

Supplementary Materials

VEHICLE REBALANCING MODEL

At each time step t , the model determines the rebalancing flows $\{y_{i,j,o}^t\}$ for operator o that minimize the total rebalancing cost while satisfying the desired vehicle distribution specified by the actor network. The optimization problem is formulated as:

$$\min \sum_{(i,j) \in \mathcal{E}} c_{i,j,o}^t y_{i,j,o}^t \quad (1)$$

$$\text{s.t.} \quad \sum_{j \neq i} (y_{j,i,o}^t - y_{i,j,o}^t) + m_{i,o}^t \geq \tilde{m}_{i,o}^t, \quad i \in \mathcal{V} \quad (2)$$

$$\sum_{j \neq i} y_{i,j,o}^t \leq m_{i,o}^t, \quad i \in \mathcal{V} \quad (3)$$

$$y_{i,j,o}^t \geq 0, \quad (i,j) \in \mathcal{E} \quad (4)$$

where $\tilde{m}_{i,o}^t$ denotes the number of desired vehicles at region i for operator o at time t . Objective (1) minimizes the rebalancing cost. Constraint (2) ensures that the desired vehicle number is satisfied, accounting for the current idle vehicles $m_{i,o}^t$ and the net inflow of rebalanced vehicles. Constraint (3) limits the rebalancing flow from each region by the number of available idle vehicles. The desired vehicle distribution is calculated by $\tilde{m}_{i,o}^t = \lfloor w_{i,o}^t \sum_{i \in \mathcal{V}} m_{i,o}^t \rfloor$, where $w_{i,o}^t$ is the rebalancing weight output by the actor network for region i . Note that the constraint matrix of this network flow problem is totally unimodular, and since both $m_{i,o}^t$ and $\tilde{m}_{i,o}^t$ are integer-valued, the optimal solution is guaranteed to be integral.

PARAMETERS USED IN TRAINING

Category	Hyperparameter	Value
Training	Actor learning rate (α_π)	2×10^{-4}
	Critic learning rate (α_V)	4×10^{-4}
	Discount factor (γ)	0.97
	Reward scaling factor	4,000
	Actor gradient clip	1,000
	Critic gradient clip	1,000
	Critic warmup episodes	50
	Training episodes	150,000
Network	Hidden layer size	256
	Look-ahead horizon (T)	6
	Scale factor	0.01
Price Scalars	Observe OD-prices	Yes
	OD-price scalars	No

CHOICE MODEL SENSITIVITY

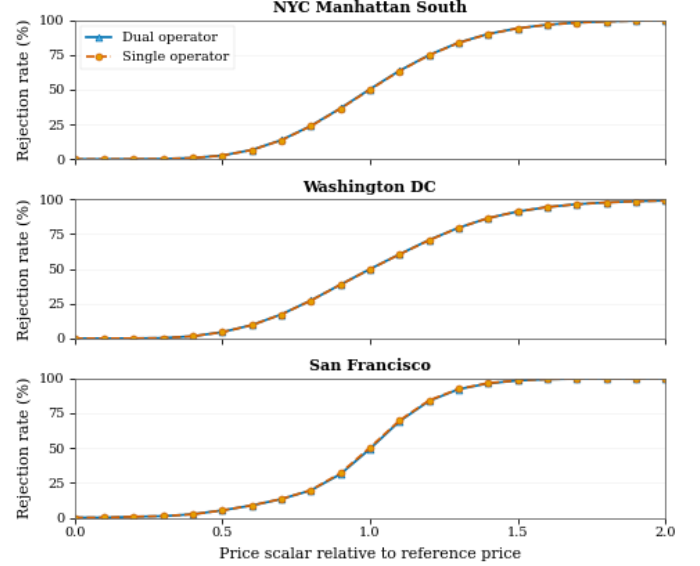


Fig. 1: Rejection rate versus price scalar relative to the historical reference price across studied datasets for both single and dual-operator setups, with the model calibrated to a 50% rejection rate at the historical reference price.

THREE-STEP CONTROL ARCHITECTURE

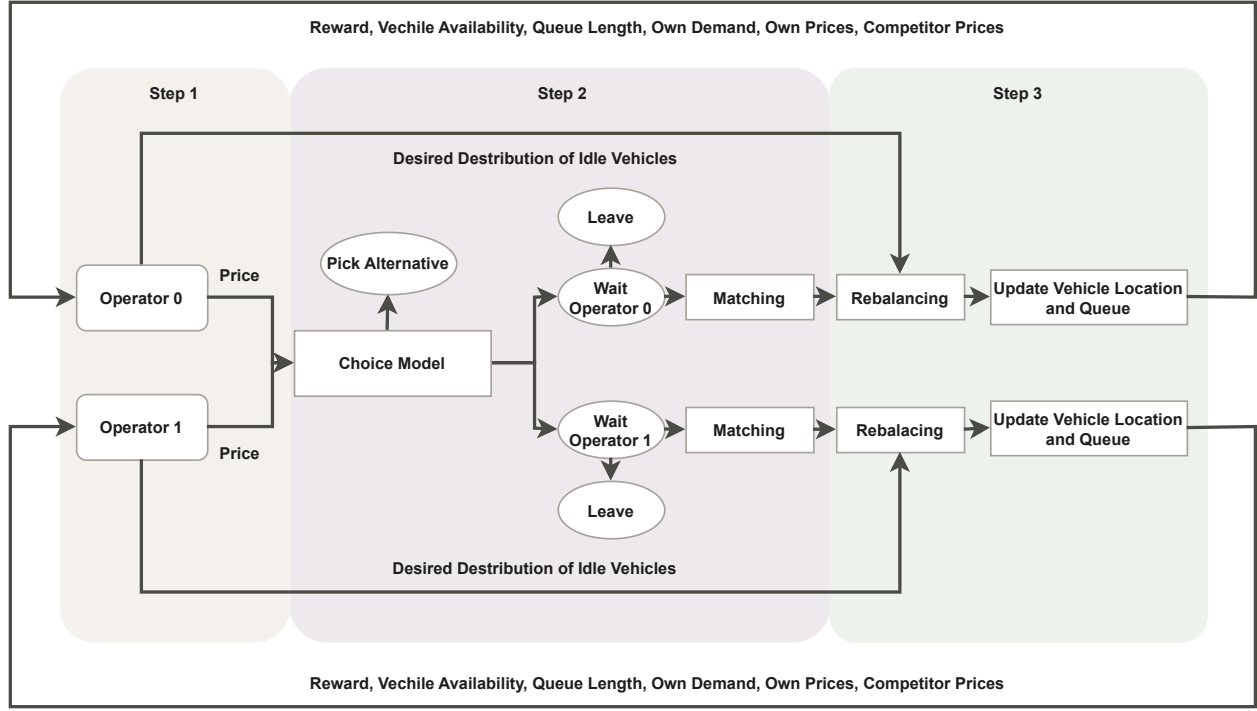


Fig. 2: Three-step control architecture for dual-operator AMoD control. Step 1: operators formulate pricing and desired idle-vehicle distribution policies. Step 2: passenger assignment via choice model, queueing, and matching. Step 3: idle-vehicle rebalancing and update of vehicle positions and queues.

OPERATOR ARCHITECTURE

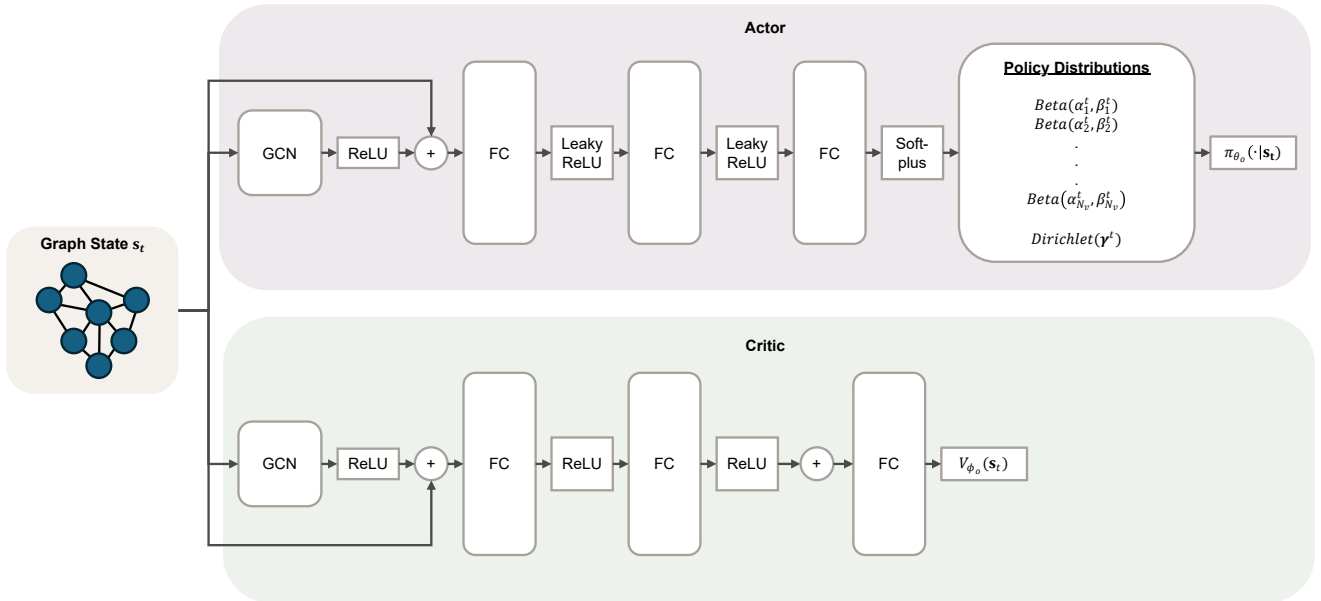


Fig. 3: The Actor-Critic architecture employed by the operators. Each operator maintains independent actor and critic networks.

PRICING SCALARS UNDER PRICING-ONLY POLICY NYC

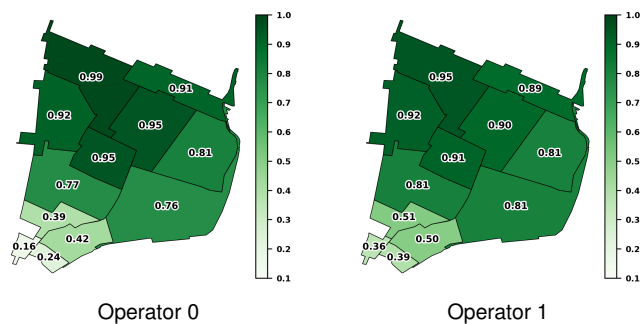


Fig. 4: Initial pricing policies at timestep 0 for the pricing-only policy.

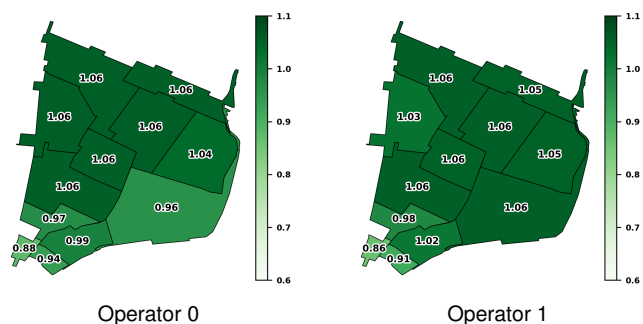


Fig. 5: Final pricing policies at time step 19 for the pricing-only policy.

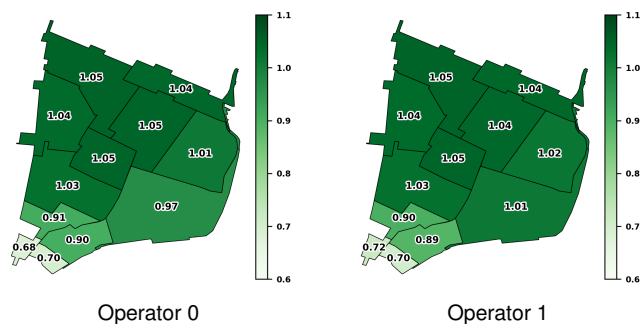


Fig. 6: Time-averaged pricing scalars for the pricing-only policy, computed across all 20 time steps.

FLEET SIZE SENSITIVITY

Fleet Size	Joint Policy						NC			UD				
	Reward	Served	Rebal. Costs	Rebal. Trips	Price A0	Price A1	Reward	Served	Price Both	Reward	Served	Rebal. Costs	Rebal. Trips	Price Both
450	14170.5 (251.7)	2563.9 (28.4)	865.6 (66.9)	177.3 (16.0)	1.04 (0.00)	1.04 (0.00)	12270.7 (251.8)	2550.6 (15.9)	1.00 (0.00)	13217.4 (153.7)	2550.6 (15.9)	989.0 (70.0)	200.6 (16.6)	1.00 (0.00)
650	17481.1 (258.7)	3561.1 (29.0)	1379.0 (83.7)	284.7 (19.0)	0.96 (0.00)	0.96 (0.00)	16048.4 (416.6)	2894.2 (73.9)	1.00 (0.00)	17652.0 (277.5)	3496.5 (32.2)	1865.1 (98.3)	387.9 (19.6)	1.00 (0.00)
850	19513.1 (288.1)	4538.7 (38.9)	1871.0 (79.9)	386.3 (16.7)	0.88 (0.00)	0.88 (0.00)	18683.9 (461.8)	3367.4 (83.3)	1.00 (0.00)	19941.6 (334.8)	4171.5 (44.7)	3401.8 (128.6)	761.5 (27.2)	1.00 (0.00)
1050	20373.0 (198.7)	5512.9 (30.1)	2152.1 (77.1)	446.8 (19.7)	0.79 (0.00)	0.79 (0.00)	20667.8 (467.6)	3714.9 (83.3)	1.00 (0.00)	20450.0 (543.4)	4538.1 (72.0)	5016.9 (155.3)	1187.7 (41.2)	1.00 (0.00)
1250	20305.6 (235.2)	6440.1 (32.2)	2509.8 (117.9)	526.9 (29.9)	0.71 (0.00)	0.70 (0.00)	21955.0 (397.2)	3934.1 (71.2)	1.00 (0.00)	19798.1 (526.1)	4639.2 (77.2)	6259.8 (192.8)	1501.7 (46.1)	1.00 (0.00)

FLEET SPLIT SENSITIVITY

Fleet Split (O0:O1)	Total Reward	O0 Reward	O1 Reward	Total Rebal. Trips	O0 Rebal. Trips	O1 Rebal. Trips	O0 Price	O1 Price	Total Served	O0 Served	O1 Served
5:5	17481.1 (258.7)	8717.0 (249.1)	8764.1 (223.5)	284.7 (19.0)	144.9 (17.8)	139.8 (14.3)	0.96 (0.00)	0.96 (0.00)	3561.1 (29.0)	1773.9 (26.1)	1787.2 (24.0)
4:6	17567.0 (231.0)	7295.2 (136.2)	10271.7 (243.7)	279.2 (17.8)	116.1 (12.6)	163.1 (14.6)	0.98 (0.00)	0.95 (0.00)	3577.2 (27.9)	1437.2 (14.7)	2140.0 (30.1)
3:7	17434.2 (241.7)	5772.6 (143.7)	11661.6 (259.2)	285.3 (19.8)	79.6 (11.5)	205.7 (18.9)	1.01 (0.00)	0.94 (0.00)	3562.5 (32.1)	1090.1 (17.4)	2472.4 (33.0)
2:8	16968.2 (263.3)	4140.0 (88.8)	12828.2 (244.2)	294.5 (21.2)	41.3 (4.9)	253.2 (19.8)	1.04 (0.00)	0.93 (0.00)	3530.7 (33.5)	742.3 (13.6)	2788.4 (31.9)
1:9	16655.4 (291.8)	2348.7 (77.9)	14306.7 (299.7)	311.6 (22.3)	14.1 (3.0)	297.5 (22.9)	1.10 (0.00)	0.93 (0.00)	3499.9 (35.8)	380.7 (8.7)	3119.2 (38.1)

IMPACT OF PRICING INFORMATION

Policy	No Information Sharing			Information Sharing		
	Reb.	Pricing	Joint	Reb.	Pricing	Joint
Total Reward	18096.6 (296.6)	18983.6 (267.4)	17689.4 (242.4)	18133.7 (330.8)	18631.6 (376.7)	17481.1 (258.7)
Reward Operator 0	8981.1 (227.9)	9413.0 (111.9)	8951.6 (212.4)	9010.8 (242.6)	9351.6 (161.4)	8717.0 (249.1)
Reward Operator 1	9115.5 (317.0)	9570.6 (192.9)	8737.8 (190.5)	9123.0 (341.4)	9280.1 (265.1)	8764.1 (223.5)
Total Rebalancing Costs	1544.4 (110.7)	—	1349.7 (76.5)	1600.2 (95.4)	—	1379.0 (83.7)
Rebal. Costs Operator 0	806.0 (78.6)	—	662.5 (65.8)	813.5 (73.6)	—	699.6 (76.7)
Rebal. Costs Operator 1	738.5 (103.5)	—	687.1 (59.2)	786.8 (107.0)	—	679.4 (63.1)
Total Rebalance Trips	327.8 (25.4)	—	276.5 (16.8)	338.4 (21.2)	—	284.7 (19.0)
Rebal. Trips Operator 0	172.0 (20.6)	—	134.7 (15.3)	173.5 (18.2)	—	144.9 (17.8)
Rebal. Trips Operator 1	155.8 (23.6)	—	141.8 (12.4)	164.9 (24.2)	—	139.8 (14.3)
Total Served Demand	3521.9 (34.2)	3388.9 (38.1)	3579.5 (29.0)	3535.8 (39.4)	3337.6 (52.3)	3561.1 (29.0)
Served Demand Operator 0	1751.8 (26.1)	1678.1 (13.8)	1800.4 (25.4)	1758.1 (29.4)	1674.3 (19.1)	1773.9 (26.1)
Served Demand Operator 1	1770.1 (36.8)	1710.8 (28.8)	1779.1 (21.0)	1777.7 (39.5)	1663.3 (38.7)	1787.2 (24.0)
Price Operator 0	—	0.96 (0.00)	0.96 (0.00)	—	0.95 (0.00)	0.96 (0.00)
Price Operator 1	—	0.95 (0.01)	0.96 (0.00)	—	0.95 (0.01)	0.96 (0.00)
Wait/mins Operator 0	1.82 (0.08)	1.34 (0.06)	1.97 (0.10)	1.80 (0.08)	1.51 (0.08)	2.03 (0.10)
Wait/mins Operator 1	1.81 (0.07)	1.37 (0.07)	2.06 (0.09)	1.78 (0.07)	1.45 (0.08)	2.06 (0.08)
Queue Operator 0	8.37 (1.09)	4.35 (0.70)	9.89 (1.12)	8.41 (1.12)	5.69 (0.73)	10.4 (1.3)
Queue Operator 1	7.78 (1.21)	4.62 (0.80)	10.7 (1.4)	7.75 (1.05)	4.96 (0.67)	10.8 (1.2)
Total Demand	4683.0 (73.9)	4202.3 (55.2)	5185.7 (58.2)	4683.0 (73.9)	4311.7 (65.1)	5231.3 (61.6)
Demand Operator 0	2336.2 (56.0)	2075.3 (34.3)	2552.0 (52.5)	2336.2 (56.0)	2145.4 (36.7)	2581.1 (59.5)
Demand Operator 1	2346.8 (39.4)	2127.0 (31.8)	2633.7 (34.4)	2346.8 (39.4)	2166.3 (34.7)	2650.2 (42.1)

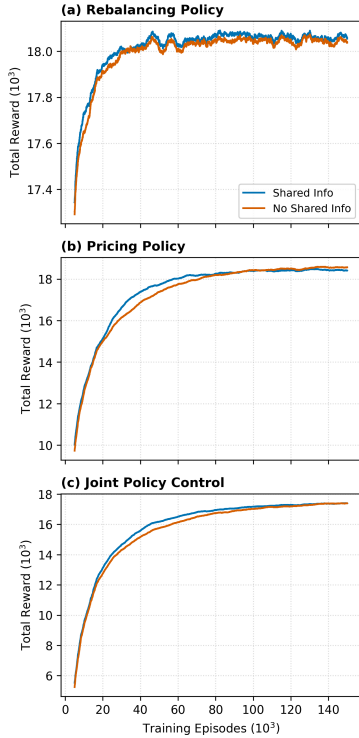


Fig. 7: Convergence dynamics comparing scenarios with and without competitor price visibility: (a) Rebalancing, (b) Pricing, (c) Joint. Curves are smoothed over 30 episodes; the first 5,000 episodes are excluded for clarity. Note that rewards are training rewards based on sampling from the policy distributions and may therefore be lower than test rewards.

REGIONAL WAGE HETEROGENEITY

Metric	Joint Policy
Total Reward	23564.6 (341.1)
Reward Operator 0	11659.3 (206.4)
Reward Operator 1	11905.3 (248.4)
Total Rebalancing Costs	2661.3 (87.2)
Rebalancing Costs Operator 0	1359.8 (52.2)
Rebalancing Costs Operator 1	1301.5 (57.0)
Total Rebalance Trips	609.7 (21.0)
Rebalance Trips Operator 0	315.0 (10.9)
Rebalance Trips Operator 1	294.7 (15.0)
Total Served Demand	2990.9 (28.1)
Served Demand Operator 0	1480.6 (18.9)
Served Demand Operator 1	1510.3 (21.1)
Price Operator 0	1.32 (0.00)
Price Operator 1	1.31 (0.00)
Wait/mins Operator 0	1.42 (0.07)
Wait/mins Operator 1	1.46 (0.09)
Queue Operator 0	6.40 (1.01)
Queue Operator 1	5.97 (1.34)
Total Demand	3567.0 (46.7)
Demand Operator 0	1762.8 (24.5)
Demand Operator 1	1804.2 (33.7)
Average Wage	25.8 (0.1)

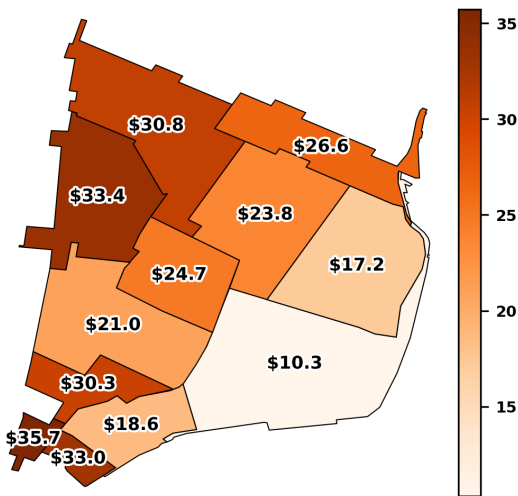


Fig. 8: Average hourly passenger wage distribution across regions in NYC Manhattan South under regional income heterogeneity.

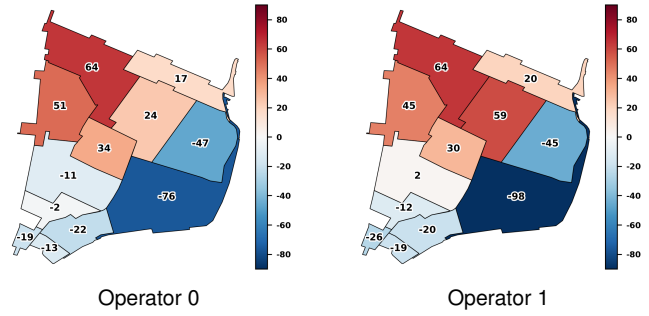


Fig. 9: Net rebalancing flows under regional income heterogeneity, showing cumulative vehicle movements across all time steps. Red = net receiver, blue = net sender.

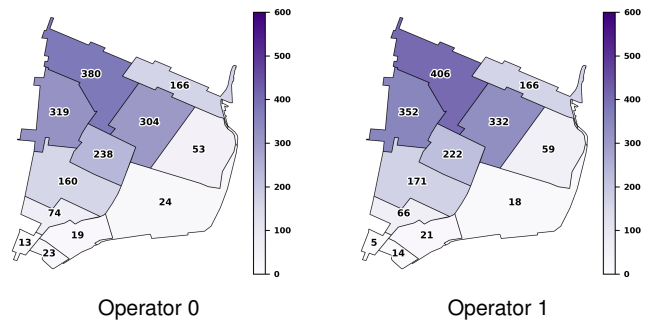


Fig. 10: Total demand originating from each region under regional income heterogeneity.

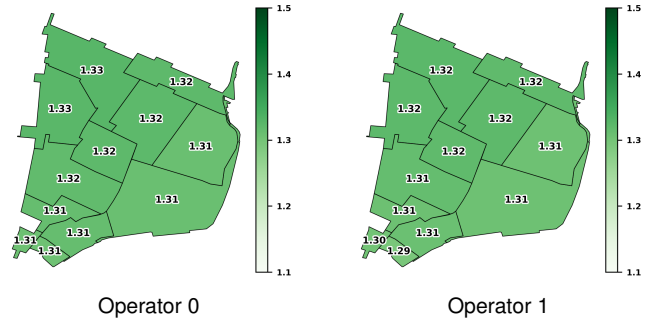


Fig. 11: Average pricing scalars per region across all time steps under regional income heterogeneity.