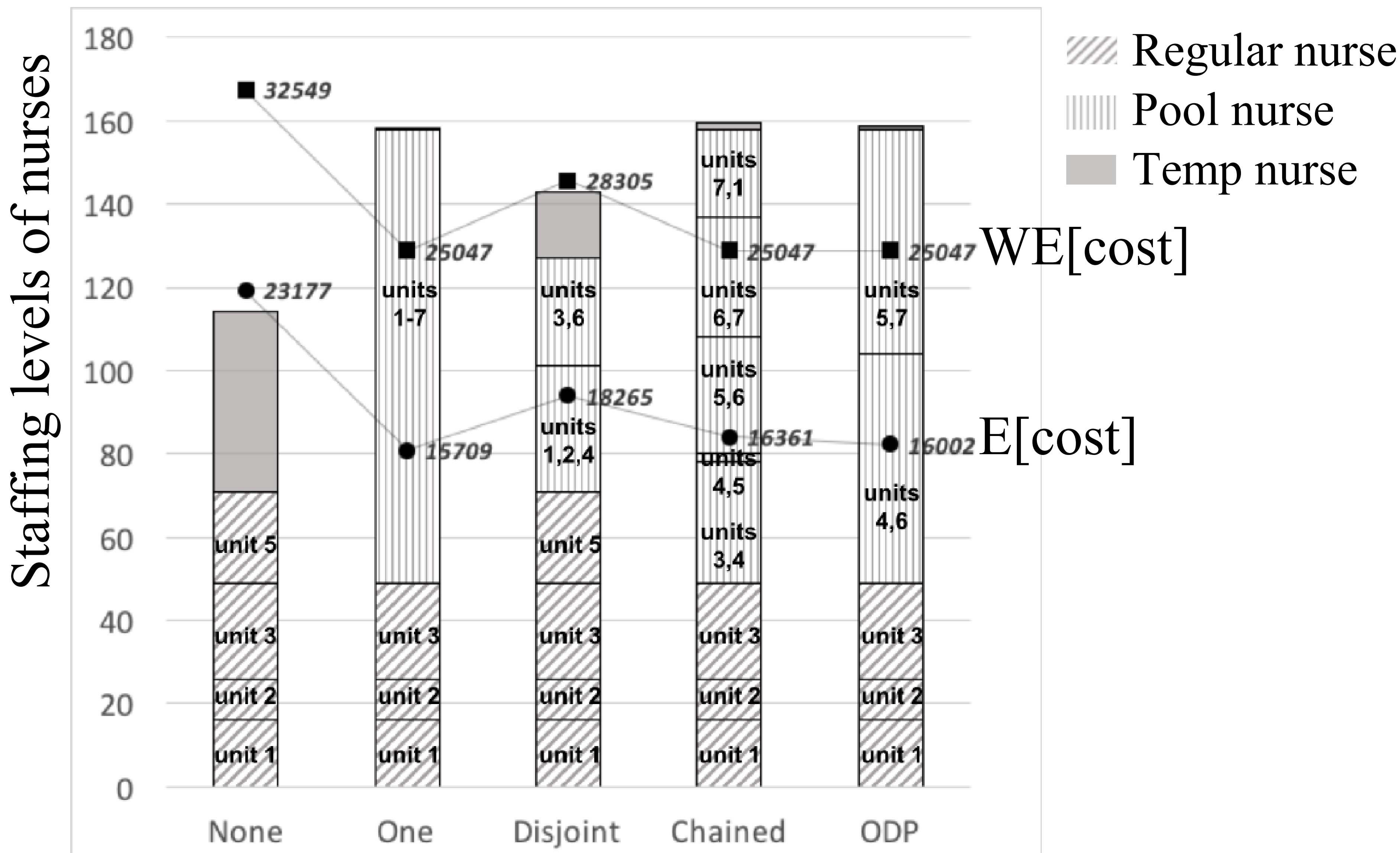
- Out-of-sample simulation
  1. Compute the staffing decision *considering / ignoring* absenteeism.
  2. Simulate the out-of-sample average costs.
- Result ($E[cost]_C$ and $E[cost]_I$ for 100 instances)



Computational Results

4. Nurse staffing decisions under various pool structures.

- Result



✓ One pool: lowest staffing cost (21 cross-training)

✓ Optimal disjoint pool : lowest staffing cost (2 cross-training)

Computational Results

5. Which units to pool together?

- Motivation
  - In hospital, each unit has different level of uncertainties

	unit 1	unit 2	unit 3	unit 4
Demand	high	low	high	low
Absenteeism	high	high	low	low

  - Which units to pool together?

Computational Results

5. Which units to pool together?

- Experiment
  - Label units with *low uncertainties* as “A”.
  - Label units with *high uncertainties* as “B”.
  - Consider 3 cases:

    (case 1) *demand only.*

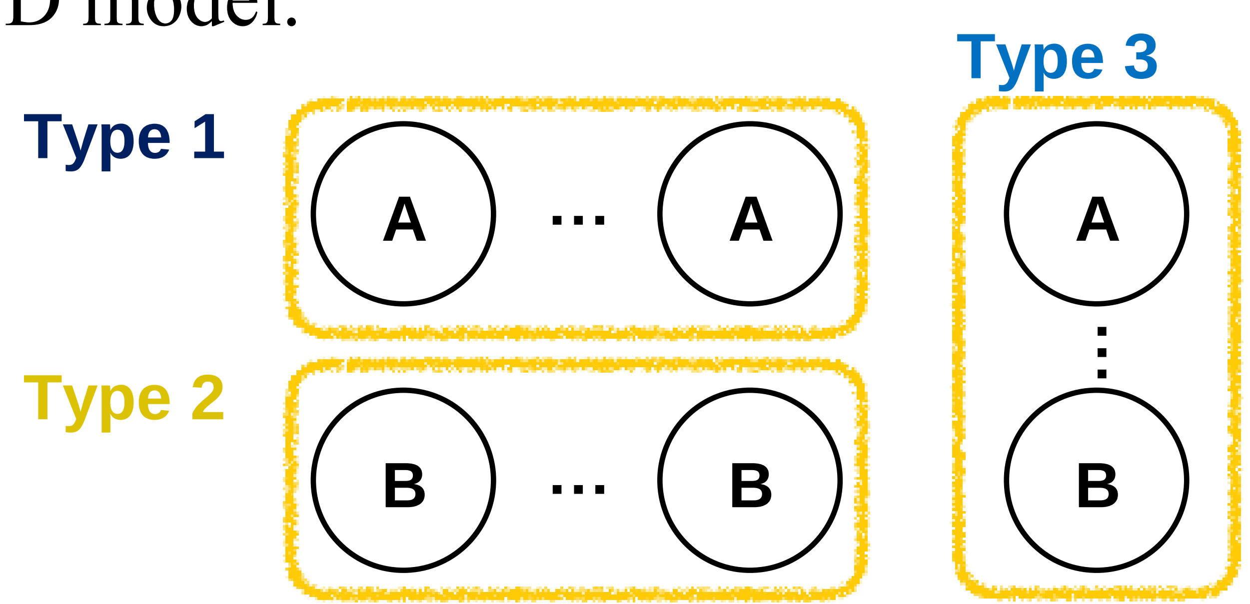
    (case 2) *absenteeism only.*

    (case 3) *both demand and absenteeism.*
  - For each case, we solve the ONPD model.
  - Resulting pools:

    (Type 1) *A units only.*

    (Type 2) *B units only.*

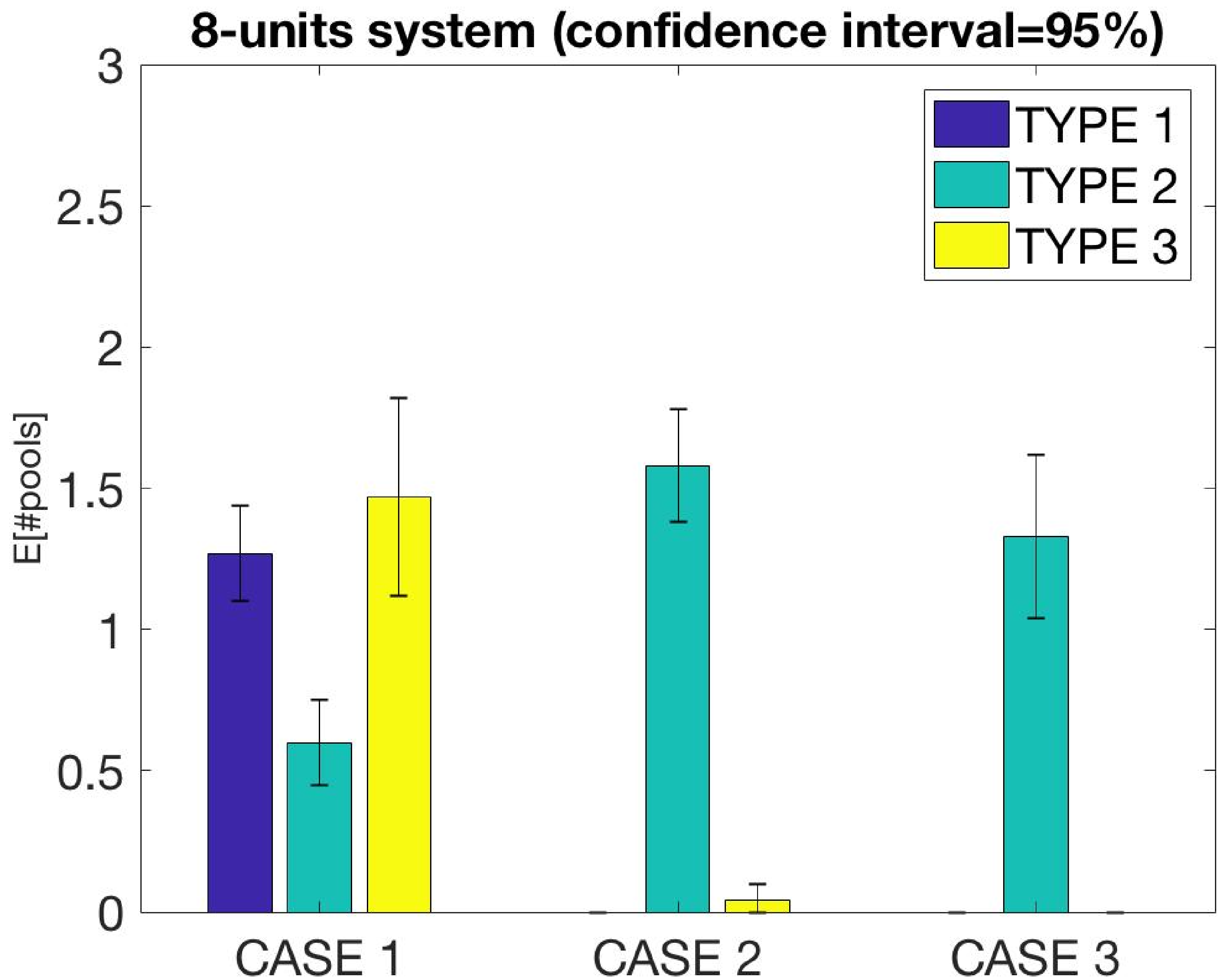
    (Type 3) *both A and B units.*



Computational Results

5. Which units to pool together?

- Result (Expected number of Type 1, 2, 3 pools)



Uncertainties	
Case 1	demand
Case 2	absenteeism
Case 3	demand + absenteeism

Pool	
Type 1	units A only
Type 2	units B only
Type 3	units A & B

✓ Units with high uncertainties are grouped together (Type 2 pool).
✓ Absenteeism plays an important role on the pooling strategy.

Thank You for Your Attention!

Questions / Comments?