Day 8: Herding Analysis v1

When does herding begin?

Herding begins after the initial stabilization phase, once volatility collapses and agent strategies become aligned.
In the simulation, the mean pairwise correlation of agent positions rises sharply from near zero values to persistently high levels (0.9 - 1.0) shortly after the early transient period. This transition occurs well before the external shock window, indicating that herding is not shock-induced but endogenously formed.
This marks a regime shift from heterogeneous, independent trading to coordinated behavior.

What agent type triggers herding?

Herding is primarily triggered by Trend Followers, with reinforcement from Market Maker volatility adaptation.
Trend Followers share similar decision rules (SMA deviation – based trading)
As price trends stabilize, many Trend Followers receive the same signal at the same time, leading to synchronized market orders.
Market Makers respond to reduced volatility by tightening spreads, lowering execution friction, which further amplifies coordinated behavior.
Noise Traders provide volume but do not initiate alignment and the RL agent adapts to the herd (instead of the causes).
Thus, herding emerges from strategy similarity and feedback through price impact and not from centralized control.

Is herding reversible?

Herding is weakly reversible, but structurally persistent.
Short lived drops in correlation occur during sudden price adjustments or volatility spikes.
However, correlations quickly revert to high levels once volatility decays again.
Even during the external shock window, the system returns to a herded state rather than fragmenting.
This indicates that once agents lock into a shared behavioral regime, temporary disturbances are insufficient to break coordination unless agent incentives or decision rules change.