Loss-Constrained Minimum Cost Flow under Arc Failure Uncertainty with Applications in Risk-Aware Kidney Exchange

Siqian Shen
University of Michigan at Ann Arbor

joint work with Qipeng Zheng & Yuhui Shi
(Univ. of Central Florida; Univ. of Michigan)

2015 INFORMS Computing Society Conference
Richmond, Virginia

Outline

Outline

The Minimum Cost Flow Problem

- $\mathcal{G} = (\mathcal{N}, \mathcal{A})$, where $\mathcal{N}$: node set, and $\mathcal{A}$: arc set.
- $\mathcal{S} \subset \mathcal{N}$: supply node set; $\mathcal{T} \subset \mathcal{N}$: demand node set.
- S_i/D_i: the absolute value of supply/demand at node i
- C_{ij}: unit flow cost; U_{ij}: arc capacity.

The Minimum Cost Flow Problem

- $\mathcal{G} = (\mathcal{N}, \mathcal{A})$, where $\mathcal{N}$: node set, and $\mathcal{A}$: arc set.
- $\mathcal{S} \subset \mathcal{N}$: supply node set; $\mathcal{T} \subset \mathcal{N}$: demand node set.
- S_i/D_i: the absolute value of supply/demand at node i
- C_{ij}: unit flow cost; U_{ij}: arc capacity.

A Minimum Cost Flow problem is:

$$\textbf{[MCF]}: \quad \min \quad \sum_{(i,j)\in\mathcal{A}} C_{ij}x_{ij} \tag{1a}$$

$$\text{s.t.} \quad \sum_{j:(i,j)\in\mathcal{A}} x_{ij} - \sum_{j:(j,i)\in\mathcal{A}} x_{ji} = \begin{cases} S_i & \forall i \in \mathcal{S}, \\ 0 & \forall i \in \mathcal{N} \setminus \mathcal{S} \setminus \mathcal{T}, \\ -D_i & \forall i \in \mathcal{T}, \end{cases} \tag{1b}$$

$$0 \le x_{ij} \le U_{ij}, \quad \forall (i,j) \in \mathcal{A}, \tag{1c}$$

Literature Review

1. Various Stochastic shortest path: Loui (1983), Eiger et al. (1985), Fan et al. (2005), Hutson and Shier (2009)
2. MCF under uncertain demand, capacity, and/or traveling cost (Glockner et al. (2001), Peraki and Servetto (2004), Powell and Frantzeskakis (1994), Prékopa and Boros (1991))
3. "Last-mile delivery" in humanitarian relief: Balcik et al. (2008), Salmeron and Apte (2010), Ozdamar et al. (2004)
4. VaR: Miller and Wagner (1965), Prékopa (1970))
5. CVaR: Rockafellar and Uryasev (2000;2002)

Loss-Constrained MCF and Applications

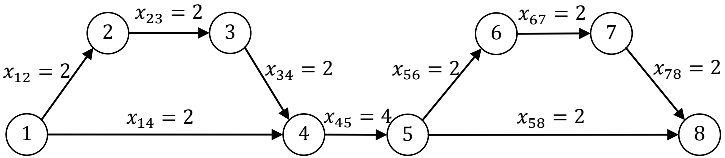
Goal: minimize the arc flow cost, while under random 0-1 arc failures, the VaR/CVaR of random path-flow losses is bounded.

- Applications: Logistics, telecommunication, humanitarian relief...
- We test a class of **stochastic kidney exchange** problems, in which we maximize the utility of pairing kidneys subject to constrained risk of utility losses, under random match failure of paired kidneys.

Assumptions

1. the failure of an arc will cause flow losses on all paths using that arc;
2. for any path carrying positive flows, the failure of one or multiple arcs on the path will lead to losing the whole amount of flows it carries
3. The total loss of an arc flow solution is the summation of path flows on all paths that have arc failures.

Motivating Example



If destroy arcs $(2, 3)$ and $(6, 7)$:

- Solution 1: two units of flow via path "1–2–3–4–5–6–7–8," and two units via path "1–4–5–8"; will lose two units.
- Solution 2: two units of flow via path "1–2–3–4–5–8," and the other two via path "1–4–5–6–7–8"; will lose four units.

Constrained "*maximum*" flow losses ⇒ being robust
Constrained "*minimum*" flow losses ⇒ being opportunistic

Outline

Notation

We formulate an LP model to compute possible flow losses.

- Let $Y_{ij} = \begin{cases} 1 & \text{if arc } (i,j) \in \mathcal{A} \text{ fails,} \\ 0 & \text{otherwise.} \end{cases}$
- Recall that the original network is $\mathcal{G}(\mathcal{N}, \mathcal{A})$
- Given an MCF solution $\hat{x}$, build a residual graph $\mathcal{G}(\hat{x})$:
  - Disconnect all arcs (i,j) having $Y_{ij} = 1$ in graph $\mathcal{G}(\mathcal{N}, \mathcal{A})$.

Notation

We formulate an LP model to compute possible flow losses.

- Let $Y_{ij} = \begin{cases} 1 & \text{if arc } (i,j) \in \mathcal{A} \text{ fails,} \\ 0 & \text{otherwise.} \end{cases}$
- Recall that the original network is $\mathcal{G}(\mathcal{N}, \mathcal{A})$
- Given an MCF solution $\hat{x}$, build a residual graph $\mathcal{G}(\hat{x})$:
  - Disconnect all arcs (i,j) having $Y_{ij} = 1$ in graph $\mathcal{G}(\mathcal{N}, \mathcal{A})$.
  - Add a fixed demand of $\hat{x}_{ij}$ at node i (where $Y_{ij} = 1$ for some j).

Notation

We formulate an LP model to compute possible flow losses.

- Let $Y_{ij} = \begin{cases} 1 & \text{if arc } (i,j) \in \mathcal{A} \text{ fails,} \\ 0 & \text{otherwise.} \end{cases}$
- Recall that the original network is $\mathcal{G}(\mathcal{N}, \mathcal{A})$
- Given an MCF solution $\hat{x}$, build a residual graph $\mathcal{G}(\hat{x})$:
    - Disconnect all arcs (i,j) having $Y_{ij} = 1$ in graph $\mathcal{G}(\mathcal{N}, \mathcal{A})$.
    - Add a fixed demand of $\hat{x}_{ij}$ at node i (where $Y_{ij} = 1$ for some j).
    - Add a demand variable ρ_j at node j (where $Y_{ij} = 1$ for some j).

Notation

We formulate an LP model to compute possible flow losses.

- Let $Y_{ij} = \begin{cases} 1 & \text{if arc } (i,j) \in \mathcal{A} \text{ fails,} \\ 0 & \text{otherwise.} \end{cases}$
- Recall that the original network is $\mathcal{G}(\mathcal{N}, \mathcal{A})$
- Given an MCF solution $\hat{x}$, build a residual graph $\mathcal{G}(\hat{x})$:
  - Disconnect all arcs (i,j) having $Y_{ij} = 1$ in graph $\mathcal{G}(\mathcal{N}, \mathcal{A})$.
  - Add a fixed demand of $\hat{x}_{ij}$ at node i (where $Y_{ij} = 1$ for some j).
  - Add a demand variable ρ_j at node j (where $Y_{ij} = 1$ for some j).
  - Add a variable λ_s representing accumulated losses at each supply node $s \in \mathcal{S}$

An Example of Constructing $\mathcal{G}(\hat{x})$

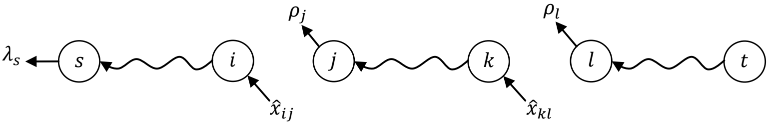
Given $Y_{ij} = Y_{kl} = 1$:

(a) Original graph and solution $\hat{x}$

An Example of Constructing $\mathcal{G}(\hat{x})$

Given $Y_{ij} = Y_{kl} = 1$:

(c) Original graph and solution $\hat{x}$



(d) The corresponding $\mathcal{G}(\hat{x})$

An LP Model for Computing Flow Losses

Theorem

Denote $L(x, Y)$ as some flow loss. For given x and Y, $L(x, Y) = \sum_{i \in \mathcal{S}} \lambda_i$, where (f, ρ, λ) satisfy:

[Flow Loss LP]:

$$\sum_{j:(i,j)\in\overline{\mathcal{A}}} f_{ij} - \sum_{j:(j,i)\in\overline{\mathcal{A}}} f_{ji} = \begin{cases} -\lambda_i + \sum_{j:(i,j)\in\mathcal{A}} Y_{ij}x_{ij} - \rho_i & \forall i \in \mathcal{S} \\ \sum_{j:(i,j)\in\mathcal{A}} Y_{ij}x_{ij} - \rho_i, & \forall i \in \mathcal{N} \setminus \mathcal{S} \end{cases} \tag{2a}$$

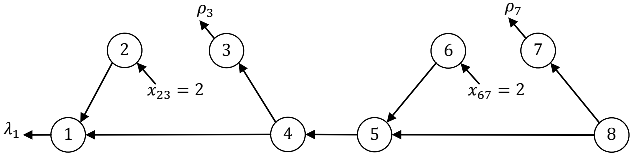
$$0 \leq f_{ij} \leq (1 - Y_{ji})x_{ji} \quad \forall (i,j) \in \overline{\mathcal{A}} \tag{2b}$$

$$0 \leq \lambda_i \leq S_i \quad \forall i \in \mathcal{S} \tag{2c}$$

$$0 \leq \rho_i \leq \sum_{j:(j,i)\in\mathcal{A}} Y_{ji}x_{ji} \quad \forall i \in \mathcal{N}. \tag{2d}$$

- (2a) is the flow balance constraint in the residual network $\mathcal{G}(x)$.
- It includes withdraw demand variable λ_i only at each supply node i in $\mathcal{S}$ if it is associated with a failed path.

An Example: Computing possible $L(x, Y)$-values



For the previous case of flowing 4 total units from node 1 to node 8, with arcs $(2,3)$ and $(6,7)$ failed, two feasible solutions to the LP correspond to the two possible path solutions:

1. $f_{87}^1 = f_{76}^1 = f_{32}^1 = f_{41}^1 = f_{85}^1 = 0$, $f_{65}^1 = f_{54}^1 = f_{43}^1 = f_{21}^1 = 2$, $\rho_7^1 = 0$, $\rho_3^1 = 2$, $\lambda_1^1 = 2$;
2. $f_{87}^2 = f_{76}^2 = f_{32}^2 = f_{85}^2 = f_{43}^2 = 0$, $f_{65}^2 = f_{54}^2 = f_{41}^2 = f_{21}^2 = 2$, $\rho_7^2 = 0$, $\rho_3^2 = 0$, $\lambda_1^2 = 4$.

Outline

Flow Losses and Risk Measures

- $\underline{L}(x, Y) = \min_{f,\lambda,\rho}\{L(x, Y)|[\text{Flow Loss LP}]\}$: The least amount of flow losses among all possible path-flow solutions
- $\overline{L}(x, Y) = \max_{f,\lambda,\rho}\{L(x, Y)|[\text{Flow Loss LP}]\}$: The largest amount of flow losses among all possible path-flow solutions

Flow Losses and Risk Measures

- $\underline{L}(x, Y) = \min_{f,\lambda,\rho}\{L(x, Y)|[\text{Flow Loss LP}]\}$: The least amount of flow losses among all possible path-flow solutions
- $\overline{L}(x, Y) = \max_{f,\lambda,\rho}\{L(x, Y)|[\text{Flow Loss LP}]\}$: The largest amount of flow losses among all possible path-flow solutions

We solve problems of bounding $\underline{L}(x, Y)$ or $\overline{L}(x, Y)$ by CVaR or VaR:

- For (VaR, $\underline{L}(x, Y)$), reformulate the problem as an MIP with logic binary variables, named SMCF-VaR; apply a cutting-plane algorithm.
- For (CVaR, $\underline{L}(x, Y)$), reformulate the problem as an LP, named SMCF-CVaR
- Both (CVaR, $\overline{L}(x, Y)$) and (VaR, $\overline{L}(x, Y)$) are intractable bilevel non-convex programs; not investigated in this talk

Constraining the VaR of $\underline{L}(x, Y)$

Replace $\underline{L}(x, \xi)$ with $L(x, \xi)$, eliminate the minimization, and solve:

$$\textbf{SMCF-VaR}: \ \min \left\{ \sum_{(i,j)\in\mathcal{A}} C_{ij} x_{ij} : \text{(1b), (1c)}, \mathbb{P}\{L(x, Y_\xi) \leq \eta\} \geq 1 - \theta \right\}, \qquad (3)$$

- Denote the random form of parameter Y by Y_ξ.
- Ω: a set of realizations of Y_ξ, denoted by Y_{ξ^s}, $\forall s \in \Omega$.
- η: flow loss threshold, i.e., $\text{VaR}_{1-\theta}$ of $L(x, Y_\xi)$.
- θ is a given risk tolerance parameter.

An MIP Model of SMCF-VaR

SMCF-VaR-D:

$$\min \quad \sum_{(i,j)\in\mathcal{A}} C_{ij}x_{ij}$$

$$\text{s.t.} \quad \text{(1b)–(1c)}$$

(2a)–(2d) with inputs Y_{ξ^s} and variables f^s, λ^s, and ρ^s, $\forall s \in \Omega$

$$L(x, Y_{\xi^s}) = \sum_{i\in\mathcal{S}} \lambda_i^s \leq Mz^s + \eta \quad \forall s \in \Omega \tag{4a}$$

$$\sum_{s\in\Omega} \mathrm{Prob}_{\xi^s} z^s \leq \theta \tag{4b}$$

$$z^s \in \{0, 1\} \quad \forall s \in \Omega, \tag{4c}$$

- Prob_{ξ^s}: the probability of realizing ξ^s with $\sum_{s\in\Omega} \mathrm{Prob}_{\xi^s} = 1$.

An LP Model of SMCF-CVaR

We formulate SMCF-CVaR for given risk parameter θ as

$$
\begin{aligned}
\min \quad & \sum_{(i,j)\in\mathcal{A}} C_{ij} x_{ij} \\
\text{s.t.} \quad & \text{(1b)–(1c)} \\
& \text{(2a)–(2d) with inputs } Y_{\xi^s} \text{ and variables } f^s,\ \lambda^s, \text{ and } \rho^s,\ \forall s \in \Omega \\
& \alpha + \sum_{s\in\Omega} \frac{\mathrm{Prob}_{\xi^s} b_{\xi^s}}{\theta} \le \eta && \text{(5a)} \\
& \sum_{i\in\mathcal{S}} \lambda_i^s \le b_{\xi^s} + \alpha \quad \forall s \in \Omega && \text{(5b)} \\
& \alpha \ge 0, \quad b_{\xi^s} \ge 0 \quad \forall s \in \Omega && \text{(5c)}
\end{aligned}
$$

where α represents the corresponding VaR_θ, and is enforced to be nonnegative. Continuous variable b_{ξ^s} denotes the amount of loss (i.e., $L(x, Y_{\xi^s}) = \sum_{i\in\mathcal{S}} \lambda_i^s$) larger than VaR_θ in scenario s

Decomposition for SMCF-VaR and SMCF-CVaR

- 1^{st} Stage: Decide an MCF solution x.
- 2^{nd} Stage: Check whether the risk constraint is satisfied based on outcomes of each scenario.
- As needed, generate cutting planes by using LP-based dual information.

Decomposition for SMCF-VaR and SMCF-CVaR

- 1^{st} Stage: Decide an MCF solution x.
- 2^{nd} Stage: Check whether the risk constraint is satisfied based on outcomes of each scenario.
- As needed, generate cutting planes by using LP-based dual information.

Additional steps before generating a cut:

- Given an MCF solution $\hat{x}$, with the knowledge of $\underline{L}(\hat{x}, Y^s)$ or a $\overline{L}(\hat{x}, Y^s)$, we can quickly decide whether a cut is needed, rather than directly solve the [Flow Loss LP].
- Derive an algorithm ALG($\mathcal{M}$) (uses an augmenting path idea) that reroutes flows in the residual graph $\mathcal{G}(\hat{x})$ and compute the max/min flow losses.
- Complexity of the algorithm: $O(n^2 \times$ the complexity of MAXIMUM-FLOW).

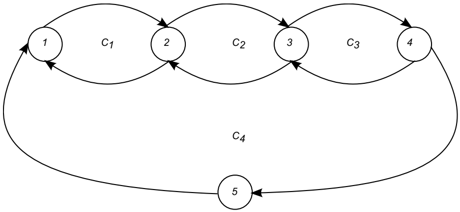
Outline

Kidney Exchange Problem

- Focus on pairing kidneys given by living donors who may be incompatible, with their target patients.
- The method has recently emerged to enable willing but incompatible donor-patient pairs to swap donors.
- Roth et al. (2004) initially propose to organize kidney exchange on a large scale, with the formation of the New England Program for Kidney Exchange (NEPKE).
- Idea: each incompatible donor-patient pair seeks to swap their donors with other pairs to obtain a compatible kidney.

Encoding Kidney Exchange to an MCF in graph $\mathcal{G}(\mathcal{N}, \mathcal{A})$

- A node for each donor-patient pair in $\mathcal{N}$
- An arc from one pair i to another pair j if the donor of pair i is compatible with the patient of pair j.
- Assign weight w_{ij} with each arc (i, j) in $\mathcal{A}$, representing the utility or social welfare attained if the transplant from i to j is implemented.
- A *cycle* in this graph represents a possible swap among multiple pairs, with each pair in the cycle receiving the kidney from the next pair.
- A feasible *exchange* solution is a collection of node-disjoint cycles since each pair can give at most one kidney.
- $\{c_1, c_3\}$ and $\{c_4\}$ are both feasible and maximal exchanges.



MCF Formulation of Deterministic Kidney Exchange

- Seek a set of node-disjoint cycles with the maximum total weights.
- Define binary variables x'_{ij}:

$$x'_{ij} = \begin{cases} 1 & \text{arc } (i,j) \in \mathcal{A} \text{ is contained in the exchange solution} \\ 0 & \text{otherwise} \end{cases}.$$

- A network optimization model:

$$\max \quad \sum_{(i,j)\in\mathcal{A}} w_{ij} x'_{ij} \tag{6a}$$

$$\text{s.t.} \quad \sum_{j:(i,j)\in\mathcal{A}} x'_{ij} - \sum_{j:(j,i)\in\mathcal{A}} x'_{ji} = 0 \quad \forall i \in \mathcal{N} \tag{6b}$$

$$\sum_{j:(i,j)\in\mathcal{A}} x'_{ij} \le 1 \quad \forall i \in \mathcal{N} \tag{6c}$$

$$x'_{ij} \in \{0,1\} \quad \forall (i,j) \in \mathcal{A}, \tag{6d}$$

Risk-Aware Kidney Exchange

A loss-constrained stochastic kidney exchange problem:

- Arc failure: previously compatible donors and receivers may be found incompatible after pairing all the exchanges
- Consequences: all pairs in those cycles containing incompatible pairs are affected since a planned transplant operation is no longer possible
- Current studies: keep the size of kidney-exchange cycles small, e.g., only allow ≤ 3 pairing arcs in each cycle.
- We optimize risk-aware kidney exchange solutions by using SMCF-VaR and SMCF-CVaR models

Example Illustration

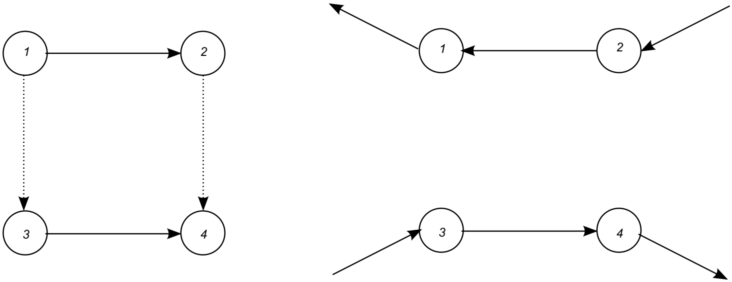


Figure: A 4-way exchange cycle with the failure of arc $(2, 1)$ and arc $(4, 3)$, meaning that the donor of pair 2 (pair 4) cannot give the kidney to the patient of pair 1 (pair 3), and therefore all exchanges involved in the cycle cannot be implemented.

Generalized SMCF-VaR/CVaR Formulation

- Consider random 0-1 match failure of arc (i,j), denoted by a Bernoulli variable Y'_{ij} such that $Y'_{ij} = 1$ if it fails and Y'_{ij} otherwise, for all $(i,j) \in \mathcal{A}$.
- $Y' = [Y'_{ij},\ (i,j) \in \mathcal{A}]^{\mathsf{T}}$

$$L(x', Y') = \sum_{(j,i)\in\mathcal{A}} Y'_{ji} w_{ji} + \min_{f} \sum_{(i,j)\in\overline{\mathcal{A}}} w_{ij} f_{ij} \tag{7a}$$

$$s.t. \sum_{j:(i,j)\in\overline{\mathcal{A}}} f_{ij} - \sum_{j:(j,i)\in\overline{\mathcal{A}}} f_{ji} = \sum_{j:(i,j)\in\mathcal{A}} Y'_{ij} x'_{ij} - \sum_{j:(j,i)\in\mathcal{A}} Y'_{ji} x'_{ji}, \quad \forall i \in \mathcal{N} \tag{7b}$$

$$0 \le f_{ij} \le (1 - Y'_{ji}) x'_{ji}, \quad \forall (i,j) \in \overline{\mathcal{A}}. \tag{7c}$$

Theorem

The value of $L(x', Y')$, given x' and Y', measures exactly the total utility losses of affected exchanges due to match failure.

Generalized SMCF-VaR/CVaR Formulation

- Different from general SMCF models, the utility loss $L(x', Y')$ has a unique value given fixed x' and Y' because a feasible exchange only consists of node-disjoint cycles.
- Denote the random failure by Y'_ξ, we formulate and solve

$$\max\left\{\sum_{(i,j)\in\mathcal{A}} w_{ij}x'_{ij} : \text{(6b)–(6d)},\ \mathbb{P}\left\{L(x', Y'_\xi) \le \eta\right\} \ge 1-\theta\right\}. \quad (8)$$

- An SMCF-CVaR model can be established in a similar way.

Outline

Risk Averse Kidney Exchange (RAKE): Experimental Design

- The number of donor-patient pairs: $|\mathcal{N}| = 50, 100, 200$.
- Each node having an outgoing degree between $0.04|\mathcal{N}|$ and $0.12|\mathcal{N}|$
- Unit utility $w_{ij} = 1$ for all $(i, j) \in \mathcal{A}$.
- Follow the literature (Dickerson et al. (2013)) to set the failure probabilities of each match in $\mathcal{A}$. We sample randomly from a bimodal distribution with 30% of arcs having a low failure rate in (0, 0.2] while 70% arcs having a high failure rate between [0.8, 1), and thus the fail percentage is 66% (verified by the literature).
- 200 scenarios with equal probability 0.5% of realizing each scenario according to these arc-failure rates.

Computational Procedures and Benchmark

- **MaxU**: solve a deterministic kidney exchange problem that maximizes the total exchange utility without any arc failure (yielding an optimal exchange solution x'_{MaxU}).
- Computing the expected utility losses caused by x'_{MaxU}:

$$L_{\mathrm{Max}} = \mathbb{E}_{\xi}\left[L\left(x'_{\mathrm{MaxU}}, Y'_{\xi}\right)\right] = \frac{1}{|\Omega|}\sum_{s\in\Omega} L\left(x'_{\mathrm{MaxU}}, Y'_{\xi^s}\right). \tag{9}$$

- **MinEL**: maximize the total utility of exchanges and meanwhile minimize the expected losses due to the uncertain compatibility, i.e., $\max_x \sum_{s\in\Omega} w_{ij}x'_{ij} - \frac{1}{|\Omega|}\sum_{s\in\Omega} L\left(x', Y'_{\xi^s}\right)$.
- Denote its optimal objective by L_{Min}.
- Set the threshold loss η as the middle point in $[L_{\mathrm{Min}}, L_{\mathrm{Max}}]$, for both SMCF-VaR or SMCF-CVaR.

Average Cycle Length Given by Different Approaches

	MaxU	MinEL	SMCF-VaR $(1-\theta)$				SMCF-CVaR $(1-\theta)$			
			70%	80%	90%	99%	70%	80%	90%	99%
Avg	10.4	2.7	4.8	3.8	3.2	2.2	4.2	3.6	2.6	2.2
Max	18.2	4.2	6.8	5.0	4.8	2.8	6.4	4.4	4.0	2.4
Min	5.4	2.2	3.2	2.2	2.0	2.0	3.0	2.0	2.0	2.0

- MaxU yields the least conservative solutions, reflected by significantly longer cycles
- MinEL is the most conservative and yields relatively small cycles.
- Using SMCF-VaR, we balance the total utility yielded by large-cycle exchanges and potential utility losses due to match failure
- For the same $1-\theta$ reliability, SMCF-CVaR tends to result in more conservative and thus shorter cycles for exchanging kidneys.
- Both SMCF-VaR and SMCF-CVaR yield shorter-cycle exchanges as we increase the reliability $1-\theta$.