

The figure is a line plot titled 'Average Return vs. Iterations'. The x-axis is labeled 'Iterations' and ranges from 0 to 10,000. The y-axis is labeled 'Average Return' and ranges from 0.7 to 1.0. The plot displays the performance of various policies over time. The 'Random Policy' (blue line) starts at an average return of approximately 0.75 and remains relatively stable. The 'Warm Start' policies (orange, green, red, purple, brown, pink, grey, yellow) all start at an average return of approximately 0.75 and show a rapid increase in performance, eventually plateauing at higher returns. The 'Warm Start (70%)' policy (yellow line) achieves the highest final average return of approximately 0.95, while the 'Random Policy' (blue line) achieves the lowest final average return of approximately 0.75.

Policy	Initial Average Return (Iter 0)	Final Average Return (Iter 10,000)
Random Policy	~0.75	~0.75
Warm Start (0%)	~0.75	~0.85
Warm Start (10%)	~0.75	~0.88
Warm Start (20%)	~0.75	~0.90
Warm Start (30%)	~0.75	~0.92
Warm Start (40%)	~0.75	~0.93
Warm Start (50%)	~0.75	~0.94
Warm Start (60%)	~0.75	~0.95
Warm Start (70%)	~0.75	~0.95

