information

MDPI

Article

Is the Taiwan Stock Market (Swarm) Intelligent?

Ren-Raw Chen [1,2]

[1] Gabelli School of Business, Fordham University, 113 W 60th Street, New York, NY 10023, USA; rchen@fordham.edu
[2] College of Management, Chang Gung University, 259, Wenhua 1st Rd., Guishan Dist., Taoyuan City 33302, Taiwan

Abstract: It is well-believed that most trading activities tend to herd. Herding is an important topic in finance. It implies a violation of efficient markets and hence, suggests possibly predictable trading profits. However, it is hard to test such a hypothesis using aggregated data (as in the literature). In this paper, we obtain a proprietary data set that contains detailed trading information, and as a result, for the first time it allows us to validate this hypothesis. The data set contains all trades transacted in 2019 by all the brokers/dealers across all locations in Taiwan of all the equities (stocks, warrants, and ETFs). Given such data, in this paper, we use swarm intelligence to identify such herding behavior. In particular, we use two versions of swarm intelligence—Boids and PSO (particle swarm optimization)—to study the herding behavior. Our results indicate weak swarm among brokers/dealers.

Keywords: swarm intelligence; Taiwan stock market; herding

1. Introduction

It is well-believed that most trading activities tend to herd. Herding is an important topic in finance (see a recent work by Mavruk [1], who adopts various machine learning algorithms). Ukpong, Tan, and Yarovaya [2] and Rahayu et al. [3] provide an excellent review of herding in the financial markets (An earlier review can be found in Bikhchandani and Sharma [4]). However, it is hard to test such a hypothesis using aggregated data (as in the literature). In this paper, we obtain a propriety data set that contains detailed trading information of herding, and as a result, for the first time it allows us to validate this hypothesis. The data set contains all trades transacted by all the brokers/dealers across all locations in Taiwan of all the equities (stocks, warrants, and ETFs).

Technical analysts follow patterns drawn from prices and volume and develop various trading indices. Other than the popular indices like 50-day and 200-day moving averages, there are also sophisticated ones like on-balance volume, stochastic oscillator, and relative strength index, among numerous others. No matter what indices these analysts follow, they are not based upon any fundamental information of the companies. Instead, they are based upon behaviors of the market participants—known as herding.

For pretty much the entire history of studying stock returns, such indices have been ridiculed and looked down upon by economists. Traditional economists criticize such trading indicators as being lacking theory, completely ad-hoc, and highly judgmental (i.e., different analysts can come to different conclusions from the same data) . . . until now. However, we believe that all those technical indicators, although not supported by any economic theory, are supported by psychological behaviors of market participants—herding.

Not just in stock markets, people in general herd. Often, we vote for a political candidate not because how much we know about this person, or how much we are familiar with his past record, but because our friends vote for him. We listen to our friends' suggestions to make decisions all the time, without doing actual research ourselves. This is herding.

check for updates

Citation: Chen, R.-R. Is the Taiwan Stock Market (Swarm) Intelligent? *Information* **2024**, *15*, 707. https://doi.org/10.3390/info15110707

Academic Editor: Heming Jia

Received: 23 August 2024
Revised: 26 October 2024
Accepted: 29 October 2024
Published: 5 November 2024

Not only humans herd, but animals herd as well. Ants and bees rely on one another to find food and avoid prey. Computer scientists observe this and translate their behavior into an algorithm so that we can solve complex problems related to herding—swarm intelligence.

In this paper, we use swarm intelligence to study the herding behavior in stock markets. This is the first time an artificial intelligence model is applied to herding in financial markets. Swarm intelligence is perfectly suitable for herding in that herding is regarded as a psychological behavior that is unrelated to firms' fundamentals. In order to test if herding exists, we estimate the hyperparameters in the swarm model. Closed-form solutions are derived for special cases, while the general case can be solved numerically.

The rest of the paper is organized as follows. In the next section, we discuss the basic idea and math of swarm intelligence. In Section 3, we also derive the closed-form formulas for the parameters in swarm. We then do not need to numerically solve for the parameters, as many machine learning models do. Section 4 is the main section of the paper. We present data and empirical results. In Section 5, we discuss the major shortcoming of this paper—noise, insufficient granularity, and length of data. And the paper is concluded.

2. A Quick Glance on Herding

Herding is the behavior of individuals in a group acting collectively without centralized direction. Herding is originally observed in animals in herds, packs, bird flocks, fish schools, and so on, as well as in humans. Shiller [5] seems to be the first author who studied herding in the financial market. He describes investors influenced by their peers in making investment decisions as follows:

> "Investing in speculative assets is a social activity. Investors spend a substantial part of their leisure time discussing investments, reading about investments, or gossiping about others' successes or failures in investing. It is thus plausible that investors' behavior (and hence, prices of speculative assets) would be influenced by social movements."

As we can all agree, although humans are highly intelligent, they can behave foolishly when they are in a crowd. Hence, with no surprise, the literature has predominantly regarded herding as an irrational behavior. This is because herding does not generate better returns (see Mavruk [1] for a review of such results). The literature documents the following reasons for herding, all of which demonstrate a certain behavioral bias (See Mavruk [1] for further details):

- fads
- fear
- greed
- reputation
- noise

All of the above causes can easily result in market disturbances and destabilization. Hence, herding can move the market away from its fundamentals. The conclusion that herding is irrational is consistent with the common wisdom raised by Shiller [5].

As a result, it can be expected that herding is not persistent. It happens sometimes, ceases to happen sometimes, and furthermore contradicts itself sometimes. This is extremely similar to the performances of technical analyses that herding is time-varying and situation-dependent. For this reason, a major effort in the literature is to identify the determinants of herding. Summarized nicely by Rahayu et al. [3], herding is stronger (Please see the citations of relevant papers in Rahayu [3]):

- when volatility is higher
- in a crisis
- for small stocks
- when there are large price movements
- in a declining market

- when rates rise
- in a poor information environment

It is obvious that such situations are not sustainable situations. They happen only temporarily and are naturally not persistent.

Herding is in general classified as stock herding or investor herding. As their names suggest, the former refers to how investors rush in and out of a stock, and the latter refers to how investors follow their peers. Understandably, the methodology to study both types of herding is the same, while the difference is in data. For stock herding, the data are usually more price related (i.e., in order to measure trading profits) and more frequent, while for investor herding, the data are more shares related (i.e., investors' holdings) and less frequent.

The measures of herding vary. Yet they can be traced back to Lakonishok, Shleifer, and Vishney [6], who define herding as the difference between the actual change in "purchase intensity" and the expected change of "purchase intensity" (For other measures of herding, see Bikhchandani and Sharma [4] and Mavruk [1] for their surveys):

$$|p_{t+1} - p_t| - E[|p_{t+1} - p_t|]$$

where p_t is purchase intensity. While this definition reflects common wisdom, finding a good proxy for purchase intensity is a challenge. Various authors, given data availability and research focus, use different proxies. Lakonishok, Shleifer, and Vishney [6] use changes in share holdings. While this is a sensible proxy for purchase intensity, such data are only available at the institutional level and only four times a year (i.e., quarterly). As a result, not only does it fail to measure more frequent herding, which is how herding can meaningfully impact trading profits, it also fails to measure retail trading, which is believed to be more interesting (because it is more tightly related to trading profits). In other words, their results are limited to fund managers and are in low frequency. Strictly speaking, it is a result of how information is transferred among fund managers, not herding, which is more understood as trading related.

The other extreme is to use very frequent data (daily), and yet the quality of the proxy is lowered. For example, Ukpong et al. [2] use excess returns (the difference between a stock's return and the market return) as the proxy. Contrary to Lakonishok, Shleifer, and Vishny [6], they capture the dynamics of herding more perfectly and yet the measurement errors in their results are larger. Using swarm, we are free of the above criticisms in that we directly measure how individual investors follow or not their peers.

A small portion of the literature argues that herding can be rational, mainly caused by information asymmetry. Some individuals in a group have superior information than others. As one can understand, such evidence can only exist in investor herding (not stock herding) and is only limited to institutional investors. These institutions mimic one another due to information asymmetry, which is distinctly different from and hard to reconcile with the aforementioned reasoning of herding. In swarm intelligence, however, these two types of herding can be easily integrated. The former is the usual notion of swarm via separation, alignment, and cohesion (see the next section for details), and the latter is leader following (either via a landscape or a predefined leader or a group of leaders).

Evidence on herding also indicates anti-herding—known as contrarian. Similar to the standard contrarian notion, anti-herding refers to a situation where certain investors tend to diverge from the rest of the group. Again, this can be captured more faithfully by swarm intelligence via alignment, cohesion, and separation.

Some researchers relate herding to momentum (The pioneering academic research on momentum investing can be traced back to Jegadeesh and Titman [7]. They documented how strategies of buying recent stock winners and selling recent losers generated significantly higher near-term returns than the U.S. market overall from 1965 to 1989. They established the basic time frame for momentum-investing success as a 3-to-12-month window on either side. Since then, it has been booming into one of the largest research areas in finance. Recently, they (Jegadeesh and Titman [8]) wrote an excellent review piece of the past 30 years of momentum research). As the first authors systemically documenting herding, Grinblatt, itman, and Wermers [9] define herding as how a group of investors move in and out of the market simultaneously—like a herd. They do not study why they herd. Recently, Demirer, Lien, and Zhang [10] evaluate the impact of industry herding on return momentum. They find that the profitability of industry momentum strategies depends on the level of herding in an industry. Lin, Wu, and Zhang [11] investigate the impact of herding behavior on the momentum effect. Using a new firm-level herding measurement, they find that investors require higher returns in high herding stocks, and they require even higher returns in high herding stocks among previous losers, indicating that investors herd against the previous losers while they herd toward the winners. Chen [12], using intra-day volume data of 2016 over 62 countries, finds that uninformed country-level herding is highly related to momentum (Note that although the data used by Chen (2021) are intra-day, he aggregates information to daily).

Finally, herding is observed internationally. Besides Chen's work on 62 countries, Rahayu et al. [3] provide an excellent review of literature. They have documented strong evidence of herding internationally (over 20 nations).

3. Swarm Intelligence

Wikipedia describes swarm intelligence as "the collective behavior of decentralized, self-organized systems". The basic idea of swarm intelligence is derived from those animals (such as birds, ants, bees, and fish) that rely on group effort to achieve their basic survival needs—seek food and avoid prey. The intelligence behind this collective behavior is how they communicate among one another.

There are two versions of swarm intelligence. They are related and yet applied differently on different problems. The first is Boids, and the second is particle swarm optimization. In an Appendix A, we provide a simple numerical, step-by-step example with five fish and two dimensions. The example demonstrates how migration of fish in each iteration is calculated.

3.1. Boids

Reynolds [13] was the first to "artificialize" such natural intelligence and create a computer algorithm named Boids (for bird-oid object). Reynolds' algorithm is amazingly simple. For any given bird, Reynolds devises a set of linear equations (vectors) combining which determines how the bird should fly to its next destination.

> The factors that determine how various vectors are combined are: separation, alignment, and cohesion. As their names suggest, "separation" is to avoid collision with other birds, "alignment" decides how a particular bird should fly in a direction by referencing to its fellow birds, and "cohesion" decides how fast (speed) a particular bird should fly to the center of its fellow birds.

There are countless versions of Boids. One can add obstacles. One can add an objective destination (swim to target). One can do Boids in a maze. The basic Boids as described in Figure 1 can be described by the following algorithm.

In the swarm model, let $i = 1, \cdots, m$ be firm (bird) and $j = 1, \cdots, n$ be stock (dimension). At each given point in time, a map of the locations of bird is taken. The map is a 15-dimensional hypercube image with each axis bounded between 0 and 1.

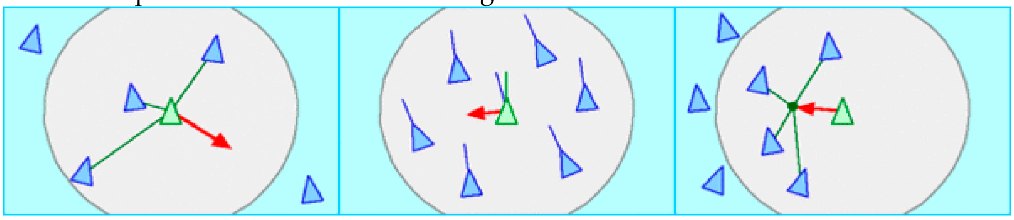


Figure 1. Swarm Intelligence (source: https://en.wikipedia.org/wiki/Boids, accessed on 28 October 2024). The green triangle is the target fish, while the blue triangles are its neighbors. The circle defines how far the target fish makes its references. Lines (within the circle) show how the target fish is connected to its neighbors. The red arrow then is the final vector the target fish will move.

In the swarm model, let $\vec{x}_t^{(i)}$ be a vector of n coordinates (dimensions) where $x_{j,t}^{(i)}$ and $j = 1, \cdots, n$ is an element. In other words, $\vec{x}_t^{(i)}$ is the position of bird i at time t. For the sake of completeness, define F as a flock of birds $\{f^{(1)}, \cdots, f^{(m)}\}$ whose positions at time t are collected in set $X_t = \{\vec{x}_t^{(1)}, \cdots, \vec{x}_t^{(m)}\}$. In other words, X_t is an $n \times m$ matrix. The velocity of each bird is defined as a weighted average of various forces:

$$\vec{v}_t^{(i)} = w_A \vec{v}_{A,t}^{(i)} + w_C \vec{v}_{C,t}^{(i)} + w_S \vec{v}_{S,t}^{(i)} \tag{1}$$

where weights sum to 1 and

$$\begin{aligned} \vec{v}_{A,t}^{(i)} &= \text{avg}\left(\vec{v}_{t-1}^{(j \neq i)} \middle| f^{(j)} \in F\right) - \vec{v}_{t-1}^{(i)} \\ \vec{v}_{C,t}^{(i)} &= \text{avg}\left(\vec{x}_{t-1}^{(j \neq i)} \middle| f^{(j)} \in F\right) - \vec{x}_{t-1}^{(i)} \\ \vec{v}_{S,t}^{(i)} &= \max\left\{\left|\vec{x}_t^{(i)} - \vec{x}_{t-1}^{(i)}\right|, \varepsilon\right\} \end{aligned} \tag{2}$$

representing alignment, cohesion, and separation respectively.

The update of position is:

$$\vec{x}_t^{(i)} = \vec{x}_{t-1}^{(i)} + \vec{v}_t^{(i)} \tag{3}$$

In terms of data, each position is a snapshot of the locations of all birds at a given time. From one snapshot to another, it represents a migration (i.e., velocity).

As it can be easily seen, the main purpose of the swarm here is to describe how birds move. For example, if cohesion is dominant, then eventually all birds will line up in a straight line. Similarly, if alignment is dominant, then they all want to move in the same direction parallelly. Finally, if separation is dominant, then they all move randomly. Since there is no stopping time, these birds will keep moving indefinitely.

3.2. Particle Swarm Optimization

Particle swarm optimization (PSO), from its name, is an optimization tool using swarm (See Eberhart and Kennedy [14] and Shi and Eberhart [15]). In PSO, an objective function (or penalty function) is given so birds can all reach the optimal location. The communication mechanism among the birds is the same as Boids, and yet how they move is different.

One can think of PSO is swarm with a landscape (i.e., the objective function). Birds now will try to reach either the peak or bottom (i.e., global optimum) of the landscape. In this case, they will stop moving once their objective is met.

In order to achieve convergence, the velocity of a PSO is given as:

$$\vec{v}_t^{(i)} = w_t \vec{v}_{t-1}^{(i)} + c_1 r_1 (\vec{p}_{t-1}^{(i)} - \vec{x}_{t-1}^{(i)}) + c_2 r_2 (\vec{g}_{t-1} - \vec{x}_{t-1}^{(i)}) \tag{4}$$

where $w_t < 1$ is a decaying weight (e.g., we can let $w_t = \delta^t$ and $\delta < 1$), c_1 and c_2 are two constants, r_1 and r_2 are two random variables, and

$$\vec{p}_t^{(i)} = \left\{ \vec{x}_{\tau \le t}^{(i)} \middle| \max_{\tau} \phi\left(\vec{x}_{\tau}^{(i)}\right) \right\} \text{where } \phi(\cdot)\text{isthefitnessfunction.} \tag{5}$$

is the personal best and $\vec{g}_t = \max_i \vec{p}_t^{(i)}$ is the global best.

In (4), c_1 and c_2 are learning rates. These are standard parameters in artificial intelligence to attain most effective learning. The two random variables r_1 and r_2 are "exploration", meaning that the birds do not follow what they "learn" exactly. This is key to artificial intelligence to avoid local optima. Finally, birds follow the instruction from $\vec{g}_{t-1} - \vec{x}_{t-1}^{(i)}$ and $\vec{p}_{t-1}^{(i)} - \vec{x}_{t-1}^{(i)}$ which are regarded as "exploitation" since they would like to learn from given information.

The update of position is the same as (3) and repeated here:

$$\vec{x}_t^{(i)} = \vec{x}_{t-1}^{(i)} + \vec{v}_t^{(i)} \tag{6}$$

A graphical depiction of the particle swarm optimization is given in Figure 2 (with the replacement of $\vec{p}_t^{(i)}$ by $\hat{x}_i(t)$).

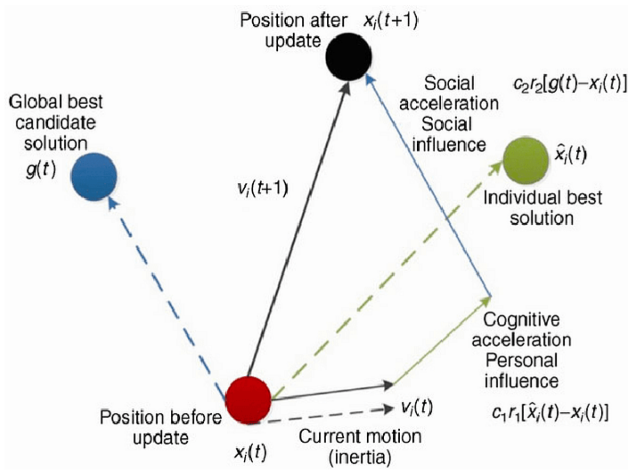


Figure 2. Particle Swarm Optimization (source: https://medium.com/analytics-vidhya/implementing-particle-swarm-optimization-pso-algorithm-in-python-9efc2eb179a6).

Note that swarm is very flexible, and we can put exploration separately (i.e., not mingled with exploitation). In this paper, we alter (4) as follows:

$$\vec{v}_t^{(i)} = w_L \vec{v}_{L,t}^{(i)} + w_P \vec{v}_{P,t}^{(i)} \tag{4a}$$

where $w_{P,t} + w_{L,t} = 1$ and

$$\vec{v}_{P,t}^{(i)} = (1 - \alpha_P^{(i)}) \vec{u}_t^{(i)} + \alpha_P^{(i)} (\vec{p}_{t-1}^{(i)} - \vec{x}_{t-1}^{(i)}) \tag{7}$$

and

$$\vec{v}_{L,t}^{(i)} = \alpha_L (\vec{g}_{t-1} - \vec{x}_{t-1}^{(i)}) \text{ or } = \alpha_L^{(i)} (\vec{g}_{t-1} - \vec{x}_{t-1}^{(i)}) \tag{8}$$

Note that in this paper, we do not have a landscape. The global best (which is the best position of the landscape) is replaced by a "leader", who is the one with the position $\vec{g}_t$. As a result, there is no convergence or stopping time in our empirical work.

3.3. Swarm Intelligence and Herding

Swarm intelligence is particularly suitable for studying human herding in that humans, although highly intelligent, behave rather foolishly (i.e., by instinct) when they are in a group. They tend to follow their peers in taking certain actions as opposed to relying on their own intelligence (which includes independent research, judgment, and analyses). In trading stocks, the finance literature has long documented herding in human behavior. It is rather obvious if we just look around how people choose their investment targets. Very few actually use their intelligence but rather make hasty decisions based upon the suggestions of their peers. It is not just trading; we also observe a similar phenomenon in political elections. Reference groups are essential to how people vote. It is this observation that motivates this research. As a result, instead of using the traditional methods (i.e., statistical methods), which mainly rely upon covariances to study herding (Regressions in the herding literature are based upon covariances (between dependent and explanatory variables)), it is more natural and common-sense to use swarm to study herding. In swarm, herding is not determined by covariances of the performances of stocks. Rather, it detects if a trader actually follows the other traders.

As seen in the above subsections, the hyperparameters of a swarm determine its behavior (and in turn, determines the performance of the market). In a Boid, alignment, cohesion, and separation (and leader-following) are the main parameters. Alignment is how an investor follows the same trend of his peers, cohesion is how an investor would like to hold the same position of stocks as his peers, and finally, separation is how an investor would like to be a contrarian (anti-herding). Although one can add other behavioral parameters (e.g., leader-following, or grouping), Reynolds [13] argues that these are the main parameters that a swarm needs. Indeed, the "social movements" recognized by Shiller [5] are quite consistent with Reynolds' design of swarm. Any individual either wants to "keep up with the Joneses" (i.e., cohesion), takes a successful story and learns from it (i.e., follows its path, alignment), or takes a failure story and avoids its path (i.e., separation). In other words, Reynolds' swarm is a direct model for an individual who wants to herd. The statistical methods used in herding are at best an indirect detection of herding, while swarm intelligence is a direct recognition of herding.

Swarm intelligence has been widely used in scientific research (which is not relevant to this paper). (Note that there is a wide variety of different swarm models (in both boids and particle swarm optimization)). In finance, unfortunately, there have been very limited number of applications. Recent work includes Chen [16] on stock picking, Chen and Behrndt [17] on commodity options, Chen et al. [18] on stock options, Chen, Huang, and Yeh [19] on portfolio construction, and Chen, Miller and Toh [20,21] on firm search (See Huang [22] for a survey).

4. Empirical Methodology

We can estimate the parameters of swarm (Equation (1)) using data. Empirical (i.e., data) positions can be labeled as $\vec{\xi}_t^{(i)}$ and velocity as $\vec{\nu}_t^{(i)}$ (to substitute for $\vec{x}_t^{(i)}$ and $\vec{v}_t^{(i)}$ in (3)). Hence, similar to (3):

$$\vec{\nu}_t^{(i)} = \vec{\xi}_t^{(i)} - \vec{\xi}_{t-1}^{(i)} \tag{9}$$

Our objective function is to minimize the sum of squared errors between $\vec{\nu}_t^{(i)}$ and $\vec{v}_t^{(i)}$:

$$\min_{\alpha_{P,t}^{(i)}} (\vec{v}_t^{(i)} - \vec{\nu}_t^{(i)})'(\vec{v}_t^{(i)} - \vec{\nu}_t^{(i)}) = \min_{\alpha_{P,t}^{(i)}} \sum_{j=1}^{n} (v_{j,t}^{(i)} - \nu_{j,t}^{(i)})^2 \tag{10}$$

Taking partial derivative and setting it to 0:

$$\sum (v_{j,t}^{(i)} - \nu_{j,t}^{(i)}) \frac{\partial v_{j,t}^{(i)}}{\partial \theta} = 0 \tag{11}$$

where θ is a chosen parameter.

Clearly, (11) in general has no closed-form solution and needs to be solved numerically. However, if we focus on one parameter at a time (i.e., holding other parameters constant), then there is a closed-form solution, which is what we implement in the empirical section.

4.1. Boids

Rewrite (1) slightly as follows:

$$\begin{aligned}\vec{v}_t^{(i)} &= w_{A,t}^{(i)}\vec{v}_{A,t}^{(i)} + w_{C,t}^{(i)}\vec{v}_{C,t}^{(i)} + (1 - w_{A,t}^{(i)} - w_{C,t}^{(i)})\vec{v}_{S,t}^{(i)} \\ &= w_{A,t}^{(i)}\left(\vec{v}_{A,t}^{(i)} - \vec{v}_{S,t}^{(i)}\right) + w_{C,t}^{(i)}\left(\vec{v}_{C,t}^{(i)} - \vec{v}_{S,t}^{(i)}\right) + \vec{v}_{S,t}^{(i)} \\ &= w_{A,t}^{(i)}\vec{A}_t^{(i)} + w_{C,t}^{(i)}\vec{C}_t^{(i)} + \vec{S}_t^{(i)}\end{aligned} \tag{1a}$$

Then we follow (11) and take partial derivatives with respect to alignment w_A and cohesion w_C parameters, respectively. This leads to a simultaneous equation system (the derivation is in Appendix C), and the solution is:

$$\begin{bmatrix} w_{A,t}^{(i)} \\ w_{C,t}^{(i)} \end{bmatrix} = \frac{1}{x_{11}x_{22} - x_{12}^2}\begin{bmatrix} x_{22} & -x_{12} \\ -x_{12} & x_{11} \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \tag{12}$$

where (note that i, j, and t are dropped from x for easy expression)

$x_{AC} = \sum_{j=1}^{n} A_{j,t}^{(i)} C_{j,t}^{(i)}$,

$y_X = \sum_{j=1}^{n} \nu_{j,t}^{(i)} X_{j,t}^{(i)} - \sum_{j=1}^{n} X_{j,t}^{(i)} S_{j,t}^{(i)}$ (in which X is either A or C),

And finally, $\nu_{j,t}^{(i)} = \xi_{j,t}^{(i)} - \xi_{j,t-1}^{(i)}$ is velocity computed from data.

For the reason that is clear later, we also run the estimation without separation. In that case, (1) can be rewritten as:

$$\vec{v}_t^{(i)} = w_t^{(i)}\vec{v}_{A,t}^{(i)} + (1 - w_t^{(i)})\vec{v}_{C,t}^{(i)} \tag{13}$$

Similar to the process in solving (12), we can solve for the weight (details in Appendix C) as follows:

$$w_t^{(i)} = \frac{\sum_{j=1}^{n}(\nu_{j,t}^{(i)} - v_{C,j,t}^{(i)})(v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)})}{\sum_{j=1}^{n}(v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)})^2} \tag{14}$$

Note that in this case, we force the brokers/dealers to swarm. If in reality brokers/dealers do not swarm, then the weight should be random and present no patterns.

4.2. Particle Swarm Optimization

We have several tests (for our hypotheses). The first is to see if a broker/dealer explores. In this test, we rewrite (7) as follows ($w_P = 1$):

$$\begin{aligned}\vec{v}_t^{(i)} &= \vec{v}_{P,t}^{(i)} \\ &= (1 - \alpha_P^{(i)})\vec{u}_t^{(i)} + \alpha_P^{(i)}(\vec{p}_t^{(i)} - \vec{x}_t^{(i)})\end{aligned} \tag{7a}$$

From the Appendix, we solve the coefficient analytically as:

$$\alpha_{P,t}^{(i)} = \frac{\sum_{j=1}^{n}(\nu_{j,t}^{(i)} - u_{j,t}^{(i)})(p_{j,t}^{(i)} - x_{j,t}^{(i)} - u_{j,t}^{(i)})}{\sum_{j=1}^{n}(p_{j,t}^{(i)} - x_{j,t}^{(i)} - u_{j,t}^{(i)})^2} \tag{15}$$

Alternatively, we could let $w_L = 1$ and

$$\begin{aligned} \vec{v}^{(i)}_{L,t} &= \alpha^{(i)}_{L,t}(\vec{g}_{t-1} - \vec{x}^{(i)}_{t-1}) \\ \vec{v}^{(i)}_{L,t} &= \alpha_{L,t}(\vec{g}_{t-1} - \vec{x}^{(i)}_{t-1}) \end{aligned} \tag{16}$$

where the former is each firm has its own $\alpha^{(i)}_{L,t}$ and the latter is all firms share the same $\alpha_{L,t}$. The solution (see Appendix B) is:

$$\alpha^{(i)}_{L,t} = \frac{\sum_{j=1}^{n} \nu^{(i)}_{j,t}(g_{j,t} - x^{(i)}_{j,t})}{\sum_{j=1}^{n} (g_{j,t} - x^{(i)}_{j,t})^2} \tag{17}$$

and the average across individual brokers/dealers is:

$$\begin{aligned} \overline{\alpha}_{L,t} &= \frac{1}{m}\sum_{i=1}^{m} \alpha^{(i)}_{L,t} \\ &= \frac{1}{m}\sum_{i=1}^{m} \frac{\sum_{j=1}^{n} \nu^{(i)}_{j,t}(g_{j,t} - x^{(i)}_{j,t})}{\sum_{j=1}^{n} (g_{j,t} - x^{(i)}_{j,t})^2} \end{aligned} \tag{18}$$

This can be compared to the other solutions where all firms adopt the exact same parameter value, which is:

$$\alpha_{L,t} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} \nu^{(i)}_{j,t}(g_{j,t} - x^{(i)}_{j,t})}{\sum_{i=1}^{m}\sum_{j=1}^{n} (g_{j,t} - x^{(i)}_{j,t})^2} \tag{19}$$

Finally, we can estimate both parameters jointly (see Appendix B).

$$\vec{v}^{(i)}_{t} = w^{(i)}_t \vec{v}^{(i)}_{P,t} + (1 - w^{(i)}_t)\vec{v}^{(i)}_{L,t} \tag{20}$$

However, we end up with three simultaneous equations:

$$\begin{aligned} &\sum_{j=1}^{n}\left\{ w^{(i)}_t u^{(i)}_{j,t} z^{(i)}_{j,t} + w_t \alpha^{(i)}_{P,t}(z^{(i)}_{j,t})^2 + (1 - w^{(i)}_t) v^{(i)}_{L,j,t} z^{(i)}_{j,t} - \nu_{j,t} z^{(i)}_{j,t} \right\} = 0 \\ &\sum_{j=1}^{n}[w^{(i)}_t v^{(i)}_{P,j,t} + (1 - w^{(i)}_t)\alpha^{(i)}_{L,t}(g_{j,t-1} - x^{(i)}_{j,t-1}) - \nu_{j,t}][g_{j,t-1} - x^{(i)}_{j,t-1}] = 0 \\ &\sum_{j=1}^{n}(v^{(i)}_{t-1} - \nu^{(i)}_{t-1})\left\{ (u^{(i)}_{t-1} + \alpha^{(i)}_{P,t}(p^{(i)}_{t-1} - x^{(i)}_{t-1} - u^{(i)}_{t-1})) - \alpha^{(i)}_{L,t}(g_{j,t} - x^{(i)}_{t-1}) \right\} = 0 \end{aligned} \tag{21}$$

Instead of solving the system of simultaneous equations, in the empirical work, we simplify the problem by holding the weight constant (0.5 and 0.6) and then solve for $\alpha^{(i)}_{L,t}$ and $\alpha^{(i)}_{P,t}$ respectively. See Appendix B for details.

5. Empirical Findings

As mentioned in the introduction, herding is widely observed in financial markets, yet it has been only studied with low-frequency data. Furthermore, such empirical evidence is indirect and could be subject to large measurement errors (due to proxies for herding). In this study, a benefit from a proprietary data set, we are for the first time able to examine at the broker/dealer level the herding behavior. We use two swarm intelligence models to examine such a behavior.

5.1. Data

We use a proprietary data obtained from HiHedge. (We are highly grateful for the CEO of HiHedge, Dr. Gu Jiaqi, for generously providing the data at no cost). The data contain the whole year of 2019 of all the trading activities (prices and volumes of buy, sell, and trade) of all the locations of all the securities firms in Taiwan. The data therefore include all (80) broker/dealer trades in all (294) locations in Taiwan. Consequently, the data contain a total 205,678,058 transactions. To my knowledge, such granularity of data has not been possible in the literature of swarm intelligence.

Each transaction is labeled as buy or sell, its price (NT$), and its volume (shares). The data do not contain time stamps, and hence, they are aggregated within a day. In 2019, there are a total of 223 trading days. In other words, for the same broker/dealer and location, all buy transactions and sell transactions (separately) are summed up into one transaction within a day.

The data include stocks, warrants, and ETFs. In this study, we limit our focus on only TSE (967) stocks. In particular, we first focus on only the top 20 stocks, which account for 50% of the market size in TSE (Note that TSMC (#2330) alone accounts for over 25% of the TSE). The distribution is given in Figure 3.

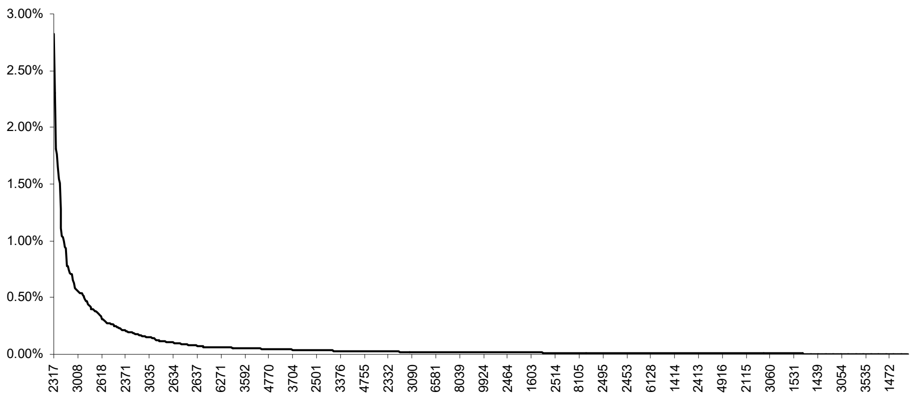


Figure 3. Stock Weights (excluding TSMC #2330).

As we can see, the size of companies becomes exponentially smaller. For this reason, in our empirical work, we study swarm using all stocks as well as the top 20 stocks.

The top 20 stocks are listed in Table 1. These companies allocate across 11 industries (financial 6, semi-conductor 4, plastic 2, auto 1, energy 1, telecom 1, steel 1, electronic parts 1, computer 1, food 1, other electronics 1.

Table 1. Trading Volume of the Top 20 Firms in TSE.

Rank	Ticker	Name	Name	Share wt	Dollar wt	Industry	Industry
1	2330	台積電	TSMC	5.31%	0.59%	semi-conductor	半導體
2	3406	玉晶光	Genius Electronic Optical	2.70%	0.20%	photoelectric	光電
3	2317	鴻海	Hon Hai Precision	1.98%	0.72%	other electronics	其他電子
4	2327	國巨	Yageo Electronics	1.95%	0.18%	electronics	電子零組件
5	3008	大立光	Largan Precision	1.90%	0.01%	photoelectric	光電
6	2454	聯發科	MediaTek	1.75%	0.15%	semi-conductor	半導體
7	3150	穩懋	WIN Semiconductors	1.70%	0.22%	semi-conductor	半導體
8	2492	華新科	Walsin Technology	1.66%	2.92%	electronic parts	電子零組件
9	6488	環球晶	Global Wafer Corporation	1.50%	0.24%	semi-conductor	半導體
10	2337	旺宏	Macronix	1.45%	0.13%	semi-conductor	半導體
11	6462	神盾	Egis Technology	1.03%	1.07%	semi-conductor	半導體
12	2023	欣興	Unimicron	1.02%	0.13%	electronic parts	電子零組件
13	3034	聯詠	Novatek Microelectronics	0.99%	0.80%	semi-conductor	半導體
14	2474	可成	Catcher Technology	0.88%	0.14%	other electronics	其他電子

Table 1. *Cont.*

Rank	Ticker	Name	Name	Share wt	Dollar wt	Industry	Industry
15	2408	南亞科	Nanya Technology	0.84%	0.10%	semi-conductor	半導體
16	3324	雙鴻	Auras Technology	0.81%	0.55%	other electronics	其他電子
17	2308	台達電	Delta Electronics	0.79%	0.34%	electronic parts	電子零組件
18	2345	智邦	Accton Technology	0.76%	0.14%	Telecommunication	通信網路
19	2313	華通	Compeq Manufacturing	0.65%	0.13%	electronic parts	電子零組件
20	2379	瑞昱	Realtek	0.64%	0.13%	semi-conductor	半導體

As we can see, TSMC has the most share trading volume (5.31%), and yet Hon Hai Precision has the most dollar trading volume (0.72%). However, Hon Hai Precision is ranked #3 in share volume (1.98%), which is only one-third of TSMC.

Both TSMC and Hon Hai Precision are Taiwan's most valuable companies (TSMC (ADR) is also traded on the NYSE under the ticker TSM. The market cap as of May 24 is $826 billion. In 2019, the market cap of TSM was roughly $220 billion). TSMC is the world's largest chip manufacturer (According to SemiWiki, TSMC occupies 28% of the chip market, followed by Samsung of 10%), and Hon Hai Precision (who owns Foxconn in China), is the most important manufacturer for Apple's iPhones. Hence, we also provide the list of the top 20 stocks by market capitalization in Table 2.

Table 2. Market Caps of the Top 20 Firms in TSE.

Rank	Ticker	Name	Name	Weight	Industry	Industry
1	2330	台積電	TSMC	26.58%	semi-conductor	半導體
2	2317	鴻海	Hon Hai Precision	2.83%	other electronics	其他電子
3	2454	聯發科	MediaTek	2.30%	semi-conductor	半導體
4	2382	廣達	Quanta Group	1.81%	computer	電腦週邊
5	2412	中華電	China Telecom	1.76%	telecom	通訊網路
6	2308	台達電	Delta Electronics	1.65%	electronic parts	電子零組件
7	2881	富邦金	Fubon Financial	1.55%	financial	金融業
8	6505	台塑化	Formosa Petrochemical	1.50%	energy	油電燃氣
9	2882	國泰金	Cathay Financial	1.28%	financial	金融業
10	2303	聯電	United Microelectronics	1.11%	semi-conductor	半導體
11	2886	兆豐金	Mega Financial	1.04%	financial	金融業
12	1303	南亞	Nan Ya Plastics	1.04%	plastic	塑膠
13	1301	台塑	Formosa Plastics	1.00%	plastic	塑膠
14	2891	中信金	China Trust	0.94%	financial	金融業
15	3711	日月光投控	ASE Technology	0.94%	semi-conductor	半導體
16	1216	統一	Uni-President Enterprises	0.78%	food	食品
17	2002	中鋼	China Steel	0.78%	steel	鋼鐵
18	2884	玉山金	E.sun Financial	0.74%	financial	金融業
19	5880	合庫金	Taiwan Cooperative Financial	0.74%	financial	金融業
20	2207	和泰車	Hotai Motor	0.72%	auto	汽車

Now we can see that TSMC has a dominant market share in TSE of 26.58%, with Hon Hai Precision being the second of 2.83% (not counting Foxconn in China). For our study, we use Table 1 for our top 20 firms, so we are internally consistent.

Among the total of 80 brokers/dealers, three brokers/dealers are data error (i.e., no such brokers/dealers in Taiwan), and two brokers/dealers trade only futures and options but are mistakenly listed as stockbrokers/dealers. They are removed from our study. As a result, we are left with 75 stockbrokers/dealers, which are given in Table 3.

Table 3. List of all Brokers/Dealers in Taiwan.

1	合庫	Taiwan Coop Bank	26	新百王	NHK Securities	51	華南永昌	Hua Nan Sec's
2	土銀	Land Bank	27	光和	Kuanz Ho Sec's	52	富邦	Fuban Bank
3	臺銀證券	Bank Taiwan Sec's	28	永全	Yun Chuan Sec's *	53	元大	Yuanta Comm Bank
4	臺銀	Bank of Taiwan	29	大昌	Dah Chang Sec's	54	永豐金	SinoPac Sec's
5	台灣企銀	Taiwan Bsns Bank	30	德信	Reliance Sec's	55	日進	J Zin Sec's
6	臺灣企銀	Taiwan Bsns Bank	31	福勝	Fushan Sec's	56	聯邦商銀	Union Bank
7	日盛	Jih Sun Bank	32	兆豐	Mega Int'l Bank	57	奔亞證券	Primasia Sec's
8	彰銀	Chang Hwa Bank	33	致和	Z Ho Securities *	58	台灣匯立	Hui Li Sec's *
9	宏遠	Horizon Securities	34	豐農	Feng Nun Sec's *	59	美林	Merrill Lynch
10	港商麥格理	Macquarie Group	35	石橋	Bridge Stone Sec's *	60	港商野村	Nomura Sec's
11	台灣摩根士丹利	Morgan Stanley	36	金港	Kovack Securities	61	富隆	Fullong Sec's
12	亞東	Oriental Securities	37	北城	Pei Cheng Sec's	62	寶盛	Monex Boom Sec's
13	大展	Tachan Securities	38	國票	Waterland Sec's	63	福邦	Grand Fortune Sec's
14	大慶	Ta Ching Sec's	39	台新	Taishin Int'l Bank	64	萬泰	KGI Bank
15	高橋	Mizuho Securities	40	安泰	Entie Comm Bank	65	中農	Chung Nourn Sec's
16	第一金	First Financial	41	摩根大通	JP Morgan	66	全泰	Chuan Tai Sec's *
17	永興	YS Securities	42	康和	Concord Int'l Sec's	67	瑞士信貸	Credit Swiss
18	統一	President Sec's	43	新光	Shin Kong Bank	68	大鼎	Da-Din Securities
19	盈溢	I Win Securities	44	聯邦	Union Securities	69	美商高盛	Goldman Sachs
20	光隆	Kuang Long Sec's	45	陽信	Sunny Securities	70	港商德意志	Dueche Bank
21	元富	Masterlink Sec's	46	玉山	E.SUN Comm Bank	71	港商法國興業	Société Générale
22	日茂	Jee Mach Sec's	47	國泰	Cathay Securities	72	花旗環球	Citi Bank
23	奔亞	Primasia Sec's	48	大和國泰	Daiwa Cathay Sec's	73	新加坡商瑞銀	Swiss Bank
24	台中銀	Taichung Comm Bank	49	群益金鼎	Capital Sec's	74	法銀巴黎	Bank
25	中國信託	China Trust	50	凱基	KGI Securities	75	香港上海匯豐	HSBC

* translated (phonetically) by author.

In our empirical work, we use volume data to detect swarm. There are both share volume and dollar volume (shares multiplied by price) bought and sold. To have a quick glance of such data, we plot them in Figure 4. In Figure 4, we plot daily trading volumes of all the stocks in 2019. They are shares bought and dollar volumes bought.

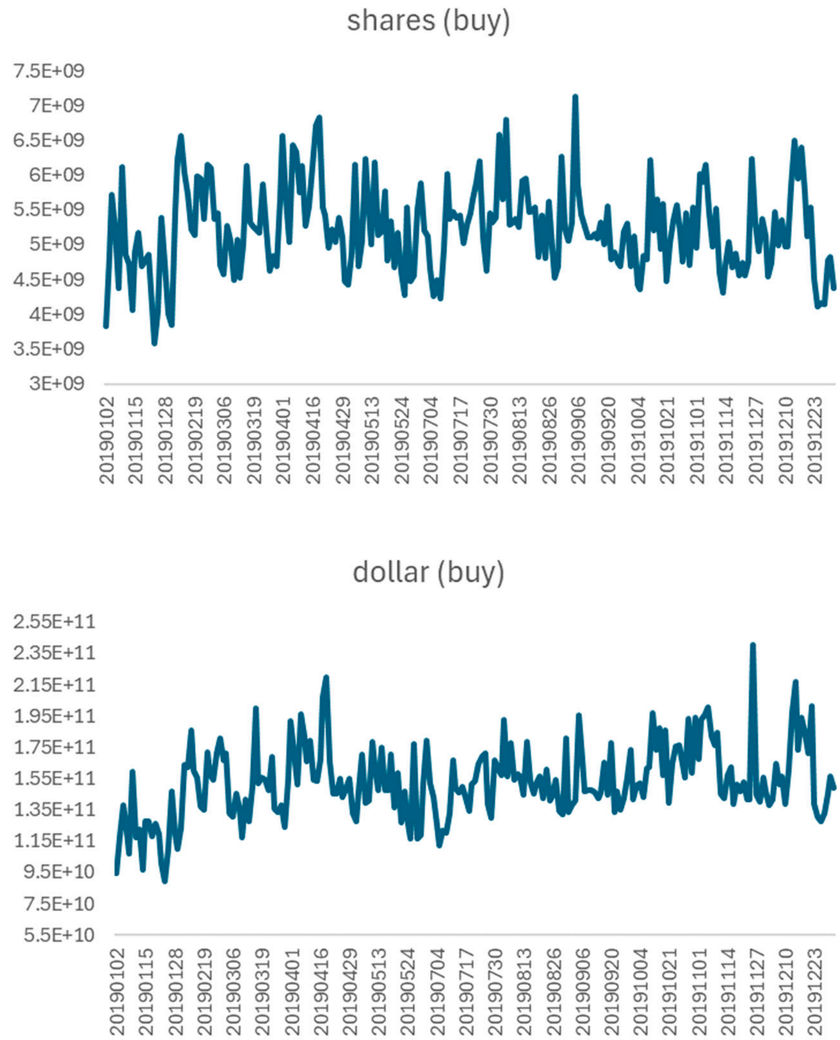


Figure 4. Daily Trading Share and Dollar Volumes of 2019 (223 days).

It is clear that volume data are noisy, and hence, difficult to see patterns. Overall, dollar and share volumes are similar, indicating that the variability of shares dominates that of prices.

The average share trading volume is 5.22 billion shares per day and NT$15.3 trillion (about $5 billion). Hence, the average price per share is NT$29.28 (about 97 cents). (Note that the largest stock in Taiwan—Taiwan Semi-conductor Manufacturing Company, or TSMC (ticker = 2330), shows an average price of NT$263.5 (about $8.5) in the data. TSMC closed at NT$867 (about $28.5) as of 24 May 2024).

We also note that there is no growth in share volume (a linear regression fit has a slightly negative slope near 0 R^2) and yet a noticeable 3.12% growth in dollar volume (R^2 is 11%). Clearly, the visible growth in dollar volume is a result of price growth in 2019 of the overall stock market in Taiwan.

The data have the following shortcomings:

- It is a one-time collection. The data are proprietary, and hence, there is no subsequent effort of such data collection. The year 2019 is not chosen for any particular reason.
- It does not have time stamps, which makes it impossible to study intraday swarm, which is believed to be more evident.

5.2. Boids Results

We present two estimation results in this subsection. One is a joint estimation of all three hyperparameters: alignment, cohesion, and separation. The other is an estimation of only alignment and cohesion. The purpose of eliminating separation is to directly compare the relative strengths of alignment and cohesion, given that separation dominates.

5.2.1. Joint Estimation of Alignment, Cohesion, and Separation

The first set of results using (12) refer to Boids and are plotted in Figure 5. In Figure 5, we set $\vec{v}^{(i)}_{S,t} = \vec{0}$. Note that the idea behind separation is a random move (similar to exploration), and as a result of that, its velocity $\vec{v}^{(i)}_{S,t}$ is a random number. This would unpleasantly cause the solution to the weights $w^{(i)}_{A,t}$ and $w^{(i)}_{C,t}$ to be random. In such a case, we must take expected values. If the effect of Jensen's inequality is small, the expected values of $w^{(i)}_{A,t}$ and $w^{(i)}_{C,t}$ would be similar to (12) when the expected value of $\vec{v}^{(i)}_{S,t}$ is 0. $\mathbb{E}[\vec{v}^{(i)}_{S,t}] = 0$ is not a bad assumption because it is a white-noise move, and it is reasonable for its expected value to be 0. In other words, by assuming $\mathbb{E}[\vec{v}^{(i)}_{S,t}] = 0$, we only suffer from the bias of Jensen's inequality.

We plot both alignment and cohesion weights in Figure 5. We fit the model (12) to shares bought, shares sold, dollar volume bought, and dollar volume sold. These values are median values across all 75 brokers/dealers. While not reported here, mean values across 75 brokers/dealers are way more volatile. For instance, the maximum $w^{(i)}_{A,t}$ is 6241.57 and the minimum is -4744.30 using shares bought. Similarly, using shares bought, the maximum $w^{(i)}_{C,t}$ is 650,133.46 and the minimum is $-186{,}886.23$.

We first observe that, in all four cases, cohesion weight $w_{C,t}$ (median across brokers/dealers, hence, superscript (i) disappears) is higher than alignment weight $w_{A,t}$. Furthermore, while $w_{C,t}$ is mostly positive, $w^{(i)}_{A}$ is mostly negative. Secondly, there is higher volatility of $w_{C,t}$ than of $w_{A,t}$. The median of medians for alignment weight is around -0.6 in all four cases, and for cohesion it is 0 in all four cases. This implies that separation is around 1.6, given that the three parameters must sum to 1 by construction.

The above result indicates that there is not much swarm (or anti-swarm) among these firms. Note that 0 cohesion represents that firms do not move towards the positions of other firms, and negative alignment represents that firms move in an opposite direction of other firms. However, these median values are not significant (although there is no formal statistical test, we can see that the standard deviation is over 10 times of the median).

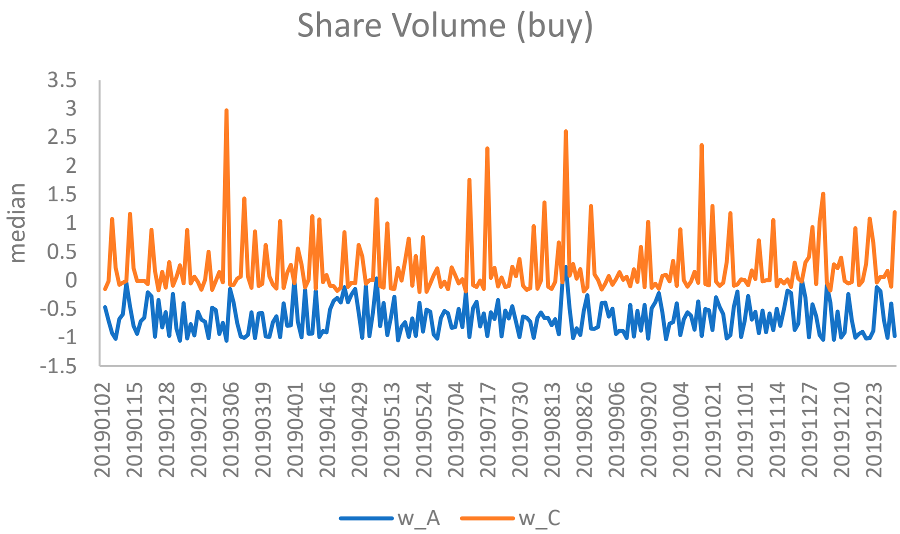


Figure 5. *Cont.*

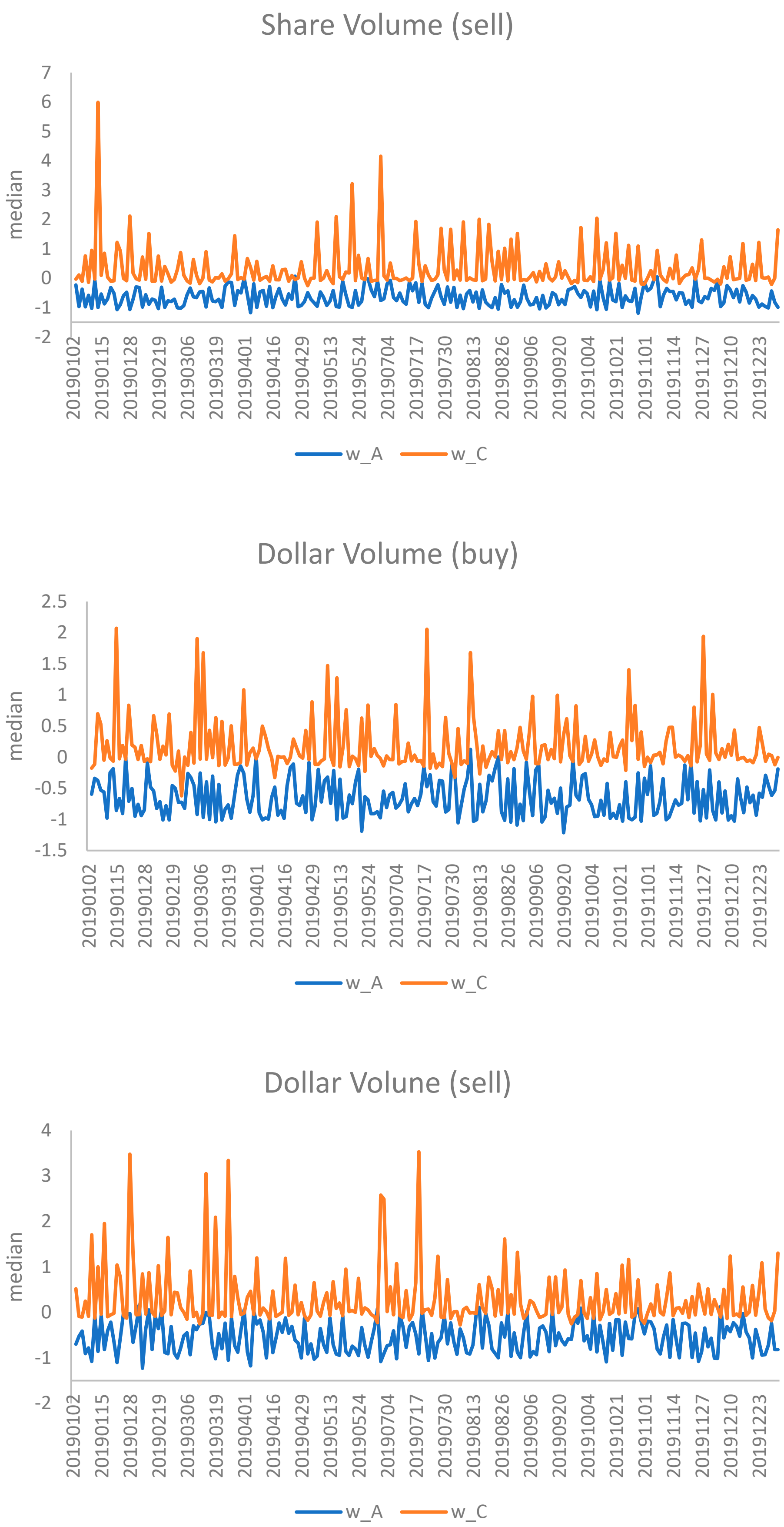


Figure 5. Time Series of Alignment and Cohesion weights (median).

5.2.2. Joint Estimation of Alignment and Cohesion Only (Without Separation)

In lieu of the above result, we remove separation and only compare cohesion against alignment. In other words, we solve for the alignment weight in (14). The results are presented in Figure 6.

The purpose to remove separation and only concentrate on alignment and cohesion is to distinguish the relative strengths of the two swarm behaviors. In estimating (12), we do not constrain on alignment or cohesion. Here, we constrain that $w_{A,t}^{(i)} + w_{C,t}^{(i)} = 1$. Similar to Figure 5, Figure 6 also contains four time series plots. Both means and medians of $w_{A,t}^{(i)}$ across 75 brokers/dealers are plotted in Figure 6. $w_{C,t}^{(i)}$ is simply $1 - w_{A,t}^{(i)}$.

Now, we find that the means and medians are closer to each other, indicating no extreme values (across 75 brokers/dealers). A quick glance of Figure 6 leads to a general observation that cohesion is stronger than alignment. Buy share volume has an overall median of 0.1507 (i.e., average of medians while the median of medians is 0.1324), and sell share volume average of medians is 0.1053 and median of medians is 0.0882. This indicates that, overall, brokers/dealers have a much weaker alignment than cohesion, since sum of the two weights is 1.

This result confirms the relative magnitude of the two parameters, that cohesion has a higher value than alignment. However, the interpretation of the result is quite different. Since now both are positive (since we remove separation, hence anti-swarm), we can more easily see that firms tend to move toward the positions of their competitors rather than follow where they want to go.

Dollar volumes tell the same story. The median of medians for the buy volume is 0.1689 and for the sell volume is 0.1604 (while the averages of the median are 0.1776 and 0.1900, respectively). The average (i.e., average of averages) buy dollar volume yields an even more negligent alignment of -0.0220, and the dollar sell volume yields similarly of 0.0184. This is even a stronger cohesion tendency than share volumes.

Secondly, it is clear that alignment (hence cohesion) is volatile, for both median and mean. This prevents us from identifying a time series pattern of swarming by Taiwanese brokers/dealers (at least in our sample period 2019). There are some possible explanations. The most obvious one is that the volume data are known to be extremely noisy. As a result, it is translated to model parameters—cohesion and alignment. In such a case, the remedy is usually to take moving average. However, given that we have only a very short time series (and worse, the data are daily), taking moving averages is no more insightful than simply looking at the overall averages.

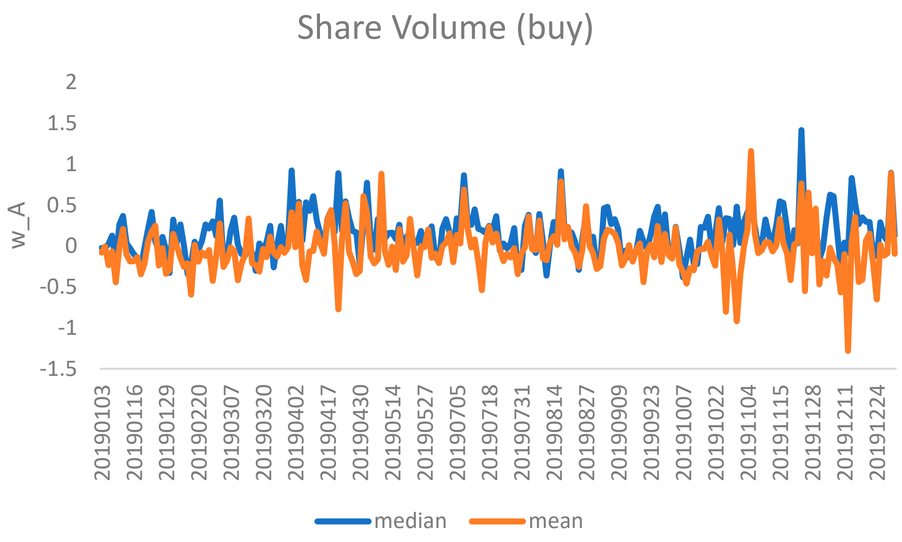


Figure 6. *Cont.*

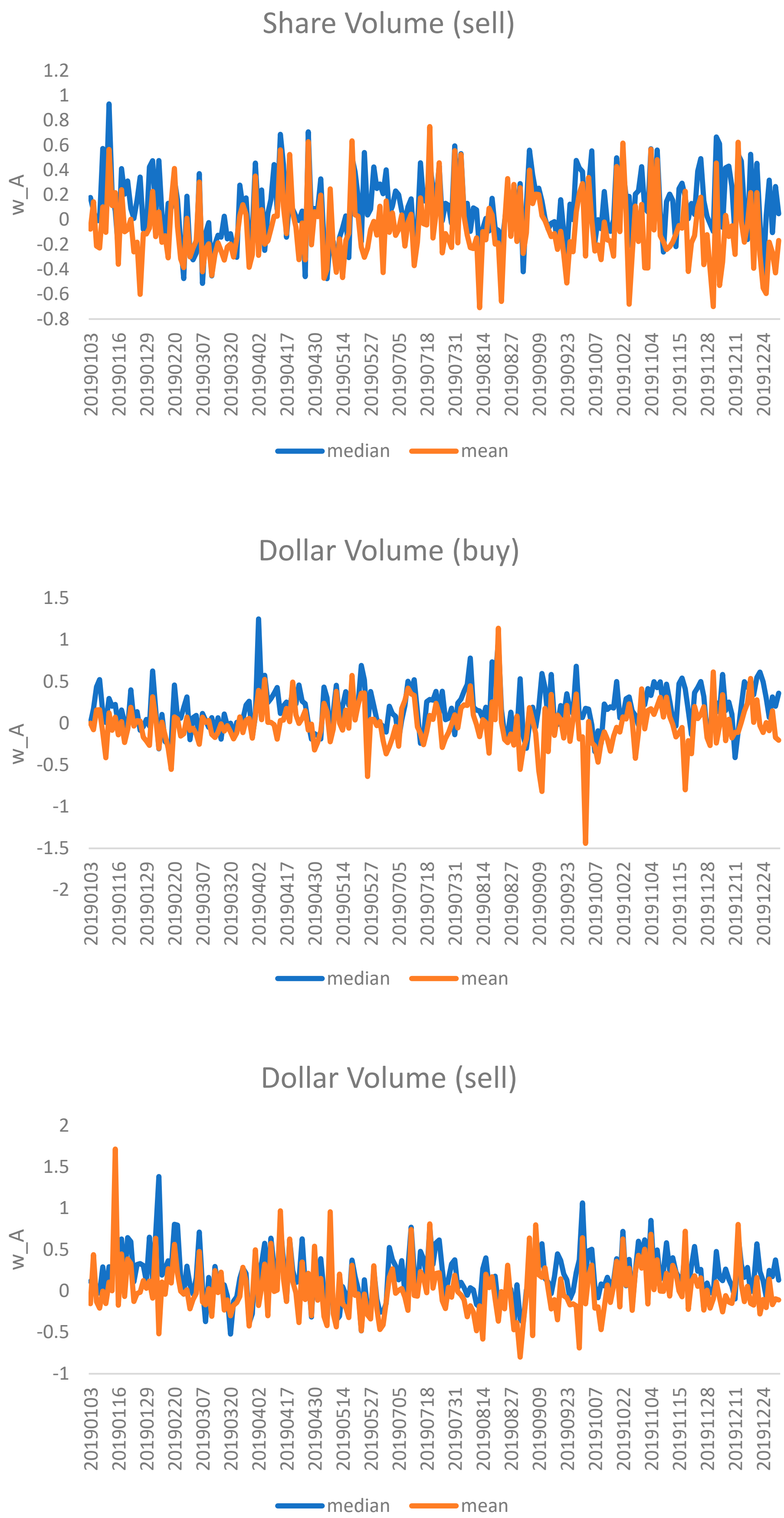


Figure 6. Time Series of Alignment Weights w_A (note: Each result is the mean/median across 75 brokers/dealers).

To see that, we calculate a series of correlation coefficients and report them in Table 4. Table 4 reports correlation coefficients between alignment and cohesion weights and volume data plotted in Figure 4.

Table 4. Correlation Between Alignment, Cohesion and General Volume.

	Buy Share	Sell Share	Buy Dollar	Sell Dollar
alignment	12.30%	24.93%	1.30%	25.82%
p value	0.1027	0.0307	0.4320	0.0287
cohesion	−23.27%	−22.92%	−22.75%	−22.36%
p value	0.0351	0.0361	0.0366	0.0378

We can see that in general, the correlation values are quite substantial. The cohesion weight has over 22% correlation with volume data in all four categories. The alignment weight can be as highly correlated with shares sold as almost 25% (although with shares bought, the correlation values are low). This confirms the suspicion above that noisy data impacts the estimates of parameters. With this observation in mind, we can roughly conclude that the brokers/dealers in the Taiwan stock market present a weak but non-trivial swarm behavior.

The only outlier is how alignment is correlated with dollar volume bought (1.30%). Also note that the second lowest correlation is between alignment and share volume bought (12.30%), although it is barely significant at the 10% level. Both of these low correlation results indicate that buy volumes do not significantly impact alignment of brokers/dealers. In contrast, cohesion seems to reflect more closely of data (correlation is consistently at −22~23%). If we look at the equation that computes the alignment parameter, it is clear that it is dominated by how alignment velocity is different from the cohesion velocity. It is the randomness in this quantity that results in low correlation between alignment buy volume data.

5.2.3. Who Swarms?

As in Lakonishok, Shleifer, and Vishny [6], we study investor herding as opposed to stock herding. Yet more insightful than Lakonishok, Shleifer, and Vishny [6], the swarm result can identify the tendency of herding of each individual firm. This is the key advantage of using swarm intelligence as a traditional statistical method.

In this subsection, we examine which brokers/dealers are the strongest followers in the swarm. The results are plotted in Figure 7. Like previous figures, we present the results in four volume panels. We sort these brokers/dealers by $w_{A,t}^{(i)}$ the alignment weight (y-axis). Note that the cohesion weight is $w_{C,t}^{(i)} = 1 - w_{A,t}^{(i)}$. The x-axis is dealer/broker (by their numbers listed in Table 3).

Except for the buy dollar volume, the other three panels have no higher than one alignment weight. In the case of share volumes (buy and sell), the value of $w_{A,t}$ (average across brokers/dealers and hence, subscript (i) disappears) is 0.7 and −0.4. The decrease of $w_{A,t}$ is roughly linear, indicating that there are no particular outliers (however, we do notice two outliers in buy dollar volume and one outlier in sell dollar volume).

We selectively look at the top 20 brokers/dealers who show strong $w_A^{(i)}$ (average across time and hence, subscript t disappears). Their names are given in Table 5. There are 32 non-repeating names in Table 5.

Table 5. Top 20 Highest Cohesion Brokers/Dealers.

Share Bought	Share Sold	Dollar Bought	Dollar Sold
16	16	39	34
56	56	42	56
35	34	22	19
34	36	8	16
36	35	72	33
18	31	70	17
7	19	1	20
65	65	14	18
19	18	19	6
33	17	51	36
2	33	33	65
20	2	13	14
31	20	73	72
54	7	16	31
17	54	35	7
30	12	4	30
12	71	56	50
71	30	2	12
51	6	52	51
14	50	30	35

In Table 5, we first observe that there are six brokers/dealers who are on the list of all four panels: #16 (First Financial), #19 (I Win Securities), #30 (Reliance Securities), # 33 (Z Ho Securities), #35 (Bridge Stone Securities), and #56 (Union Bank). There is no coincidence that these are all small brokers/dealers. Small firms follow large firms in that they do not have the same research capabilities as well as market power to lead the market. They are more of a follower as a result. For the remaining 26 firms, 12 of which are on three of the four panels. This indicates that regardless which volume measure is used, a follower in one tends to be also a follower in another.

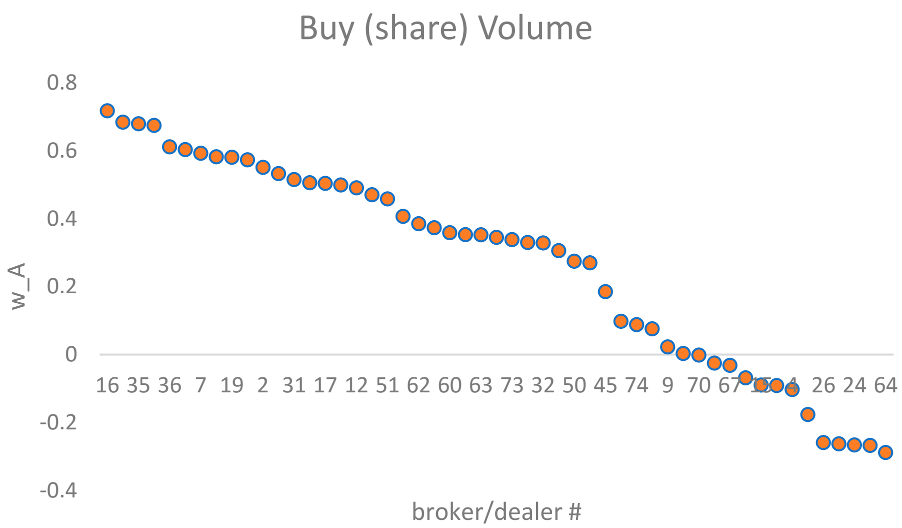


Figure 7. *Cont.*

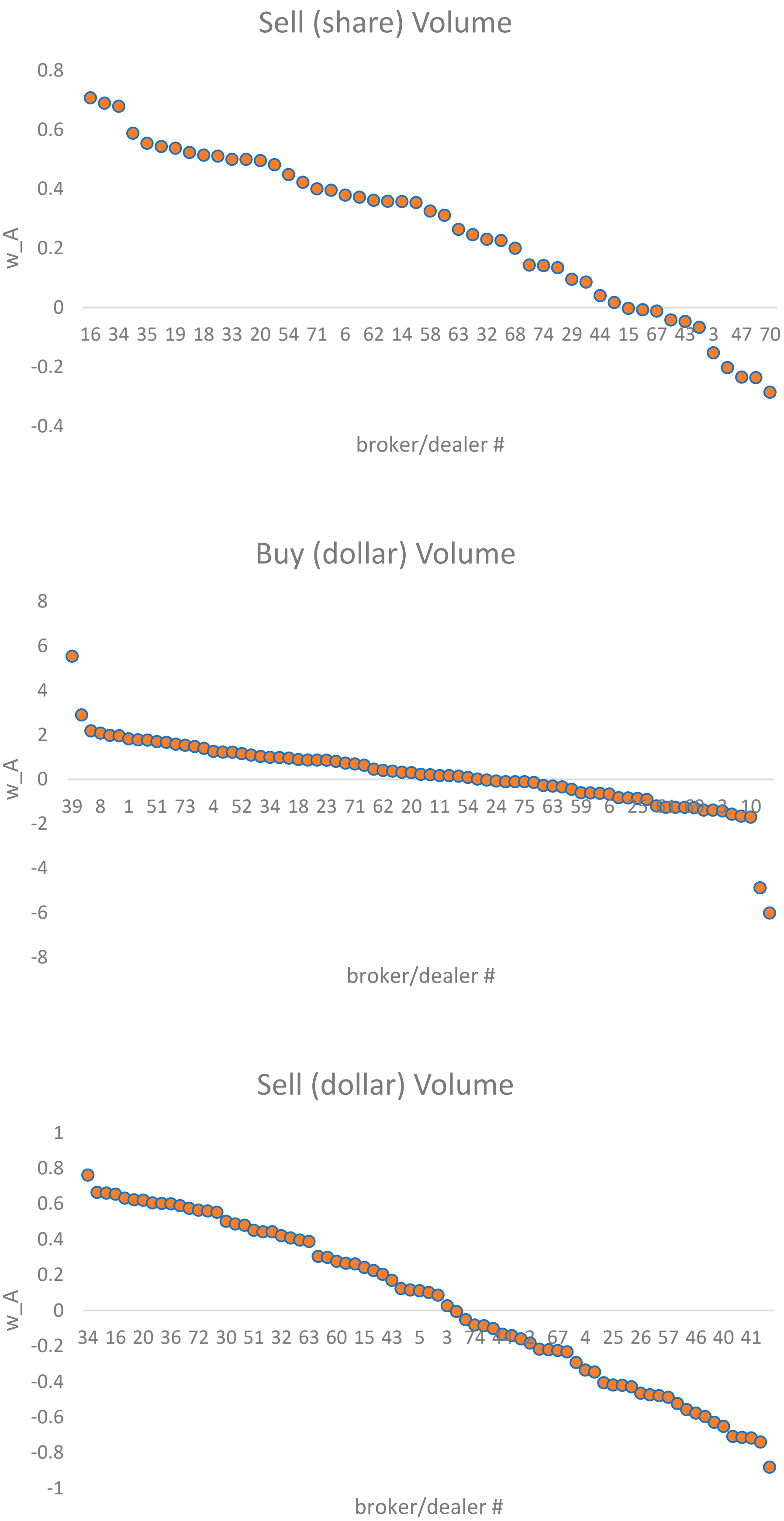


Figure 7. Alignment Weight w_A of 75 Brokers/Dealers.

5.3. PSO Results

In addition to estimating the Boids swarm model, we also estimate the particle swarm optimization model. Particle swarm optimization (PSO) is a swarm model but is mainly used to seek the optimum of an objective function (known as landscape). As a result, PSO can be regarded as a heuristic search (or smart grid search). Boids now are assigned a target to meet.

Because of this particular purpose, while the intelligence of swarm is reserved, the information sought by these birds is different. Now they look for a leader to follow. Hence, instead of following the whole crowd (i.e., align with the crowd in directions or seek to move to toward the crowd), now each bird will identify who the leader (which is the one closest to the target) is and move toward the leader (known as "exploitation"), but at the same time, it is necessary for each bird to "explore" its current neighborhood to see if the true optimum is simply just nearby. Note that there are a number of ways to structure exploration. Equation (4) is the most common expression where exploitation and exploration are intertwined. However, this is not necessarily the case. For the purpose of our empirical work, we follow a modified PSO as in (4a).

We first, as in the previous section, use only the top 20 stocks (listed in Table 2) to fit a PSO model. Then, as a robust check, we use the entire stock market in TSE, which has 967 stocks in total. A flow chart is provided in Figure 8.

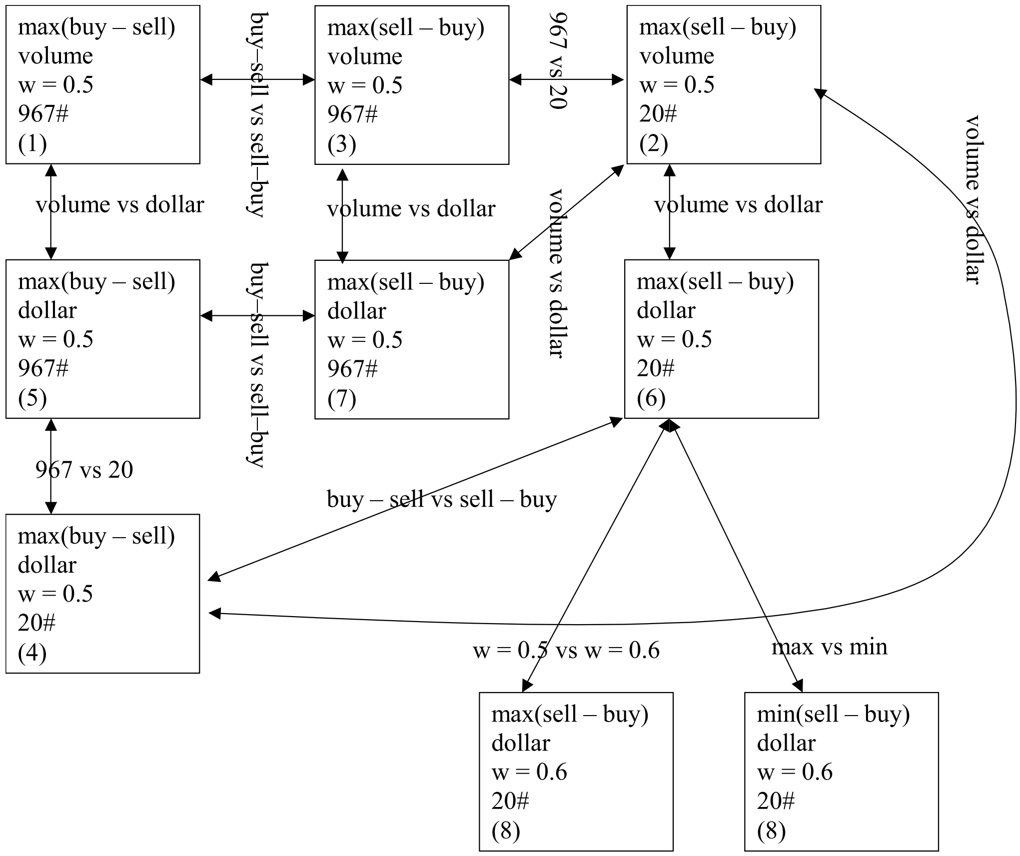


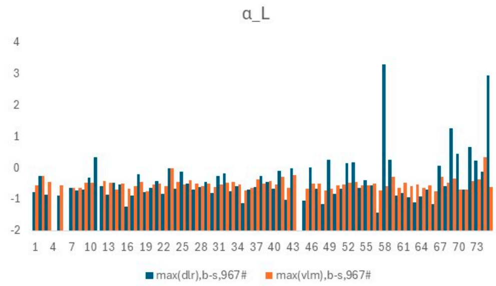
Figure 8. Flow Chart of Various Tests of PSO.

In Figure 8, the relationships of various tests are provided. From Figure 8, one can see how parameters can differ if a different setup of swarm is used. As we can see, we fit the swarm model to the entire data set (976 stocks) and also the top stocks (20 stocks). We also fit the model to share volume (both buy volume and sell volume) as well as dollar volume. We experiment various possibilities such as net buy (which is buy volume subtracting sell volume) and net sell (which is sell volume subtracting buy volume). As a result, there are a large number of pair comparisons. To conserve space, we only represent those pairs that show a substantial difference. Other results are available upon request.

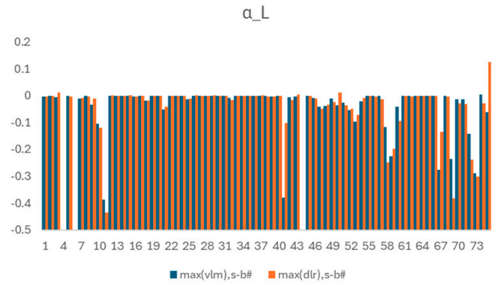
It seems that the swarm behaviors between different volume measures and different sample sizes matter the most. And between the two, sample size matters more than volume measure. As a result, we select four comparisons (other results are available on request) in Figure 9:

(**A**) $\alpha_L^{(i)}$ where $i = 1, \cdots, 75$

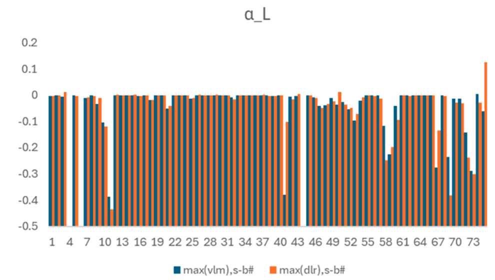
(1)-(5) share vs. dollar (all 967 stocks)



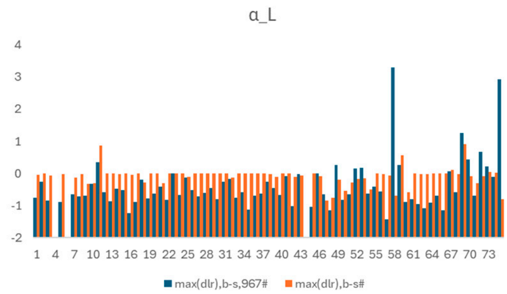
(2)-(6) share vs. dollar (20 stocks)



(3)-(2) 967 vs. 20 stocks (share volume)

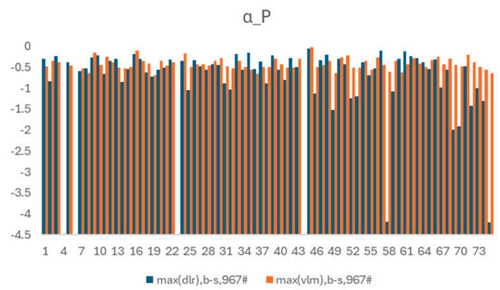


(5)-(4) 967 vs. 20 stocks (dollar volume)

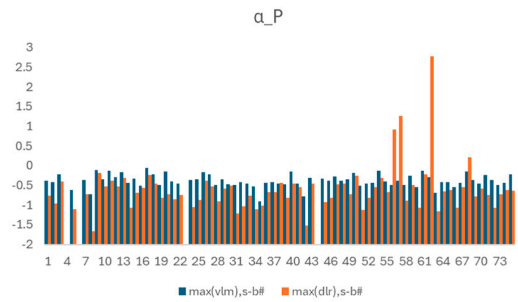


(**B**) $\alpha_P^{(i)}$ where $i = 1, \cdots, 75$

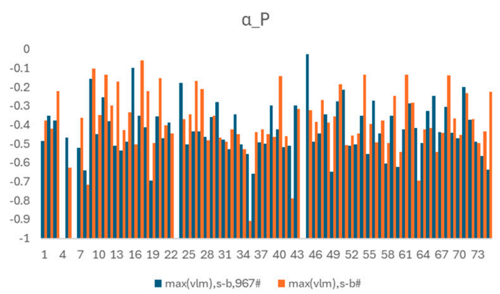
(1)-(5) share vs. dollar (all 967 stocks)



(2)-(6) share vs. dollar (20 stocks)



(3)-(2) 967 vs. 20 stocks (share volume)



(5)-(4) 967 vs. 20 stocks (dollar volume)

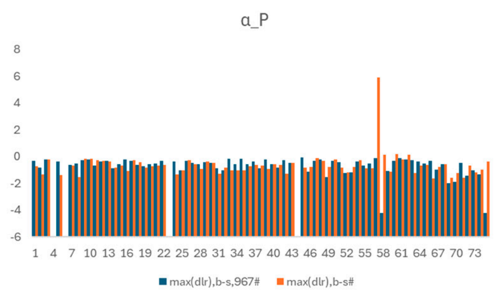


Figure 9. Joint α_L and α_P Estimates $w\alpha_L + (1 - w)\alpha_P$ (average across time).

(1)-(5) share volume versus dollar volume under all (967) stocks.
(2)-(6) share volume versus dollar volume under top (20) stocks
(5)-(4) all (967) stocks versus top (20) stock under dollar volume
(3)-(2) all (967) stocks versus top (20) stock under share volume

Figure 9 plots a joint estimate of exploitation (leader-following) $\alpha_L^{(i)}$ and exploration (personal-freedom) $\alpha_P^{(i)}$ in panels (A) and (B), respectively. Along the x-axis are 75 brokers/dealers ($i = 1, \cdots, 75$ and the names of the brokers/dealers are given in Table 3). Each $\alpha_L^{(i)}$ value is an average across time, and hence, subscript t drops out) for each broker/dealer.

The first comparison, (1)-(5), is share volume versus dollar volume, where the leader is defined as holding the highest net buy position and all 976 stocks are considered. $\alpha_L^{(i)}$ in both cases are largely negative but less negative for share volume than for dollar volume. Furthermore, the two sets of estimates are marginally similar to each other. The correlation between them is 23.77% (significant at the 6.5% level).

We do observe some exceptionally large positive values but only very few. The grand averages are still negative even with these exceptionally large positive values. They are -0.39 and -0.51, respectively. (The medians are -0.61 and -0.53, respectively).

In contrast, (2)-(6), is the same comparison (share volume versus dollar volume), but only the top 20 stocks are considered (hence, a much smaller sample). Similar to the case (1)-(5), both sets of $\alpha_L^{(i)}$ values are similar to each other. However, they are much more similar than the case (1)-(5). The correlation now is much higher 82.60% (significant at the 0.5% level).

This indicates that in a smaller sample, we observe less difference between share volume and dollar volume. Note that the smaller sample is also more dominant in that the top 20 stocks account for a large portion of the market. In other words, except for the top stocks, the majority of the stocks in TSE provide a lot of noise trading.

Now we turn to comparing directly whole same (967 stocks) and subsample (20 stocks). We first compare the two under the share volume, i.e., (3)-(2). In comparison (3)-(2), we observe drastic differences. $\alpha_L^{(i)}$ values are substantially less negative in the case of 20 stocks than in the case of all stocks. The averages are -0.50 and -0.04, respectively, for all stocks and 20 stocks. The correlation is -27.31% (significant at the 5% level), implying that the two set of estimates are rather opposite of each other. This indicates that not only the top 20 stocks do not dominate the market, but the other stocks also move in an opposite direction from the top 20 stocks.

In contrast, the dollar volume comparison (5)-(4) presents a different result from which of (3)-(2). The grand averages are -0.39 and -0.08, respectively, for all stocks and 20 stocks. The correlation is only -8.36% (insignificant). This indicates that the effect that the other stocks move against the top 20 stocks is significantly less. Again, this is due to the price impact. Alternatively speaking, the general price trend negates the negative correlation between the whole market and the top submarket.

The same set of graphs for $\alpha_P^{(i)}$ is provided in Panel (B) of Figure 9. They are quite different from the results of following the leader $\alpha_L^{(i)}$. Except that the values of $\alpha_P^{(i)}$ are mostly negative, similar to the values of $\alpha_L^{(i)}$, the difference between whole market and submarket (i.e., (3)-(2) for share volume, and (5)-(4) for dollar volume) is less substantial. Now the two markets both show positive correlation (25.12% and 8.35%, respectively).

In terms of difference between share and dollar volumes (i.e., (1)-(5) for all stocks and (2)-(6) for 20 stocks), the $\alpha_P^{(i)}$ values are very similar (between -0.4 and -0.6). The correlation is 40.51% and 42.49%, respectively (significant at 2% level). This is different from the case of $\alpha_L^{(i)}$ (see above).

To further investigate the difference, we turn to the raw data (that are used to compute $\alpha_L^{(i)}$ and $\alpha_P^{(i)}$) of all 967 stocks and only the top 20 stocks. We compute the daily net shares bought and net dollars bought of all stocks and the top 20 stocks. To conserve space, these results are available on request. The correlation between the two of share volume is 2.82% and the correlation of dollar volume is 46.48%. It is clear that high correlation in price movements boosts the correlation of the two-dollar volume measurements.

Note that the net dollar volume of the stock is the net purchase of the stock. Hence, it is reasonable to assume that a broker/dealer will swarm according to the net purchases of other brokers/dealers. As a result, this is a more reliable metric to measure how brokers/dealers swarm.

To grasp a general sense of the magnitudes of the results, we report their summary statistics in Table 6. In Table 6, the columns correspond to Figure 8. However, different from Figure 9 that show swarm by brokers/dealers (i.e., taking averages over time for each broker/dealer), Table 6 first takes averages across brokers/dealers, and then the summary statistics are taken over 223 days.

Table 6. Summary Statistics of Swarm Parameters.

(A) $\alpha_{L,t}^{(i)}$							
	(5)	(1)	(4)	(7)	(3)	(2)	(6)
min	−1.1039	−6.2376	−0.4679	−1.1039	−6.2376	−0.4633	−0.4679
max	2.0728	1.6416	1.2320	2.0728	1.6416	0.0379	1.2320
avg	−0.4001	−0.6777	−0.0739	−0.4001	−0.6777	−0.0568	−0.0739
std	0.3936	0.4920	0.1792	0.3936	0.4920	0.0642	0.1792
med	−0.4573	−0.6897	−0.0968	−0.4573	−0.6897	−0.0428	−0.0968
(B) $\alpha_{P,t}^{(i)}$							
	(5)	(1)	(4)	(7)	(3)	(2)	(6)
min	−2.4921	−2.0772	−2.0000	−2.4921	−2.0772	−3.3800	−2.0000
max	0.0000	0.0000	5.5925	0.0000	0.0000	2.5114	9.5754
avg	−0.6893	−0.4554	−0.6372	−0.6893	−0.4554	−0.6572	−0.5466
std	0.2897	0.1992	0.7322	0.2897	0.1992	0.5287	1.0786
med	−0.6413	−0.4340	−0.7388	−0.6413	−0.4340	−0.7256	−0.7542

We can see that while various versions of PSO show differences in how brokers/dealers swarm in Figure 9, the overall averages are quite similar. Regardless of different versions of PSO, they are all negative and roughly at a level about −0.4 (with cases #2, #4, and #6 at about −0.07). We notice that the minimum and maximum values are mild, mainly due to each day there is already an average taken across all brokers/dealers. We also notice that the difference between α_L (average over both brokers/dealers and time, and hence (i) and t both drop out) and α_P is much smaller, although α_P is slightly more negative than α_L. Besides, α_P is more uniform across different versions of PSO, ranging from −0.43 to −0.74.

The next empirical work is to estimate $\alpha_{L,t}^{(i)}$ and $\alpha_{P,t}^{(i)}$ by themselves separately (not jointly). We obtain the same number of results for $\alpha_{L,t}^{(i)}$ and $\alpha_{P,t}^{(i)}$ estimated separately on their own. To conserve space, we only choose a few to compare to the results that they are estimated jointly. The results are reported in Figure 10.

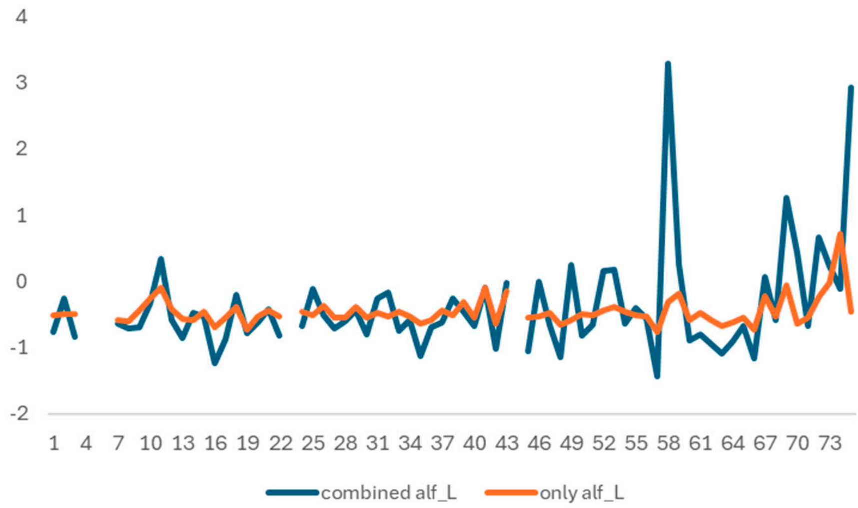
The first is to use case #6 where all (967) stocks and dollar volume are used to estimate $\alpha_{L,t}^{(i)}$ and $\alpha_{P,t}^{(i)}$, as seen in the Panel (A) of Figure 10. On the left, we compare $\alpha_L^{(i)}$ (average across time, and hence, subscript t drops out) when it is estimated on its own and with $\alpha_P^{(i)}$. As we can see, the two results are similar in direction and yet different in magnitude. The correlation between the two is 40.64%, which is significant. Interestingly the magnitude (and also variation) of $\alpha_L^{(i)}$ estimated jointly is so much larger than if estimated in separation.

In the middle of Panel (A) is the same comparison for $\alpha_P^{(i)}$ (average across time, and hence, subscript t drops out). We do see that when $\alpha_P^{(i)}$ has a larger magnitude when it is estimated jointly with $\alpha_L^{(i)}$ than by itself. This is similar to the result of $\alpha_L^{(i)}$. Also similar is the high correlation between the two, which is 41.92%.

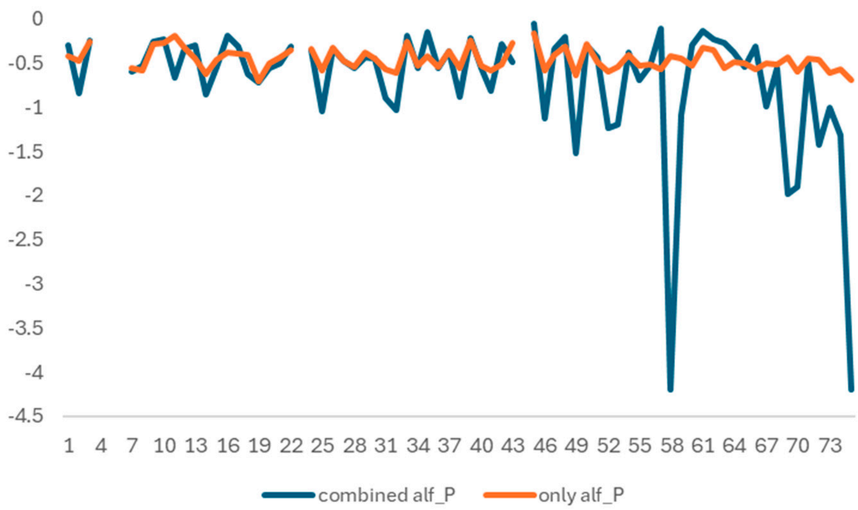
While $\alpha_L^{(i)}$ and $\alpha_P^{(i)}$ (average across time, and hence, subscript t drops out) both have higher magnitudes when they are estimated together jointly, they tend to compensate each other. To see that, we plot the two on the right of Panel (A). Now, it is clear that $\alpha_L^{(i)}$ and $\alpha_P^{(i)}$ move in exactly the opposite direction. This is expected in that they split the total velocity. The correlation here is –95.36%.

(**A**) Case (5) 967 stocks, dollar volume

α_L Only versus Combined

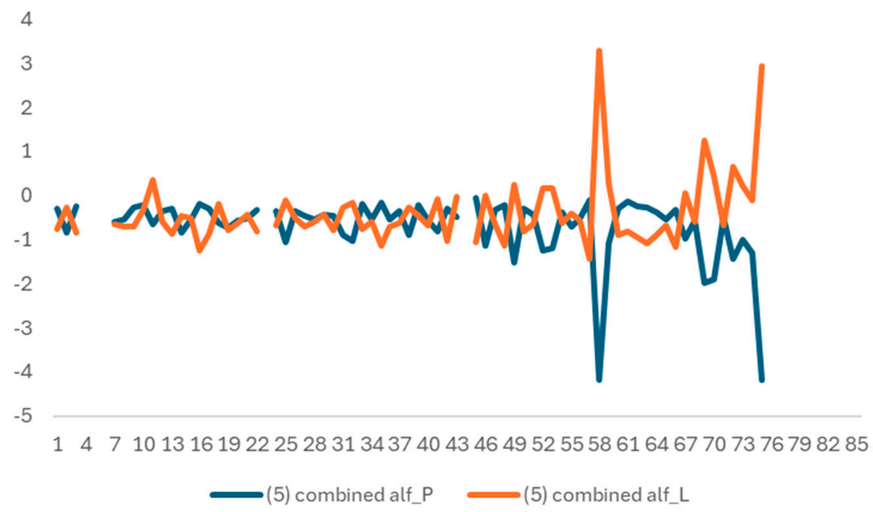


α_P Only versus Combined



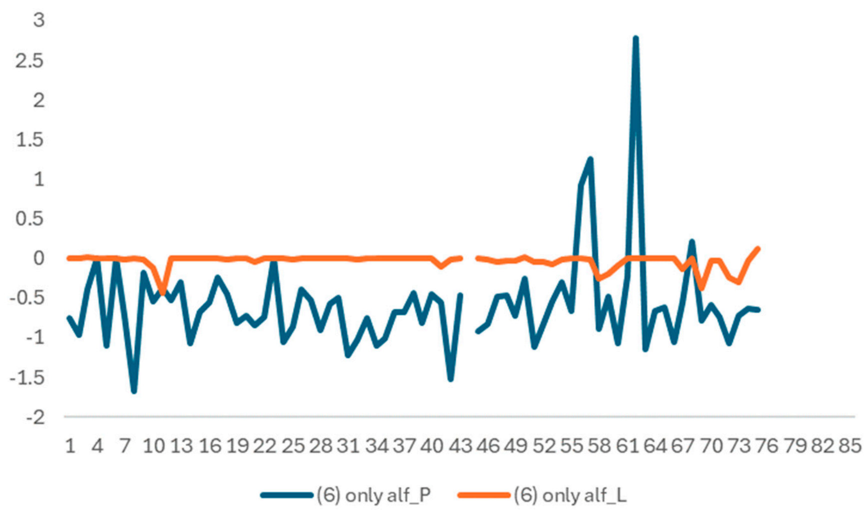
α_L versus α_P (combined)

Figure 10. *Cont.*

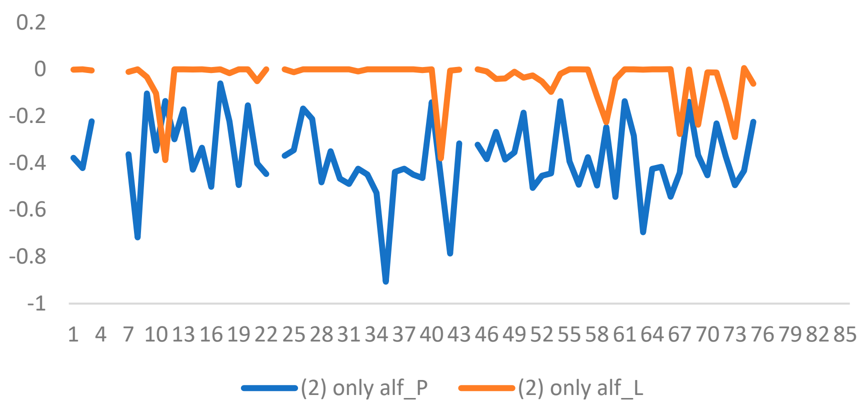


(**B**) α_L versus α_P (only)

(6) 20 stocks, dollar volume



(2) 20 stocks, share volume



(6) 20 stocks, dollar volume

Figure 10. *Cont.*

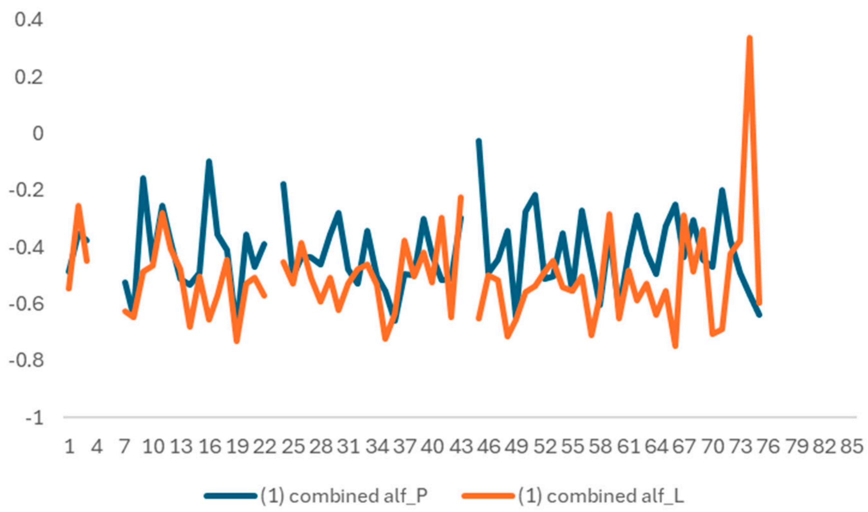


Figure 10. Swarm Intelligence (source: https://en.wikipedia.org/wiki/Boids).

Now we turn to compare $\alpha_L^{(i)}$ and $\alpha_P^{(i)}$ (average across time, and hence, subscript t drops out) in Panel (B) when they are estimated on their own separately. For the first comparisons, we choose the case of the top (20) stocks. On the very left of Panel (B) is dollar volume, and in the middle of Panel (B) is share volume. As we can see, the two results have very similar conclusions. $\alpha_L^{(i)}$ and $\alpha_P^{(i)}$ move in the same direction in both cases. The correlation is about 12%, which is not as high as those in Panel (A) and yet is still quite noticeable, especially for those last several brokers/dealers. Also, we observe that $\alpha_P^{(i)}$ has a much larger magnitude than $\alpha_L^{(i)}$. This is expected as the former contains a random number in each iteration.

Lastly, we examine the similarity between $\alpha_L^{(i)}$ and $\alpha_P^{(i)}$ (average across time, and hence, subscript t drops out) when they are estimated jointly. This is analogous to the right-most case in Panel (A). The difference is here we have dollar volume and in Panel (A) it is share volume. We see that some brokers/dealers are similar, but some others are different. For example, we can see that the last few brokers/dealers whose $\alpha_L^{(i)}$ and $\alpha_P^{(i)}$ move in oppose directions in both cases. Yet, those brokers/dealers in the middle tend to move in opposite directions in Panel (A) but in the same direction in Panel (B). There are two possible sources that can cause such a result. First is the number of stocks considered. In Panel (B), only 20 stocks are used as opposed to all 967 stocks in Panel (A). Furthermore, dollar volume has a price influence. As prices move higher as a general trend, it will cause positive correlation (or negate negative correlation).

To grasp a general sense of the magnitude of the results, we report their summary statistics in Table 7. Columns of Table 7 correspond to Figure 8, reflecting different versions of PSO.

The values of α_L and α_P (average across both brokers/dealers and time, and hence, (i) and t drop out) in Table 7 are quite similar to those in Table 6. The averages of α_L and α_P in particular are similar to those in Table 6. However, the standard deviations are smaller. In the case of α_L, they are at the magnitudes of 0.1 and 0.2 now as opposed to 0.3 and 0.4 in Table 6. This is same in the case of α_P. Minimum and maximum values are also milder when they are estimated separately than jointly.

Table 7. Separation Estimation of $\alpha_{L,t}^{(i)}$ and $\alpha_{P,t}^{(i)}$.

(A) $\alpha_{L,t}^{(i)}$ 75 Brokers Under Various Scenarios (α_L only)							
	(5)	(1)	(4)	(7)	(3)	(2)	(6)
min	−0.7659	−0.7500	−0.4550	−0.7659	−0.7500	−0.3874	−0.2692
max	0.7199	0.3377	0.2741	0.7199	0.3377	0.0054	0.0471
avg	−0.4598	−0.5109	−0.0940	−0.4598	−0.5109	−0.0391	−0.0106
std	0.2106	0.1571	0.1533	0.2106	0.1571	0.0860	0.0444
med	−0.5148	−0.5266	−0.0121	−0.5148	−0.5266	−0.0017	0.0000
(B) $\alpha_{P,t}^{(i)}$ 75 Brokers Under Various Scenarios (α_L only)							
	(5)	(1)	(4)	(7)	(3)	(2)	(6)
min	−0.7022	−0.6947	−0.8377	−0.7022	−0.6947	−0.9061	−0.8377
max	−0.1680	−0.0272	0.7602	−0.1680	−0.0272	−0.0599	0.7602
avg	−0.4603	−0.4213	−0.3849	−0.4603	−0.4213	−0.3822	−0.3849
std	0.1196	0.1335	0.2376	0.1196	0.1335	0.1594	0.2376
med	−0.4799	−0.4422	−0.3979	−0.4799	−0.4422	−0.3932	−0.3979

6. Limitations of Data (We Thank an Anonymous Referee for His/er Excellent Insight and Suggestions Toward This Limitation of This Paper)

There are two major limitations of the data used in this study. The first is its lack of intraday time stamps. The second is the noisy nature of volume data (compared to institutional holdings).

6.1. Lack of Time Stamps

The proprietary data given by HiHedge contain intraday trading by all brokers/dealers across all locations in Taiwan. Theoretically, this would be an ideal data set for studying swarm. Unfortunately, these trades lack time stamps during the day. As a result, we must aggregate the trades into daily volumes. In a swarm, each fish is moving according to its neighbor fish. With time stamps, each broker/dealer then can observe, at every moment, all the other brokers/dealers' tractions and decide if it wants to align with them, cohere to the same holdings as them, or separate from them. With estimation Equations (14)~(17), we can then use data to back out what a particular movement of a particular broker/dealer means. Does he align, cohere, or separate (or follow a leader). This is superior to using an indirect proxy to measure herding and a linear regression model to determine its determinants.

Unfortunately, without time stamps, we cannot actually detect such swarm. Instead, we can only observe between each day how holdings of each broker/dealer change. This is similar to Lakonish, Shleifer, and Vishney [6], yet we do have one advantage—daily as opposed to quarterly.

Not only do we lose information by aggregating, we also mitigate the effects of swarm due to not able to precisely observe the time of the movements. In daily data, we implicitly assume that the movement of today is based upon the movement of others of yesterday. As we clearly see, this is not accurate at all if the true swarm is based upon intra-day reactions. For example, a trade in the middle of day is based upon the movements of the others on the same day, not the previous day.

6.2. Volume Too Noisy

While the deficiency of intra-day information is data collection limitation, the volume data themselves are simply noisy (even noisier for intra-day). It is natural that noisy volume data can result in noisy parameter estimates. Such a limitation cannot be fixed via better data collection methods. It must be handled via noise reduction methods.

Fortunately there are many powerful noise-reduction (smoothing) methods in statistics such as Kalman filter or Savitzky-Golay filter. One can also employ image processing techniques for pattern recognition to screen out noise. Although we can use these techniques on our volume data (in fact, we did play with Savitzky-Golay filter), simple eyeballing the data clearly reveals that the data presents no pattern. This is because we have too few points (232 only) to show any pattern. A longer dataset (either via a higher frequency such as intra-day or a longer time period such as multi-year) is necessary.

However, unfortunately, with just one year of data (our major empirical limitation), we would not be able to fully benefit from a filter. Without a visible pattern or trend, filtered data would generate result very similar to just a simple average of current estimates. (Imagine an extreme case where raw data contain only noise but no pattern or trend. Then a filter will simply produce a straight line as the result—at the average). Waggle and Agrrawal [23], when discussion the "sell-in-May" anomaly, study the seasonal effects in trading profits. When such an effect (or any patterns) is observed, it is possible that swarm is existent. Then filters like Kalman or Savitzky-Golay would then be helpful in reducing the noise in raw data.

7. Conclusions and Future Research

In this paper, we examine if the stock brokers/dealers in Taiwan swarm. The contributions of this paper are as follows: (i) It is the first paper to use swarm intelligence to model herding. Swarm is a perfect choice for herding in that it describes a group behavior. The goal-seeking incentive (e.g. bees seeking honey) that underlies swarm is the same as profit-seeking investors. (ii) We explicitly estimate hyper-parameters of swarm (via a set of closed-form solutions derived in this paper) which allow readers (different from all previous herding research) to see the exact behavior of herding. (iii) We luckily acquired a proprietary dataset from HiHedge. The data contain all transactions happened in all the branches of all the brokers/dealers in Taiwan in 2019. Without such data the swarm model could not be estimated.

Two swarm models are employed in the study. One is Boids which describes how birds swarm and the other is particle swarm optimization (PSO) which use swarm to find the global optimum. We adapt both models to fit in our situation here. In particular, we modify the PSO algorithm to capture if there is a leader that all firms follow.

The results show a weak swarm behavior in Taiwan's stock market in 2019. The PSO result is slightly stronger than the Boids result. This weak result can be largely attributed to noisy volume data in the first place. Noisy data translate naturally to noisy parameter estimates. We find high co-movements between parameter estimates and raw volume data. For future research, we must reduce noise in the volume data.

Not only do we detect swarm in Taiwan's stock market, we also observe some econometric issues. For example, we find that parameters in swarm tend to offset each other. This is frequently observed in financial studies as models are forced to fit data.

Finally, as mentioned earlier, Mavruk [1] uses various machine learning algorithms to retrieve features that explain herding. However, his study separates stock herding from investor herding. Using swarm, we jointly study which investor herds which stock. As a result, we can fully benefit his feature-selection algorithms in improving the current paper (provided that better data (i.e., longer and more granular) can be accessible).

There are limitations to this study. First, there is no time stamp attached to each transaction. As a result, although there are over 200 million transactions, they are aggregated to less than a million observations on a daily basis. Such a loss of information drastically reduces the validity of the result. One obvious problem, as mentioned above, is how information is transmitted in a swarm. Note that in a swarm, each fish examines where other fish are located and how they want to move, and then decides how it will make its own move. Such an assumption requires at any time s on day t, positions of all fish at time $s-1$ on day t are known. Without intraday time stamps, each fish makes its move at time s (e.g., 10:30 a.m.) on day t (e.g., 2 June 2019) based upon information on day $t-1$ (i.e., 1 June 2019), as opposed to time $s-1$ (i.e., 10:30 a.m.) on day t (i.e., 2 June 2019). This clearly defeats the validity of the result. Another problem is the noisy nature of the volume data. With a long enough time series of data, one can adopt various statistical tools, such as filters and pattern-recognition tools in image processing to reduce noises. Due to the aggregation of intraday data into daily data, we lose a large majority of data. This prohibits us from identifying any pattern in the volume data. This is clearly present in Figure 4 in which it is indistinguishable between herding and noise.

For future research, the methodologies proposed in this paper can be applied when more granular data are available in order to obtain more insightful results. Furthermore, one can utilize textual data to obtain insightful results. This is because herding is often triggered by news. The development of models used to retrieve information from textual data has been exploding, most notably "generative pre-trained transformers" (or GPT). The methodologies proposed in this paper can be applied to these transformers to obtain more insightful results.

Funding: This research received no external funding.

Institutional Review Board Statement: The study did not require ethical approval.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is unavailable due to privacy restrictions.

Acknowledgments: We are grateful for HiHedge for the proprietary data.

Conflicts of Interest: The author declares no conflicts of interest.

Appendix A

This appendix provides an example of a simple swarm algorithm with five fish in two dimensions. We first randomly set the initial positions of the five fish in a unit square (i.e., between 0 and 1) and also initial velocity (between 0 and 1), as follows:

	position		velocity	
	x1	x2	v1	v2
1	0.2038	0.6760	0.5824	0.6446
2	0.1787	0.4814	0.6943	0.0557
3	0.7132	0.8629	0.0372	0.9300
4	0.2838	0.2329	0.9947	0.3429
5	0.7271	0.2468	0.8690	0.6400

From the position values, we can calculate cohesion. From the velocity values, we can calculate alignment. Finally, we can add separation by prohibiting fish from colliding into one another. Note that each fish has a tendency to swim toward the center of other fish. Hence, for each fish, we need to calculate the averages of the coordinates of other fish (i.e., center). For example, for fish 1, the center of the other 4 fish (fish 2, 3, 4, and 5) is, following Equation (2):

x1-axis: 0.1787 + 0.7132 + 0.2838 + 0.7271 = 0.4757

x2-axis: 0.4814 + 0.8629 + 0.2329 + 0.2468 = 0.4560

The distance of fish 1 from this center is simply to subtract the position of fish 1 from the position of the center:

x1-axis: 0.4757 − 0.2038 = 0.2719

x2-axis: 0.4560 − 0.6760 = −0.2200

This is cohesion. Fish 1 will take a portion of this to make its next move (weighted by a fraction). Now, one can repeat the same process for other fish, and the final result is given below:

	Cohesion	
	x1	x2
1	0.2719	−0.2200
2	0.3033	0.0232
3	−0.3649	−0.4536
4	0.1719	0.3339
5	−0.3822	0.3164

The alignment is calculated the same way. Instead of position center, now we need to find the center of velocity for every fish. Again, for fish 1, we need to calculate the averages of the coordinates of the other four fish (fish 2~fish 5):

x1-axis: 0.6943 + 0.0372 + 0.9947 + 0.8690 = 0.6488

x2-axis: 0.0557 + 0.9300 + 0.3429 + 0.6400 = 0.4921

and then similarly calculate the difference of the velocity of fish 1 and this center in order to obtain alignment.

x1-axis: 0.6488 − 0.5824 = 0.0664

x2-axis: 0.4921 − 0.6446 = −0.1524

Now we simply repeat the process to calculate the alignment amounts for the other fish, shown in the table below:

	Alignment	
	x1	x2
1	0.0664	−0.1524
2	−0.0734	0.5837
3	0.7479	−0.5092
4	−0.4490	0.2247
5	−0.2919	−0.1467

The last step is the weighted sum up of the two forces to arrive at the total velocity as in Equation (1): $\vec{v}_t^{(i)} = w_A \vec{v}_{A,t}^{(i)} + w_C \vec{v}_{C,t}^{(i)}$. With the hyperparameters for alignment and cohesion to be 1.5 and 0.8, respectively.

x1-axis: 1.5 × 0.0664 + 0.8 × 0.2719 = 0.3170

x2-axis: 1.5 × (−0.1524) + 0.8 × 0.2200 = −0.4046

Now, we simply repeat the same process for all the other fish and obtain the velocity for all the fish as in Equation (2), listed as follows:

	total velocity	
	x1	x2
1	0.3170	−0.4046
2	0.1325	0.8941
3	0.8299	−1.1268
4	−0.5359	0.6042
5	−0.7436	0.0331

Adding the velocity to the current position, we obtain the position of the next iteration, as in Equation (3) $\vec{x}_t^{(i)} = \vec{x}_{t-1}^{(i)} + \vec{v}_t^{(i)}$, shown as follows:

	position in the next iteration	
	x1	x2
1	0.5209	0.2713
2	0.3112	1.3755
3	1.5431	−0.2639
4	−0.2522	0.8370
5	−0.0165	0.2800

In our empirical study, we do not use Equation (1) $\vec{v}_t^{(i)} = w_A \vec{v}_{A,t}^{(i)} + w_C \vec{v}_{C,t}^{(i)} + w_S \vec{v}_{S,t}^{(i)}$ but rather calculate separation separately. Note that Equation (1) does not guarantee the minimum distance. It is merely a way to incorporate a certain degree of separation. In a Boids example, this is fine. And yet in our work, we do not use actual fish to collide with one another. Separation is activated if a fish is too close to any other fish; otherwise, separation is ignored. We implement separation in the order of fish numbers. That is, we examine if separation is needed for fish 1, then fish 2, and so on. Certainly, if we use a different order, the result will be different. For example, we can randomly select fish for separation (e.g., 3, 4, 2, 1, 5 as opposed to 1, 2, 3, 4, 5).

Appendix B

This appendix derives separate estimation of α_P and α_L.

Note that, for α_P only

$$\vec{v}_t^{(i)} = \vec{v}_{P,t}^{(i)} \tag{A1}$$

where

$$\begin{aligned} \vec{v}_{P,t}^{(i)} &= (1-\alpha_{P,t}^{(i)})\vec{u}_t^{(i)} + \alpha_{P,t}^{(i)}(\vec{p}_t^{(i)} - \vec{x}_t^{(i)}) \\ &= \vec{u}_t^{(i)} + \alpha_{P,t}^{(i)}(\vec{p}_t^{(i)} - \vec{x}_t^{(i)} - \vec{u}_t^{(i)}) \end{aligned} \tag{A2}$$

where $\vec{p}_t^{(i)}$ is bird i's personal best and $\vec{g}_{t-1}$ is the global best. Taking derivative w.r.t. $\alpha_P^{(i)}$:

$$\frac{\partial v_{P,j,t}^{(i)}}{\partial \alpha_P^{(i)}} = p_{j,t}^{(i)} - x_{j,t}^{(i)} - u_{j,t}^{(i)} \tag{A3}$$

The objective is to solve for cohesion across all stocks:

$$\frac{\partial}{\partial \alpha_{P,t}^{(i)}} \sum_{j=1}^{n} (v_{j,t}^{(i)} - \nu_{j,t}^{(i)})^2 = \sum_{j=1}^{n} 2(v_{j,t}^{(i)} - \nu_{j,t}^{(i)}) \frac{\partial v_{j,t}^{(i)}}{\partial \alpha_{P,t}^{(i)}} = 0 \tag{A4}$$

Hence, the solution to $\alpha_{P,t}^{(i)}$ is:

$$\begin{aligned} &0 = \textstyle\sum_{j=1}^{n}(v_{j,t}^{(i)} - \nu_{j,t}^{(i)})(p_{j,t}^{(i)} - x_{j,t}^{(i)} - u_{j,t}^{(i)}) \\ &\textstyle\sum_{j=1}^{n} v_{j,t}^{(i)}(p_{j,t}^{(i)} - x_{j,t}^{(i)} - u_{j,t}^{(i)}) = \sum_{j=1}^{n} \nu_{j,t}^{(i)}(p_{j,t}^{(i)} - x_{j,t}^{(i)} - u_{j,t}^{(i)}) \\ &\textstyle\sum_{j=1}^{n}(u_{j,t}^{(i)} + \alpha_{P,t}^{(i)}(p_{j,t}^{(i)} - x_{j,t}^{(i)} - u_{j,t}^{(i)}))(p_{j,t}^{(i)} - x_{j,t}^{(i)} - u_{j,t}^{(i)}) = \sum_{j=1}^{n} \nu_{j,t}^{(i)}(p_{j,t}^{(i)} - x_{j,t}^{(i)} - u_{j,t}^{(i)}) \\ &\alpha_{P,t}^{(i)} \textstyle\sum_{j=1}^{n}(p_{j,t}^{(i)} - x_{j,t}^{(i)} - u_{j,t}^{(i)})(p_{j,t}^{(i)} - x_{j,t}^{(i)}) = \sum_{j=1}^{n}(\nu_{j,t}^{(i)} - u_{j,t}^{(i)})(p_{j,t}^{(i)} - x_{j,t}^{(i)} - u_{j,t}^{(i)}) \\ &\alpha_{P,t}^{(i)} = \frac{\sum_{j=1}^{n}(\nu_{j,t}^{(i)} - u_{j,t}^{(i)})(p_{j,t}^{(i)} - x_{j,t}^{(i)} - u_{j,t}^{(i)})}{\sum_{j=1}^{n}(p_{j,t}^{(i)} - x_{j,t}^{(i)} - u_{j,t}^{(i)})^2} \end{aligned} \tag{A5}$$

Because $u_{j,t}^{(i)}$ is random, $\alpha_{P,t}^{(i)}$ is random. To calculate $\mathbb{E}[\alpha_{P,t}^{(i)}]$, we must simulate $u_{j,t}^{(i)}$ a number of times.

To estimate α_L only,

$$\vec{v}_t^{(i)} = \begin{cases} \vec{v}_{L,t}^{(i)} \\ \vec{v}_{L,t} \end{cases}$$

where the former is each firm has its own $\alpha_{L,t}^{(i)}$ and the latter is all firms share the same $\alpha_{L,t}$.

In the first case, $\alpha_{L,t}^{(i)}$. Taking derivative w.r.t. $\alpha_{L,t}^{(i)}$:

$$\frac{\partial \vec{v}_{j,t}^{(i)}}{\partial \alpha_{L,t}^{(i)}} = \vec{g}_{j,t}^{(i)} - \vec{x}_{j,t}^{(i)} \tag{A6}$$

Differentiate the objective function with respect to $\alpha_{L,t}^{(i)}$:

$$\frac{\partial}{\partial \alpha_{L,t}^{(i)}} \sum_{j=1}^{n} (v_{j,t}^{(i)} - \nu_{j,t}^{(i)})^2 = \sum_{j=1}^{n} 2(v_{j,t}^{(i)} - \nu_{j,t}^{(i)}) \frac{\partial v_{j,t}^{(i)}}{\partial \alpha_{L,t}^{(i)}} = 0 \tag{A7}$$

The solution to $\alpha_{L,t}^{(i)}$ is:

$$\begin{aligned}
&0 = \sum_{j=1}^{n} 2(v_{j,t}^{(i)} - \nu_{j,t}^{(i)})(g_{j,t-1} - x_{j,t-1}^{(i)}) \\
&\sum_{j=1}^{n} v_{j,t}^{(i)}(g_{j,t-1} - x_{j,t-1}^{(i)}) = \sum_{j=1}^{n} \nu_{j,t}^{(i)}(g_{j,t-1} - x_{j,t-1}^{(i)}) \\
&\sum_{j=1}^{n} \alpha_{L,t}^{(i)}(g_{j,t-1} - x_{j,t-1}^{(i)})(g_{j,t-1} - x_{j,t-1}^{(i)}) = \sum_{j=1}^{n} \nu_{j,t-1}^{(i)}(g_{j,t-1} - x_{j,t-1}^{(i)}) \\
&\alpha_{L,t}^{(i)} = \frac{\sum_{j=1}^{n} \nu_{j,t}^{(i)}(g_{j,t-1} - x_{j,t-1}^{(i)})}{\sum_{j=1}^{n} (g_{j,t-1} - x_{j,t-1}^{(i)})^2}
\end{aligned} \tag{A8}$$

In the second case $\alpha_{L,t}$.

$$\begin{aligned}
&0 = \sum_{i=1}^{m} \sum_{j=1}^{n} (v_{j,t}^{(i)} - \nu_{j,t}^{(i)}) \frac{\partial v_{j,t}}{\partial \alpha_{L,t}} \\
&\sum_{i=1}^{m} \sum_{j=1}^{n} v_{j,t}^{(i)} \frac{\partial v_{j,t}}{\partial \alpha_{L,t}} = \sum_{i=1}^{m} \sum_{j=1}^{n} \nu_{j,t}^{(i)} \frac{\partial v_{j,t}}{\partial \alpha_{L,t}} \\
&\sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{L,t}(g_{j,t-1} - x_{j,t-1}^{(i)})(g_{j,t-1} - x_{j,t-1}^{(i)}) = \sum_{i=1}^{m} \sum_{j=1}^{n} \nu_{j,t}^{(i)}(g_{j,t-1} - x_{j,t-1}^{(i)}) \\
&\alpha_{L,t} = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} \nu_{j,t}^{(i)}(g_{j,t-1} - x_{j,t-1}^{(i)})}{\sum_{i=1}^{m} \sum_{j=1}^{n} (g_{j,t-1} - x_{j,t-1}^{(i)})^2}
\end{aligned} \tag{A9}$$

Compare to the average of $\alpha_{L,t}^{(i)}$.

Appendix C

This appendix derives joint estimation of α_P and α_L.
Now,

$$\begin{aligned}
&\vec{v}_t^{(i)} = w\vec{v}_{P,t}^{(i)} + (1-w)\vec{v}_{L,t}^{(i)} \\
&\begin{cases} \frac{\partial}{\partial \alpha_P} \vec{v}_t^{(i)} = 0 \\ \frac{\partial}{\partial \alpha_L} \vec{v}_t^{(i)} = 0 \end{cases}
\end{aligned} \tag{A10}$$

We have α_L

$$\begin{aligned}
&\vec{v}_{L,t}^{(i)} = \alpha_{L,t}^{(i)}(\vec{g}_{t-1} - \vec{x}_{t-1}^{(i)}) \\
&\vec{v}_{P,t}^{(i)} = \vec{u}_t^{(i)} + \alpha_P^{(i)}(\vec{p}_t^{(i)} - \vec{x}_{t-1}^{(i)} - \vec{u}_t^{(i)}) \\
&\frac{\partial v_{P,j,t}^{(i)}}{\partial \alpha_{P,t}^{(i)}} = p_{j,t}^{(i)} - x_{j,t}^{(i)} - u_{j,t}^{(i)};\ \frac{\partial v_{L,j,t}^{(i)}}{\partial \alpha_{P,t}^{(i)}} = 0
\end{aligned}$$

and then α_P

$$\begin{aligned}
&\vec{v}_{L,t}^{(i)} = \alpha_{L,t}^{(i)}(\vec{g}_{t-1} - \vec{x}_{t-1}^{(i)}) \\
&\vec{v}_{P,t}^{(i)} = \vec{u}_t^{(i)} + \alpha_P^{(i)}(\vec{p}_t^{(i)} - \vec{x}_{t-1}^{(i)} - \vec{u}_t^{(i)}) \\
&\frac{\partial v_{L,j,t}^{(i)}}{\partial \alpha_{L,t}^{(i)}} = g_{j,t-1} - x_{j,t-1}^{(i)};\ \frac{\partial v_{P,j,t}^{(i)}}{\partial \alpha_{L,t}^{(i)}} = 0
\end{aligned}$$

Solving the equation leads to:

$$\begin{aligned}
&\frac{\partial}{\partial \alpha_{P,t}^{(i)}} \sum_{j=1}^{n} [wv_{P,j,t}^{(i)} + (1-w)v_{L,j,t}^{(i)} - \nu_{j,t}]^2 \\
&= \sum_{j=1}^{n} 2[wv_{P,j,t}^{(i)} + (1-w)v_{L,j,t}^{(i)} - \nu_{j,t}] \frac{\partial v_{P,j,t}^{(i)}}{\partial \alpha_{P,t}^{(i)}} \\
&= \sum_{j=1}^{n} 2[wv_{P,j,t}^{(i)} + (1-w)v_{L,j,t}^{(i)} - \nu_{j,t}][p_{j,t}^{(i)} - x_{j,t}^{(i)} - u_{j,t}^{(i)}] = 0
\end{aligned}$$

and

$$\frac{\partial}{\partial \alpha_{L,t}^{(i)}} \sum_{j=1}^{n} \left[w v_{P,j,t}^{(i)} + (1-w) v_{L,j,t}^{(i)} - \nu_{j,t} \right]^2$$
$$= \sum_{j=1}^{n} 2\left[w v_{P,j,t}^{(i)} + (1-w) v_{L,j,t}^{(i)} - \nu_{j,t} \right] \frac{\partial v_{L,j,t}^{(i)}}{\partial \alpha_{L,t}^{(i)}}$$
$$= \sum_{j=1}^{n} 2\left[w v_{P,j,t}^{(i)} + (1-w) v_{L,j,t}^{(i)} - \nu_{j,t} \right] \left[g_{j,t-1} - x_{j,t-1}^{(i)} \right] = 0$$

Let $z_{j,t}^{(i)} = p_{j,t}^{(i)} - x_{j,t}^{(i)} - u_{j,t}^{(i)}$ (to simplify notation). Then, we can set up the first equation as:

$$(1)\ \sum_{j=1}^{n} \left\{ w u_{j,t}^{(i)} z_{j,t}^{(i)} + w \alpha_{P,t}^{(i)} \left(z_{j,t}^{(i)}\right)^2 + (1-w) \vec{v}_{L,j,t}^{(i)} z_{j,t}^{(i)} - \nu_{j,t} z_{j,t}^{(i)} \right\} = 0$$

Expanding the equation, we have:

$$w \alpha_{P,t}^{(i)} \sum_{j=1}^{n} \left(z_{j,t}^{(i)}\right)^2 = \sum_{j=1}^{n} \left\{ \nu_{j,t} z_{j,t}^{(i)} - w u_{j,t}^{(i)} z_{j,t}^{(i)} - (1-w) v_{L,j,t}^{(i)} z_{j,t}^{(i)} \right\}$$

From which we can then solve for the parameter $\alpha_{P,t}^{(i)}$ as follows:

$$\alpha_{P,t}^{(i)} = \frac{\sum_{j=1}^{n} \left\{ \nu_{j,t} z_{j,t}^{(i)} - w u_{j,t}^{(i)} z_{j,t}^{(i)} - (1-w) v_{L,j,t}^{(i)} z_{j,t}^{(i)} \right\}}{w \sum_{j=1}^{n} \left(z_{j,t}^{(i)}\right)^2}$$

where

$$z_{j,t}^{(i)} = p_{j,t}^{(i)} - x_{j,t}^{(i)} - u_{j,t}^{(i)}$$

Now, we can set up the second equation as follows:

$$(2)\ \sum_{j=1}^{n} \left[w v_{P,j,t}^{(i)} + (1-w) \alpha_{L,t}^{(i)} \left(\vec{g}_{t-1} - \vec{x}_{t-1}^{(i)} \right) - \nu_{j,t} \right] \left(g_{j,t-1} - x_{j,t-1}^{(i)} \right) = 0$$

Similarly, we can expand it and get:

$$\sum_{j=1}^{n} (1-w) \alpha_{L,t}^{(i)} \left(g_{j,t-1} - x_{j,t-1}^{(i)} \right) \left(g_{j,t-1} - x_{j,t-1}^{(i)} \right)$$
$$= \sum_{j=1}^{n} \left[\nu_{j,t} - w v_{P,j,t}^{(i)} \right] \left(g_{j,t-1} - x_{j,t-1}^{(i)} \right)$$

Solving for $\alpha_{L,t}^{(i)}$ to get:

$$\alpha_{L,t}^{(i)} = \frac{\sum_{j=1}^{n} \left(\nu_{j,t} - w \left(u_{j,t}^{(i)} + \alpha_{P}^{(i)} \left(p_{j,t}^{(i)} - x_{j,t}^{(i)} - u_{j,t}^{(i)} \right) \right) \right) \left(g_{j,t-1} - x_{j,t-1}^{(i)} \right)}{(1-w) \sum_{j=1}^{n} \left(\vec{g}_{t-1} - \vec{x}_{t-1}^{(i)} \right)^2}$$

Putting (1) and (2) together, we then solve for both parameters simultaneously as follows:

$$\alpha_{P,t}^{(i)} = \frac{\sum_{j=1}^{n} \left\{ \nu_{j,t} \left(p_{j,t}^{(i)} - x_{j,t}^{(i)} \right) - (1-w) \alpha_{L,t}^{(i)} \left(g_{j,t-1} - x_{j,t-1}^{(i)} \right) \left(p_{j,t}^{(i)} - x_{j,t}^{(i)} \right) \right\}}{w \sum_{j=1}^{n} \left(z_{j,t}^{(i)}\right)^2}$$
$$= \frac{\sum_{j=1}^{n} \nu_{j,t} \left(p_{j,t}^{(i)} - x_{j,t}^{(i)} \right)}{w \sum_{j=1}^{n} \left(p_{j,t}^{(i)} - x_{j,t}^{(i)} \right)^2} - \alpha_{L,t}^{(i)} \frac{(1-w) \sum_{j=1}^{n} \left(g_{j,t-1} - x_{j,t-1}^{(i)} \right) \left(p_{j,t}^{(i)} - x_{j,t}^{(i)} \right)}{w \sum_{j=1}^{n} \left(p_{j,t}^{(i)} - x_{j,t}^{(i)} \right)^2}$$
$$= \pi_1 - \pi_2 \alpha_{L,t}^{(i)}$$

and

$$
\begin{aligned}
\alpha_{L,t}^{(i)} &= \frac{\sum_{j=1}^{n}(\nu_{j,t} - w\alpha_P^{(i)}(p_{j,t}^{(i)} - x_{j,t}^{(i)}))(g_{j,t-1} - x_{j,t-1}^{(i)})}{(1-w)\sum_{j=1}^{n}\left(g_{j,t-1} - x_{j,t-1}^{(i)}\right)^2} \\
&= \frac{\sum_{j=1}^{n}\nu_{j,t}(g_{j,t-1} - x_{j,t-1}^{(i)})}{(1-w)\sum_{j=1}^{n}\left(g_{j,t-1} - x_{j,t-1}^{(i)}\right)^2} - \alpha_P^{(i)}\frac{w\sum_{j=1}^{n}(p_{j,t}^{(i)} - x_{j,t}^{(i)})(g_{j,t-1} - x_{j,t-1}^{(i)})}{(1-w)\sum_{j=1}^{n}\left(g_{j,t-1} - x_{j,t-1}^{(i)}\right)^2} \\
&= \rho_1 - \rho_2\alpha_P^{(i)}
\end{aligned}
$$

Finally,

$$
\begin{aligned}
\alpha_{L,t}^{(i)} &= \tfrac{\rho_1 - \rho_2\pi_1}{1-\rho_2\pi_2} \\
\alpha_{P,t}^{(i)} &= \tfrac{\pi_1 - \pi_2\rho_1}{1-\pi_2\rho_2}
\end{aligned}
\tag{A11}
$$

Appendix D

This appendix derives estimation of alignment, cohesion and separation. Rewrite (1) slightly as follows

$$
\begin{aligned}
\vec{v}_t^{(i)} &= w_{A,t}\vec{v}_{A,t}^{(i)} + w_{C,t}\vec{v}_{C,t}^{(i)} + (1 - w_{A,t} - w_{C,t})\vec{v}_{S,t}^{(i)} \\
&= w_{A,t}\left(\vec{v}_{A,t}^{(i)} - \vec{v}_{S,t}^{(i)}\right) + w_{C,t}\left(\vec{v}_{C,t}^{(i)} - \vec{v}_{S,t}^{(i)}\right) + \vec{v}_{S,t}^{(i)} \\
&= w_{A,t}\vec{A}_t^{(i)} + w_{C,t}\vec{C}_t^{(i)} + \vec{S}_t^{(i)}
\end{aligned}
\tag{1a}
$$

Then we follow (11) and take partial derivatives with respect to alignment w_A and cohesion w_C parameters respectively.

$$
\begin{aligned}
\sum_{j=1}^{n} 2(v_{j,t}^{(i)} - \nu_{j,t}^{(i)})\frac{\partial v_{A,j,t}^{(i)}}{\partial w_{A,t}} = 0 \\
\sum_{j=1}^{n} 2(v_{j,t}^{(i)} - \nu_{j,t}^{(i)})\frac{\partial v_{C,j,t}^{(i)}}{\partial w_{C,t}} = 0
\end{aligned}
\tag{A12}
$$

This leads to the following simultaneous equation system:

$$
\begin{bmatrix} \sum_{j=1}^{n}\left(A_{j,t}^{(i)}\right)^2 & \sum_{j=1}^{n} A_{j,t}^{(i)}C_{j,t}^{(i)} \\ \sum_{j=1}^{n} A_{j,t}^{(i)}C_{j,t}^{(i)} & \sum_{j=1}^{n}\left(C_{j,t}^{(i)}\right)^2 \end{bmatrix} \begin{bmatrix} w_{A,t} \\ w_{C,t} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n}\nu_{j,t}^{(i)}A_{j,t}^{(i)} - \sum_{j=1}^{n}A_{j,t}^{(i)}S_{j,t}^{(i)} \\ \sum_{j=1}^{n}\nu_{j,t}^{(i)}C_{j,t}^{(i)} - \sum_{j=1}^{n}A_{j,t}^{(i)}S_{j,t}^{(i)} \end{bmatrix} \tag{A13}
$$

and the solution is:

$$
\begin{aligned}
\begin{bmatrix} w_{A,t} \\ w_{C,t} \end{bmatrix} &= \begin{bmatrix} \sum_{j=1}^{n}\left(A_{j,t}^{(i)}\right)^2 & \sum_{j=1}^{n} A_{j,t}^{(i)}C_{j,t}^{(i)} \\ \sum_{j=1}^{n} A_{j,t}^{(i)}C_{j,t}^{(i)} & \sum_{j=1}^{n}\left(C_{j,t}^{(i)}\right)^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{j=1}^{n}\nu_{j,t}^{(i)}A_{j,t}^{(i)} - \sum_{j=1}^{n}A_{j,t}^{(i)}S_{j,t}^{(i)} \\ \sum_{j=1}^{n}\nu_{j,t}^{(i)}C_{j,t}^{(i)} - \sum_{j=1}^{n}A_{j,t}^{(i)}S_{j,t}^{(i)} \end{bmatrix} \\
&= \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix}^{-1} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\
&= \frac{1}{x_{11}x_{22} - x_{12}^2}\begin{bmatrix} x_{22} & -x_{12} \\ -x_{12} & x_{11} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}
\end{aligned}
$$

We also estimate another setup of the model where both alignment and cohesion weights are equal, i.e., $w_{A,t}^{(i)} = w_{C,t}^{(i)} = w_t^{(i)}$, as follows:

$$\min \sum_{j=1}^{n} \left(v_{j,t}^{(i)} - \nu_{j,t}^{(i)}\right)^2$$
$$\sum_{j=1}^{n} 2\left(v_{j,t}^{(i)} - \nu_{j,t}^{(i)}\right)\frac{\partial v_{j,t}^{(i)}}{\partial w_t} = 0$$
$$\sum_{j=1}^{n} v_{j,t}^{(i)}\frac{\partial v_{j,t}^{(i)}}{\partial w_t} = \sum \nu_{j,t}^{(i)}\frac{\partial v_{j,t}^{(i)}}{\partial w_t}$$
$$\sum_{j=1}^{n}\left(w_t\left(v_{S,j,t}^{(i)} - v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right) + v_{A,j,t}^{(i)} + v_{C,j,t}^{(i)}\right)\left(v_{S,j,t}^{(i)} - v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right) = \sum_{j=1}^{n} \nu_{j,t}^{(i)}\left(v_{S,j,t}^{(i)} - v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right)$$
$$w_t^{(i)} \sum_{j=1}^{n}\left(v_{S,j,t}^{(i)} - v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right)\left(v_{S,j,t}^{(i)} - v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right) = \sum_{j=1}^{n}\left(\nu_{j,t}^{(i)} - v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right)\left(v_{S,j,t}^{(i)} - v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right)$$
$$w_t^{(i)} = \frac{\sum_{j=1}^{n}\left(\nu_{j,t}^{(i)} - v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right)\left(v_{S,j,t}^{(i)} - v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right)}{\sum_{j=1}^{n}\left(v_{S,j,t}^{(i)} - v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right)^2}$$

Since $v_{S,j,t}^{(i)}$ is random, $w_t^{(i)}$ is random. Take expectation

$$\mathbb{E}[w_t] \approx \frac{\sum_{j=1}^{n}\left(\nu_{j,t}^{(i)} - v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right)\left(c - v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right)}{\sum_{j=1}^{n}\left(c - v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right)^2} = \frac{\sum_{j=1}^{n}\left(\nu_{j,t}^{(i)} - v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right)\left(-v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right)}{\sum_{j=1}^{n}\left(-v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right)^2}$$

Lastly (This result is not presented in the paper but can be available upon request), we can estimate alignment and cohesion without separation as follows:

$$\vec{v}_t^{(i)} = w_t^{(i)}\vec{v}_{A,t}^{(i)} + \left(1 - w_t^{(i)}\right)\vec{v}_{C,t}^{(i)}$$

The solution is:

$$\sum_{j=1}^{n} 2\left(v_{j,t}^{(i)} - \nu_{j,t}^{(i)}\right)\frac{\partial v_{j,t}^{(i)}}{\partial w_t} = 0$$
$$\sum_{j=1}^{n} v_{j,t}^{(i)}\frac{\partial v_{j,t}^{(i)}}{\partial w_t} = \sum \nu_{j,t}^{(i)}\frac{\partial v_{j,t}^{(i)}}{\partial w_t}$$
$$\sum_{j=1}^{n}\left(w_t\left(v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right) + v_{C,j,t}^{(i)}\right)\left(v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right) = \sum_{j=1}^{n} \nu_{j,t}^{(i)}\left(v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right)0.3$$
$$w_t^{(i)} \sum_{j=1}^{n}\left(v_{A,j,t}^{(i)} -{}_{C,j,t}^{(i)}\right)\left(v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right) = \sum_{j=1}^{n}\left(\nu_{j,t}^{(i)} -{}_{C,j,t}^{(i)}\right)\left(v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right)$$
$$w_t^{(i)} = \frac{\sum_{j=1}^{n}\left(\nu_{j,t}^{(i)} - v_{C,j,t}^{(i)}\right)\left(v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right)}{\sum_{j=1}^{n}\left(v_{A,j,t}^{(i)} - v_{C,j,t}^{(i)}\right)^2}$$

References

1. Mavruk, T. Analysis of Herding Behavior in Individual Investor Portfolios using Machine Learning Algorithms. *Res. Int. Bus. Financ.* **2022**, *62*, 101740. [CrossRef]
2. Ukpong, I.; Tan, H.; Yarovaya, L. Determinants of industry herding in the US stock market. *Financ. Res. Lett.* **2021**, *43*, 101953. [CrossRef]
3. Rahayu, A.D.; Putra, A.; Oktaverina, C.; Ningtyas, R.A. Herding Behavior in The Stock Market: A Literature Review. *Int. J. Soc. Sci. Rev.* **2020**, *1*, 8–25.
4. Bikhchandani, S.; Sunil, S. Herd Behavior in Financial Markets. *IMF Staff. Pap.* **2001**, *47*, 279–310. [CrossRef]
5. Shiller, R. Stock Prices and Social Dynamics. *Brook. Pap. Econ. Act.* **1984**, *15*, 457–510. [CrossRef]
6. Lakonishok, J.; Shleifer, A.; Vishny, R.W. The Impact of Institutional Trading on Stock Prices. *J. Financ. Econ.* **1992**, *32*, 23–43. [CrossRef]
7. Jegadeesh, N.; Titman, S. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *J. Financ.* **1993**, *48*, 65–91. [CrossRef]
8. Jegadeesh, N.; Titman, S. Momentum: Evidence and insights 30 years later. *Pac.-Basin Financ. J.* **2023**, *82*, 102202. [CrossRef]

9. Grinblatt, M.; Titman, S.; Wermers, R. Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior. *Am. Econ. Rev.* **1995**, *85*, 1088–1105.
10. Demirer, R.; Lien, D.; Zhang, H. Industry herding and momentum strategies. *Pac.-Basin Financ. J.* **2015**, *32*, 95–110. [CrossRef]
11. Lin, Z.; Wu, W.; Zhang, H. Momentum, Information, and Herding. *J. Behav. Financ.* **2023**, *24*, 219–237. [CrossRef]
12. Chen, T. Does Country Matter to Investor Herding? Evidence from an Intraday Analysis. *J. Behav. Financ.* **2020**, *22*, 56–64. [CrossRef]
13. Reynolds, C. Flocks, herds and schools: A distributed behavioral model. In Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques, Association for Computing Machinery, Anaheim, CA, USA, 27–31 July 1987; pp. 25–34.
14. Eberhart, R.C.; Kennedy, J. A New Optimizer Using Particle Swarm Theory. In Proceedings of the Sixth International Symposium on Micro Machine and Human Science, Nagoya, Japan, 4–6 December 1995.
15. Shi, Y.; Eberhart, R.C. A Modified Particle Swarm Optimizer. In Proceedings of the IEEE International Conference on Evolutionary Computation, Anchorage, AK, USA, 4–9 May 1998; pp. 69–73.
16. Chen, R.-R. Index Tracking: A Stock Selection Model using Particle Swarm Optimization. *J. Investig.* **2023**, *32*, 53–73. [CrossRef]
17. Chen, R.-R.; Miller, C.D.; Toh, P.K. Modeling Firm Search and Innovation Trajectory Using Swarm Intelligence. *Algorithms* **2023**, *16*, 72. [CrossRef]
18. Chen, R.-R.; Miller, C.; Toh, P.K. Search on a NK Landscape with Swarm Intelligence: Limitations and Future Research Opportunities. *Algorithms* **2023**, *16*, 527. [CrossRef]
19. Chen, R.-R.; Huang, K.; Yeh, S.-K. Particle Swarm Optimization Approach to Portfolio Construction. *Intell. Syst. Account. Financ. Manag.* **2021**, *28*, 182–194. [CrossRef]
20. Chen, R.-R.; Behrndt, T. A New Look at the Swing Contract: From Linear Programming to Particle Swarm Optimization. *J. Risk Financ. Manag.* **2022**, *15*, 246.
21. Chen, R.-R.; Huang, W.; Huang, J.; Yu, R. An Artificial Intelligence Approach to the Valuation of American-style Derivatives: A Use of Particle Swarm Optimization. *J. Risk Financ. Manag.* **2021**, *14*, 57. [CrossRef]
22. Huang, K. *Particle Swarm Optimization Central Mass on Portfolio Construction*; Gabelli School of Business, Fordham University: Bronx, NY, USA, 2019.
23. Waggle, D.; Agrrawal, P. Is the "sell in May and go away" adage the result of an election-year effect? *Manag. Financ.* **2018**, *44*, 1070–1082. [CrossRef]