Online Supplement

Appendix B: Decision-making process for network G_2 in Figure 3

- **Step 0**: $I^{0,\lambda} = \emptyset$, $z^0 = 2$;
- **Step 1**: $I^{1,\lambda} = \{(1,3),(1,4)\}$, and the interdictor would expect the evader to traverse path $1-2-5$.

This implies that $z_R^{1,*} = 7$ and that the evader would go through path $1-2-5$, implying that $z^{1,\lambda} = 2$.

- **Step 2**: $I^{2,\lambda} = \{(1,3),(4,5)\}$, $z_R^{2,*} = 7$, $z^{2,\lambda} = 2$;
- **Step 3**: $I^{3,\lambda} = \{(1,4),(3,5)\}$, $z_R^{3,*} = 7$, $z^{3,\lambda} = 2$;
- **Step 4**: $I^{4,\lambda} = \{(3,5),(4,5)\}$, $z_R^{4,*} = 7$, $z^{4,\lambda} = 2$;
- **Step 5**: $I^{5,\lambda} = \{(1,2),(1,4)\}$, $z_R^{5,*} = 6$, $z^{5,\lambda} = 3$;
- **Step 6**: $I^{6,\lambda} = \{(1,2),(4,5)\}$, $z_R^{6,*} = 6$, $z^{6,\lambda} = 3$;
- **Step 7**: $I^{7,\lambda} = \{(1,4),(2,5)\}$, $z_R^{7,*} = 6$, $z^{7,\lambda} = 3$;
- **Step 8**: $I^{8,\lambda} = \{(2,5),(4,5)\}$, $z_R^{8,*} = 6$, $z^{8,\lambda} = 3$;
- **Step 9**: $I^{9,\lambda} = \{(1,4),(4,5)\}$, $z_R^{9,*} = 6$, $z^{9,\lambda} = 2$;
- **Step 10**: $I^{10,\lambda} = \{(1,2),(1,3)\}$, $z_R^{10,*} = 5$, $z^{10,\lambda} = 4$;
- **Step 11**: $I^{11,\lambda} = \{(1,2),(3,5)\}$, $z_R^{11,*} = 5$, $z^{11,\lambda} = 4$;
- **Step 12**: $I^{12,\lambda} = \{(1,3),(2,5)\}$, $z_R^{12,*} = 5$, $z^{12,\lambda} = 4$;
- **Step 13**: $I^{13,\lambda} = \{(2,5),(3,5)\}$, $z_R^{13,*} = 5$, $z^{13,\lambda} = 4$;
- **Step 14**: $I^{14,\lambda} = \{(1,2),(2,5)\}$, $z_R^{14,*} = 5$, $z^{14,\lambda} = 3$;
- **Step 15**: $I^{15,\lambda} = \{(1,3),(3,5)\}$, $z_R^{15,*} = 5$, $z^{15,\lambda} = 2$;
- **Step 16**: $I^{16,\lambda} = \{(1,2),(2,3)\}$, $z_R^{16,*} = 4$, $z^{16,\lambda} = 4$.

We can see that after 16 steps, the interdictor finally identifies an optimal solution for the full information problem.

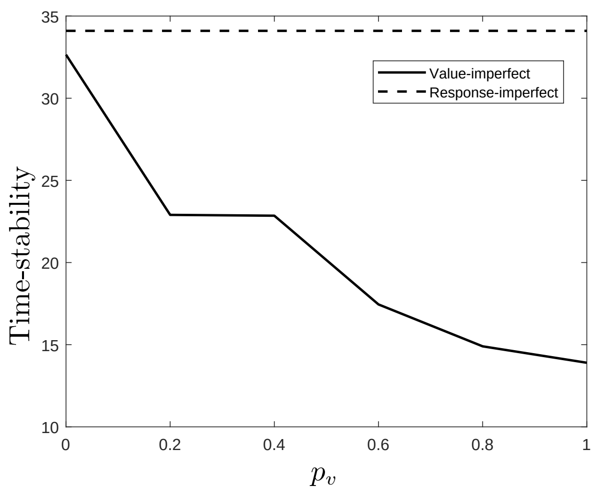
Appendix C: Supplementary Computational Results for Uniform Random Graphs

C.1. Policy Performance: Sensitivity with Respect to p_r and p_v

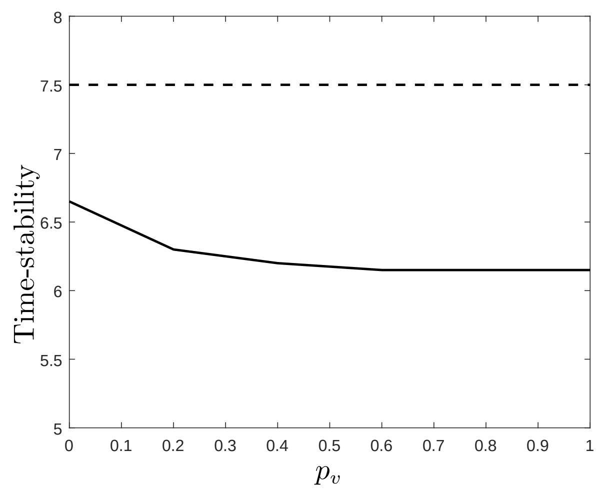
In Section 7.4, we test the performance of the approximate GRN policies on uniform random graphs with right-skewed costs. For left-skewed and symmetric cost structures, we use the same graph size as in Section 7.4, i.e. $n = 50$ and probability of having an arc between any two nodes is 0.5. We set $T = 50$.

The results for symmetric and left-skewed costs are shown in Figure 7 and 8, respectively. We observe that for both response-imperfect and value-imperfect feedback, GRN policies obtain optimal solutions in less time steps as p_r or p_v increases. Moreover, under value-imperfect feedback, the policies converge faster than under response-imperfect feedback.

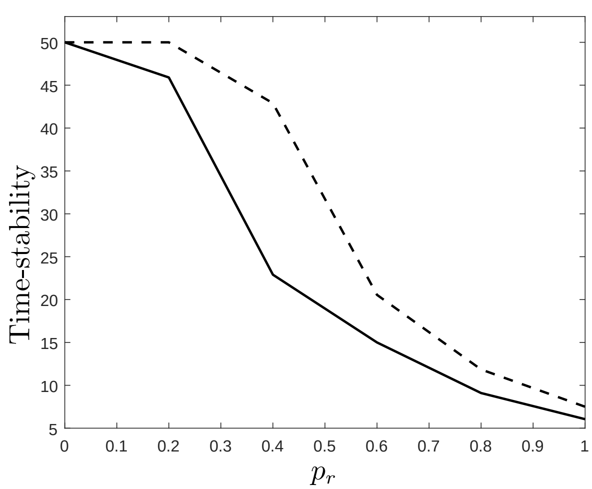
Figure 7 Average time-stability for policies in $\underline{\Lambda}$ for different types of feedback as p_r and p_v increase for $k = 6$ and the symmetric costs ($T = 50$, uniform random graphs, greedy evader).



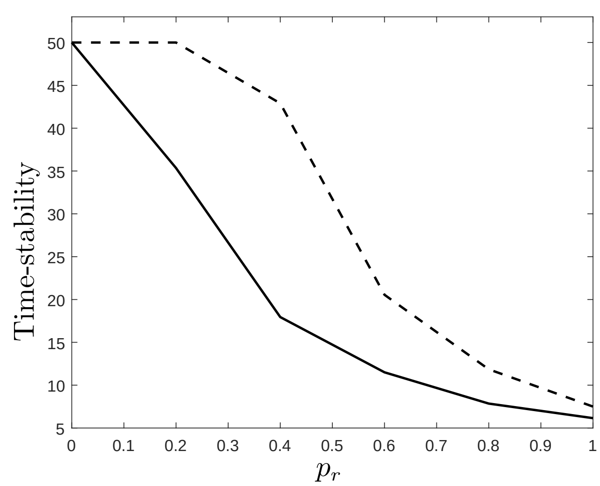
$p_r = 0.5$, p_v vs. time-stability, symmetric



$p_r = 1.0$, p_v vs. time-stability, symmetric

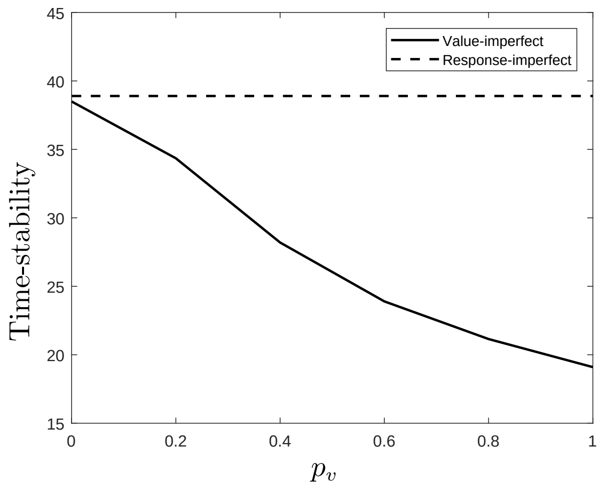


$p_v = 0.5$, p_r vs. time-stability, symmetric

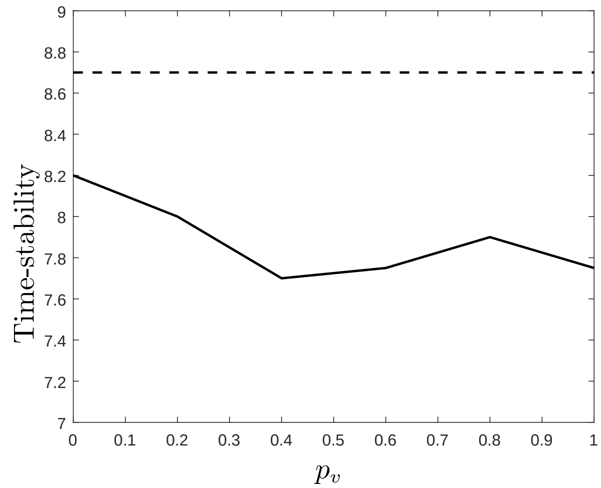


$p_v = 1.0$, p_r vs. time-stability, symmetric

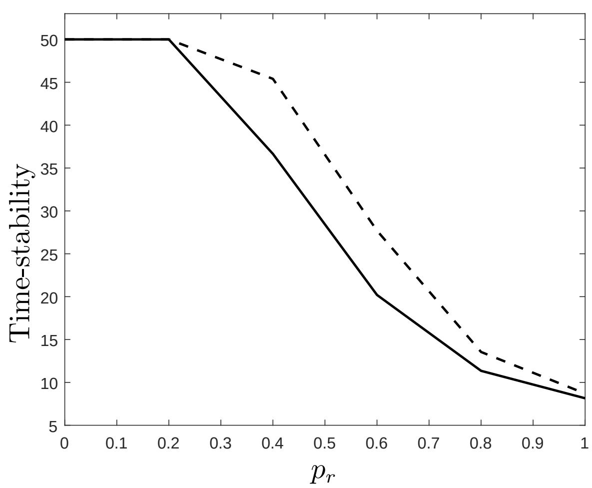
Figure 8 **Average time-stability for policies in $\underline{\Lambda}$ for different types of feedback as p_r and p_v increase for $k = 6$ and the left-skewed costs ($T = 50$, uniform random graphs, greedy evader).**



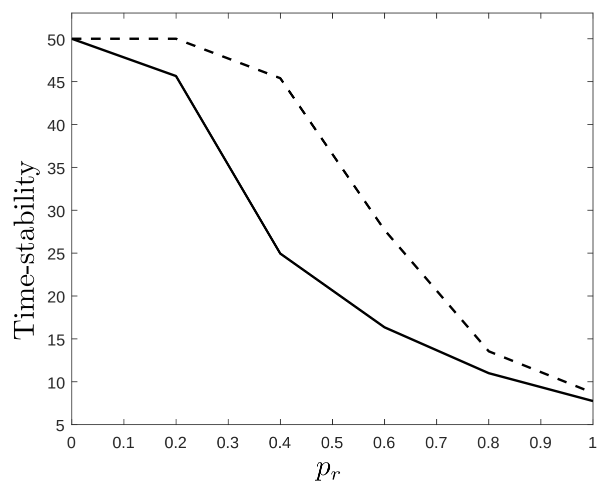
$p_r = 0.5$, p_v vs. time-stability, left-skewed



$p_r = 1.0$, p_v vs. time-stability, left-skewed



$p_v = 0.5$, p_r vs. time-stability, left-skewed



$p_v = 1.0$, p_r vs. time-stability, left-skewed

C.2. Performance under a General Evader

In Section 7.6, we test policy performance when we relax the assumption on the greedy nature of the evader and consider the setting introduced in Section 6, where the evader's response is constrained to a $1-n$ (not necessarily shortest) path on the interdicted graph. Policy performance under Random–0.9 and Random–0.5 in depicted in Tables 6 and 7, respectively. The definitions of Random–0.9 and Random–0.5 evader can be found in Section 7.6. Observe that with *less randomness*, for example under Random–0.9 evader, there are more cases where the approximate GRN policies outperform the mean-bound policies.

Table 6 Average time-stability and MAD (in parenthesis) for $\lambda_v^G \in \underline{\Lambda}_v^G$ and $\pi_M^G \in \Pi_M^G$ policies (with value-imperfect feedback) when evader type is Random–0.9 on uniform random graphs with $k=6$ and $T=100$. Note that a Random-0.9 evader chooses the shortest path with probability 0.9 and the second shortest path with probability 0.1.

n		Right-skewed		Symmetric		Left-skewed		Random	
		λ_v^G	$\pi_M^G(+\mathcal{V})$	λ_v^G	$\pi_M^G(+\mathcal{V})$	λ_v^G	$\pi_M^G(+\mathcal{V})$	λ_v^G	$\pi_M^G(+\mathcal{V})$
15	Time-stability ($\tilde{\tau}$)	**20.2**	70.9	**13.7**	23.5	20.0	**19.0**	**20.6**	59.7
	MAD	23.9	35.0	5.5	31.7	6.9	8.0	5.2	41.6
	Total regret ($\tilde{R}^T$)	580.8	806.8	106.6	62.5	166.5	140.3	221.0	310.6
	MAD	895.8	850.1	35.1	29.2	47.3	45.7	61.5	154.7
20	Time-stability ($\tilde{\tau}$)	**32.9**	71.3	**19.8**	21.6	**18.9**	24.4	**19.5**	34.7
	MAD	40.1	35.2	7.7	28.7	7.1	8.0	6.5	30.5
	Total regret ($\tilde{R}^T$)	1065.5	1189.9	132.4	54.8	118.8	130.5	164.8	181.1
	MAD	1438.0	1334.1	50.5	36.6	39.0	43.4	67.6	117.1
25	Time-stability ($\tilde{\tau}$)	**15.3**	70.6	15.2	**12.9**	25.3	**22.3**	**21.8**	60.8
	MAD	17.0	34.8	4.9	16.4	7.1	7.7	4.4	38.8
	Total regret ($\tilde{R}^T$)	315.4	549.0	96.3	42.0	149.3	115.2	168.7	314.5
	MAD	491.4	529.7	28.4	26.1	39.0	35.9	53.9	210.8
30	Time-stability ($\tilde{\tau}$)	**34.8**	77.0	**19.8**	35.9	**24.9**	28.0	**20.3**	38.8
	MAD	39.2	30.4	9.7	41.3	5.8	12.4	4.4	25.7
	Total regret ($\tilde{R}^T$)	965.0	1091.0	122.3	253.4	118.9	118.6	144.8	178.2
	MAD	1290.7	1190.3	68.3	332.5	36.7	49.0	39.1	76.4
35	Time-stability ($\tilde{\tau}$)	**24.6**	78.2	**13.9**	21.2	**20.1**	28.0	**28.7**	57.4
	MAD	30.2	30.8	5.8	29.1	5.4	7.7	10.5	37.3
	Total regret ($\tilde{R}^T$)	676.4	824.5	87.4	44.9	103.7	140.8	225.7	216.4
	MAD	1023.0	965.9	31.6	34.1	47.2	76.2	128.0	112.6
40	Time-stability ($\tilde{\tau}$)	**24.6**	86.5	**15.3**	22.3	28.6	**28.4**	**22.8**	42.9
	MAD	30.2	17.6	4.0	29.0	8.8	6.3	7.8	31.9
	Total regret ($\tilde{R}^T$)	473.8	754.8	80.7	53.8	118.9	101.8	137.8	165.0
	MAD	699.4	631.7	20.5	50.3	46.3	27.9	45.6	77.6
45	Time-stability ($\tilde{\tau}$)	**36.1**	86.6	**15.5**	35.8	**23.0**	30.5	**25.3**	46.8
	MAD	38.4	17.0	5.2	40.8	7.4	11.0	6.9	30.6
	Total regret ($\tilde{R}^T$)	739.2	942.7	74.5	69.4	91.1	104.9	163.1	206.0
	MAD	979.6	873.2	24.0	60.9	31.9	43.3	62.1	125.1
50	Time-stability ($\tilde{\tau}$)	**29.4**	77.7	**24.3**	34.8	**24.6**	28.5	**22.4**	53.1
	MAD	34.3	30.7	15.5	40.9	7.4	9.4	5.6	32.9
	Total regret ($\tilde{R}^T$)	506.4	636.6	222.4	197.1	74.9	96.5	122.0	179.4
	MAD	708.0	595.1	263.5	265.9	25.5	39.2	41.0	68.9

Table 7 Average time-stability and MAD (in parenthesis) for $\lambda_v^G \in \underline{\Lambda}_v^G$ and $\pi_M^G \in \Pi_M^G$ policies (with value-imperfect feedback) when evader type is Random–0.5 on uniform random graphs with $k=6$ and $T=100$. Note that a Random–0.5 evader randomly chooses from the shortest path and the second shortest path.

n		Right-skewed		Symmetric		Left-skewed		Random	
		λ_v^G	$\pi_M^G(+\mathcal{V})$	λ_v^G	$\pi_M^G(+\mathcal{V})$	λ_v^G	$\pi_M^G(+\mathcal{V})$	λ_v^G	$\pi_M^G(+\mathcal{V})$
15	Time-stability ($\tilde{\tau}$)	**20.9**	57.0	**13.5**	14.5	20.1	19.1	**17.7**	45.3
	MAD	23.7	30.9	6.3	19.1	5.6	5.8	5.2	33.2
	Total regret ($\tilde{R}^T$)	570.5	738.6	99.1	49.0	163.4	134.4	177.1	241.2
	MAD	855.2	820.1	36.2	18.1	56.9	42.4	66.7	120.2
20	Time-stability ($\tilde{\tau}$)	**30.7**	58.5	17.1	**16.6**	**20.9**	26.6	**19.0**	28.2
	MAD	34.7	29.1	6.8	17.5	5.6	7.9	5.4	22.2
	Total regret ($\tilde{R}^T$)	912.0	1022.5	113.0	55.1	124.0	135.0	153.3	163.1
	MAD	1285.0	1223.6	42.9	31.3	36.7	44.1	61.8	97.3
25	Time-stability ($\tilde{\tau}$)	**16.3**	47.6	17.0	**8.0**	**23.8**	26.9	**23.4**	42.9
	MAD	16.8	24.8	5.5	8.6	6.1	8.2	5.6	27.9
	Total regret ($\tilde{R}^T$)	327.6	421.3	109.9	36.8	132.7	146.6	165.1	268.4
	MAD	474.2	468.4	42.5	19.8	34.2	47.5	44.3	184.2
30	Time-stability ($\tilde{\tau}$)	**35.2**	68.8	**24.9**	28.8	**25.4**	30.7	**20.6**	36.2
	MAD	38.9	34.4	16.6	32.1	6.4	7.6	4.8	21.4
	Total regret ($\tilde{R}^T$)	844.7	932.0	232.5	214.1	109.9	118.3	136.8	165.3
	MAD	1096.3	1056.8	261.8	274.0	31.6	35.2	49.7	53.3
35	Time-stability ($\tilde{\tau}$)	**25.7**	67.9	16.0	**15.3**	**24.9**	31.0	**22.2**	45.5
	MAD	29.7	35.3	5.0	20.0	7.0	9.4	5.6	23.2
	Total regret ($\tilde{R}^T$)	594.5	703.0	86.2	44.9	122.2	138.4	158.3	195.6
	MAD	871.4	829.1	28.7	36.8	56.4	52.7	65.6	80.0
40	Time-stability ($\tilde{\tau}$)	**25.1**	68.2	**15.3**	20.0	**24.1**	27.2	**21.7**	45.9
	MAD	30.0	31.8	4.3	23.3	5.4	10.1	7.0	25.7
	Total regret ($\tilde{R}^T$)	435.2	602.3	68.8	51.4	98.2	108.1	127.9	172.0
	MAD	632.7	574.8	17.7	42.5	34.2	46.4	48.2	80.0
45	Time-stability ($\tilde{\tau}$)	**37.3**	71.3	**17.5**	25.6	**25.3**	33.0	**24.7**	39.5
	MAD	37.7	31.6	5.2	25.6	8.5	9.9	5.7	21.4
	Total regret ($\tilde{R}^T$)	643.5	768.0	73.5	57.2	104.2	118.5	135.9	180.1
	MAD	832.5	771.3	18.1	38.2	47.4	45.3	54.1	98.2
50	Time-stability ($\tilde{\tau}$)	**26.2**	68.1	29.9	**28.9**	**23.3**	28.2	**24.6**	37.5
	MAD	29.5	35.1	14.3	31.4	7.7	6.1	7.6	20.1
	Total regret ($\tilde{R}^T$)	372.2	482.7	239.5	192.3	77.6	89.2	131.8	159.3
	MAD	518.8	459.0	240.2	258.5	31.5	27.3	49.0	76.3

Appendix D: Computational Results for Layered Graphs

D.1. Graph generation

We generate layered graphs using parameters (θ, ϕ), where ϕ and θ denote the number of layers and nodes in each layer, respectively. We add a source node before the first layer and a destination node after the last layer. Thus, the total number of nodes is $\theta \times \phi + 2$. There is an arc from source node to all the nodes in the first layer and from all the nodes in the last layer to the destination node. Moreover, there is an arc with probability 0.5 from every node in layer i to nodes in layers $i+1, i+2, \ldots, \phi$.

D.2. Comparison of Policies without Information Updates

We compare policy performance when no information updates are used, i.e., $\mathcal{C}^t = \mathcal{C}^0$ for all periods t. We test policy performance using the right-skewed, symmetric, left-skewed and random cost structures on layered graphs. For each structure we randomly generate 20 instances with $(\theta, \phi) = (7, 3)$. Finally, we use $k \in \{2, 4\}$, and set either $T = 500$ or $T = 1000$, respectively. Table 8 summarizes the results, which are similar to those for uniform random graphs, see Table 2.

D.3. Improvements When Information Updates Are Applied

We consider layered graphs with 7 nodes ($\theta = 7$) in each layer and number of layers $\phi \in \{3, 4, \ldots, 10\}$ for all four cost structures. We set $T = 500$ and for each cost structure we randomly generate 20 instances. For response-imperfect feedback we set $p_r = 0.5$, and for value-imperfect feedback we set $p_r = p_v = 0.5$. Note that in this section and the following two sections, since we have polyhedral cost update mechanism, the approximate GRN policies are implemented. We measure policy performance using the average time-stability and MAD, see Table 9. As in Section 7.3, we observe that policy performance improves as more information revealed in the feedback. However, note that there is no significant improvement for policies in Π, even if the interdictor learns more information. On the other hand, GRN policies outperform mean bound policies under response- and value-imperfect feedback.

D.4. Policy Performance: Sensitivity with Respect to p_r and p_v

We set $(\theta, \phi) = (7, 10)$ for all the experiments, and generate 20 instances with different cost structures (right-skewed, symmetric and left-skewed). Interdictor's resource limit and time horizon are set as $k = 6$ and $\mathcal{T} = \{0, 1, \ldots, 50\}$. Figure 9, 10 and 11 depict the behaviour of time-stability for response- and value-imperfect updates with the right-skewed, symmetric and left-skewed costs, respectively. As in the case of uniform random graphs in Section 7.4, performance of policies $\underline{\Lambda}_r$ and $\underline{\Lambda}_v$ improves as p_r and p_v increase. We observe that time-stability of policies Λ_v decreases faster than that of Λ_r.

D.5. Policy Performance: Sensitivity with Respect to Quality of Arc Cost Bounds

We generate 20 instances for I.1, I.2 and I.3 as in Section 7.5, and set $k = 6$. Test instances are layered graphs with $(\theta, \phi) = (7, 10)$ and we set $T = 200$. We test policy performance for various values of probabilities p_r and p_v. Table 10 summarizes the results obtained. There, we observe similar results to those presented in Section 7.5, which indicate that the quality of the initial bounds has a significant affect on policy performance.

Table 8 **Performance of policies in Λ and benchmark policies without information updates (7×3, layered graphs, greedy evader).**

Graph		$k=2$ ($T=500$)				$k=4$ ($T=1000$)			
		λ	π_L	π_M	π_R	λ	π_L	π_M	π_R
Right-skewed	Time-stability	**82.0**^{1}	450.2^{18}	201.2^{8}	450.2^{18}	**445.3**^{7}	900.2^{18}	600.8^{12}	950.2^{18}
	MAD	116.1	89.6	239.0	89.6	420.3	179.6	479.0	94.7
	Relative difference	**0.2%**	18.5%	3.7%	18.2%	**2.9%**	19.9%	5.6%	17.0%
	MAD	0.3%	11.9%	4.4%	12.6%	2.6%	9.1%	5.6%	9.5%
	Total regret	**242.3**	4925.0	1025.0	5082.5	**2002.4**	12700.0	3650.0	11103.4
	MAD	360.4	3167.5	1237.5	3575.8	1973.0	5140.0	3680.0	6306.9
Symmetric	Time-stability	381.6^{12}	350.6^{14}	**133.4**^{4}	378.5^{15}	946.4^{17}	900.2^{18}	**153.2**^{3}	691.2^{14}
	MAD	142.1	209.2	167.2	182.3	91.1	179.6	246.5	401.5
	Relative difference	10.6%	11.9%	**0.5%**	10.6%	17.4%	15.3%	**0.3%**	10.5%
	MAD	9.5%	12.2%	0.8%	10.8%	8.4%	11.3%	0.6%	9.5%
	Total regret	2343.3	2450.0	**146.8**	2259.4	10055.2	7750.0	**382.7**	5945.2
	MAD	1516.7	2580.0	185.5	2249.0	3969.8	6025.0	641.3	4763.7
Left-skewed	Time-stability	411.9^{15}	**151.4**^{6}	337.5^{9}	324.4^{9}	978.6^{19}	**351.3**^{7}	867.6^{15}	876.0^{16}
	MAD	132.2	209.2	146.3	165.7	40.8	454.1	206.9	198.5
	Relative difference	27.4%	**3.5%**	14.1%	9.1%	34.4%	**3.2%**	19.4%	22.2%
	MAD	19.2%	4.9%	15.8%	10.0%	13.5%	4.3%	15.7%	18.8%
	Total regret	4085.8	**425.0**	2247.8	2334.9	10635.1	**1050.0**	7860.9	8042.5
	MAD	2592.7	595.0	1377.0	1811.9	3858.9	1370.0	3827.3	5773.4
Random	Time-stability	425.2^{15}	**226.1**^{9}	290.6^{9}	350.9^{14}	974.1^{19}	850.3^{17}	**781.2**^{14}	950.5^{18}
	MAD	115.3	246.5	189.9	208.7	49.3	254.5	306.3	94.1
	Relative difference	29.5%	9.0%	**8.1%**	14.8%	32.0%	22.3%	**18.6%**	25.9%
	MAD	18.9%	10.4%	9.2%	12.6%	15.4%	13.6%	14.3%	17.5%
	Total regret	4912.9	1850.0	**1719.5**	2895.8	16014.8	11600.0	**9979.1**	14468.1
	MAD	2889.4	2170.0	1587.4	2365.4	7935.9	7500.0	6205.8	9943.2

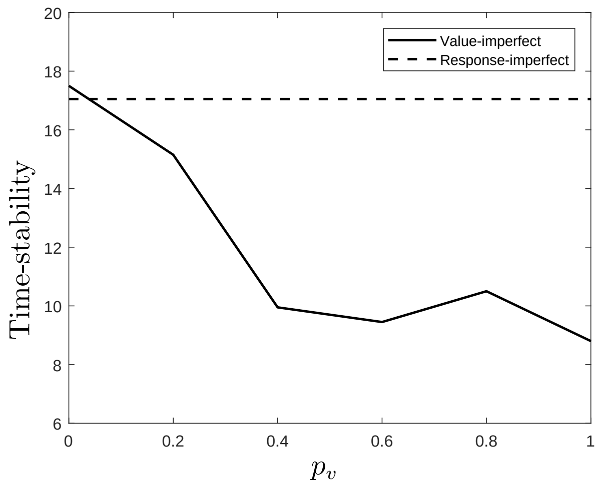
Notes: Entries in bold denote the best policy in each setting; the numbers in superscript of time-stability denote the number of instances out of 20 for which the corresponding policy failed to converge within T time periods.

Table 9 Average time-stability and MAD (in parenthesis) for $\lambda \in \Delta$ and $\pi_M \in \Pi_M$ policies when information updates are used ($k=6$, $T=500$, layered graphs, greedy evader).

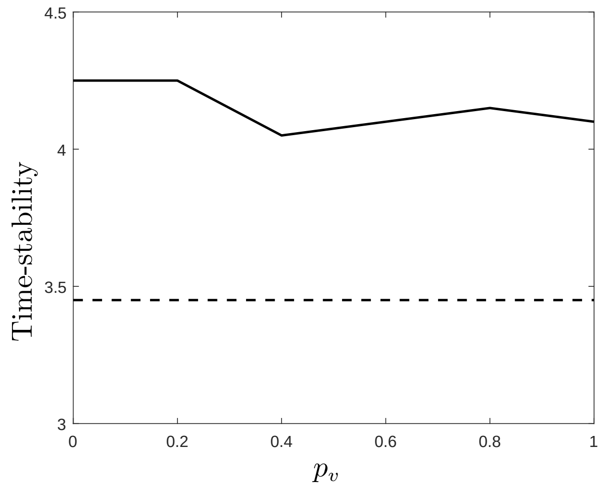
Size	Right-skewed					Symmetric					Left-skewed					Random				
	λ	λ_r	λ_v	π_M	$\pi_M(+)$	λ	λ_r	λ_v	π_M	$\pi_M(+)$	λ	λ_r	λ_v	π_M	$\pi_M(+)$	λ	λ_r	λ_v	π_M	$\pi_M(+)$
7×3	341.6^{11}	12.7	**7.3**	325.7^{13}	325.7^{13}	500^{20}	31.3	**19.2**	226.9^{7}	128.5^{5}	500^{20}	43.8	**24.7**	489.4^{19}	27.0	500^{20}	37.5	**23.0**	330.1^{17}	329.16^{13}
	(195.8)	(5.6)	(2.4)	(226.6)	(226.6)	(0.0)	(12.0)	(6.2)	(232.7)	(185.8)	(0.0)	(19.9)	(6.7)	(20.2)	(6.0)	(0.0)	(14.9)	(5.3)	(208.6)	(222.2)
7×4	383.2^{15}	11.4	**8.5**	375.5^{15}	375.5^{15}	500^{20}	33.1	**16.7**	155.0^{6}	176.6^{7}	500^{20}	39.1	33.6	500^{20}	31.6	500^{20}	40.9	**22.6**	500^{20}	282.3^{11}
	(175.2)	(5.6)	(3.3)	(186.8)	(186.8)	(0.0)	(12.0)	(6.0)	(207.0)	(226.4)	(0.0)	(17.1)	(15.7)	(0.0)	(7.7)	(0.0)	(13.8)	(4.9)	(0.0)	(239.5)
7×5	383.6^{15}	15.4	**8.4**	300.8^{12}	300.8^{12}	500^{20}	29.8	$\mathbf{43.4}^{1}$	281.9^{11}	153.3^{6}	500^{20}	43.2	50.0^{1}	500^{20}	29.7	500^{20}	35.2	**24.4**	500^{20}	212.8^{8}
	(174.7)	(8.5)	(3.4)	(239.04)	(239.0)	(0.0)	(15.4)	(45.7)	(239.91)	(208.0)	(0.0)	(16.7)	(45.0)	(0.0)	(8.9)	(0.0)	(9.7)	(6.4)	(0.0)	(229.8)
7×6	345.7^{13}	13.5	**11.6**	350.6^{13}	350.6^{14}	500^{20}	60.1^{1}	$\mathbf{67.9}^{3}$	308.5^{11}	228.1^{9}	500^{20}	68.9^{1}	36.1	500^{20}	34.1	500^{20}	74.6^{1}	$\mathbf{52.1}^{1}$	500^{20}	283.3^{11}
	(200.7)	(7.6)	(7.2)	(209.2)	(209.2)	(0.0)	(45.1)	(86.4)	(229.9)	(244.8)	(0.0)	(44.5)	(10.4)	(0.0)	(9.5)	(0.0)	(44.9)	(47.5)	(0.0)	(238.4)
7×7	430.2^{16}	17.9	**14.7**	400.4^{16}	400.4^{16}	500^{20}	55.9^{1}	$\mathbf{45.7}^{1}$	259.8^{9}	106.4^{4}	500^{20}	77.2^{1}	76.8^{2}	500^{20}	37.5	500^{20}	73.2^{1}	**29.1**	500^{20}	307.8^{12}
	(117.3)	(6.9)	(9.1)	(159.4)	(159.4)	(0.0)	(47.0)	(45.8)	(240.2)	(157.4)	(0.0)	(53.5)	(84.7)	(0.0)	(10.0)	(0.0)	(46.7)	(5.5)	(0.0)	(230.7)
7×8	382.5^{14}	38.9	**9.2**	350.6^{15}	400.4^{16}	500^{20}	122.8^{4}	$\mathbf{93.5}^{3}$	335.7^{12}	231.4^{9}	500^{20}	53.6	**28.1**	500^{20}	36.0	500^{20}	128.9^{3}	**30.6**	500^{20}	190.6^{7}
	(176.3)	(46.1)	(3.4)	(209.2)	(159.4)	(0.0)	(150.9)	(122.0)	(213.6)	(241.7)	(0.0)	(15.8)	(9.1)	(0.0)	(11.7)	(0.0)	(117.2)	(6.3)	(0.0)	(216.6)
7×9	453.2^{17}	22.8	**11.5**	325.7^{13}	300.8^{12}	500^{20}	101.4^{3}	**46.2**	295.1^{11}	82.4^{3}	500^{20}	83.8^{1}	**31.6**	500^{20}	38.1	500^{20}	41.8	**53.0**	500^{20}	214.5^{8}
	(84.2)	(11.0)	(3.3)	(226.6)	(239.0)	(0.0)	(119.6)	(45.4)	(225.4)	(125.3)	(0.0)	(50.3)	(7.3)	(0.0)	(9.0)	(0.0)	(10.5)	(44.7)	(0.0)	(228.4)
7×10	379.0^{15}	11.1	**8.4**	400.4^{16}	400.4^{16}	500^{20}	114.9^{3}	$\mathbf{69.2}^{2}$	299.0^{10}	155.6^{6}	500^{20}	138.4^{4}	$\mathbf{77.6}^{2}$	500^{20}	107.5^{3}	500^{20}	46.5	$\mathbf{52.0}^{1}$	500^{20}	166.9^{8}
	(181.6)	(4.7)	(4.5)	(159.4)	(159.4)	(0.0)	(115.5)	(86.2)	(221.2)	(206.7)	(0.0)	(144.6)	(84.5)	(0.0)	(117.8)	(0.0)	(9.7)	(44.8)	(0.0)	(199.9)

Note. The numbers in superscript of time-stability denote the number of instances out of 20 for which the corresponding policy failed to converge within T time periods.

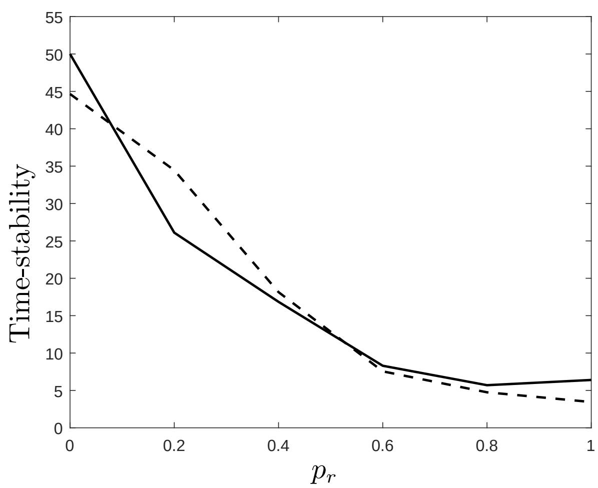
Figure 9 **Average time-stability for policies in $\underline{\Lambda}$ for different types of feedback as p_r and p_v increase for $k = 6$ and the right-skewed costs ($T = 50$, layered graphs, greedy evader).**



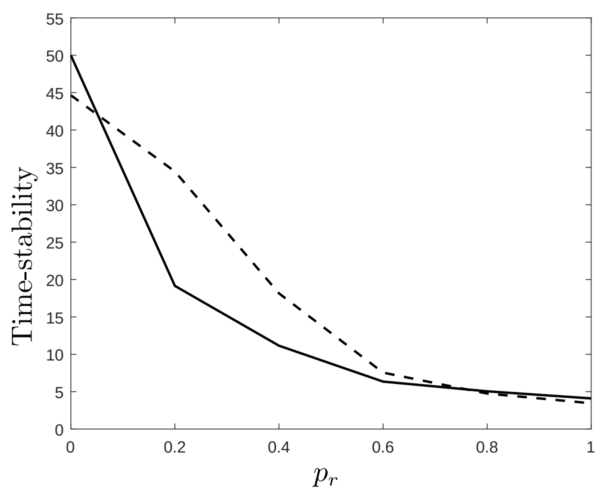
$p_r = 0.5$, p_v vs. time-stability, right-skewed



$p_r = 1.0$, p_v vs. time-stability, right-skewed

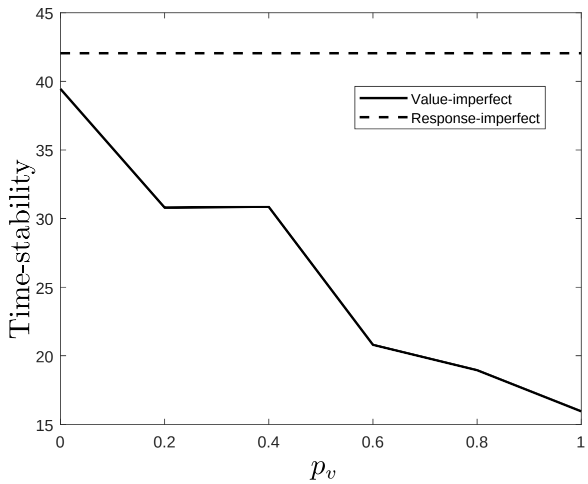


$p_v = 0.5$, p_r vs. time-stability, right-skewed

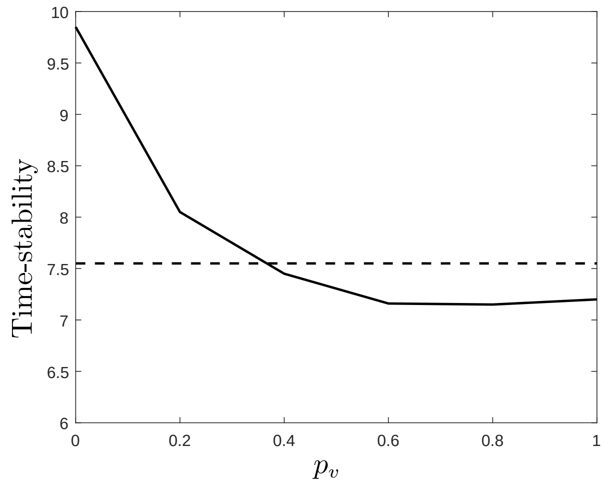


$p_v = 1.0$, p_r vs. time-stability, right-skewed

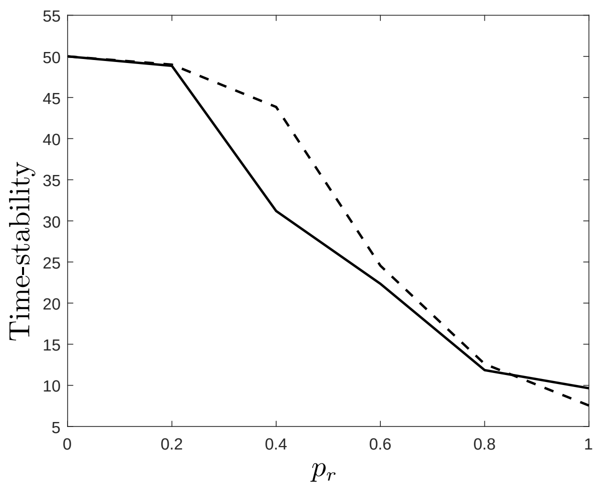
Figure 10 **Average time-stability for policies in $\underline{\Lambda}$ for different types of feedback as p_r and p_v increase for $k = 6$ and the symmetric costs ($T = 50$, layered graphs, greedy evader).**



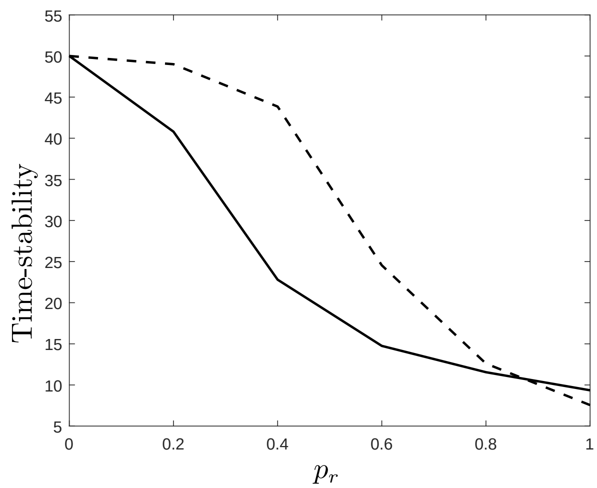
$p_r = 0.5$, p_v vs. time-stability, symmetric



$p_r = 1.0$, p_v vs. time-stability, symmetric

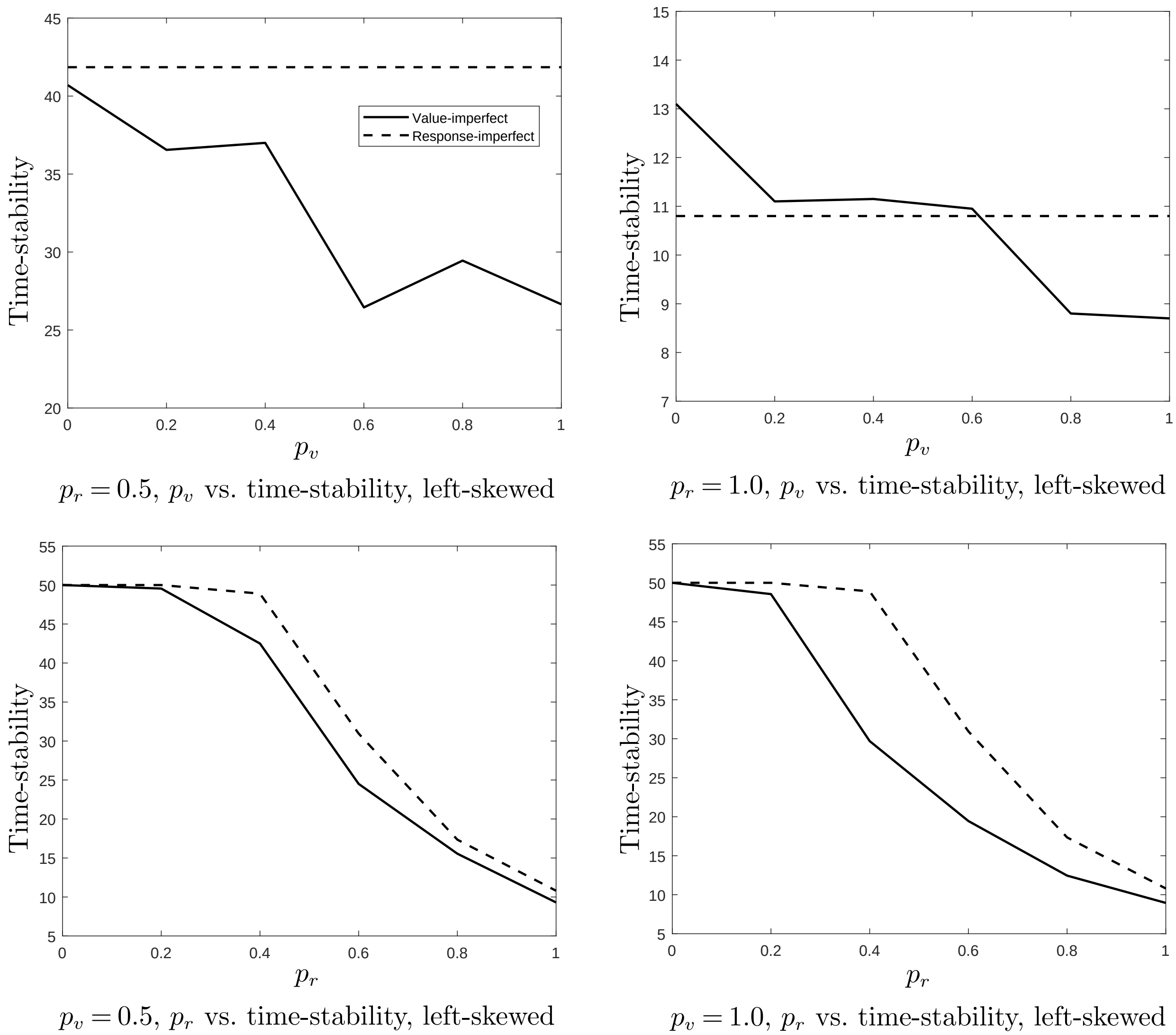


$p_v = 0.5$, p_r vs. time-stability, symmetric



$p_v = 1.0$, p_r vs. time-stability, symmetric

Figure 11 **Average time-stability for policies in $\underline{\Lambda}$ for different types of feedback as p_r and p_v increase for $k=6$ and the left-skewed costs ($T=50$, layered graphs, greedy evader).**



$p_r = 0.5$, p_v vs. time-stability, left-skewed

$p_r = 1.0$, p_v vs. time-stability, left-skewed

$p_v = 0.5$, p_r vs. time-stability, left-skewed

$p_v = 1.0$, p_r vs. time-stability, left-skewed

Table 10 Behaviour of policies in $\underline{\Lambda}$ with respect to the cost bound quality ($k = 6$, $T = 200$, layered graphs, greedy evader).

	$p_r = p_v$		0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Policy λ_r	I.1	Time-stability	**150.3**	**107.1**	**44.4**	**19.0**	**10.5**	**8.2**	**6.1**	**5.0**	**4.1**	**7.8**	**2.7**
		MAD	54.7	60.8	36.8	8.3	3.9	3.8	2.9	2.0	1.7	9.3	0.6
		Total regret	**6102.0**	**5447.0**	**3063.7**	**126.4**	**56.3**	**32.7**	**26.6**	**17.3**	**19.7**	**43.2**	**7.5**
		MAD	5112.9	5995.8	4908.5	142.8	52.5	34.3	25.9	17.0	23.1	65.1	7.2
	I.2	Time-stability	193.3	181.0	112.7	53.6	31.3	15.3	12.4	14.0	11.8	15.6	4.9
		MAD	12.8	30.5	58.0	26.9	14.4	6.6	4.0	10.3	9.7	18.0	1.1
		Total regret	13827.2	12969.3	8212.0	2822.1	1632.5	658.1	575.2	582.5	610.3	567.9	183.2
		MAD	9674.3	7633.7	6920.5	2336.6	1251.9	460.8	446.1	422.9	601.8	621.1	113.1
	I.3	Time-stability	200.0	198.1	155.1	70.2	54.4	30.4	27.6	13.7	14.4	9.0	7.1
		MAD	0.0	3.6	44.3	29.3	30.2	8.8	17.5	3.7	9.5	1.3	0.1
		Total regret	33018.8	21630.7	11753.0	5899.6	4605.6	2515.5	2334.1	1109.2	981.0	772.8	546.4
		MAD	12001.0	7696.5	5683.3	2858.8	2406.5	1113.8	1575.9	412.2	648.2	360.3	208.9
Policy λ_v	I.1	Time-stability	**161.5**	**75.4**	**28.8**	**13.1**	**6.6**	**6.2**	**4.9**	**4.3**	**4.4**	**3.5**	**3.2**
		MAD	57.8	55.2	16.5	7.0	2.7	2.7	2.0	1.7	1.3	1.3	1.1
		Total regret	**12643.3**	**3461.0**	**361.0**	**53.2**	**38.5**	**28.6**	**17.0**	**18.7**	**19.7**	**17.7**	**16.6**
		MAD	12171.1	4513.4	417.7	58.4	39.6	27.1	16.6	19.8	20.0	19.4	17.8
	I.2	Time-stability	193.3	163.9	81.6	30.8	17.6	20.3	9.5	8.1	15.7	5.4	4.9
		MAD	12.8	49.0	51.2	15.5	6.5	18.0	3.4	2.9	18.4	1.3	1.2
		Total regret	13894.5	9782.2	5299.3	1174.7	691.5	968.0	317.4	327.8	573.2	189.0	171.3
		MAD	8683.5	5049.2	5479.6	934.9	380.5	1181.5	196.9	238.3	667.4	119.4	102.0
	I.3	Time-stability	200.0	197.3	123.3	49.2	29.1	30.6	15.7	21.9	9.9	8.6	7.1
		MAD	0.0	5.0	29.8	12.0	6.0	18.0	3.5	17.8	1.4	0.9	0.1
		Total regret	33589.3	22066.9	11660.3	4200.6	2533.7	2043.9	1333.5	1285.0	795.4	637.6	546.4
		MAD	13519.1	9613.6	4940.6	1440.0	1143.8	956.6	611.9	809.8	345.0	262.4	208.9

Note. The numbers in superscript of time-stability denote the number of instances out of 20 for which the corresponding policy failed to converge within T time periods

Appendix E: Computational Results for Watts-Strogatz Graphs

E.1. Graph generation

We generate Watts-Strogatz graphs following the model in Watts and Strogatz (1998), using parameters (n, d, β). Mean degree of nodes and rewiring probability are denoted as d and β, respectively. Note that β denotes the graph instances' degree of randomness. Through all the experiments, we set $\beta = 1$.

E.2. Comparison of Policies without Information Updates

For each structure we randomly generate 20 instances with $(n, d) = (15, 8)$. Finally, we have $k \in \{2, 4\}$, and use $T \in \{500, 1000\}$. Table 11 summarizes the results, which are similar to those for uniform random and layered graphs, see Table 2 and Table 8, respectively.

E.3. Improvements When Information Updates Are Applied

We consider Watts-Strogatz graphs with

$$(n, d) \in \{(15, 8), (20, 15), (25, 14), (30, 16), (35, 22), (40, 22), (45, 28), (50, 30)\},$$

for all four cost structures. We set $T = 500$ and for each cost structure we randomly generate 20 instances. For response-imperfect feedback we set $p_r = 0.5$, and in value-imperfect feedback we set $p_r = p_v = 0.5$. Note that similar in Appendix D in the Online Supplement, in this section and the following two sections, since we have polyhedral cost update mechanism, the approximate GRN policies are implemented. Table 12 depicts our results, which are similar to those obtained for uniform random and layered graphs.

E.4. Policy Performance: Sensitivity with Respect to p_r and p_v

We set $(n, d) = (50, 30)$ for all the experiments, and generate 20 instances with different cost structures (right-skewed, symmetric and left-skewed). Interdictor's resource limit and time horizon are set as $k = 6$ and $\mathcal{T} = \{0, 1, \ldots, 50\}$. Figure 12, 13 and 14 depict the behaviour of time-stability for response- and value-imperfect updates with the right-skewed, symmetric and left-skewed costs, respectively.

E.5. Policy Performance: Sensitivity with Respect to Quality of Arc Cost Bounds

We generate 20 instances for I.1, I.2 and I.3 as in Section 7.5, and set $k = 6$. We set $(n, d) = (50, 30)$ and $T = 200$. We test policy performance for various values of probabilities p_r and p_v. Table 13 summarizes the results obtained, which are similar results to those in Section 7.5 and Appendix D.5 in the Online Supplement.

Table 11 Performance of policies in Λ and benchmark policies without information updates ($n=15$, Watts-Strogatz graphs, greedy evader).

Graph		$k=2$ ($T=500$)				$k=4$ ($T=1000$)			
		λ	π_L	π_M	π_R	λ	π_L	π_M	π_R
Right-skewed	Time-stability	**28.4**	325.7^{13}	101.6^{5}	400.4^{16}	$\mathbf{280.0}^{3}$	850.3^{17}	311.8^{8}	850.6^{17}
	MAD	37.8	226.6	159.4	159.4	303.2	254.5	412.9	254.1
	Relative difference	**0.0%**	13.8%	3.4%	17.9%	**1.4%**	23.2%	1.9%	21.0%
	MAD	0.0%	12.6%	5.3%	11.8%	2.4%	11.3%	2.1%	12.4%
	Total regret	**90.9**	3775.0	875.0	4978.4	**1629.1**	17050.0	1772.4	15650.2
	MAD	145.2	3425.0	1350.0	3328.4	1997.1	8255.0	1949.6	9414.9
Symmetric	Time-stability	182.0^{4}	375.5^{15}	$\mathbf{87.3}^{3}$	195.2^{7}	843.9^{15}	800.4^{16}	$\mathbf{178.5}^{3}$	690.0^{12}
	MAD	143.2	186.8	135.3	220.7	234.2	319.4	267.5	412.4
	Relative difference	2.6%	14.1%	**0.6%**	4.7%	14.9%	13.9%	**1.0%**	10.2%
	MAD	4.2%	13.2%	1.0%	6.1%	11.2%	9.0%	1.7%	8.9%
	Total regret	749.2	3025.0	**155.4**	1048.9	8347.9	7350.0	**655.9**	5715.5
	MAD	622.0	2935.0	247.1	1263.6	4912.7	4985.0	973.4	4588.0
Left-skewed	Time-stability	315.4^{10}	176.3^{7}	246.2^{7}	$\mathbf{88.6}^{1}$	1000.0^{20}	$\mathbf{401.2}^{8}$	702.2^{13}	723.7^{13}
	MAD	184.6	226.6	191.8	99.4	0.0	479.0	387.2	359.2
	Relative difference	12.4%	4.9%	7.6%	**0.5%**	33.9%	**5.2%**	18.6%	13.1%
	MAD	13.2%	6.3%	9.9%	1.0%	13.2%	6.6%	14.3%	13.3%
	Total regret	2362.7	850.0	1543.6	**646.0**	14373.8	**2150.0**	7452.3	6734.9
	MAD	1687.4	1105.0	1373.7	798.4	6177.3	2710.0	4383.1	4523.3
Random	Time-stability	281.4^{9}	350.6^{17}	$\mathbf{216.6}^{7}$	301.6^{12}	940.9^{18}	900.2^{18}	$\mathbf{342.2}^{6}$	751.1^{15}
	MAD	196.8	209.2	206.0	238.1	106.4	179.6	404.1	373.4
	Relative difference	10.7%	15.8%	**4.5%**	11.4%	31.6%	20.3%	**5.0%**	15.1%
	MAD	12.8%	8.8%	6.1%	12.1%	10.9%	10.9%	7.1%	13.9%
	Total regret	2715.5	3250.0	**1293.3**	2540.2	16614.3	11900.0	**2985.4**	9675.5
	MAD	2411.5	1925.0	1442.2	2825.9	4588.5	7190.0	3706.1	8355.2

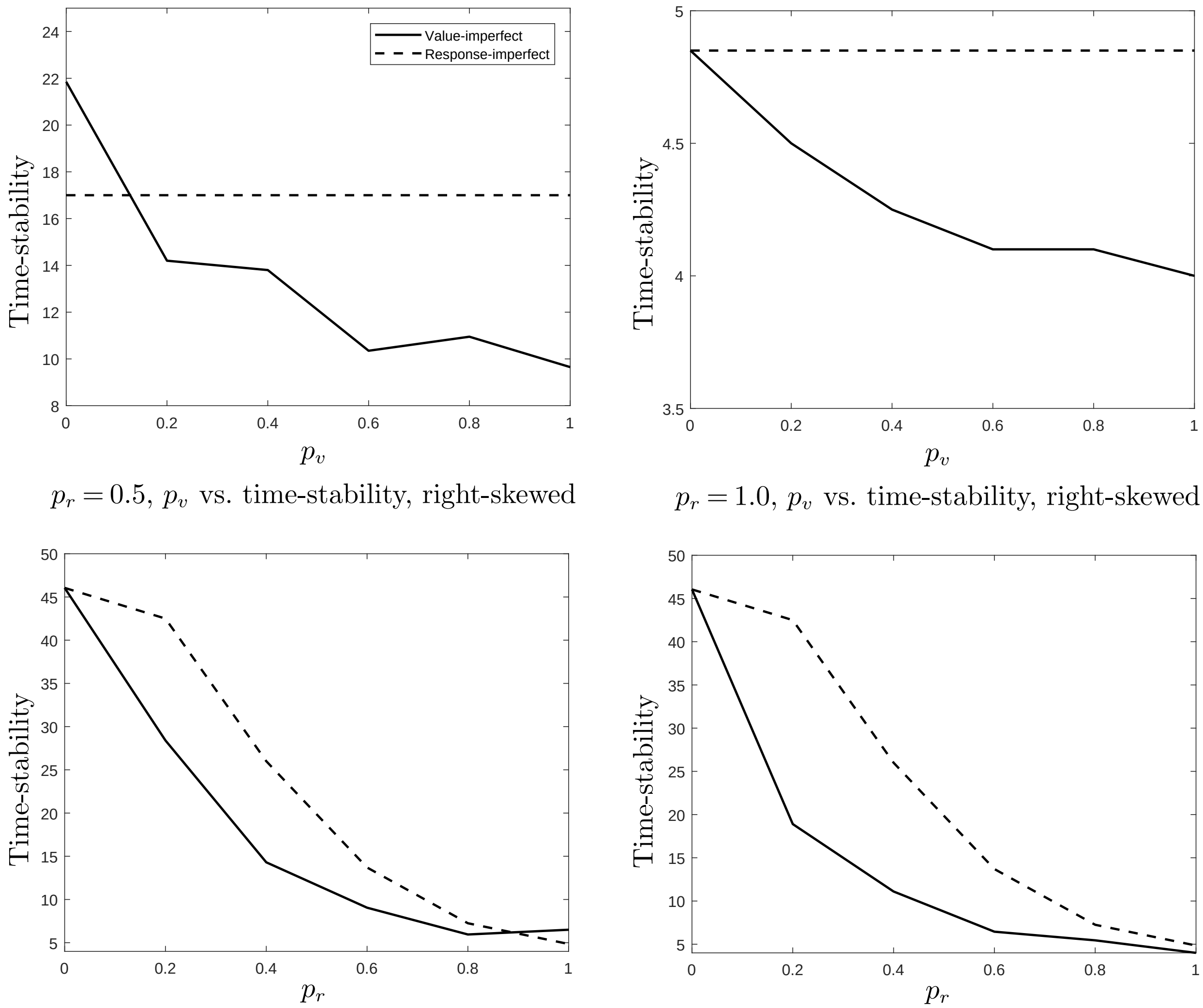
Notes: Entries in bold denote the best policy in each setting; the numbers in superscript of time-stability denote the number of instances out of 20 for which the corresponding policy failed to converge within T time periods.

Table 12 **Average time-stability and MAD (in parenthesis) for $\lambda \in \underline{\Lambda}$ and $\pi_M \in \Pi_M$ policies when information updates are used ($k=6$, $T=500$, Watts-Strogatz graphs, greedy evader).**

n	Right-skewed					Symmetric					Left-skewed					Random				
	λ	λ_r	λ_v	π_M	$\pi_M(+\mathcal{V})$	λ	λ_r	λ_v	π_M	$\pi_M(+\mathcal{V})$	λ	λ_r	λ_v	π_M	$\pi_M(+\mathcal{V})$	λ	λ_r	λ_v	π_M	$\pi_M(+\mathcal{V})$
15	218.6	36.6	**20.3**	215.4	251.0	500.0	66.3	**15.1**	74.2	123.0	500.0	82.1	**68.6**	500.0	157.0	500.0	89.8	**20.8**	364.7	337.6
	(211.2)	(11.0)	(26.5)	(243.9)	(249.0)	(0.0)	(8.5)	(3.6)	(121.7)	(170.0)	(0.0)	(4.8)	(79.5)	(0.0)	(196.0)	(0.0)	(9.1)	(8.4)	(193.3)	(208.8)
20	363.4	41.5	**32.8**	251.0	251.0	500.0	123.5	**18.8**	152.0	103.5	500.0	72.5	**48.9**	500.0	202.4	500.0	94.6	**53.0**	410.1	354.5
	(191.3)	(54.3)	(46.7)	(249.0)	(249.0)	(0.0)	(149.2)	(5.1)	(208.8)	(158.6)	(0.0)	(47.6)	(45.1)	(0.0)	(208.3)	(0.0)	(84.0)	(54.8)	(143.9)	(203.7)
25	347.2	43.1	**11.1**	251.0	251.0	500.0	59.5	**45.0**	220.9	61.1	500.0	61.6	**26.8**	500.0	143.3	500.0	58.4	**52.6**	476.8	333.3
	(200.3)	(45.8)	(4.1)	(249.0)	(249.0)	(0.0)	(45.9)	(45.5)	(231.2)	(87.8)	(0.0)	(21.4)	(4.9)	(0.0)	(156.3)	(0.0)	(22.3)	(44.7)	(44.2)	(216.7)
30	475.8	41.3	**9.9**	275.9	275.9	500.0	43.4	**19.7**	129.7	153.1	500.0	63.4	**26.7**	500.0	128.5	500.0	61.7	**28.2**	452.1	354.0
	(46.1)	(47.9)	(2.9)	(246.5)	(246.5)	(0.0)	(23.1)	(5.6)	(185.2)	(208.1)	(0.0)	(27.2)	(6.0)	(0.0)	(148.6)	(0.0)	(21.7)	(7.7)	(81.5)	(204.4)
35	428.9	21.0	**10.1**	350.6	375.5	500.0	61.4	**21.2**	257.1	179.4	500.0	72.8	**37.4**	500.0	180.8	500.0	50.5	**29.6**	479.6	253.3
	(121.0)	(10.7)	(4.0)	(209.2)	(186.8)	(0.0)	(44.1)	(5.1)	(242.9)	(224.5)	(0.0)	(34.3)	(9.1)	(0.0)	(191.5)	(0.0)	(18.9)	(7.2)	(38.9)	(233.8)
40	427.4	73.7	**60.1**	375.5	350.6	500.0	38.2	**25.6**	159.5	80.9	500.0	90.1	**56.4**	500.0	223.8	500.0	98.0	**75.9**	500.0	236.2
	(123.5)	(85.3)	(88.0)	(186.8)	(209.2)	(0.0)	(9.8)	(6.4)	(204.3)	(125.7)	(0.0)	(53.2)	(44.4)	(0.0)	(221.0)	(0.0)	(82.5)	(84.8)	(0.0)	(237.4)
45	377.9	46.7	**10.3**	350.6	350.6	500.0	65.6	**48.3**	304.0	83.2	500.0	93.6	**62.7**	500.0	271.1	500.0	94.9	**35.5**	500.0	288.1
	(183.2)	(47.0)	(3.4)	(209.2)	(209.2)	(0.0)	(48.8)	(45.4)	(235.3)	(125.0)	(0.0)	(53.7)	(45.8)	(0.0)	(228.9)	(0.0)	(49.4)	(6.7)	(0.0)	(233.1)
50	403.0	17.0	**11.5**	400.4	400.4	500.0	70.3	**21.4**	205.6	276.9	500.0	93.1	**34.0**	500.0	270.9	500.0	111.7	**31.2**	500.0	358.0
	(155.2)	(5.7)	(3.1)	(159.4)	(159.4)	(0.0)	(50.4)	(4.8)	(235.5)	(245.5)	(0.0)	(51.7)	(7.6)	(0.0)	(229.2)	(0.0)	(78.5)	(7.9)	(0.0)	(198.9)

Note. The numbers in superscript of time-stability denote the number of instances out of 20 for which the corresponding policy failed to converge within T time periods.

Figure 12 **Average time-stability for policies in $\underline{\Lambda}$ for different types of feedback as p_r and p_v increase for $k=6$ and the right-skewed costs ($T=50$, Watts-Strogatz graphs, greedy evader).**



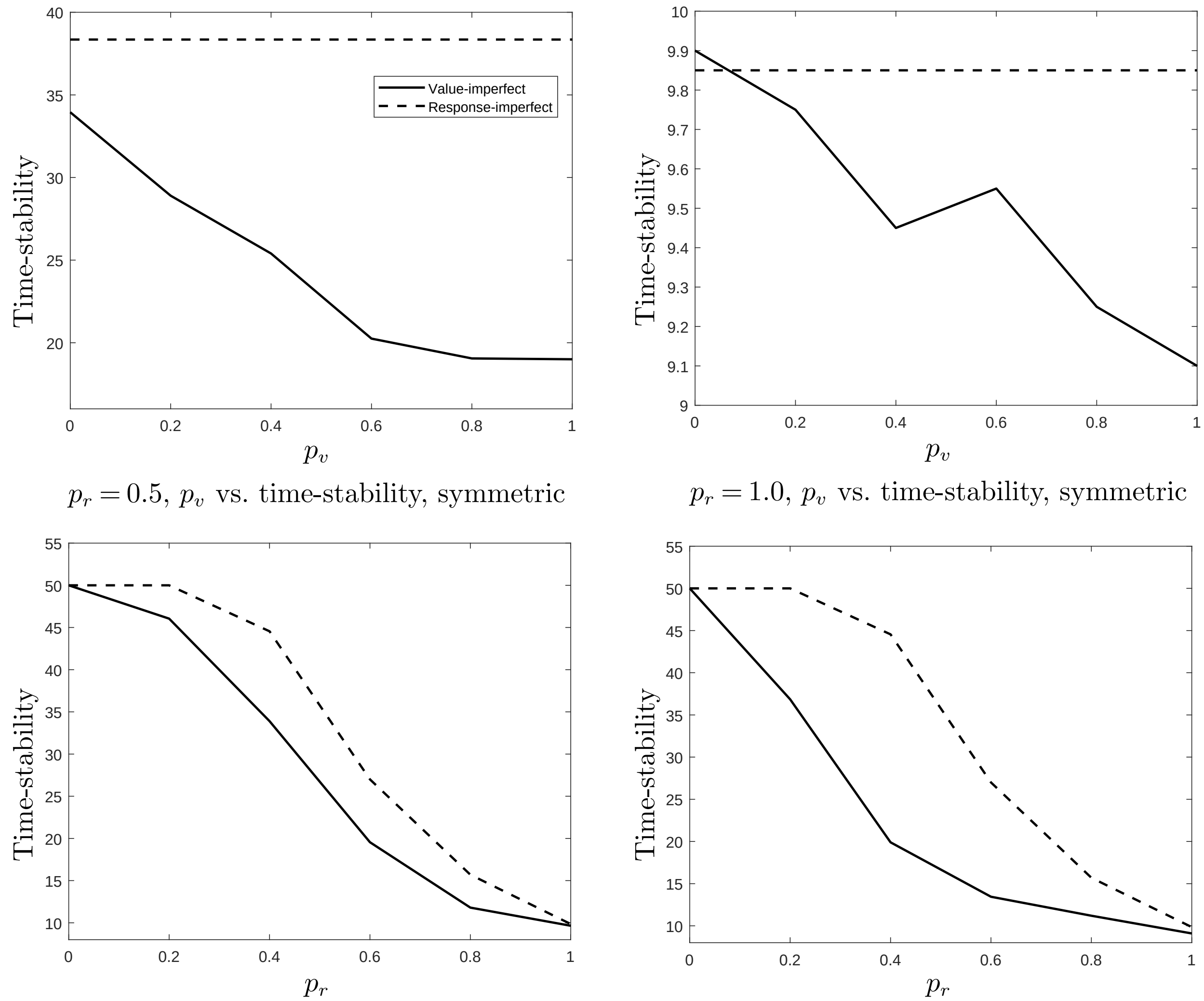
$p_r=0.5$, p_v vs. time-stability, right-skewed

$p_r=1.0$, p_v vs. time-stability, right-skewed

$p_v=0.5$, p_r vs. time-stability, right-skewed

$p_v=1.0$, p_r vs. time-stability, right-skewed

Figure 13 **Average time-stability for policies in $\underline{\Lambda}$ for different types of feedback as p_r and p_v increase for $k = 6$ and the symmetric costs ($T = 50$, Watts-Strogatz graphs, greedy evader).**



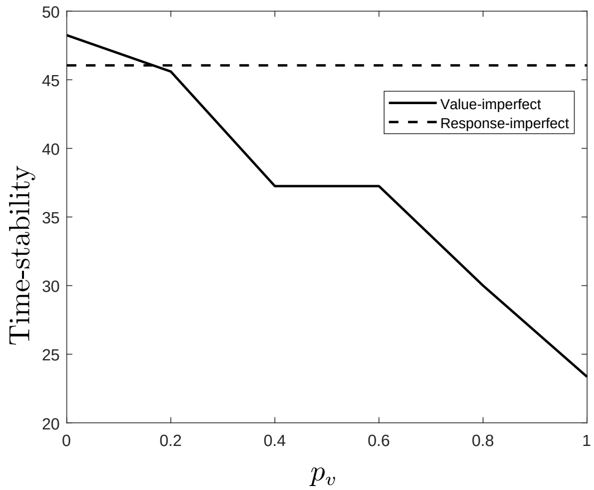
$p_r = 0.5$, p_v vs. time-stability, symmetric

$p_r = 1.0$, p_v vs. time-stability, symmetric

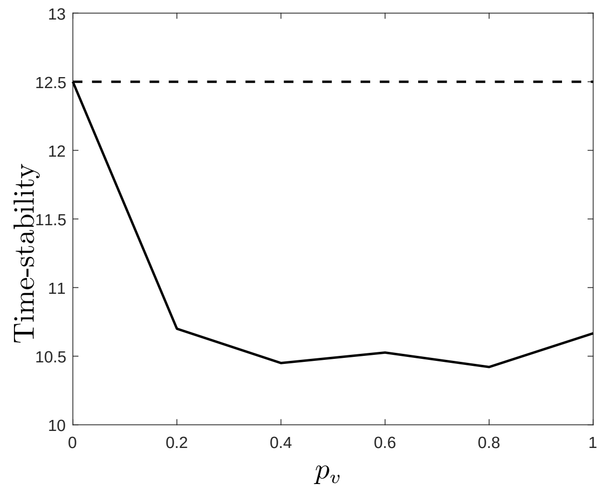
$p_v = 0.5$, p_r vs. time-stability, symmetric

$p_v = 1.0$, p_r vs. time-stability, symmetric

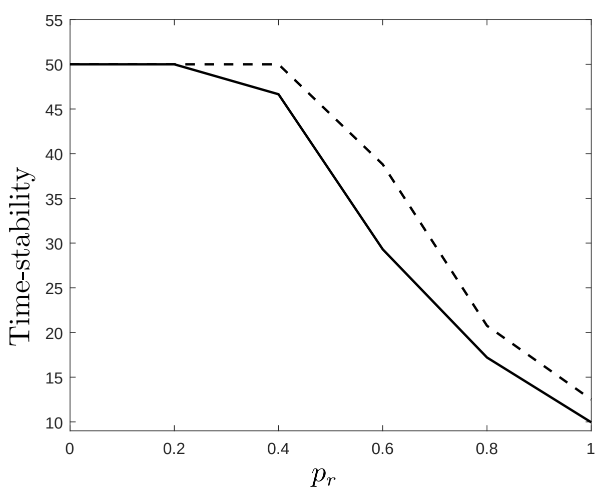
Figure 14 **Average time-stability for policies in $\underline{\Lambda}$ for different types of feedback as p_r and p_v increase for $k = 6$ and the left-skewed costs ($T = 50$, Watts-Strogatz graphs, greedy evader).**



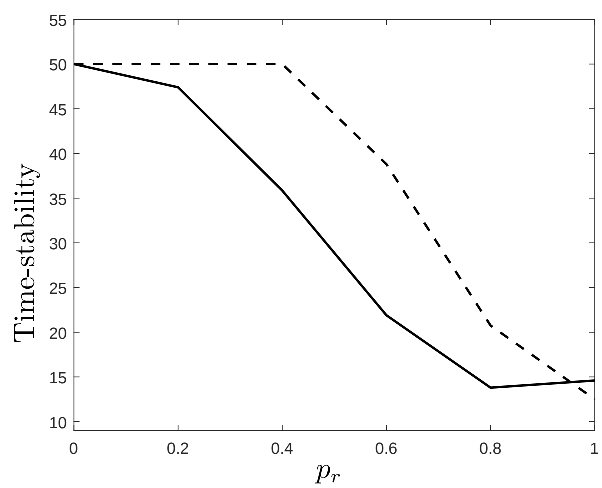
$p_r = 0.5$, p_v vs. time-stability, left-skewed



$p_r = 1.0$, p_v vs. time-stability, left-skewed



$p_v = 0.5$, p_r vs. time-stability, left-skewed



$p_v = 1.0$, p_r vs. time-stability, left-skewed

Table 13 **Behaviour of policies in $\underline{\Lambda}$ with respect to the cost bound quality ($k=6$, $T=200$, Watts-Strogatz graphs, greedy evader).**

	$p_r = p_v$		0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Policy λ_r	I.1	Time-stability	**105.5**	**77.2**	**31.6**	**16.3**	**12.4**	**6.4**	**5.8**	**3.6**	**3.6**	**2.7**	**2.5**
		MAD	70.1	65.3	17.3	7.7	7.8	3.0	2.5	1.4	1.4	0.8	0.6
		Total regret	**1572.6**	**909.9**	**390.3**	**236.4**	**151.6**	**19.6**	**18.0**	**10.7**	**6.7**	**7.1**	**5.2**
		MAD	2193.9	1306.2	600.4	338.4	217.5	27.0	25.3	15.4	9.1	10.1	7.1
											7.7		
	I.2	Time-stability	200.0	151.6	84.6	41.5	21.3	17.3	9.0	7.8	6.3	5.1	4.3
		MAD	0.0	48.2	44.2	21.7	6.5	12.3	2.7	2.4	1.6	1.2	0.8
		Total regret	10297.9	9986.2	5490.7	1881.5	832.8	852.7	350.9	268.7	202.4	158.5	126.2
		MAD	6342.5	7238.2	4813.5	1731.2	490.5	946.4	265.6	190.0	139.6	98.0	76.3
											7.7		
	I.3	Time-stability	200.0	200.0	138.9	81.6	46.2	29.7	22.6	20.1	11.6	8.1	7.0
		MAD	0.0	0.0	39.7	31.4	18.7	13.4	11.9	10.4	2.1	0.7	0.0
		Total regret	27668.4	20968.8	14129.3	7069.2	3950.6	2552.3	2177.5	1744.6	797.9	591.7	502.2
		MAD	10027.8	10213.1	5741.8	3264.7	1774.7	1528.0	1648.3	1211.4	253.4	173.5	129.0
Policy λ_v	I.1	Time-stability	**109.1**	**60.5**	**22.3**	**15.4**	**6.2**	**6.1**	**3.4**	**3.8**	**3.1**	**2.6**	**2.5**
		MAD	74.1	36.5	8.5	9.7	2.6	2.7	1.0	1.5	0.6	0.6	0.5
		Total regret	**2244.2**	**1163.8**	**46.1**	**23.3**	**15.4**	**13.3**	**9.0**	**10.0**	**5.7**	**5.2**	**5.2**
		MAD	3268.4	1874.0	64.4	32.9	22.0	19.0	12.8	13.9	7.7	7.1	7.1
											7.7		
	I.2	Time-stability	200.0	144.5	58.4	22.4	15.1	9.7	8.3	6.0	5.6	4.7	4.3
		MAD	0.0	38.7	28.6	7.7	6.2	3.8	3.1	1.6	1.4	1.1	0.7
		Total regret	10785.2	6594.4	3361.4	742.1	517.2	341.1	228.6	180.5	156.7	142.6	132.9
		MAD	6942.5	4936.1	3504.3	549.0	409.9	263.9	138.1	106.9	95.1	89.4	84.3
											7.7		
	I.3	Time-stability	200.0	198.0	108.6	68.7	29.1	20.3	15.0	11.7	10.4	8.4	7.3
		MAD	0.0	3.5	31.4	33.2	8.0	3.6	2.4	1.9	1.3	0.9	0.5
		Total regret	27200.2	20596.5	9600.6	5408.1	2264.0	1610.5	1214.1	844.4	753.8	578.3	513.7
		MAD	11068.9	6487.9	4343.4	2991.2	1072.0	372.5	414.4	300.8	235.1	133.8	129.9

Note. The numbers in superscript of time-stability denote the number of instances out of 20 for which the corresponding policy failed to converge within T time periods