CENTER FOR TAX AND ACCOUNTING RESEARCH

Center for Risk Management

TAXATION, ACCOUNTING, AND FINANCE

TAF WORKING PAPER

No. 74 / July 2022 / revised June 2023

Social trading, communication, and networks

Deng, Jiaying / Yang, Mingwen / Pelster, Matthias / Tan, Yong

PADERBORN UNIVERSITY

www.taf-wps.cetar.org.

Social trading, communication, and networks*

Jiaying Deng[1], Mingwen Yang[2], Matthias Pelster[3], Yong Tan[4]
Gabelli School of Business, Fordham University, New York, New York, United States[1]
Michael G. Foster School of Business, University of Washington, Seattle, Washington, United States[2,4]
Mercator School of Management, Duisburg-Essen University & european center for financial services (ecfs), Duisburg, Germany[3]

Social trading is an emerging market in the sharing economy, allowing inexperienced investors (followers) to automatically follow the trades of experts (leaders) in real time. We use a separable temporal exponential random graph model (STERGM) to analyze the formation and dissolution of links in a large social trading network. In contrast to traditional social networks, social trading networks are characterized by the rapid dissolution of links, thereby increasing the importance of studying network dissolution. We investigate how social communication, along with financial performance and demographics, affects dynamic network evolution and address the existing dependence in the leader-follower links. The determinants of link formation and dissolution are asymmetric. Different types of social communication, such as posts and comments, have different implications for link formation and dissolution. Our results show that financial performance and demographic characteristics also become important determinants of link formation. However, once a link is formed, followers focus mainly on financial performance, in addition to social communication, and not on demographic characteristics. Thus, our findings have important implications for both investors and social trading platforms.

Key words: Social Trading, Copy Trading, Social Communication, Network Dynamics

1. Introduction

Online engagement among individual investors has grown significantly in recent years. The recent Reddit hype (Pedersen 2022) vividly shows that social media plays an important role in financial markets and in transmitting relevant information to potential investors. Social trading platforms incorporate elements from the worlds of social media *and* online trading and have recently garnered tremendous attention in both research and practice (e.g., Ammann and Schaub 2021, Apesteguia et al. 2020, Yang et al. 2022). Social trading is a novel form of investment that allows retail investors to observe the trading behavior of

* *Information Systems Research*, forthcoming. Data were obtained from a large social trading platform. We thank our data provider for use and explanations of their data. Matthias Pelster gratefully acknowledges financial support by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) – Individual Research Grants project number 434732045. We are grateful for excellent comments and suggestions from the editors and three anonymous referees. Any errors, misrepresentations, and omissions are our own.

other investors and to automatically follow their investment strategies through so-called "copy trading" or "mirror trading" (Apesteguia et al. 2020). An autocopy service enables novice investors (followers) to link their trading accounts with those of expert investors (leaders) and thereby delegate their trading activities (Doering et al. 2015). Investors are able to earn additional income by sharing their trading knowledge with a group of followers. Large social trading platforms, such as eToro, Zulutrade, and FX Junction, have gained popularity, as evidenced by the growing pool of investors on such platforms.[1]

Social trading platforms offer several unique features. First, such platforms offer a very transparent information flow, as (potential) followers are able to see the details of the transactions completed by other investors and track their gains and losses in real time. Second, these platforms allow for straightforward and transparent communication among investors. Investors can share their opinions, publish posts, and leave comments in a news feed that is publicly available to all users. Third, different from mutual fund managers, most participants of social trading platforms are individual traders who lack institutional endorsement.

Social trading platforms also require a new perspective regarding the evolution of social networks, as their network structure follows a different dynamic than that of traditional social networks, such as Facebook and Twitter, which have been extensively studied (Li et al. 2017, Kim et al. 2018). In social trading networks, the links between users are directly tied to cash flows. Individual investors may become leaders who share their trading strategies or may become followers who copy the trading strategies of leaders. Platforms typically share some of their revenue with leaders. As a result of this monetary incentive—compared to other traditional social networks such as Facebook or Twitter—link dissolutions are more frequent in social trading networks. While a link in a traditional social network is commonly characterized by stability and longevity, that in a social trading network is short lived and volatile (Pelster 2017). Thus, not only the process of link formation but also that of link dissolution is crucial. Considering the increasing spread of social trading networks and their economic implications, an extensive understanding of network evolution is important. However, network evolution with a frequent dissolution of links has not yet been studied in detail. Although a large number of studies have focused on the preformation process, i.e., how a social network is formed, none have analyzed the postformation process. Thus, our study fills this void.

[1] https://www.coindesk.com/company/etoro, last accessed Jan. 8, 2021.

We study how the directed leader-follower network on one of the largest social trading platforms evolves over time. Building on the theory of soft and hard information (Liberti and Petersen 2019), we investigate the determinants of link formation and dissolution. Social trading platforms provide a transparent environment in which two types of information with which potential followers can evaluate leaders are released: their *trading activities* (financial performance, i.e., hard information) and their *social activities* (social communication, i.e., soft information). The combination of these data is typically difficult to obtain. For example, in traditional mutual funds, researchers can observe the financial performance of a mutual fund manager but typically lack soft communication information. While some mutual fund managers may have social media channels, for example, YouTube (e.g., Cathie Wood), this is not the case for all managers. In addition, the mutual fund industry allows for private communication between managers and investors that is unobservable to other investors, which may affect investment decisions. In contrast, (most) social trading platforms do not allow for private communication between leaders and followers.[2] Access to various kinds of information on social trading platforms allows us to examine the role of soft and hard information in this innovative form of delegated investment management.

We place a particular emphasis on social network features and study the impact of social communication on link formation and dissolution. While prior studies document that investors chase past financial performance (see, e.g., Barber et al. 2016, Huang et al. 2007), the role of social communication is not clear. Financial performance signals a trader's trading ability. The platform summarizes such information in a highly transparent manner and does not allow users to modify or manipulate the data, thus making them be seen as trustworthy. Social communication provides an additional channel through which leaders can convince potential followers to follow their investment strategies. Given that the primary goal of investors is to make money, followers may focus mainly on financial performance, which provides an objective measure of investment skill, instead of nonmonetary soft information—in particular since, in contrast to the objective features of financial performance, the textual information involved in communication is more complex to interpret and evaluate. Such information requires more time for followers to read through text messages and filter out irrelevant information. The limited attention of (potential) followers

[2] The exception is those rare cases in which leaders and followers may know each other in real life.

may make social communication less effective. In addition, the reliability of social communication is questionable, given that individually disclosed information is not screened by the platform and may constitute "cheap talk". In other words, social communication may not be as trustworthy as financial performance, and the role of social communication may thus remain unclear.

In this study, we build a separable temporal exponential random graph model (STERGM) to disentangle the reasons why a follower follows *and* why she or he unfollows a leader in the social trading context. To capture unobserved heterogeneity and address potential endogeneity concerns, we incorporate Chamberlain correlated random effects (Chamberlain 1980) into the STERGM. We find that financial performance, social communication, and demographic characteristics are important determinants of link formation. However, once a link is formed, followers focus mainly on financial performance and social communication (instead of demographic characteristics) when deciding who to unfollow. We also find that the impact of these factors is asymmetric in the link formation and dissolution processes. Different types of social communication such as posts and comments have different implications for link formation and dissolution. Both the quality and quantity of a leader's posts increase the follower's probability of forming a new link and maintaining an existing link. Followers also rely on "peer reviews": leaders who receive more positive comments are more likely to attract new followers and maintain existing followers. Followers are less likely to form new links or sustain existing links if leaders receive more negative comments. Moreover, the impacts of negative and positive comments are asymmetric; negative comments have a larger impact than do positive comments in both link formation and link dissolution. Overall, we find that social communication plays an important role in leaders' ability to convince potential followers to follow their trading strategies and existing followers to sustain their links.

Our work makes several contributions to the extant literature. First, this study is the first to model leader-follower network evolution on social trading platforms. Unlike those in traditional social networks, relations among investors on social trading platforms involve a monetary dimension, and therefore, social trading features frequent link formation *and* dissolution as investors adjust their investment strategies. Our findings enrich the literature on the determinants of social networks by providing empirical evidence of the evolution process of an innovative network structure. Second, our study contributes to the literature on financial advice. Recent developments in financial technology (FinTech) have made it easier and more convenient for investors to share their trading knowledge and turn to other investors for advice. With the increasing importance of social interactions on financial markets, our results contribute to this stream of the literature by showing that social communication, especially from leaders, can generate economic impacts (i.e., leaders can attract or maintain more followers to earn higher compensation). Third, our study contributes to the literature on individual investor behavior. While financial performance is an important signal of traders' trading skills, we find that followers also rely on communication when evaluating peer traders. We find evidence that both hard information (i.e., financial performance) and soft information (i.e., social communication) play important roles in the link formation and dissolution processes on social trading platforms. Finally, from a methodological perspective, we incorporate Chamberlain random effects into the STERGM to alleviate concerns about confounding effects from individual-level unobserved heterogeneity in the network analysis.

Our paper has important managerial implications. While social trading has some features that are comparable to mutual funds in the sense of "delegated portfolio management" (Doering et al. 2015),[3] the extreme flexibility of followers in dissolving links and thereby terminating their relationship instantaneously brings about large income uncertainty for the leader. Thus, for a leader, a thorough understanding of network evolution and its determinants is crucial. Here, social communication can mitigate information asymmetries and help build trust. Our results on the impact of social communication can provide important guidance for leaders on when and how to communicate with followers. Second, our findings also provide implications for platform providers. Understanding the role of soft information in network dynamics helps platforms understand the importance of their features to enhance their business models. Third, as most recently demonstrated by the GameStop frenzy, vocal leaders on social media may exert a significant influence on financial markets (see, e.g., Pedersen 2022). Thus, a better understanding of the evolution of social networks with an investment focus is important for regulators.

[3] In particular, investors who invest in mutual funds entrust their money to a third party who then makes specific investment decisions for them. This situation is the same in copy trading on social trading platforms: investors entrust their money to a leader, and the leader makes specific investment decisions for them. Due to these similarities, the Markets in Financial Instruments Directive (MiFID) characterizes social trading as a form of portfolio management.

The remainder of this paper is structured as follows. Section 2 summarizes the relevant literature and presents the theoretical background. Section 3 introduces our model. Section 4 describes the data. Section 5 presents the results, and Section 6 presents various robustness checks. Finally, Section 7 discusses the implications and conclusions of the study.

2. Theoretical background

While the literature has studied social networks in detail (Kane et al. 2014), it is unclear how insights from other social networks can be applied to social trading platforms. Social networks on social trading platforms differ from traditional social networks. The relations between users in social trading networks are directly tied to cash flows. The platform provides a service that allows for "copy trading"—duplicating investment strategies from other investors with one's own money, without approving each individual transaction—by its customers. Followers who make use of the social features of the platform and form a copy trading link with other investors can delegate their trading; at the same time, leaders can earn additional income, receiving compensation from the platform. As a result, compared to other traditional social networks such as Facebook or Twitter, link dissolution occurs more frequently (Pelster 2017). As the prior literature has focused mainly on link formation, link dissolution has received little attention in the literature—partially because link dissolution is a relatively rare event in many traditional social networks.

2.1. Social trading

Our study contributes to the fast-growing literature on social trading, which, generally, can be divided into three streams. The first stream addresses the institutional aspects of social trading (see, e.g., Doering et al. 2015). A key feature of social trading platforms is a high level of information transparency (Gemayel and Preda 2018a). Investors can observe the trading behavior of their peers at the trade level and in real time. Considering that some investor's trades may contain valuable information, making these trades available in real time potentially undercuts platforms' payoff potential. To resolve this issue, Yang et al. (2022) propose a personalized trade-level information release policy that allows platforms to optimize their revenue.

A large second group of studies examines how the information transparency that allows investors to observe other investors in real time may affect their trading behavior and the performance implications of social trading. Gemayel and Preda (2018a,b) label the state of permanent reciprocal observation and scrutiny that are typical of social trading platforms as a "scopic regime". Trading in a scopic regime alters investors' behavioral biases such as through the disposition effect (Heimer 2016, Pelster and Hofmann 2018).[4] Focusing on the copy trading feature, the experimental study by Apesteguia et al. (2020) suggests that having the option to directly follow other investors significantly increases the risk-taking behavior of investors. This increased risk taking does not, however, yield superior investment returns (e.g., Pan et al. 2012). Several studies find that on average, traders on social trading platforms do not outperform the market in the long term (Dorfleitner et al. 2018, Oehler et al. 2016). Only a few investors can earn significant short-term excess returns (Dorfleitner et al. 2018, Oehler et al. 2016).

Most closely related to our paper, the last stream of literature studies network relationships on social trading platforms. Ammann and Schaub (2021) find that social traders are more likely to have their investment strategies duplicated within three weeks of making (positive) posts on trading platforms. These posts do not, however, seem to contain valuable information, as they do not have any predictive power over future performance. While Ammann and Schaub (2021) focus on fund flows in their analysis, in contrast, we focus on individual activity-based relationships between traders, i.e., links within the network. Additionally, focusing on fund flows, Röder and Walter (2019) document a positive relationship between flow and performance in social trading portfolios, which is limited to the top past performers (i.e., investment flows chase past performance). To optimally exploit the copy trading function, Lee and Ma (2015) develop a system called "whom to follow (W2F)" that enables the users of social trading platforms to "discover expert traders" who consistently realize high risk-adjusted performance.

2.2. Link formation and dissolution in (social) networks

The evolution of leader-follower networks or, in other words, when and why an investor (follower) follows or unfollows another investor (leader) on a social trading platform is one of the fundamental questions involving social trading. Prior research has mostly focused on the trading behaviors of leaders in this context but has largely ignored followers' decisions

[4]The disposition effect is an anomaly discovered in behavioral finance that is related to the tendency of investors to sell assets that have increased in value while holding on to assets that have dropped in value (Shefrin and Statman 1985).

about whom to follow or unfollow. In this study, we aim to shed light on this question by focusing on the factors that drive the following and unfollowing decisions of followers on social trading platforms. First, we review the related literature that discusses the determinants of tie (link) formation and dissolution in other (related) settings and discuss social-trading-specific features that may influence tie formation and dissolution.

Trust. It is widely acknowledged that a good relationship between managers and investors is beneficial for both parties. Trust is important for managing relationships and regulating their quality (Kaiser and Berger 2021). Trust is relevant when a person (trustor) has specific expectations of another person (trustee) and is vulnerable to whether the trustee fulfills those expectations, regardless of the degree of control (Mayer et al. 1995). Various factors can influence the existence of trust in a relationship. As noted by Mayer et al. (1995), trustworthiness must be established before any factor can lead to trust. Trustworthiness refers to the trustee and relates to his or her specific characteristics, for example, benevolence, integrity, or ability. The most common factor that can establish trustworthiness and ultimate trust is communication (Kaiser and Berger 2021). Communication and timely feedback increase trust. Moreover, other factors that establish trust are reputation, quality, and partner fit (Kaiser and Berger 2021).

Recent developments in FinTech have led to new challenges in managing relationships, particularly because investors cannot establish a personal one-on-one trust relationship and must instead seek to build a relationship with a more or less anonymous mass (Kaiser and Berger 2021). Trust among online community members plays an even more important role in the online trading context, where investors can automatically, without further evaluation, duplicate the investment strategies of their peers (Wohlgemuth et al. 2016).

Social communication. Social communication can help build trust and reduce information asymmetries (Xu and Chau 2018). Duarte et al. (2012) show that borrowers who appear to be more trustworthy based on their pictures have a higher probability of having their loans funded. Moreover, Xu and Chau (2018) find that both credit grades and lender-borrower communication affect funding outcomes on peer-to-peer lending platforms. In particular, social communication can be a tool that particularly allows listers with a poor degree of hard information (e.g., credit grades) to improve their chances of being funded (Xu and Chau 2018).

Social communication can be regarded as soft information, which refers to qualitative information such as media press, communication texts, or market commentary, whereas hard information refers to quantitative information such as stock returns or credit ratings. Even though soft information is usually qualitative, it can be "hardened" using information technologies such as text mining and converted into a quantitative measure. Thus, the main difference between hard and soft information is that the former can be objectively verified and is independent of context, while the quantification of the latter makes use of various degrees of freedom (Liberti and Petersen 2019). An immediate consequence of the nature of soft information is that its assessment (e.g., how trustworthy another agent is or other informational cues are) depends on each agent's personal standards. When evaluating information from others, the quality of information is important for building trust (Xu and Chau 2018). We hypothesize that social communication is important for investors in establishing trust on social trading platforms. Thus, we incorporate social communication variables, including a proxy for the quality of social communication, and explore how they affect link formation and dissolution.

Financial performance. As noted above, hard information is also an important determinant of whether investors decide to fund a project, sell a stock, or provide a loan (Liberti and Petersen 2019). Social trading platforms are comparable to mutual and hedge funds, as they allow for some form of delegated portfolio management (Doering et al. 2015). Both mutual and hedge funds have received considerable attention in the financial literature, with a particular focus on the determinants of their performance (e.g., Agarwal et al. 2009, Grinblatt et al. 2020) and the relationship between fund performance and (net) fund flows (Sirri and Tufano 1998, Goetzmann et al. 2003). As investors can infer the skills, at least to some degree, of mutual and hedge fund managers from their past performance, (net) fund flows should thus be explained by past performance (Barber et al. 2016). This stream of literature also documents not only that mean performance is important but also that the volatility of performance is negatively related to fund flows (Sirri and Tufano 1998, Huang et al. 2007). Importantly, investors seem to determine their inflows and outflows in different ways (Ivković and Weisbenner 2009). The importance of returns has also been documented in nonprofessional settings, such as online crowdfunding markets (Lin and Viswanathan 2016). Based on this stream of literature, we examine how the formation and dissolution of leader-follower links are affected by financial performance.

Demographic factors and homophily. A large stream of the social networks literature documents that similarity (homophily) breeds connections (McPherson et al. 2001). Homophily describes the preference of people to favor others who are similar to them rather than those who are dissimilar to them. In financial markets, so-called home bias is a prominent example of a preference for similarity (Coval and Moskowitz 1999, Lin and Viswanathan 2016). In the social trading context, potential cultural differences or language barriers in online communications may also contribute to homophily. In addition, the profile picture (image) and biography on an investor's public profile page can affect how he or she is perceived by others (Duarte et al. 2012).

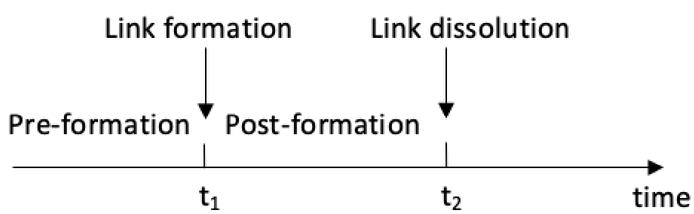
Dissolution of ties. Even though link dissolution happens less frequently in traditional social networks such as Facebook or Twitter, the dissolution of ties in interfirm networks or financial markets is a rather frequent event. Studying interfirm networks, Greve et al. (2010) argue that the dissolution of ties may happen particularly when embeddedness is low. Polidoro et al. (2011) further study the importance of embeddedness for tie dissolution and argue that network centrality, i.e., positional embeddedness, does not promote stability but that having common partners, i.e., structural embeddedness, does. Translated to social trading, where common partners do not play a meaningful role and where there are almost no costs to dissolve a tie and engage in a new relationship (aside from transaction costs, which are charged via the spread), these findings suggest that ties could be dissolved rather quickly. Furthermore, Baker et al. (1998) argue that unsatisfactory performance is a main driver of tie dissolution—an argument that can be translated to social trading at face value.

Shafi et al. (2020) study the dissolution of ties in early entrepreneurial finance and argue that tie discontinuation can have important ripple effects on other ties. In particular, once well-established investors cut their ties with a start-up, smaller investors may follow suit. Thus, discontinuation may have important repercussions for start-ups (Shafi et al. 2020). In social trading, despite the high level of transparency, investors face some uncertainty when deciding to follow other investors. As commonly stated regarding delegated investment opportunities, "Past performance is no guarantee of future results". Consequently, investors may not be satisfied with the outcomes of a given tie and decide to discontinue the relationship, or in the words of Shafi et al. (2020), "The decision to withdraw financial support may be primarily related to a venture's underperformance". Based on this stream of literature, we hypothesize that financial performance, or hard information, is a key factor determining link dissolution. In addition, the role of soft information in link dissolution is unclear. We thus investigate how hard and soft information affect link dissolution.

3. Model

Due to the nature of social trading networks, particularly the frequent link formation and dissolution in such networks, it is important to study both link formation and dissolution. Figure 1 illustrates a typical link formation and dissolution process on social trading platforms. A link between a follower and a leader is formed at t_1 and dissolved at t_2. Following a leader is equivalent to (automatically) copying the trading strategy of the leader. A link is formed when a follower follows a leader, and the link is dissolved when the follower stops following the leader. The network is constructed through leader-follower links. We use the word "follower" to maintain consistency with the prior literature (Ammann and Schaub 2021, Yang et al. 2022).

Figure 1 **Illustration of link formation and dissolution**



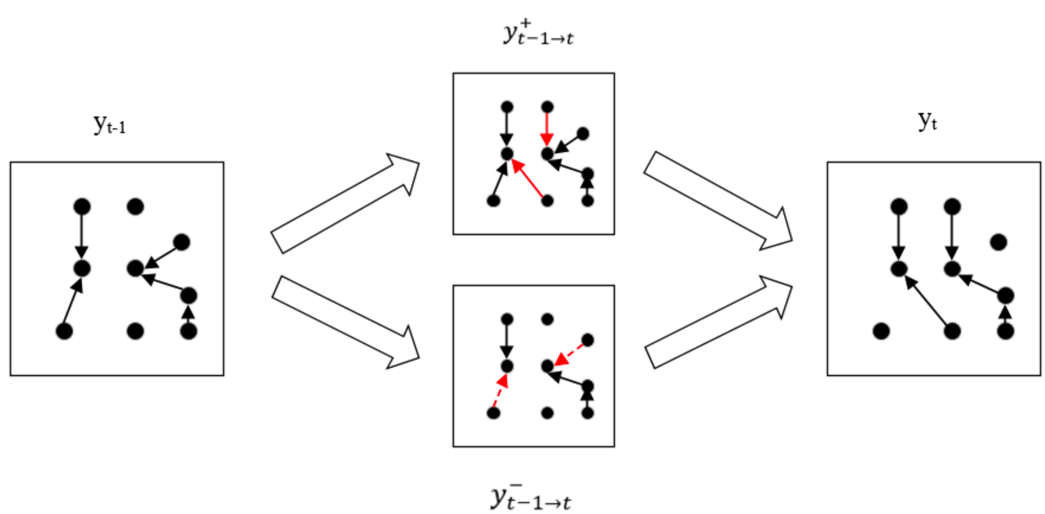
3.1. STERGM

We use extensions of the exponential random graph model (ERGM) (Snijders et al. 2006, Robins et al. 2007) to model network evolution. ERGMs represent a general class of models based on exponential family theory that can be used to specify the probability distribution underlying a set of random graphs or networks (Robins et al. 2007, Snijders et al. 2006) and are widely used for network analyses in the field of information systems (Yan et al. 2015, Hwang et al. 2022). An ERGM aims to identify the factors that affect link formation in a network by comparing the probability of the realized network structure with all alternative network configurations. However, the conventional ERGM neither accounts for the intertemporal dependence in longitudinally observed networks nor models the link dissolution process. In this study, we adopt the STERGM (Krivitsky and Handcock 2014), an extension of the ERGM, to model the network dynamics that allow us to capture both intertemporal dependence and the link dissolution process.

We consider dynamic leader-follower networks with a total of T time periods. At time period t, suppose that there are N_t nodes, and let Y_t be an $N_t \times N_t$ adjacency matrix for a random network. $y_{ijt} = 1$ indicates a link between nodes i and j at time t, and $y_{ijt} = 0$ indicates that there is no link between these nodes at time t. We define $\mathcal{Y}_t$ as the set of all possible networks among the nodes and y_t as a realized network for $y_t \in \mathcal{Y}_t$ at time t.

Figure 2 illustrates a visualization of directed network changes from time $t-1$ to t. Realized networks at times $t-1$ and t are denoted as y_{t-1} and y_t, respectively. We define two networks to track the network evolution: *formation network* y^+ and *dissolution network* y^-. $y^+_{t-1\to t}$ is defined as network y_{t-1} plus the links established from time $t-1$ to t. Similarly, $y^-_{t-1\to t}$ is defined as network y_{t-1} minus the links dissolved from time $t-1$ to t. In our illustration, two new links are added (denoted by red solid arrows), and two existing links are removed (denoted by red dashed arrows). Thus, we are able to track the network evolution in terms of links from time $t-1$ to t. Although we observe only networks y_{t-1} and y_t, we can recover $y^+_{t-1\to t}$ and $y^-_{t-1\to t}$ since $y^+_{t-1\to t} = y_{t-1} \cup y_t$ and $y^-_{t-1\to t} = y_{t-1} \cap y_t$, respectively. Appendix A presents a detailed description of how we track network evolution.

Figure 2 **Visualization of network changes from time $t-1$ to t**



Mathematically, the formation process is modeled as

$$\mathrm{P}(Y^+_{t-1\to t} = y^+_{t-1\to t} | Y_{t-1} = y_{t-1}; \theta^+) = \frac{e^{(\theta^{+\prime} g^+(y^+_{t-1\to t}, X_{t-1}))}}{\kappa(\theta^+, X_{t-1}, \mathcal{Y}^+(y_{t-1}))}, \tag{1}$$

and the dissolution process is modeled as

$$\mathrm{P}(Y^-_{t-1\to t} = y^-_{t-1\to t} | Y_{t-1} = y_{t-1}; \theta^-) = \frac{e^{(\theta^{-\prime} g^-(y^-_{t-1\to t}, X_{t-1}))}}{\kappa(\theta^-, X_{t-1}, \mathcal{Y}^-(y_{t-1}))}, \tag{2}$$

where $g^+(y^+_{t-1\to t}, X_{t-1})$ $(g^-(y^-_{t-1\to t}, X_{t-1}))$ is the vector of model covariates for formation network $y^+_{t-1\to t}$ (dissolution network $y^-_{t-1\to t}$) and θ^+ (θ^-) is the vector of coefficients for network $y^+_{t-1\to t}$ $(y^-_{t-1\to t})$. The denominators in Equations (1) and (2) are normalizing factors that represent the sum of the numerator over all possible networks to ensure that the probability of observing the realized formation (dissolution) network is between 0 and 1. Mathematically, the factor is defined as follows:

$$\kappa(\theta^+, X_{t-1}, \mathcal{Y}^+(y_{t-1})) = \sum_{z^+ \in \mathcal{Y}^+(y_{t-1})} e^{(\theta^{+\prime} g^+(z^+, X_{t-1}))} \tag{3}$$

and

$$\kappa(\theta^-, X_{t-1}, \mathcal{Y}^-(y_{t-1})) = \sum_{z^- \in \mathcal{Y}^-(y_{t-1})} e^{(\theta^{-\prime} g^-(z^-, X_{t-1}))}, \tag{4}$$

where z^+ (z^-) denotes a possible formation (dissolution) network from time $t-1$ to t.

3.2. Identification

A dynamic network analysis of thousands of nodes requires significant computing resources and is computationally intractable (Yan et al. 2015). Thus, we adopt a degenerate statistical model to estimate the coefficients in the link formation and dissolution processes, similar to maximum pseudolikelihood estimation (Strauss and Ikeda 1990).

A common issue in network analysis is endogeneity. First, we use lagged independent variables to mitigate potential reverse causality. Second, in our context, the information provided on the platform is highly transparent. We observe the information that is observed by followers on the platform, which may affect link formation and dissolution. We have access to the complete transaction history and social communications of each investor, as well as rich demographic information; the platform does not allow for a private chat channel. We construct various covariates, including follower characteristics, leader characteristics, homophily, and network structure, as elaborated in a later section. However, some determinants that explain link formation and dissolution may still be unobserved, at least by researchers. For example, when followers make their decisions, their investment goals on the platform and their intrinsic trust in others might affect their link formation and dissolution. Hence, to mitigate the concern of omitted variables, we control for follower-specific unobservables (η_i) in the link formation and dissolution model. Specifically, for the link formation process,[5] we define

$$y_{ijt} = \begin{cases} 1, & y^*_{ijt} > 0; \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

y_{ijt} is a binary variable that is equal to 1 if follower i forms a link with leader j from period $t-1$ to t, and y^*_{ijt} is the corresponding latent utility. The utility of follower i due to forming a link with leader j at time t is defined as follows:

$$y^*_{ijt} = \boldsymbol{\alpha} X_{it-1} + \boldsymbol{\beta} W_{jt-1} + \boldsymbol{\lambda} V_{ij} + \boldsymbol{\tau} C_i + \eta_i + \epsilon_{ijt}, \tag{6}$$

where X_{it-1} is a vector of follower i's time-variant covariates at period $t-1$, W_{jt-1} is a vector of leader j's time-variant covariates at period $t-1$, V_{ij} is a set of dummies indicating whether follower i and leader j share the same demographics (homophily), C_i is a set of follower-specific time-invariant observable controls, and η_i is the follower-specific fixed effects. $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\lambda}$, and $\boldsymbol{\tau}$ are the corresponding vectors of the coefficients to be estimated.

A conventional approach to estimating fixed effects is to treat η_i as a parameter and use maximum likelihood estimation. However, such estimation is inconsistent when the number of nodes is large and the number of periods is finite, which denotes the incidental parameter problem (Neyman and Scott 1948). To correct this problem, Chamberlain (1980) proposes a correlated random effects model (Wooldridge 2010). We control for follower-specific Chamberlain correlated random effects in our model.[6] Chamberlain (1980) allows for follower-specific unobservables to be correlated with independent variables. Specifically, we implement Chamberlain correlated random effects following Mundlak (1978). η_i is defined as

$$\eta_i = \psi + \boldsymbol{\xi} \bar{X}_i + a_i, \tag{7}$$

where a_i follows a normal distribution with a mean of zero and variance of σ_a^2, ψ is a constant, and $\bar{X}_i$ is the time average of the follower's time-variant observables. $\bar{X}_i =$

[5]The link dissolution process is defined in the same fashion. In the dissolution process, y_{ijt} is equal to 1 if follower i maintains the link with leader j in period t.

[6]We also estimate link formation and dissolution using an alternative estimation approach, the conditional logit estimator. The estimation results remain consistent with the results from the main model (see Appendix G).

$(\Gamma_i)^{-1}\sum_{t=1}^{\Gamma_i} X_{it}$, where Γ_i equals the number of periods in which follower i exists on the platform multiplied by the number of leaders that follower i follows in each period.

In Equation (7), a_i is independent of X_i, and the model allows for dependence between η_i and X_i by adding $\bar{X}_i$ to the equation. From Equation (7), we see that η_i follows a conditional normal distribution, that is, $\eta_i|X_i \sim Normal(\psi + \bar{X}_i\xi, \sigma_a^2)$. Thus, unlike the conventional fixed effects model, the coefficients on the follower-specific time-invariant controls C_i in Equation (6) can be identified.

Although no significant regulatory changes are made to social trading during our sample period in general, link formation and dissolution may be affected by some other time-dependent events, for example, some policy changes on the platform. To mitigate this concern, we estimate a model with time-level fixed effects, which allow us to control for time-specific peculiarities. The estimation results remain consistent with the results from the main model (see Appendix B).

Finally, despite the leader characteristics included in the model, there may still be some unobservables that affect link formation and dissolution. For example, it is possible for leaders on the platform to advertise themselves via other social media platforms. To mitigate omitted variables on the leader's side, we also include leader-specific Chamberlain correlated random effects in the link formation model. The estimation results remain generally consistent with the results from the main model (see Appendix C).

4. Data and variables

In this section, we first introduce the data and then describe how we construct the variables used in the analysis.

4.1. Data

We obtain our data from eToro, the largest social trading platform. Similar to other online trading brokerage services, this platform allows its customers to trade stocks, commodities, currency pairs, and crypto-assets. For each trade, the platform charges transaction fees as a portion of the bid-ask spread. In addition, the platform incorporates various features that are typical of social media. Specifically, the platform contains a news feed in which investors can disclose their trading activities (*open book trading*) and publish posts. Here, investors can conveniently discuss their trading strategies, like and comment on others' trading activities, and automatically copy those trades of other investors. Investors who have their trades copied receive monetary compensation from the brokerage service in relation to their number of followers, the amount of their assets under management, and their degree of investment performance, similar to professional fund managers. Each investor has a public profile page, which shows detailed and transparent information on his or her past trading activities, including financial performance, social activities (e.g., posts, comments, likes, and replies), and number of followers.

Our data cover the complete social and trading activity histories of all investors in 2016 and 2017. Social activity histories include all posts, comments, replies, and likes, together with the exact timestamp of each activity. Trading activity histories include detailed information on each trade. In addition, the data include the dynamics in the leader-follower networks. For each link in the network, we know the exact timestamp of the formation and dissolution of the link between the follower and the leader. Finally, the data include each investor's nationality, age, gender, use of a profile image, publication of a biography, trading experience before joining the platform, and desired risk level upon registration.

We consider each investor as a node in the network. If an investor (follower) follows or autocopies another investor (leader), then this relation is modeled as a directed link between the follower and the leader. We set $y_{ijt} = 1$ if a link exists between nodes i and j in period t. We use the data from 2016 to proxy for historical trading performance (e.g., average profit and standard deviation of profit) to guarantee a long-term horizon on which followers can evaluate leaders. We examine link formation and dissolution using the leader-follower network in 2017. One period denotes one month.

We first illustrate how we sample investors in period 1. We select all leaders who have at least 5 followers (to alleviate the sparsity of the network and exclude some casual investors) in period 1 and stay on the platform for two successive months (i.e., periods 1 and 2), ending with a total of 462 leaders.[7] We then obtain information on all the followers of these leaders, ending with a total of 13,533 unique followers who exist during these two successive periods. As a large number of nodes can cause computational intractability issues in network analyses, we randomly sample 600 followers out of 13,533, resulting in 1,057 unique investors (because some investors may be both followers and leaders, the total

[7]Existence for two successive months is the minimum requirement because the formation network (dissolution network) is constructed by tracking the links added (removed) between two successive periods.

number of investors is less than the sum of leaders and followers).[8] In period 2, as some investors sampled from period 1 may exit the platform and some new investors may join the platform, we first keep those investors who remain on the platform for two successive months (i.e., periods 2 and 3). Then, using the number of investors in period 1 as an anchor, we add investors by randomly sampling from the new investors who join the platform in period 2 and stay on the platform during both periods 2 and 3. We repeat this procedure across all periods. Table 1 summarizes the network statistics: the number of nodes, number of links, and network density (the proportion of links in a network relative to the total number possible).

Table 1 Network dynamics

Period	Nodes	Links	Density
1	1,057	1,595	0.0014
2	1,053	1,588	0.0014
3	1,053	1,658	0.0015
4	1,053	1,703	0.0015
5	1,053	2,025	0.0018
6	1,054	2,012	0.0018
7	1,055	1,923	0.0017
8	1,055	1,857	0.0017
9	1,056	1,832	0.0016
10	1,056	1,809	0.0016
11	1,057	1,742	0.0016
12	1,057	1,250	0.0011

From Table 1, we see that the number of unique nodes varies slightly across different periods because the numbers of traders who are both followers and leaders are different in each period. Note that although the number of nodes is approximately 1,057 in our network analysis, the leader-follower relations between the nodes are described by a 1,057×1,057-dimensional matrix in each period, and there are 12 periods in total. In addition, our network sampling process accounts for the scenario in which new investors join and existing investors quit the platform since we resample the nodes every period. Therefore, the total number of unique nodes across all periods is 2,737. To mitigate concerns about our sample

[8] It is a common practice to sample a smaller set of nodes to achieve computational feasibility when estimating network analysis models (Yan et al. 2015, Lee et al. 2016). For example, Lee et al. (2016) study strategic network formation in a location-based social network. Their network analyses are conducted on three city-level subsamples consisting of 336, 129, and 146 users. Moreover, Yan et al. (2015) examine the driving forces behind patients' social network formation and evolution using a subsample consisting of 1,322 individuals.

selection, we take a second random sample and reestimate our model in Appendix D. Our results are generally consistent with those for the alternative sample.

4.2. Variables

Based on the theoretical background presented in Section 2, we consider different variables that may affect the dynamics of leader-follower networks.

4.2.1. Variables of interest

Social communication An important feature of social trading platforms is that investors are able to conveniently interact with other investors. For example, eToro allows its users to publish posts, comment on posts, and distribute likes. Investors can publish posts to broadcast their recent achievements, explain their trading strategies, share their financial advice, or simply communicate with others about recent events. Other investors can comment on these posts to voice their opinions, request additional information, or ask for clarification regarding comments.[9] Investors can leave another comment to a comment that is made on their original post; we label these types of comments as replies to distinguish them from original comments. All social interactions are shown on the platform news feed and in the investor's public profile, similar to typical social media platforms such as Facebook or Twitter. eToro does not provide its users with the ability to chat privately. Consequently, the news feed is the only way that users can communicate with each other, and all social activities on the platform are public. We show examples of different types of social communication (posts, comments, and replies) in Appendix E.

We measure investors' social activities using the following variables. For each investor, we use the total number of posts over period t to measure the intensity of posting (*post quantity*). Following the literature (Khern-am-nuai et al. 2018, Cao et al. 2011), we use the number of likes that a post receives to measure its quality. We take the average over all posts in period t as a proxy for an investor's *post quality*. Considering prior evidence that sentiments in user-generated content play an important role in agents' decision-making processes (Xu and Chau 2018), we do not simply examine the number of comments but instead focus on the sentiments expressed in those comments. Due to the international

[9] We do not observe automated comments in our data and are confident that investors manually post these comments. In addition, we do not work under the assumption that all users read all comments. In fact, given the substantial number of comments and the limited attention of investors (Hirshleifer and Teoh 2003), we believe this situation to be very unlikely. However, if comments are not being read, then their impact on relationships should be zero and nonsignificant, which would be reflected in our estimation results.

customer base of the platform, comments are posted in different languages. Thus, to conduct sentiment analysis, we first use a Google Cloud translation application programming interface (API) to translate all comments into English.[10] We then remove stop words, perform word stemming, and use lexicon-based content analysis to perform our sentiment analysis. We implement the Valence Aware Dictionary and sEntiment Reasoner (VADER), specifically attuned to sentiments expressed on social media (Hutto and Gilbert 2014). The VADER has recently been applied in finance and trading and performs as well as individual human raters in terms of matching ground truth (Hutto and Gilbert 2014). The package enables us to label each comment with positive and negative sentiment scores by calculating the percentage of the text that falls into each category. Then, we average the positive and negative sentiment scores for all comments in period t and obtain the variables *comment positive* and *comment negative*, respectively.[11] A higher positive (negative) score indicates that a comment contains a larger percentage of positive (negative) words. Finally, we count the replies provided by investors to their received comments in period t using *reply*.

Financial performance To measure an investor's financial performance, we first calculate his or her daily profit. Because statistics from only one period (month) may not reflect the investor's overall performance, we use their historical average daily profit until period t to measure their *average profit*. Similarly, we calculate the standard deviation of historical daily returns until period t as a proxy for investors' volatility (in line with, e.g., Sirri and Tufano 1998, Huang et al. 2007).[12]

4.2.2. Control variables

Trading strategies We construct the following variables to proxy for investors' trading strategies. *Holding time* measures the duration from the opening to the closing of a particular position (with the unit of the day), reflecting the extent to which a trader prefers "day trading" versus a buy-and-hold strategy. We account for investors' portfolio features using

[10] This procedure is consistent with the practices of the platform, which provides a "translate" icon for all posts and comments, allowing users to view them in English.

[11] As Tirunillai and Tellis (2012) finds that the effect of negative and positive user-generated content is asymmetric, we include both positive and negative comments in the model.

[12] The standard deviation is the most common measure of risk in the field of finance (Holzmeister et al. 2020), as it is widely used in textbooks to teach financial basics or in regulation. As such, the standard deviation is part of many basic models. For example, it is used in Markowitz's famous portfolio theory. Hence, the standard deviation or volatility provides an easily interpretable and widely understood measure of risk. Therefore, we use the standard deviation to measure the risk of financial performance.

the Herfindahl–Hirschman index (*HHI*), a measure of diversification based on the sum of squared portfolio weights (Dorn et al. 2008). A smaller HHI indicates a more diversified portfolio. We include a measure of investors' preferences for investing in lottery-type stocks following Kumar (2009), based on the observation that retail investors are attracted to lottery stocks (i.e., stocks with positively skewed returns) and that this level of attraction to lottery stocks can increase as a result of social interactions, even if investors do not have inherent preferences for skewness (Han et al. 2022). In this vein, Bali et al. (2021) show that social interactions aggravate the lottery anomaly. We define *lottery preference* as the fraction of trades executed by a given investor in lottery-type stocks relative to all of his or her trades.

Demographics We include some variables to control for homophily based on demographic characteristics. First, we use a dummy variable that takes a value of one if investors come from the same country and zero otherwise (*Nationality*). In a similar fashion, we control for homophily along the investor gender dimension (*Gender*). We also construct a dummy variable to indicate whether investors are in the same age range (*Age*). Social investors may also generate trust by having a detailed profile page that includes their image and/or a biography (Wohlgemuth et al. 2016). Consequently, we include two dummy variables for leaders, *Image* and *Bio*, to denote whether a profile picture[13] or a biography is provided on the investor's profile page. Finally, we incorporate investor characteristics upon registration, including trading experience in years before joining the platform (*Experience*), total wealth in dollars (*Wealth*), annual income in dollars (*Income*), and the reported desired risk level (*Risk*), to capture potential heterogeneity.

Network structure In addition to node characteristics and dyadic covariates, the network structure may affect network evolution through reciprocity and transitivity (Wasserman et al. 1994, Holland and Leinhardt 1971). In our data, the number of mutual links (i.e., $i \rightarrow j$ and $j \rightarrow i$) is zero, as it is unlikely that leaders will follow their followers' trading strategies in a social trading network. Therefore, we do not consider reciprocity in our study. However, we incorporate a triadic term to capture potential transitivity. When links $i \rightarrow j$ and $j \rightarrow k$

[13]Our dataset does not contain more detailed information, for example, whether it is a symbolic image or a photo showing a real person, on the picture. Due to the anonymous nature of the data, we are not able to collect this information and merge it with our dataset. We study a random sample of the profile pictures of eToro users to analyze how many fantasy pictures (i.e., symbolic images), on average, are used. Our analysis of slightly over 500 randomly selected profile pages shows that approximately 80% of them contain photos showing a real person (who is not a well-known celebrity).

Table 2 Data description and statistics

Variable	Description	Mean	Std. Dev.	Min.	Max.
Post quantity	Number of posts made by investors	23.77	146.39	0	5406
Post quality	Number of likes that are received by an investor's posts	4.28	9.59	0	236
Comment positive	Percentage of positive words in those comments received by an investor	0.31	0.21	0	1
Comment negative	Percentage of negative words in those comments received by an investor	0.06	0.08	0	1
Reply	Number of replies to comments	20.57	62.49	1	1551
Average profit	Average profit	0.0008	0.04	-0.55	0.97
Std. dev. profit	Standard deviation of profit	0.05	0.11	0	3.11
MDD	Maximum drawdown (MDD) of profit	16.23	60.54	0	450.88
Holding time	Duration of a particular position from opening to closing in days	14.00	38.23	0	749.21
HHI	HHI of portfolio diversification	0.19	0.31	0	1
Lottery preference	Fraction of trades in lottery-type stocks	0.02	0.06	0	1
Gender	Dummy =1 if both investors are females or both are males and 0 otherwise	0.83	0.38	0	1
Age	Dummy =1 if both investors are in the same age range and 0 otherwise	0.29	0.45	0	1
Nationality	Dummy =1 if both investors are from the same country and 0 otherwise	0.07	0.26	0	1
Image	Dummy =1 if the investor uploads a profile picture and 0 otherwise	0.48	0.50	0	1
Bio	Dummy =1 if the investor uploads a biography to his or her profile and 0 otherwise	0.22	0.41	0	1
Experience	Trading experience before joining the platform at the time of registration	1.53	1.07	0	3
Wealth	Reported wealth at the time of registration	105,541	251,475.2	10,000	2,000,000
Income	Reported annual income at the time of registration	113,644	209,495.7	10,000	2,000,000
Risk	Reported risk at the time of registration	24.35	15.41	3	48
In-degree popularity	Number of incoming ties in an investor's social network	1.70	10.68	0	371
Out-degree activity	Number of outgoing ties initiated by an investor in his or her social network	1.70	2.18	0	43
Transitivity	Number of triadic closures for each node	0.21	1.42	0	42

exist, the likelihood that a new link, $i \to k$, will be formed may increase. While triadic effects represent the local hierarchy within the network, we also incorporate the global hierarchy among all nodes within the network—node-level in-degree-related popularity and out-degree-related activity (Hunter et al. 2008). Because it is possible that the effects of popularity and activity are different for leaders than for followers, we distinguish between leader and follower nodes.

We provide brief variable definitions and summary statistics in Table 2. For in-degree popularity and transitivity, over 80% of observations are found to be zero, and for out-degree activity, approximately 25% of observations are zero. As a result of the unique social trading context (the connection of nodes to cash flows), compared to other traditional social networks such as Twitter or Facebook, social trading networks do not have many links among nodes (as reflected in degrees and transitivity in the network structure). In other words, a social trading network is sparse. We report a correlation matrix in Appendix Q. The correlations between our variables are not strong. Hence, possible multicollinearity is less of a concern in our analysis.

5. Results

We apply the STERGM with Chamberlain correlated random effects to investigate link formation and dissolution in social trading. We study the determinants presented in Section 4.2 and estimate the coefficients that best fit our model using a maximum likelihood procedure. We provide additional robustness checks in Section 6 and in the Appendix.

Table 3 summarizes our main estimation results separately for link formation and dissolution. To account for the skewness of the data, we take the logarithm of the post quantity, post quality, number of replies, wealth, and income. We distinguish between the variables for a leader's account and those for a follower's account since they may play different roles in network evolution. Overall, we find evidence that hard information (financial performance) and soft information (social communication and demographic characteristics) play different roles in the link formation and dissolution processes.

Table 3 Estimation results

Variable	Formation		Dissolution	
Leader's post quantity	0.0425***	(0.0033)	0.0151***	(0.0042)
Leader's post quality	0.5384***	(0.0253)	0.0596***	(0.0221)
Leader's number of replies	0.1428***	(0.0208)	0.0400*	(0.0231)
Leader's comment received a positive score	1.4242***	(0.1266)	0.6735***	(0.1421)
Leader's comment received a negative score	−3.7892***	(0.7124)	−2.0273***	(0.4735)
Leader's average profit	0.1051***	(0.0138)	0.0785***	(0.0190)
Leader's std. dev. profit	−3.0988***	(0.5702)	−3.0560***	(0.7937)
Controls				
Leader's average holding time	0.3324***	(0.0445)	0.0477	(0.0657)
Leader's lottery preference	0.2924	(0.3562)	1.2306***	(0.3967)
Leader's HHI	−0.6671***	(0.0903)	−0.1384	(0.0929)
Follower's post quantity	0.0158	(0.0281)	−0.2781***	(0.0412)
Follower's post quality	−0.2556***	(0.0829)	0.0062	(0.0793)
Follower's average profit	0.0281	(0.0183)	0.0374*	(0.0193)
Follower's std. dev. profit	−3.4719***	(1.1279)	−8.3969***	(1.1207)
Nationality	0.7436***	(0.0705)	0.3422***	(0.0798)
Age	0.1161**	(0.0507)	0.0470	(0.0543)
Homophily (male)	0.9879***	(0.1360)	−0.1393	(0.1165)
Homophily (female)	−0.8602**	(0.4166)	0.2718	(0.3177)
Image	2.5061***	(0.4232)	0.4876	(0.5451)
Bio	3.0236***	(0.1871)	−0.2172	(0.2018)
Experience	−0.0689*	(0.0363)	0.1472***	(0.0406)
Wealth	0.0044	(0.0315)	0.0661*	(0.0347)
Income	−0.0271	(0.0380)	0.0219	(0.0410)
Risk	−0.0949**	(0.0471)	−0.0351	(0.0527)
Leader's popularity	0.0073***	(0.0004)	0.0011***	(0.0004)
Leader's activity	−0.0409***	(0.0123)	0.0391***	(0.0120)
Follower's popularity	−0.0591***	(0.0150)	−0.0056	(0.0093)
Follower's activity	−0.0264***	(0.0072)	−0.0846***	(0.0105)
Transitivity	0.0977***	(0.0266)	−0.0215	(0.0653)
Constant	−22.9134***	(6.9104)	1.0412	(0.9465)
Log likelihood	-13,535.87		-9,155.36	
Observations	11,000,219		19,744	

Notes: A positive coefficient in the formation process indicates that a link is more likely to form, whereas a positive coefficient in the dissolution process indicates that the link is more likely to be sustained. For better interpretation, the average profit and standard deviation of profit are scaled by a factor of 100, and the average holding time is scaled by a factor of 1/100. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

5.1. Main variables of interest

Table 3 shows the positive coefficients for a leader's post quantity for both link formation and link dissolution. Similarly, the coefficients for a leader's post quality during link formation and link dissolution are also positive, indicating that the propensity to form new links and maintain existing links increases as leaders publish a larger number higher-quality posts.[14] This observation is consistent with the notion that high-quality posts provide useful information and increase the transparency of a leader's investment strategy, which in turn increases trust in the leader, attracts more incoming links, and helps maintain existing links. In addition, the leader's number of replies can also help attract new followers. Overall, this notion is consistent with the stream of literature that argues that communication can increase trustworthiness and trust (Kaiser and Berger 2021). We further investigate the nonlinear effects of post quantity and quality on the likelihood of link formation by adding their quadratic terms to our model. We find a diminishing marginal effect of post quantity and quality in the link formation process (see Appendix P).

The coefficients on positive comments for link formation and dissolution are significantly positive, indicating a higher probability that a follower will form or maintain a link if the leader receives more positive comments. Similarly, negative and significant estimates for negative comments indicate that negative comments are associated with a lower probability of link formation or maintaining a link. The different magnitudes of the coefficients further indicate that the impacts of negative and positive comments are asymmetric and that negative comments are particularly relevant in online contexts, consistent with previous evidence from social media (Xu and Chau 2018).[15]

Comparing the estimates of the effect of social communication in link formation and dissolution, we find that social communication plays an important role in both processes. Thus, communication on social trading platforms seems to have an economic impact on leaders, given that the links in the leader-follower network are tied to cash flow and directly affect the compensation received by the leader from the platform. Communication, a type of soft information, helps leaders not only attract new links but also maintain existing links. Our coefficients indicate that the effect of communication is stronger for link formation than for link dissolution, which may be explained by the fact that posts can affect link formation through an additional channel (i.e., attention channel), which is less relevant for

[14] Also of particular interest is the interplay between post quality and quantity. A robustness check including a quantity-quality interaction term indicates that post quality has a positive moderating effect on post quantity in the link formation process. The interaction term is not significant in the link dissolution process. Please refer to Appendix F for more details.

[15] The t statistic of the difference in coefficients between negative and positive comments is 7.32 in the link formation model and 5.58 in the link dissolution model, indicating that these two types of comments are significantly different.

dissolution (see, e.g., Barber and Odean 2008, for a similar argument in the context of the financial market).

Turning to financial performance, we find that a leader's average profit tends to attract followers by forming new links and maintaining existing links. In addition, greater volatility in the leader's financial performance is negatively associated with forming new links and maintaining existing links. Overall, these observations are in line with prominent findings from the mutual fund flow literature that investors chase past performance (Barber et al. 2016) and with previous evidence on social trading (Doering et al. 2015).

5.2. Control variables

We control for the trading strategies of leaders and find that followers prefer leaders who tend to adopt diversified buy-and-hold strategies and are more likely to establish links with those leaders. Once links are established, investing in lottery-like stocks is associated with a higher probability of sustaining the link. This observation is consistent with the notion that investors may choose to accept large chances of a small loss for a small chance of a large gain (Markowitz 1952). In particular, on average, lottery-like stocks realize a low (slightly negative) return. However, they are also associated with a small potential of very large positive returns. In addition, these stocks typically have conversational features (Han et al. 2022), which may make them attractive in a social setting. Against this backdrop, it is reasonable that followers find it attractive when leaders add lottery-like positions to their diversified portfolios.

We also control for the social activities and financial performance of followers. Regarding their social activities, we find that the probability of establishing new links decreases as the quality of posts increases. Intuitively, followers with higher-quality posts may have greater financial knowledge and expertise and consequently may be more likely to trade by themselves, instead of following others. Regarding their financial performance, we find that followers are insensitive to their past profit on the platform when choosing a new link, as the coefficient on their average profit is nonsignificant in the link formation. However, followers tend to be more likely to establish and maintain links if they have a lower level of volatility.

With respect to demographic characteristics, we find—most notably—that followers tend to establish and maintain links with leaders of the same nationality, which is consistent with studies on peer-to-peer credit markets (Lin and Viswanathan 2016). These effects may be driven by language barriers or cultural differences. We also find evidence in support of age and gender homophily among male investors in link formation. However, female followers are more likely to form a link with male leaders. In line with previous findings from the literature (Wohlgemuth et al. 2016), we observe that the presence of a picture on a leader's profile page and of a biographical description significantly increases the likelihood that followers will form a new link. The disclosure of a profile picture or a biography by a leader may increase his or her perceived trustworthiness and therefore the likelihood of new links. Next, we consider a follower's experience, wealth, and risk preference. Our results show that followers with a higher risk score are less likely to form a new link, whereas those with a higher experience level tend to maintain their existing links.

Comparing the link formation and dissolution processes, we find strong differences in the impact of the demographic characteristics of leaders. This observation is intuitive in the sense that once followers have considered the demographic characteristics and established a link, there is no need to for them to be considered again, as demographics remain stable over time. Other factors such as financial performance and social communication become more relevant.

Finally, we briefly discuss the variables that capture the network structure. A leader's popularity (in-degree) increases his or her propensity to attract followers who form new links and maintain existing links, indicating preferential attachment. In contrast, the coefficient on a leader's activity (out-degree) is significantly negative in link formation, indicating that leaders who follow other investors are less attractive to potential followers. Followers with higher popularity are less likely to form new links, whereas those with higher levels of activity are less likely to form and maintain existing links. We also find a significant transitivity effect in the link formation, indicating that the presence of a link from i to j and from j to k increases the likelihood of the formation of a direct link between i and k.

6. Robustness tests

In this section, we present a series of robustness checks to make sure that our findings are not driven by a specific model setup.

6.1. Two-stage selection model

Since the total number of leaders on the platform is large and followers have a limited attention span, it is possible that some leaders are more visible than others, which may affect whether potential followers follow a particular leader. Thus, in this robustness check, we develop a two-stage selection model that attempts to accurately model the link formation process. In the first stage, to account for different exposure to leaders' profiles among followers, we model the probability of followers being aware of leader j as

$$Pr(A_{jt}=1)=\frac{e^{\gamma z_{jt}}}{1+e^{\gamma z_{jt}}}, \tag{8}$$

where z_{jt} denotes the number of followers of leader j in period t. We choose the number of followers in the first step because, first, the finance literature provides substantial evidence that herding is a relevant behavioral trait in financial markets (Devenow and Welch 1996). In the context of social trading, Gemayel and Preda (2018b) show that the scopic regime can increase herding behavior. Given that the number of followers already reflects financial performance to some degree (i.e., investors with poor performance are less likely to have many followers), we argue that investors will start their filtering decision based on fewer criteria to simplify their decision process as much as possible. Second, we refer to the concept of preferential attachment in social networks: leaders with large follower bases should be more likely to attract additional followers (Neyman and Scott 1948).

In the second stage, followers decide whether to form a link with the leader based on the hard and soft information that they see on the his or her profile page. The second stage is identical to that in the main model; the probability of follower i following leader j during period t is modeled as

$$Pr(y_{ijt}=1)=\frac{e^{(\boldsymbol{\alpha}X_{it-1}+\boldsymbol{\beta}W_{jt-1}+\boldsymbol{\lambda}V_{ij}+\boldsymbol{\tau}C_i+\eta_i)}}{1+e^{(\boldsymbol{\alpha}X_{it-1}+\boldsymbol{\beta}W_{jt-1}+\boldsymbol{\lambda}V_{ij}+\boldsymbol{\tau}C_i+\eta_i)}}. \tag{9}$$

Taking steps 1 and 2 together, we derive the overall likelihood as

$$\begin{aligned} L_{ijt} &= y_{ijt}\times(Pr(A_{jt}=1)\times Pr(y_{ijt}=1|A_{jt}=1)) \\ &+(1-y_{ijt})\times(Pr(A_{jt}=1)\times Pr(y_{ijt}=0|A_{jt}=1)+Pr(A_{jt}=0)) \end{aligned} \tag{10}$$

The overall log-likelihood value is further written as

$$TLL(\gamma,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\lambda},\boldsymbol{\tau})=\sum_{i=1}^{I_t}\sum_{t=1}^{T_i}\sum_{j=1}^{J_{it}}\ln(L_{ijt}), \tag{11}$$

where I_t is the number of followers in period t, T_i is the number of periods in which follower i exists on the platform, and J_{it} is the number of leaders that follower i follows in period

t. We estimate our proposed two-stage selection model by maximizing the overall log-likelihood value. For computational tractability, following Heckman and Singer (1984), we apply a nonparametric approach to estimate the follower's random effects after controlling for the time average of the follower's time-variant observables. The estimation results are reported in Table 4. We find that potential followers are more likely to be aware of a leader if he or she has a larger number of followers of his or her account (coefficient of 0.3498, $p < 0.01$). Overall, the results from the second stage are consistent with our findings from the main model, with the volatility of the leader's financial performance (which is no longer significant) being the exception. If followers, in general, consider the volatility of returns, then this information is already accounted for in the first stage when considering the number of followers. Thus, followers do not place weight on volatility in the second stage.

6.2. Alternative sentiment dictionary

In the main analysis, we use the widely applied VADER sentiment dictionary to calculate the sentiment scores for comments. It is possible that different sentiment dictionaries generate different sentiment scores, which might affect the estimated effects of positive and negative scores for a leader's comments on link formation and dissolution. Therefore, in this robustness check, we apply an alternative sentiment dictionary to calculate the sentiment scores for comments. In particular, we adopt the Harvard General Inquirer[16] dictionary, another widely adopted dictionary for extracting sentiment from social media, to perform sentiment analysis (Ammann and Schaub 2021). Then, we reestimate the STERGM with Chamberlain correlated random effects and show the estimation results in Table 5. The results are generally consistent with the findings in the main model.

6.3. Alternative measure of financial risk

The volatility of performance is a symmetric measure of risk that takes into account both positive and negative deviations from the mean. Investors may, however, be most concerned with extremely negative profit outcomes, i.e., those with large losses. Consequently, we consider the MDD as an alternative risk measure that accounts for large losses. The MDD measures the monthly maximum observed loss in a leader's daily profit. Similar to the standard deviation, the MDD is a widely used risk measure (Cvitanić and Karatzas

[16] `http://www.wjh.harvard.edu/~inquirer/`, last accessed Jan. 8, 2021.

Table 4 Estimation results of the two-stage selection model

Variable	Formation	
First stage		
No. of followers	0.3498***	(0.041 4)
Second stage		
Leader's post quantity	0.2164***	(0.038 1)
Leader's post quality	0.2513***	(0.036 5)
Leader's number of replies	0.1059**	(0.041 6)
Leader's received positive comment score	1.1982***	(0.194 8)
Leader's received negative comment score	−3.4817***	(0.895 9)
Leader's average profit	0.8829***	(0.180 1)
Leader's std. dev. profit	0.5282	(0.662 1)
Controls		
Leader's average holding time	0.1066	(0.073 9)
Leader's lottery preference	0.8128*	(0.483 5)
Leader's HHI	−0.8485***	(0.123 3)
Follower's post quantity	−0.0947	(0.069 6)
Follower's post quality	−0.1212	(0.137 9)
Follower's average profit	0.3988*	(0.233 5)
Follower's std. dev. profit	−4.0453***	(1.476 2)
Nationality	1.1422***	(0.106 4)
Age	0.1604**	(0.073 1)
Homophily (male)	0.9903***	(0.169 9)
Homophily (female)	−1.0283**	(0.497 7)
Image	−1.4717**	(0.574 5)
Bio	1.4668***	(0.267 0)
Experience	−1.1127***	(0.348 2)
Wealth	−48.6303	(30.310 8)
Income	−0.3570	(0.362 9)
Risk	−1.1816***	(0.453 0)
Leader's popularity	4.2868***	(0.257 6)
Leader's activity	−0.2572*	(0.147 1)
Follower's popularity	−1.3578***	(0.218 3)
Follower's activity	−1.0680***	(0.227 6)
Transitivity	0.9521***	(0.075 5)
Constant	−37.5676	(33.464 7)
Log likelihood	-13,145.14	
Observations	11,000,219	

Notes: For better interpretation, the average profit and standard deviation of profit are scaled by a factor of 100, and the average holding time is scaled by a factor of 1/100. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

1999, de Melo Mendes and Lavrado 2017). We reestimate the STERGM with Chamberlain correlated random effects and present the estimation results in Table 6. The coefficients on the main variables of interest are consistent with our main results. A higher MDD value on the leader's financial performance yields a lower likelihood of followers forming new links and sustaining existing links.

6.4. Further robustness tests

We run a series of additional robustness checks. First, we address the heterogeneity of followers' age. Prior studies have found that individuals who are younger in age are more likely to blog, visit social network sites, and rely on social media in their decision-making than are those who are older in age (Chou et al. 2009). Considering that social trading is a novel way in which to participate in financial markets that may particularly attract younger

Table 5 Estimation results using an alternative sentiment dictionary

Variable	Formation		Dissolution	
Leader's post quantity	0.0442***	(0.0032)	0.0157***	(0.0042)
Leader's post quality	0.5257***	(0.0240)	0.0476**	(0.0220)
Leader's number of replies	0.1140***	(0.0194)	0.0243	(0.0227)
Leader's received positive comment score	1.6340***	(0.1422)	0.8703***	(0.1773)
Leader's received negative comment score	−2.5631***	(0.5103)	−1.3289***	(0.3897)
Leader's average profit	0.1084***	(0.0135)	0.0795***	(0.0190)
Leader's std. dev. profit	−3.0402***	(0.5594)	−3.0657***	(0.7943)
Controls				
Leader's average holding time	0.3197***	(0.0442)	0.0518	(0.0658)
Leader's lottery preference	0.3092	(0.3553)	1.2911***	(0.3971)
Leader's HHI	−0.6483***	(0.0901)	−0.1222	(0.0931)
Follower's post quantity	0.0160	(0.0281)	−0.2755***	(0.0412)
Follower's post quality	−0.2576***	(0.0830)	0.0063	(0.0792)
Follower's average profit	0.0274	(0.0183)	0.0362*	(0.0193)
Follower's std. dev. profit	−3.4637***	(1.1280)	−8.4108***	(1.1192)
Nationality	0.7419***	(0.0704)	0.3398***	(0.0797)
Age	0.1127**	(0.0507)	0.0450	(0.0543)
Homophily (male)	0.9834***	(0.1362)	−0.1478	(0.1164)
Homophily (female)	−0.8515**	(0.4167)	0.2962	(0.3187)
Image	2.4997***	(0.4229)	0.4847	(0.5452)
Bio	3.0510***	(0.1876)	−0.2174	(0.2015)
Experience	−0.0688*	(0.0363)	0.1377***	(0.0405)
Wealth	0.0005	(0.0315)	0.0650*	(0.0347)
Income	−0.0224	(0.0380)	0.0230	(0.0409)
Risk	−0.1021**	(0.0471)	−0.0319	(0.0527)
Leader's popularity	0.0074***	(0.0004)	0.0012***	(0.0004)
Leader's activity	−0.0414***	(0.0122)	0.0389***	(0.0120)
Follower's popularity	−0.0591***	(0.0150)	−0.0055	(0.0094)
Follower's activity	−0.0260***	(0.0072)	−0.0844***	(0.0105)
Transitivity	0.0971***	(0.0267)	−0.0171	(0.0652)
Constant	−18.8462***	(7.0133)	1.0273	(0.9471)
Log likelihood	-13,550.02		-9,162.39	
Observations	11,000,219		19,744	

Notes: A positive coefficient in the formation process indicates that a link is more likely to form, whereas a positive coefficient in the dissolution process indicates that the link is more likely to be sustained. For better interpretation, the average profit and standard deviation of profit are scaled by a factor of 100, and the average holding time is scaled by a factor of 1/100. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

individuals, it is natural to ask whether our findings hold across all age groups. We thus examine whether the impact of social communication and financial performance variables varies across age groups. We split the dataset into two subsamples based on followers' age ranges. The first group includes followers between 18 and 44 years of age, and the second group includes followers who are older than 44 years. We again apply the STERGM with Chamberlain correlated random effects and summarize the estimation results in Appendix I. Younger followers are rather sensitive to positive and negative comments in the link dissolution process, whereas the effects of comments are not significant for older followers. Interestingly, on the one hand, we observe that a leader's post quantity increases the probability of younger followers maintaining existing links, while post quality becomes nonsignificant. On the other hand, for older followers, it is a leader's post quality, rather than post quantity, that increases the probability of him or her maintaining existing links.

Table 6 Estimation results with the MDD as the measure of financial risk

Variable	Formation		Dissolution	
Leader's post quantity	0.0375***	(0.0032)	0.0099**	(0.0039)
Leader's post quality	0.5354***	(0.0252)	0.0641***	(0.0221)
Leader's number of replies	0.1479***	(0.0208)	0.0462**	(0.0229)
Leader's received positive comment score	1.4298***	(0.1263)	0.6551***	(0.1415)
Leader's received negative comment score	−3.7808***	(0.7165)	−1.8946***	(0.4747)
Leader's average profit	0.0368***	(0.0044)	0.0105*	(0.0058)
Leader's MDD	−0.0022***	(0.0008)	−0.0024***	(0.0008)
Controls				
Leader's average holding time	0.3596***	(0.0426)	0.0525	(0.0654)
Leader's lottery preference	0.3055	(0.3579)	1.2747***	(0.3965)
Leader's HHI	−0.6042***	(0.0901)	−0.0969	(0.0929)
Follower's post quantity	0.0144	(0.0286)	−0.2865***	(0.0417)
Follower's post quality	−0.2444***	(0.0831)	0.0338	(0.0802)
Follower's average profit	0.0313**	(0.0156)	0.0089	(0.0198)
Follower's MDD	−0.0008*	(0.0004)	−0.0021***	(0.0003)
Nationality	0.7450***	(0.0705)	0.3277***	(0.0795)
Age	0.1126**	(0.0507)	0.0385	(0.0542)
Homophily (male)	0.9858***	(0.1361)	−0.1441	(0.1165)
Homophily (female)	−0.8582**	(0.4167)	0.2472	(0.3182)
Image	2.5688***	(0.4234)	0.5264	(0.5478)
Bio	2.9210***	(0.1827)	−0.2942	(0.2012)
Experience	−0.0602*	(0.0362)	0.1144***	(0.0414)
Wealth	0.0109	(0.0314)	0.0791**	(0.0356)
Income	−0.0304	(0.0379)	0.0167	(0.0419)
Risk	−0.0919**	(0.0468)	−0.0615	(0.0539)
Leader's popularity	0.0074***	(0.0004)	0.0011***	(0.0004)
Leader's activity	−0.0324**	(0.0128)	0.0439***	(0.0121)
Follower's popularity	−0.0598***	(0.0151)	−0.0064	(0.0095)
Follower's activity	−0.0258***	(0.0073)	−0.0802***	(0.0105)
Transitivity	0.0941***	(0.0261)	−0.0218	(0.0656)
Constant	−21.9635***	(6.8550)	0.5605	(0.9575)
Log likelihood	-13,548.55		-9,217.73	
Observations	11,000,219		19,744	

Notes: A positive coefficient in the formation process indicates that a link is more likely to form, whereas a positive coefficient in the dissolution process indicates that the link is more likely to be sustained. For better interpretation, the average profit is scaled by a factor of 100, and the average holding time is scaled by a factor of 1/100. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Next, we explore whether our main findings differ in terms of asset class. We capture the major asset classes using dummy variables and separately investigate the impact of a leader's post quantity and quality in both the link formation and dissolution model using an interaction term between asset class dummies and our variables of interest. The interaction terms are found to be nonsignificant, and the coefficients of interest do not change in a meaningful way (see Appendix J).

We address the heterogeneity of investors' nationalities by including a full set of dummy variables for a leader's country and rerun both the link formation and dissolution models. Our main results are qualitatively consistent with those of the main model (see Appendix K).

In our main model, we construct the leader-follower network using a binary variable to represent whether or not there is a fund flow. To dig deeper and shed more light on actual fund flows, we additionally use followers' portfolio weights (i.e., the fund flow relative to a follower's total portfolio value) allocated by the follower to the leader when establishing the link. This information enables us to investigate whether followers allocate a larger or smaller share of their portfolio to a given leader. We use the investment weights as the dependent variable and do not observe meaningful changes to our main conclusions (see Appendix L).

Investors who begin their social trading career as followers may become leaders over time; i.e., they may switch their roles over time, which may affect our results. To address such a concern, we create a dummy variable to indicate whether leader j is also following others at period t. This time-variant dummy variable captures various role dynamics. We incorporate the dummy into the link formation and dissolution models. Our results show that the main variables of interest, such as the leader's social communication and financial performance, remain qualitatively consistent with those of our main model. Interestingly, we find that potential followers are more likely to follow a leader who is not following others ("pure leader"), and existing followers are more likely to dissolve their links with pure leaders, which is consistent with disconfirmation theory (Oliver 1980) (see Appendix M).

Given the speed of change in social media, using a monthly horizon in our main model may raise concerns. Note that we face a tradeoff among short-lived social media timeliness, potentially longer investment horizons, and computational tractability. We reorganize the data at the bi-weekly level and rerun the link formation and dissolution models to investigate a more granular reflection of link dynamics. Our results remain qualitatively consistent with those of the main model (see Appendix N).

Investor sentiment is an important driver of investment decisions (Baker and Wurgler 2007). Market sentiment might affect the effectiveness of social communication in link formation and dissolution. To quantify sentiment, we utilize the Financial and Economic Attitudes Revealed by Search (FEARS) measure, which is among the most widely used sentiment measures in the financial literature (Birru and Young 2022). FEARS, based on Google Search Volume (GSV), is closely linked to retail investor sentiment. We then interact the index with our measures of social communication (the leader's post quantity and quality) in the link formation and dissolution models and summarize our findings in Appendix O. Both post quantity and quality are less important—have a lower effect size—when sentiment is lower (when FEARS is higher). One possible interpretation of this finding is that when sentiment is lower, people are less susceptible to peers' communication.

7. Conclusions

Social trading is a novel form of trading that combines online brokerage and traditional social media features. Social trading has attracted a large number of investors and increased attention from both practitioners and academia. Social trading allows investors to seek financial advice from their peers, observe their peers' trading strategies, and directly follow other investors in real time. Thus, inexperienced retail investors may benefit from their peers, while experienced investors are able to provide signals and earn additional income. Due to the monetary aspects involved in these leader-follower relationships, network evolution follows a distinct pattern that differs from that of traditional social media platforms. In particular, link dissolution is an important part of social trading.

We study a dynamic social trading network using the STERGM and examine how various factors affect the link formation and dissolution processes. We show that social communication, financial performance, and demographics have different implications for these processes. Followers consider financial performance, social communication, and demographics when deciding whom to follow (link formation process). However, once a link is formed, demographic characteristics become less important, as followers focus mainly on leaders' financial performance as well as social communication to decide whether or not to sustain the link (link dissolution process). Focusing on the different types of social communication, we find that the quality and quantity of a leader's posts increase the likelihood of followers forming new links and sustaining existing links. Followers are less likely to form new links or sustain existing links with leaders who receive more negative comments. Leaders who receive more positive comments are more likely to attract new followers and keep existing followers. In addition, the impacts of negative and positive comments are asymmetric. Negative comments have a larger impact than do positive comments on the link formation and dissolution processes.

Our study contributes to the growing literature on social trading by first modeling the dynamics of leader-follower networks. Our granular data allow us to thoroughly examine the implications of various factors on the link formation and dissolution processes. Our study also contributes to a better understanding of how hard information (e.g., financial performance) and, in particular, soft information (e.g., social communication) affect leader-follower network evolution in the social trading context.

Our study has practical managerial implications. We document link formation and link dissolution processes and thereby broaden and deepen our understanding of leader-follower network evolution in social trading. Social trading platforms were established in the aftermath of the global financial crisis to provide retail investors with an alternative to traditional wealth management in response to the eroding trust in financial markets following the crisis (Doering et al. 2015). Although social trading platforms provide a high level of informational transparency, investors face new challenges in building trust, particularly because most investors on the platform are individuals who lack institutional endorsements and because the relationship is online with a more and less anonymous mass. In our study, we find that social communication plays an important role in leaders' ability to convince potential followers to follow their trading strategies and existing followers to sustain their links. Social communication is effective in building trust among investors on social trading platforms. Leaders should make high-quality posts, as such posts both increase link formation and reduce link dissolution. If not used properly, posts with negative comments can backfire, reduce link formation, and increase link dissolution. As such, negative comments have a larger impact on link formation and dissolution than do positive comments. By communicating in a balanced manner, leaders can attract new followers to follow their trading strategies and encourage existing followers to sustain their links. Given the importance of social communication in the evolution of leader-follower networks with real money flow, social trading platforms should carefully regulate social communication to sustain a healthy ecosystem.

Our analysis has some caveats. First, we abstract away from the potential dependence between a follower's following decisions and a leader's subsequent trading strategies. Thus, we assume that a leader's trading strategy does not change, regardless of any given follower's decision. However, such dependencies do exist (Pelster and Hofmann 2018) and may affect network evolution. Future research may aim to study these coevolution effects in more detail. Second, we study a relatively small sample of the network compared to the actual total size. Thus, we cannot address the generalizability of our results to other (similar) networks. We invite future research to study network evolution in similar settings to provide a more complete picture. Third, our data cover a period of bullish markets

(January 2016 to December 2017). As such, we are unable to answer the impact of social communication in other market environments. It would be interesting to investigate how link formation and dissolution in social trading occur during bearish markets or when markets are in distress. Finally, we incorporate Chamberlain correlated random effects into our model to address potential omitted variable bias; for example, some leaders might advertise themselves on other social media platforms, thus affecting link formation. However, Chamberlain correlated random effects capture only the time-invariant unobservables and, for example, do not account for the possibility that the advertising activities of investors on other social platforms may change over time.

References

Agarwal, V., N. D. Daniel, N. Y. Naik. 2009. Role of managerial incentives and discretion in hedge fund performance. *The Journal of Finance* **64**(5) 2221–2256.

Ammann, M., N. Schaub. 2021. Do individual investors trade on investment-related internet postings? *Management Science* **67**(9) 5679–5702.

Apesteguia, J., J. Oechssler, S. Weidenholzer. 2020. Copy trading. *Management Science* **66**(12) 5608–5622.

Baker, M., J. Wurgler. 2007. Investor sentiment in the stock market. *Journal of Economic Perspectives* **21**(2) 129–151.

Baker, W. E., R. R. Faulkner, G. A. Fisher. 1998. Hazards of the market: The continuity and dissolution of interorganizational market relationships. *American Sociological Review* **63**(2) 147–177.

Bali, T. G., D. Hirshleifer, L. Peng, Y. Tang. 2021. Attention, social interaction, and investor attraction to lottery stocks. Tech. rep., National Bureau of Economic Research.

Barber, B., X. Huang, T. Odean. 2016. Which factors matter to investors? evidence from mutual fund flows. *The Review of Financial Studies* **29**(10) 2600–2642.

Barber, B. M., T. Odean. 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *The Review of Financial Studies* **21**(2) 785–818.

Barberis, N., M. Huang. 2008. Stocks as lotteries: The implications of probability weighting for security prices. *The American Economic Review* **98**(5) 2066–2100.

Bhattacherjee, A. 2001. Understanding information systems continuance: An expectation-confirmation model. *MIS Quarterly* **25**(3) 351–370.

Birru, J., T. Young. 2022. Sentiment and uncertainty. *Journal of Financial Economics* **146**(3) 1148–1169.

Cao, Q., W. Duan, Q. Gan. 2011. Exploring determinants of voting for the “helpfulness” of online user reviews: A text mining approach. *Decision Support Systems* **50**(2) 511–521.

Chamberlain, G. 1980. Analysis of covariance with qualitative data. *The Review of Economic Studies* **47**(1) 225–238.

Chou, W.Y.S., Y.M. Hunt, E.B. Beckjord, R.P. Moser, B.W. Hesse. 2009. Social media use in the united states: implications for health communication. *Journal of Medical Internet Research* **11**(4) e1249.

Coval, J. D., T. J. Moskowitz. 1999. Home bias at home: Local equity preference in domestic portfolios. *The Journal of Finance* **54**(6) 2045–2073.

Cvitanić, J., I. Karatzas. 1999. On dynamic measures of risk. *Finance and Stochastics* **3**(4) 451–482.

de Melo Mendes, B. V., R. C. Lavrado. 2017. Implementing and testing the maximum drawdown at risk. *Finance Research Letters* **22** 95–100.

Devenow, A., I. Welch. 1996. Rational herding in financial economics. *European Economic Review* **40**(3-5) 603–615.

Doering, P., S. Neumann, S. Paul. 2015. A primer on social trading networks–institutional aspects and empirical evidence. *EFMA Annual Meetings*.

Dorfleitner, G., L. Fischer, C. Lung, P. Willmertinger, N. Stang, N. Dietrich. 2018. To follow or not to follow–an empirical analysis of the returns of actors on social trading platforms. *The Quarterly Review of Economics and Finance* **70** 160–171.

Dorn, D., G. Huberman, P. Sengmueller. 2008. Correlated trading and returns. *The Journal of Finance* **63**(2) 885–920.

Duarte, J., S. Siegel, L. Young. 2012. Trust and credit: The role of appearance in peer-to-peer lending. *The Review of Financial Studies* **25**(8) 2455–2484.

Gemayel, R., A. Preda. 2018a. Does a scopic regime erode the disposition effect? evidence from a social trading platform. *Journal of Economic Behavior & Organization* **154** 175–190.

Gemayel, R., A. Preda. 2018b. Does a scopic regime produce conformism? herding behavior among trade leaders on social trading platforms. *The European Journal of Finance* **24**(14) 1144–1175.

Goetzmann, W. N., J. E. Ingersoll Jr., S. A. Ross. 2003. High-water marks and hedge fund management contracts. *The Journal of Finance* **58**(4) 1685–1718.

Greve, H. R., J. A. C. Baum, H. Mitsuhashi, T. J. Rowley. 2010. Built to last but falling apart: Cohesion, friction, and withdrawal from interfirm alliances. *Academy of Management Journal* **53**(2) 302–322.

Grinblatt, M., G. Jostova, L. Petrasek, A. Philipov. 2020. Style and skill: Hedge funds, mutual funds, and momentum. *Management Science* **66**(12) 5505–5531.

Han, B., D. Hirshleifer, J. Walden. 2022. Social transmission bias and investor behavior. *Journal of Financial and Quantitative Analysis* **57**(1) 390–412.

Heckman, J., B. Singer. 1984. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica: Journal of the Econometric Society* 271–320.

Heimer, R. Z. 2016. Peer pressure: Social interaction and the disposition effect. *The Review of Financial Studies* **29**(11) 3177–3209.

Hirshleifer, D., S. H. Teoh. 2003. Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics* **36** 337–386.

Holland, P. W., S. Leinhardt. 1971. Transitivity in structural models of small groups. *Comparative Group Studies* **2**(2) 107–124.

Holzmeister, F., J. Huber, M. Kirchler, F. Lindner, U. Weitzel, S. Zeisberger. 2020. What drives risk perception? a global survey with financial professionals and laypeople. *Management Science* **66**(9) 3977–4002.

Huang, J., K. D. Wei, H. Yan. 2007. Participation costs and the sensitivity of fund flows to past performance. *The Journal of Finance* **62**(3) 1273–1311.

Hunter, D. R., M. S. Handcock, C. T. Butts, S. M. Goodreau, M. Morris. 2008. ergm: A package to fit, simulate and diagnose exponential-family models for networks. *Journal of Statistical Software* **24**(3) nihpa54860.

Hutto, C., E. Gilbert. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the international AAAI conference on web and social media*, vol. 8. 216–225.

Hwang, E. H., X. Guo, Y. Tan, Y. Dang. 2022. Delivering healthcare through teleconsultations: Implications for offline healthcare disparity. *Information Systems Research* **33**(2).

Ivković, Z., S. Weisbenner. 2009. Individual investor mutual fund flows. *The Journal of Financial Economics* **92**(2) 223 – 237.

Kaiser, M., E. S. C. Berger. 2021. Trust in the investor relationship marketing of startups: a systematic literature review and research agenda. *Management Review Quarterly* **71**(2) 491–517.

Kane, G. C., G. J. Labianca, S. P. Borgatti. 2014. What's different about social media networks? a framework and research agenda. *MIS Quarterly* **38**(1) 275–304.

Khern-am-nuai, W., K. Kannan, H. Ghasemkhani. 2018. Extrinsic versus intrinsic rewards for contributing reviews in an online platform. *Information Systems Research* **29**(4) 871–892.

Kim, H., A. Kankanhalli, S. Lee. 2018. Examining gifting through social network services: A social exchange theory perspective. *Information Systems Research* **29**(4) 805–828.

Krivitsky, P. N., M. S. Handcock. 2014. A separable model for dynamic networks. *Journal of the Royal Statistical Society. Series B, Statistical Methodology* **76**(1) 29.

Kumar, A. 2009. Who gambles in the stock market? *The Journal of Finance* **64**(4) 1889–1933.

Lee, G. M., L. Qiu, A. B. Whinston. 2016. A friend like me: Modeling network formation in a location-based social network. *Journal of Management Information Systems* **33**(4) 1008–1033.

Lee, W., Q. Ma. 2015. Whom to follow on social trading services? a system to support discovering expert traders. *The Tenth International Conference on Digital Information Management* 188–193.

Li, Z., X. Fang, X. Bai, O.R.L. Sheng. 2017. Utility-based link recommendation for online social networks. *Management Science* **63**(6) 1938–1952.

Liberti, J. M., M. A. Petersen. 2019. Information: Hard and soft. *Review of Corporate Finance Studies* **8**(1) 1–41.

Lin, M., S. Viswanathan. 2016. Home bias in online investments: An empirical study of an online crowdfunding market. *Management Science* **62**(5) 1393–1414.

Markowitz, H. 1952. Portfolio selection. *The Journal of Finance* **7**(1) 77–91.

Mayer, R. C., J. H. Davis, F. D. Schoorman. 1995. An integrative model of organizational trust. *Academy of Management Review* **20**(3) 709–734.

McPherson, M., L. Smith-Lovin, J. M. Cook. 2001. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* **27**(1) 415–444.

Mundlak, Y. 1978. On the pooling of time series and cross section data. *Econometrica: journal of the Econometric Society* 69–85.

Neyman, J., E. L. Scott. 1948. Consistent estimates based on partially consistent observations. *Econometrica: Journal of the Econometric Society* 1–32.

Oehler, A., M. Horn, S. Wendt. 2016. Benefits from social trading? empirical evidence for certificates on wikifolios. *International Review of Financial Analysis* **46** 202–210.

Oliver, R. L. 1980. A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of marketing research* **17**(4) 460–469.

Pan, W., Y. Altschuler, A. S. Pentland. 2012. Decoding social influence and the wisdom of the crowd in financial trading network. IEEE, ed., *International Conference on Privacy, Security, Risk and Trust and International Conference on Social Computing*. Washington DC, 203–209.

Pedersen, L. H. 2022. Game on: Social networks and markets. *The Journal of Financial Economics* **146**(3) 1097–1119.

Pelster, M. 2017. I'll have what s/he's having: A case study of a social trading network. *Proceedings of the International Conference on Information Systems 2017* .

Pelster, M., A. Hofmann. 2018. About the fear of reputational loss: Social trading and the disposition effect. *Journal of Banking & Finance* **94** 75–88.

Polidoro, F., G. Ahuja, W. Mitchell. 2011. When the social structure overshadows competitive incentives: The effects of network embeddedness on joint venture dissolution. *Academy of Management Journal* **54**(1) 203–223.

Robins, G., P. Pattison, Y. Kalish, D. Lusher. 2007. An introduction to exponential random graph (p*) models for social networks. *Social networks* **29**(2) 173–191.

Röder, F., A. Walter. 2019. What drives investment flows into social trading portfolios? *Journal of Financial Research* **42**(2) 383–411.

Shafi, K., A. Mohammadi, S. A. Johan. 2020. Investment ties gone awry. *Academy of Management Journal* **63**(1) 295–327.

Shefrin, H., M. Statman. 1985. The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of Finance* **40**(3) 777–790.

Sirri, E. R., P. Tufano. 1998. Costly search and mutual fund flows. *The Journal of Finance* **53**(5) 1589–1622.

Snijders, T. A., P. E. Pattison, G. L. Robins, M. S. Handcock. 2006. New specifications for exponential random graph models. *Sociological methodology* **36**(1) 99–153.

Strauss, D., M. Ikeda. 1990. Pseudolikelihood estimation for social networks. *Journal of the American Statistical Association* **85**(409) 204–212.

Tirunillai, S., G. J. Tellis. 2012. Does chatter really matter? dynamics of user-generated content and stock performance. *Marketing Science* **31**(2) 198–215.

Wasserman, S., K. Faust, et al. 1994. *Social network analysis: Methods and applications*. Cambridge University Press.

Wohlgemuth, V., E. S. Berger, M. Wenzel. 2016. More than just financial performance: Trusting investors in social trading. *Journal of Business Research* **69**(11) 4970–4974.

Wooldridge, J. M. 2010. *Econometric analysis of cross section and panel data*. MIT Press.

Xu, J., M. Chau. 2018. Cheap talk? the impact of lender-borrower communication on peer-to-peer lending outcomes. *Journal of Management Information Systems* **35**(1) 53–85.

Yan, L., J. Peng, Y. Tan. 2015. Network dynamics: how can we find patients like us? *Information Systems Research* **26**(3) 496–512.

Yang, M., Z. Zheng, V. Mookerjee. 2022. How much is financial advice worth? the transparency-revenue tension in social trading. *Management Science* **68**(7) 5252–5268.

Appendix

A. Network evolution

To model network evolution over time, we track link changes by defining two networks: the formation network (y^+) and the dissolution network (y^-). $y^+_{t-1\to t}$ consists of network y_{t-1} plus the links formed from time $t-1$ to t, and $y^-_{t-1\to t}$ consists of network y_{t-1} minus the links removed from time $t-1$ to t. In the data, we observe y_{t-1} and y_t. Given the observed data, we are able to recover formation network $y^+_{t-1\to t}$ and dissolution network $y^-_{t-1\to t}$.

Table A-1 shows 4 possible transitions between nodes i and j. If there is no link between nodes i and j at both times $t-1$ and t (first row in Table A-1), then the value of $y^+_{t-1\to t}$ is set to 0 in the formation network, indicating that there is no link formation from $t-1$ to t. However, we are not able to come to any conclusion regarding link dissolution because there is no link that can be dissolved between nodes i and j from time $t-1$ to t (denoted by - in Table A-1). If the link between nodes i and j exists at time $t-1$ but no longer exists at time t (third row in Table A-1), then the value of $y^-_{t-1\to t}$ is set to 0 in the dissolution network, indicating that the link is dissolved from $t-1$ to t. If the link exists both at time $t-1$ and at time t (last row in Table A-1), then the value of $y^-_{t-1\to t}$ is set to 1 in the dissolution network, meaning that the link is sustained from time $t-1$ to t. However, we cannot infer anything about formation (denoted by - in Table A-1) when $y_{t-1}=1$, as the link between nodes i and j already exists at time $t-1$. *Note that a value of 1 in $y^-_{t-1\to t}$ means that the link is sustained and that a value of 0 means that the link is dissolved. In contrast, a value of 1 in $y^+_{t-1\to t}$ means that the link is formed.*

Table A-1 Illustration of network evolution with one link

y_{t-1}	y_t	$y^-_{t-1\to t}$	$y^+_{t-1\to t}$
0	0	-	0
0	1	-	1
1	0	0	-
1	1	1	-

B. Estimation with time fixed effects

Link formation and dissolution may be affected by events that are time dependent such as policy changes on the platform, the availability of cryptocurrencies on the platform, or a regulatory change in the social trading industry. During our sample period (2016 and 2017), no significant changes were made to social trading regulations in general (see Appendix H), and all existing regulations remained intact.

Despite the lack of regulatory changes during our sample period, it is possible that link formation and dissolution are affected by other time-dependent events. To mitigate this concern, we include time fixed effects to control for time-specific peculiarities. We estimate our extended model and present the results in Table B-1. The estimation results are generally consistent with those from the main model.

Table B-1 Estimation results with time fixed effects

Variable	Formation		Dissolution	
Leader's post quantity	0.0551***	(0.0037)	0.0159***	(0.0043)
Leader's post quality	0.5862***	(0.0254)	0.0700***	(0.0228)
Leader's number of replies	0.1398***	(0.0219)	0.0551**	(0.0236)
Leader's received positive comment score	1.2940***	(0.1344)	0.5653***	(0.1451)
Leader's received negative comment score	−3.9554***	(0.7350)	−1.8861***	(0.4815)
Leader's average profit	0.1175***	(0.0112)	0.0606***	(0.0194)
Leader's std. dev. profit	−0.8333*	(0.4574)	−2.0623**	(0.8130)
Controls				
Leader's average holding time	0.2264***	(0.0461)	0.0420	(0.0678)
Leader's lottery preference	0.6624*	(0.3547)	1.2052***	(0.4042)
Leader's HHI	−1.1364***	(0.0935)	−0.3810***	(0.0967)
Follower's post quantity	0.0042	(0.0282)	−0.2860***	(0.0414)
Follower's post quality	−0.2309***	(0.0831)	0.0197	(0.0816)
Follower's average profit	0.0140	(0.0167)	0.0434**	(0.0198)
Follower's std. dev. profit	−2.4975**	(1.0682)	−8.1191***	(1.1617)
Nationality	0.7780***	(0.0708)	0.3583***	(0.0812)
Age	0.1002**	(0.0511)	0.0523	(0.0553)
Homophily (male)	0.9440***	(0.1363)	−0.1158	(0.1192)
Homophily (female)	−0.8444**	(0.4171)	0.2376	(0.3217)
Image	2.3433***	(0.4254)	0.6696	(0.5583)
Bio	2.9788***	(0.1934)	−0.2302	(0.2056)
Experience	−0.1009***	(0.0369)	0.1246***	(0.0445)
Wealth	−0.0077	(0.0320)	0.0689*	(0.0381)
Income	−0.0128	(0.0384)	0.0247	(0.0450)
Risk	−0.0857*	(0.0478)	−0.0182	(0.0578)
Leader's popularity	0.0101***	(0.0004)	0.0015***	(0.0004)
Leader's activity	−0.0424***	(0.0125)	0.0317***	(0.0121)
Follower's popularity	−0.0561***	(0.0154)	0.0008	(0.0092)
Follower's activity	−0.0463***	(0.0089)	−0.1207***	(0.0114)
Transitivity	0.0948***	(0.0269)	−0.0198	(0.0653)
Constant	−7.2275	(7.3154)	1.6770	(1.0340)
Time fixed effects	Yes		Yes	
Log likelihood	-12,897.03		-8,996.67	
Observations	11,000,219		19,744	

Notes: A positive coefficient in the formation process indicates that a link is more likely to form, whereas a positive coefficient in the dissolution process indicates that the link is more likely to be sustained. For better interpretation, the average profit and standard deviation of profit are scaled by a factor of 100, and the average holding time is scaled by a factor of 1/100. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

The control variables (Leader's lottery preference and HHI) in Table B-1 indicate that followers look for diversified lotteries. This observation is consistent with the notion of Barberis and Huang (2008), who argue that investors in an economy may hold different portfolios, even when they have homogeneous preferences and beliefs—due to nonunique global optima. In their setting, some investors add a large, nondiversified position in a skewed security to their portfolios, while others may not take such a position. Investors may prefer the skewed security due to their overweight tails, which provide the potential for significant changes to their wealth. As a result, these investors are willing to pay a premium for the security and accept a negative average excess return. In other words, investors add a skewed asset to their diversified portfolios to obtain the chance to realize large positive returns.

C. Estimation with leader fixed effects

As mentioned in Section 3.2, in our context, the information provided on the platform is rich and highly transparent. We observe the information that is observed by followers on the platform, which may affect followers' link formation and dissolution decisions. We have access to the complete transaction and social communication histories of each trader, and the platform does not allow for a private chat channel. In addition, given the size and international reach of the platform, we believe that personal relationships between investors are very unlikely and affect, at most, only a few investors. Only approximately 14.6 percent of all links are between members of the same nationality. Considering this large size and international reach, it is unlikely that (many) investors know each other outside of the platform.

However, some investors may decide to advertise their trading via other social media channels such as YouTube and include links to their eToro profile pages. Despite the various control variables included in our network analysis, some determinants of link formation and dissolution may be unobserved, at least to researchers. For example, it is possible that leaders' advertising activities could be found on other social media sites, which may affect followers' following decisions. However, due to the anonymity of our data, we are unable to match data from other social media sites to our trading data. To mitigate this concern, we include leader-level fixed effects to capture a leader's general propensity to advertise his or her trading on social media. We believe that such general propensities are rather stable over time. While some investors are, in general, willing to advertise their investment strategies on alternative social media, others are not willing to do so.

We model only the link formation process because leader advertisements on other channels are more likely to affect the link formation process. We estimate the model using the same Chamberlain approach described in Section 3.2. To guarantee computational tractability with two random effects (follower and leader specific), we keep all leaders and randomly sample 250 followers from the sample used for the main model. The results are reported in Table C-1, and our main findings hold qualitatively in general.

Obviously, fixed effects capture only the time-invariant portion of such advertising. If the advertising activities of investors change over time, then fixed effects cannot capture these dynamics. However, in this case, we believe that both financial performance and number of posts will also capture the effect of investors advertising on other platforms. It is plausible

Table C-1 Estimation results with leader fixed effects

Variable	Formation	
Leader's post quantity	0.0134	(0.0124)
Leader's post quality	0.1822**	(0.0743)
Leader's number of replies	0.1441*	(0.0850)
Leader's received positive comment score	1.1968***	(0.3507)
Leader's received negative comment score	−6.9259***	(1.9984)
Leader's average profit	0.1756***	(0.0457)
Leader's std. dev. profit	−6.1823***	(2.2206)
Controls		
Leader's average holding time	−0.7118	(0.4848)
Leader's lottery preference	−0.4867	(1.0359)
Leader's HHI	0.5620	(0.4201)
Follower's post quantity	0.0577*	(0.0344)
Follower's post quality	−0.4086***	(0.1171)
Follower's average profit	0.0258	(0.0315)
Follower's std. dev. profit	−2.7689	(2.4210)
Nationality	1.6335***	(0.1671)
Age	0.1197	(0.1262)
Homophily (male)	0.1897	(0.3634)
Homophily (female)	−0.1767	(0.8274)
Image	1.7257***	(0.6473)
Bio	0.9630***	(0.3554)
Experience	−0.1449**	(0.0658)
Wealth	0.0828	(0.0601)
Income	0.0246	(0.0678)
Risk	−0.1533*	(0.0882)
Leader's popularity	−0.0007	(0.0016)
Leader's activity	−0.0010	(0.0572)
Follower's popularity	−0.0840***	(0.0182)
Follower's activity	0.0163	(0.0158)
Transitivity	0.1665***	(0.0340)
Constant	15.4624	(15.7955)
Log likelihood	-3,069.46	
Observations	1,592,411	

Notes: For better interpretation, the average profit and standard deviation of profit are scaled by a factor of 100, and the average holding time is scaled by a factor of 1/100. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

to assume that investors who vary their degree of communication about their trading do so on all channels to a similar degree—that is, if they use multiple channels in the first place. Ammann and Schaub (2021) show that investors are more likely to post when they achieve positive performance. Han et al. (2022) also argue that the propensity of investors to share their trading strategies increases with their improved performance. Thus, we argue that a higher performance level makes it more likely for investors to advertise their trading strategies both on social trading platforms and on other social media platforms. In a similar vein, once investors post about their trading strategies on eToro, they may also post about their strategies on other social media sites. Thus, investors' propensity to advertise their trading strategies on other social media is likely correlated with their performance level and post quantities on eToro.

D. Alternative data sample

We estimate the main model using a random subsample of all users on the platform. To mitigate the concern that our results are specific to this sample, we draw another random sample and reestimate the model. We report the results in Table D-1. Our main results are generally consistent in this exercise, indicating that they are not specific to the sample of investors drawn in our main analysis.

Table D-1 Estimation results with an alternative sample

Variable	Formation		Dissolution	
Leader's post quantity	0.0355***	(0.0036)	0.0140***	(0.0043)
Leader's post quality	0.5410***	(0.0249)	0.0787***	(0.0218)
Leader's number of replies	0.1344***	(0.0202)	0.0352	(0.0222)
Leader's received positive comment score	1.3749***	(0.1268)	0.5394***	(0.1417)
Leader's received negative comment score	−3.1217***	(0.6579)	−2.7341***	(0.4575)
Leader's average profit	0.0954***	(0.0116)	0.0482***	(0.0171)
Leader's std. dev. profit	−3.0192***	(0.4877)	−1.5855**	(0.7144)
Controls				
Leader's average holding time	0.3448***	(0.0419)	−0.0926	(0.0618)
Leader's lottery preference	0.2492	(0.3418)	0.8589**	(0.3880)
Leader's HHI	−0.5509***	(0.0862)	−0.3219***	(0.0871)
Follower's post quantity	−0.0474	(0.0295)	−0.1678***	(0.0356)
Follower's post quality	−0.2874***	(0.0881)	−0.0670	(0.0775)
Follower's average profit	0.0335	(0.0310)	0.0803***	(0.0305)
Follower's std. dev. profit	−6.5197***	(1.2552)	−11.8084***	(1.3585)
Nationality	0.6785***	(0.0670)	0.1049	(0.0784)
Age	0.0142	(0.0503)	−0.0430	(0.0539)
Homophily (male)	1.2894***	(0.1488)	−0.0653	(0.1152)
Homophily (female)	−1.6141**	(0.7136)	0.4221	(0.4180)
Image	2.6934***	(0.4250)	0.8731	(0.5382)
Bio	2.7266***	(0.1569)	−0.0283	(0.1906)
Experience	−0.0110	(0.0384)	0.0841**	(0.0403)
Wealth	−0.0338	(0.0330)	0.1197***	(0.0347)
Income	−0.0061	(0.0383)	−0.0086	(0.0408)
Risk	−0.0348	(0.0477)	−0.1263**	(0.0515)
Leader's popularity	0.0083***	(0.0004)	0.0014***	(0.0004)
Leader's activity	−0.0495***	(0.0125)	0.0285**	(0.0118)
Follower's popularity	−0.0432***	(0.0144)	−0.0086	(0.0101)
Follower's activity	−0.0143**	(0.0065)	−0.0183***	(0.0069)
Transitivity	0.0874***	(0.0241)	0.0392	(0.0605)
Constant	−27.8699***	(5.5367)	−0.0032	(1.1879)
Log likelihood	-14,168.14		-9,529.43	
Observations	11,000,237		20,796	

Notes: A positive coefficient in the formation process indicates that a link is more likely to form, whereas a positive coefficient in the dissolution process indicates that the link is more likely to be sustained. For better interpretation, the average profit and standard deviation of profit are scaled by a factor of 100, and the average holding time is scaled by a factor of 1/100. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

E. Examples of social communication texts

We provide some examples of posts, comments, and replies in Table E-1. As shown in the table, example posts are written by leaders. Leaders may want to advertise their performance, share their trading strategies or simply welcome new followers. Followers can ask for clarification in the comments (see comment 1), provide positive feedback to the leader (see comment 2), or complain about the leader's performance by leaving a negative comment (see comment 3). Finally, a leader can reply to a follower's question (see reply 1) or share his or her insights about the market by replying to the comment (see reply 2).

Table E-1 Examples of social communication

Post	**Text**
1	After being very long out I made some big mistakes. Now I trusted my instinct and made every trade without the knowledge from someone else. With this I turned back in the green and closed 2016 in green! awesome
2	The $EURUSD is a buy for me for the next few months... been at its lowest in years lately & I will definitely be looking for buy positions
3	@adelaya Hi, thank you for following. I wish you happy successful trading :)
Comment	**Text**
1	Why does it seem to fall in the after hours charts, though? Shouldn't it surge? I'm a newbie.
2	Your current investments are looking awesome!
3	I has lost so much money since i copied u and i never taste the earning sweet
Reply	**Text**
1	@Seregaomsk Additional funds will be used only when I open new deals, to distribute them to open positions while the only option is to stop copying and then copy with a new amount.
2	The beginning of the fall was undoubtedly connected with the general correction in the market, after which they are still only recovering. The current fall, in my opinion, is largely speculative, since there have been no negative indicators, news or decisions regarding YNDX lately. In this regard, I plan to keep them for now, as I look forward to recovery in the coming weeks.

Notes: Some social communications were not originally in English, and we present the translated versions.

F. Estimation with an interaction term between post quality and quantity

In this appendix, we add an interaction term between the post quantity and quality of a leader, keeping everything else the same as in the main model, to investigate the moderating effect of post quality. We estimate the extended model and show the results in Table F-1. We find that post quantity has a positive moderating effect on post quality in the link formation process, whereas the interaction term is nonsignificant in the link dissolution process.

Table F-1 Estimation results with an interaction term between post quality and quantity

Variable	Formation		Dissolution	
Leader's post quantity	0.0292***	(0.0038)	0.0137***	(0.0046)
Leader's post quality	0.4608***	(0.0278)	0.0527**	(0.0240)
Leader's post quality · leader's post quantity	0.0543***	(0.0078)	0.0060	(0.0082)
Leader's number of replies	0.0204	(0.0279)	0.0263	(0.0297)
Leader's received positive comment score	1.5178***	(0.1260)	0.6788***	(0.1423)
Leader's received negative comment score	−3.3375***	(0.7105)	−2.0007***	(0.4750)
Leader's average profit	0.1078***	(0.0136)	0.0789***	(0.0190)
Leader's std. dev. profit	−2.9316***	(0.5643)	−3.0615***	(0.7939)
Controls				
Leader's average holding time	0.3275***	(0.0439)	0.0498	(0.0658)
Leader's lottery preference	0.2642	(0.3562)	1.2250***	(0.3967)
Leader's HHI	−0.6933***	(0.0904)	−0.1408	(0.0929)
Follower's post quantity	0.0165	(0.0281)	−0.2784***	(0.0413)
Follower's post quality	−0.2531***	(0.0829)	0.0066	(0.0793)
Follower's average profit	0.0284	(0.0183)	0.0373*	(0.0193)
Follower's std. dev. profit	−3.4791***	(1.1314)	−8.3967***	(1.1205)
Nationality	0.7328***	(0.0706)	0.3396***	(0.0799)
Age	0.1186**	(0.0506)	0.0473	(0.0543)
Homophily (male)	0.9772***	(0.1360)	−0.1370	(0.1165)
Homophily (female)	−0.8536**	(0.4167)	0.2722	(0.3177)
Image	2.5040***	(0.4229)	0.4876	(0.5450)
Bio	3.0725***	(0.1879)	−0.2143	(0.2019)
Experience	−0.0687*	(0.0363)	0.1473***	(0.0405)
Wealth	0.0044	(0.0315)	0.0661*	(0.0347)
Income	−0.0275	(0.0380)	0.0218	(0.0409)
Risk	−0.0947**	(0.0471)	−0.0349	(0.0527)
Leader's popularity	0.0068***	(0.0004)	0.0011***	(0.0004)
Leader's activity	−0.0419***	(0.0124)	0.0390***	(0.0120)
Follower's popularity	−0.0595***	(0.0150)	−0.0055	(0.0093)
Follower's activity	−0.0264***	(0.0072)	−0.0844***	(0.0105)
Transitivity	0.0975***	(0.0268)	−0.0217	(0.0653)
Constant	−23.1984***	(6.9070)	1.0469	(0.9464)
Log likelihood	-13,512.34		-9,155.09	
Observations	11,000,219		19,744	

Notes: A positive coefficient in the formation process indicates that a link is more likely to form, whereas a positive coefficient in the dissolution process indicates that the link is more likely to be sustained. For better interpretation, the average profit and standard deviation of profit are scaled by a factor of 100, and the average holding time is scaled by a factor of 1/100. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

G. Alternative estimation using a conditional logit estimator

The Chamberlain correlated random effects applied in our main model require the assumption that η_i follows a conditional normal distribution depending on X_i with a constant variance, which is equivalent to a conventional random effects model that controls for the correlation function. Thus, the coefficients on the time-invariant observables (C_i) can be estimated. To correct for the incidental problem, another way to estimate follower-specific unobservables η_i is based on the conditional logit estimator (Wooldridge 2010), which allows η_i to be arbitrarily correlated with X_i. However, under the conditional logit estimator, η_i and time-invariant covariates C_i cannot be identified simultaneously. Thus, C_i should be excluded.

We use the link formation process to illustrate the model implementation. The utility gained by follower i from forming a link with leader j between periods $t-1$ and t is defined as follows:

$$y^*_{ijt} = \boldsymbol{\alpha} X_{it-1} + \boldsymbol{\beta} W_{jt-1} + \boldsymbol{\lambda} V_{ij} + \eta_i + \epsilon_{ijt}, \tag{G-1}$$

where the notation definitions are the same as those in the main model, Equation (6).

For the link formation process, we define

$$y_{ijt} = \begin{cases} 1, & y^*_{ijt} > 0 \\ 0, & \text{otherwise} \end{cases} \tag{G-2}$$

where y_{ijt} is a binary variable equal to 1 if follower i forms a link with leader j between periods $t-1$ and t. The link dissolution process is defined in the same fashion. In the dissolution process, y_{ijt} is equal to 1 if follower i maintains the link with leader j in period t and 0 otherwise.

We denote as n_i the sum of all binary outcomes for follower i's following status over all periods. That is, $n_i = \sum_{t=1}^{T_i} \sum_{j=1}^{J_{it}} y_{ijt}$, where T_i is the number of periods during which follower i exists on the platform and J_{it} is the number of leaders that follower i follows in period t. Each follower i has a corresponding vector with length $T_i \times J_{it}$. B_i is the set of all possible vectors in which n_i elements are equal to 1 and $(T_i \times J_{it} - n_i)$ elements are equal to 0. In other words, B_i represents all the possible scenarios in which follower i forms n_i links with the J_{it} leaders over T_i periods. Mathematically,

$$B_i = \{b \in \{0,1\}^{\{T_i \times J_{it}\}} | \sum_{t=1}^{T_i} \sum_{j=1}^{J_{it}} b_{jt} = n_i\}, \tag{G-3}$$

where b is one realization or scenario among all possible scenarios and b_{jt} denotes an element in vector b.

The conditional probability of y_i given n_i is defined as follows:

$$Pr(y_i|X_{it-1}, W_{jt-1}, n_i, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda}) = \frac{e^{(y_i \times (\boldsymbol{\alpha} X_{it-1} + \boldsymbol{\beta} W_{jt-1} + \boldsymbol{\lambda} V_{ij}))}}{\sum_{b \in B_i} e^{(b \times (\boldsymbol{\alpha} X_{it-1} + \boldsymbol{\beta} W_{jt-1} + \boldsymbol{\lambda} V_{ij}))}}. \quad \text{(G-4)}$$

From Equation (G-4), we observe that the conditional probability does not depend on η_i. Thus, the conditional log-likelihood is also independent of η_i and can be written as

$$CLL(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda}) = \sum_{i=1}^{I_t} \sum_{t=1}^{T_i} \sum_{j=1}^{J_{it}} \ln[Pr(y_i|X_{it-1}, W_{jt-1}, n_i, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda})], \quad \text{(G-5)}$$

where I_t is the total number of followers in period t, T_i is the number of periods in which follower i exists on the platform, and J_{it} is the number of leaders that follower i follows in period t. We estimate the model by maximizing its overall log-likelihood value. The estimation results are reported in Table G-1. We find that the results are generally consistent with those in the main model.

Table G-1 Estimation results using the conditional logit estimator

Variable	Formation		Dissolution	
Leader's post quantity	0.0423***	(0.0033)	0.0158***	(0.0043)
Leader's post quality	0.5410***	(0.0252)	0.0615***	(0.0223)
Leader's number of replies	0.1424***	(0.0207)	0.0371	(0.0235)
Leader's received positive comment score	1.4225***	(0.1265)	0.6971***	(0.1448)
Leader's received negative comment score	−3.7827***	(0.7123)	−1.9779***	(0.4737)
Leader's average profit	0.1055***	(0.0138)	0.0824***	(0.0191)
Leader's std. dev. profit	−3.0587***	(0.5681)	−3.1776***	(0.7985)
Controls				
Leader's average holding time	0.3306***	(0.0445)	0.0582	(0.0669)
Leader's lottery preference	0.2908	(0.3558)	1.2454***	(0.4001)
Leader's HHI	−0.6655***	(0.0902)	−0.1312	(0.0935)
Follower's post quantity	0.0044	(0.0318)	−0.2863***	(0.0410)
Follower's post quality	−0.2654***	(0.0835)	0.0207	(0.0782)
Follower's average profit	0.0292	(0.0197)	0.0080	(0.0197)
Follower's std. dev. profit	−3.9234***	(1.2711)	−8.4268***	(1.2890)
Nationality	0.7407***	(0.0708)	0.3411***	(0.0794)
Age	0.1122**	(0.0509)	0.0502	(0.0544)
Homophily (male)	0.9857***	(0.1359)	−0.1329	(0.1171)
Homophily (female)	−0.8746**	(0.4167)	0.2703	(0.3124)
Image	2.5029***	(0.4232)	0.5163	(0.5410)
Bio	3.0234***	(0.1872)	−0.2393	(0.2007)
Leader's popularity	0.0071***	(0.0004)	0.0012***	(0.0004)
Leader's activity	−0.0403***	(0.0123)	0.0402***	(0.0122)
Follower's popularity	−0.0372**	(0.0147)	−0.0052	(0.0084)
Follower's activity	−0.0273***	(0.0074)	−0.0819***	(0.0104)
Transitivity	0.0954***	(0.0292)	−0.0223	(0.0637)
Log likelihood	-11,343.56		-5,520.69	
Observations	11,000,219		19,744	

Notes: A positive coefficient in the formation process indicates that a link is more likely to form, whereas a positive coefficient in the dissolution process indicates that the link is more likely to be sustained. For better interpretation, the average profit and standard deviation of profit are scaled by a factor of 100, and the average holding time is scaled by a factor of 1/100. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

H. Social trading regulation

Relevant regulators classify social trading as portfolio management per the MiFID. In particular, the European Securities and Markets Authority (ESMA), the European Union (EU)'s securities market regulator, announced in 2008 that financial operators active on a social trading network could exercise "investment discretion by automatically executing the trade signals of third parties", which implied that brokers and market-makers active in that field were assimilated into the group of other financial intermediaries that need ad hoc authorization for portfolio management per the MiFID. Subsequently, the same authorization requirements were confirmed by the ESMA in 2012. Whenever a service provider makes an investment through an automated algorithm in view of trade signals coming from third parties—in relation to MiFID financial instruments—it implies that the provider has to perform some consequent duties related to a suitability assessment, the completion of business obligations, and information standards for both clients and authorities.

The corresponding directive, Directive 2004/39/EC, was first introduced in 2004. The main objective of the MiFID is to create a European financial market that encourages honest competition among participating companies and, at the same time, increasing consumer protection. The MiFID has been in force since January 31, 2007, and was superseded by the MiFID II on January 3, 2018. However, the MiFID II did not bring about meaningful changes to the regulation of social trading.

Similarly, the Financial Conduct Authority (FCA) specified that the service of social trading falls within Article 4(1)(9) of the MiFID, which defines "portfolio management" as "managing portfolios in accordance with mandates given by clients on a discretionary client-by-client basis, where such portfolios include one or more financial instruments." In copy and mirror trading, investment decisions are implemented with no intervention by the client other than an agreement ("mandate") between the service provider and the client on the discretionary service provided.[17] This interpretation has not changed over time.

[17] `https://www.fca.org.uk/firms/copy-trading`, last accessed Jul. 8, 2022

I. Heterogeneous effects across follower age

Prior studies have found that individuals who are younger in age are more likely to blog, visit social network sites, and rely on social media in their decision-making than are those who are older in age (Chou et al. 2009). Thus, we further scrutinize the implications related to age. Considering that social trading is a novel way in which to participate in financial markets that may particularly attract younger individuals, it is natural to ask whether our findings hold across all age groups. We thus examine whether the impact of the social communication and financial performance variables varies across age groups.

We split the dataset into two subsamples based on followers' age ranges. The first group includes followers between 18 and 44 years of age, and the second group includes followers who are older than 44 years. We again apply the STERGM with Chamberlain correlated random effects and summarize the estimation results in Table I-1. Younger followers are found to be rather sensitive to positive and negative comments in the link dissolution process, whereas the effects of comments are not significant for older followers. Interestingly, on the one hand, we observe that a leader's post quantity increases the probability of younger followers maintaining existing links, while post quality becomes nonsignificant. On the other hand, for older followers, it is a leader's post quality, rather than post quantity, that increases his or her probability of maintaining existing links.

Table I-1 Estimation results of heterogeneous effects across follower age

	Age 18-44				Age >45			
Variable	Formation		Dissolution		Formation		Dissolution	
Leader's post quantity	0.0401***	(0.0039)	0.0164***	(0.0049)	0.0479***	(0.0062)	0.0105	(0.0086)
Leader's post quality	0.5394***	(0.0299)	0.0378	(0.0258)	0.5360***	(0.0476)	0.1183***	(0.0435)
Leader's number of replies	0.1625***	(0.0243)	0.0362	(0.0268)	0.0980**	(0.0408)	0.0520	(0.0467)
Leader's received positive comment score	1.4521***	(0.1511)	0.7777***	(0.1677)	1.3502***	(0.2337)	0.4235	(0.2711)
Leader's received negative comment score	−3.3975***	(0.8234)	−2.2718***	(0.5437)	−4.9105***	(1.4228)	−1.4249	(0.9856)
Leader's average profit	0.1063***	(0.0168)	0.0843***	(0.0231)	0.1025***	(0.0240)	0.0644*	(0.0349)
Leader's std. dev. profit	−3.3353***	(0.6915)	−3.4721***	(0.9644)	−2.5321**	(0.9959)	−1.9930	(1.4724)
Controls								
Leader's average holding time	0.3381***	(0.0531)	0.0750	(0.0757)	0.3176***	(0.0822)	−0.0455	(0.1384)
Leader's lottery preference	0.2549	(0.4234)	1.5961***	(0.4635)	0.4184	(0.6589)	0.1769	(0.7674)
Leader's HHI	−0.6795***	(0.1067)	−0.0947	(0.1083)	−0.6742***	(0.1695)	−0.2589	(0.1839)
Follower's post quantity	0.0247	(0.0291)	−0.2958***	(0.0452)	−0.0918	(0.0914)	−0.1256	(0.1113)
Follower's post quality	−0.3982***	(0.0919)	0.0255	(0.0843)	0.3523*	(0.1911)	−0.1441	(0.2667)
Follower's average profit	0.0315	(0.0192)	0.0396*	(0.0205)	0.0485	(0.0419)	0.0188	(0.0513)
Follower's std. dev. profit	−1.9538*	(1.1476)	−7.9407***	(1.1700)	−15.3188***	(3.4464)	−12.3308***	(3.7036)
Nationality	0.7442***	(0.0835)	0.3683***	(0.0951)	0.7820***	(0.1319)	0.2821*	(0.1502)
Age	0.1678***	(0.0557)	0.0512	(0.0598)	−0.0686	(0.1378)	0.0949	(0.1458)
Homophily (male)	0.9779***	(0.1583)	−0.1980	(0.1412)	0.9981***	(0.2653)	0.0107	(0.2131)
Homophily (female)	−0.5263	(0.5133)	0.1230	(0.4684)	−1.2889*	(0.7172)	0.2633	(0.4362)
Bio	3.1826***	(0.2396)	−0.1083	(0.2663)	2.7522***	(0.2983)	−0.3716	(0.3209)
Experience	−0.0960**	(0.0424)	0.1803***	(0.0468)	0.0254	(0.0714)	0.0653	(0.0816)
Wealth	−0.0316	(0.0376)	0.0689*	(0.0406)	0.0863	(0.0587)	0.0610	(0.0660)
Income	0.0006	(0.0450)	−0.0127	(0.0476)	−0.1333*	(0.0731)	0.1288	(0.0804)
Risk	−0.0453	(0.0546)	−0.0831	(0.0602)	−0.2944***	(0.0956)	0.1110	(0.1072)
Leader's popularity	0.0071***	(0.0005)	0.0013***	(0.0004)	0.0078***	(0.0007)	0.0007	(0.0007)
Leader's activity	−0.0355**	(0.0145)	0.0398***	(0.0142)	−0.0989***	(0.0298)	0.0432*	(0.0235)
Follower's popularity	−0.0309**	(0.0139)	−0.0101	(0.0110)	−0.2332**	(0.1099)	0.3244**	(0.1585)
Follower's activity	−0.0200*	(0.0102)	−0.0995***	(0.0130)	−0.0346***	(0.0106)	−0.0533***	(0.0167)
Transitivity	0.0266	(0.0335)	0.0639	(0.0810)	0.3727***	(0.0596)	−0.3326**	(0.1346)
Constant	−23.9648***	(7.9251)	1.0044	(1.0160)	−14.7806	(15.7994)	3.5338	(4.1821)
Log likelihood	-9,739.95		-6,781.87		-3,736.63		-2,324.39	
Observations	8,015,758		14,476		2,984,461		5,268	

Notes: A positive coefficient in the formation process indicates that a link is more likely to form, whereas a positive coefficient in the dissolution process indicates that the link is more likely to be sustained. For better interpretation, the average profit and standard deviation of profit are scaled by a factor of 100, and the average holding time is scaled by a factor of 1/100. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

J. Estimation considering heterogeneous asset classes

In our main model, we investigate the effects of social communication in link formation and dissolution. Given that investors can trade assets from five different asset classes (stocks, foreign currency, cryptocurrency, commodity, and indices), this raises the question of whether followers who focus on a particular asset class consider specific factors when deciding how to allocate their funds across (potential) leaders. Across asset classes, the majority of trades are in stocks and foreign exchange. However, most (approximately 95%) leaders trade in multiple markets concurrently. Thus, it is not feasible to separately investigate investors who only trade in one asset class, as the subsample size is not large enough for reliable statistical inference. Consequently, we create two dummy variables ($Stock_j$ and FX_j) capturing the major asset classes. $Stock_j$ (FX_j) is equal to 1 if leader j trades only in the stock (foreign exchange) market, while all other scenarios are considered the reference group. To investigate whether the main effects of social communication differ by asset class, we then interact the asset class dummy with the leader's post quantity and quality in both the link formation and dissolution models.

Table J-1 summarizes the estimation results. Both interaction terms are nonsignificant. Moreover, the conclusions from the main model are not affected by the additional interaction terms in the model: that is, the higher the quantity and quality of a leader's posts are, the greater his or her likelihood of forming new links and maintaining existing links.

Table J-1 Estimation results across different asset classes

Variable	Formation		Dissolution	
Leader's post quantity	0.0418***	(0.0035)	0.0127***	(0.0045)
Leader's post quality	0.5102***	(0.0268)	0.0609***	(0.0223)
Leader's number of replies	0.1488***	(0.0213)	0.0462**	(0.0234)
Leader's received positive comment score	1.4537***	(0.1271)	0.6840***	(0.1427)
Leader's received negative comment score	−3.4831***	(0.7073)	−2.0650***	(0.4756)
Leader's average profit	0.0983***	(0.0141)	0.0685***	(0.0201)
Leader's std. dev. profit	−2.5322***	(0.5867)	−2.6632***	(0.8420)
Leader's post quantity · FX	−0.0023	(0.0069)	0.0089	(0.0084)
Leader's post quality · FX	−0.1570	(0.2532)	−0.0529	(0.2393)
Leader's post quantity · Stock	0.0473	(0.0364)	−0.1356	(0.0869)
Leader's post quality · Stock	−0.0055	(0.0506)	−0.3475	(0.2260)
FX	0.0570	(0.2529)	−0.0526	(0.2698)
Stock	0.5185***	(0.1648)	0.5227	(0.3221)
Controls				
Leader's average holding time	0.2704***	(0.0491)	0.0537	(0.0665)
Leader's lottery preference	−0.0040	(0.3591)	1.2413***	(0.3967)
Leader's HHI	−0.5531***	(0.0972)	−0.1508	(0.0955)
Follower's post quantity	0.0161	(0.0281)	−0.2816***	(0.0413)
Follower's post quality	−0.2548***	(0.0828)	0.0064	(0.0793)
Follower's average profit	0.0281	(0.0183)	0.0375*	(0.0193)
Follower's std. dev. profit	−3.4788***	(1.1295)	−8.3996***	(1.1217)
Nationality	0.7631***	(0.0705)	0.3452***	(0.0800)
Age	0.1085**	(0.0508)	0.0500	(0.0545)
Homophily (male)	0.9888***	(0.1366)	−0.1390	(0.1169)
Homophily (female)	−0.8604**	(0.4169)	0.2719	(0.3178)
Image	2.5329***	(0.4239)	0.4746	(0.5459)
Bio	2.9529***	(0.1860)	−0.2167	(0.2035)
Experience	−0.0688*	(0.0363)	0.1462***	(0.0405)
Wealth	0.0042	(0.0315)	0.0667*	(0.0347)
Income	−0.0278	(0.0380)	0.0204	(0.0409)
Risk	−0.0953**	(0.0471)	−0.0336	(0.0527)
Leader's popularity	0.0074***	(0.0004)	0.0011***	(0.0004)
Leader's activity	−0.0325***	(0.0119)	0.0403***	(0.0121)
Follower's popularity	−0.0598***	(0.0150)	−0.0058	(0.0094)
Follower's activity	−0.0262***	(0.0072)	−0.0844***	(0.0105)
Transitivity	0.1002***	(0.0265)	−0.0233	(0.0653)
Constant	−23.1621***	(6.9141)	1.0507	(0.9459)
Log likelihood	-13,507.15		-9,150.70	
Observations	11,000,219		19,744	

Notes: A positive coefficient in the formation process indicates that a link is more likely to form, whereas a positive coefficient in the dissolution process indicates that the link is more likely to be sustained. The reference group comprises all scenarios other than those only trading in the stock market or only trading in the foreign exchange market. For better interpretation, the average profit and standard deviation of profit are scaled by a factor of 100, and the average holding time is scaled by a factor of 1/100. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

K. Estimation controlling for country fixed effects

The social trading platform offers its services across a large number of countries, which raises the concern that the effects of social communication across geographical locations, for example, due to cultural differences or differences in the relevant regulation. Given the large number of countries and the diverse interaction across customers from different nationalities, it is not feasible to conduct separate analyses for each country. At the same time, it is not practical to incorporate the country as an interaction term with the variables of interest, such as the leader's post quantity and quality, as doing so would result in an excessive number of interaction terms. Therefore, we control for heterogeneity across the leader's country by controlling for his or her nationality with a set of dummies. Note that our robustness check using leader-level fixed effects also controls for heterogeneity across countries. We rerun the analysis for both the link formation and dissolution models. Table K-1 summarizes the estimation results, which are qualitatively consistent with those of the main model.

Table K-1 Estimation results with country fixed effects

Variable	Formation		Dissolution	
Leader's post quantity	0.0478***	(0.0036)	0.0166***	(0.0045)
Leader's post quality	0.4709***	(0.0266)	0.0738***	(0.0240)
Leader's number of replies	0.1042***	(0.0226)	0.0294	(0.0248)
Leader's received positive comment score	1.4637***	(0.1296)	0.7452***	(0.1501)
Leader's received negative comment score	−3.1634***	(0.6984)	−2.0067***	(0.4900)
Leader's average profit	0.1302***	(0.0151)	0.0861***	(0.0203)
Leader's std. dev. profit	−3.5845***	(0.6189)	−3.2479***	(0.8486)
Controls				
Leader's average holding time	0.2826***	(0.0489)	0.0070	(0.0749)
Leader's lottery preference	0.5930*	(0.3467)	1.2683***	(0.4152)
Leader's HHI	−0.4350***	(0.1011)	0.0327	(0.1079)
Follower's post quantity	0.0150	(0.0282)	−0.2806***	(0.0415)
Follower's post quality	−0.2550***	(0.0829)	0.0075	(0.0797)
Follower's average profit	0.0277	(0.0183)	0.0488**	(0.0200)
Follower's std. dev. profit	−3.5062***	(1.1335)	−8.8654***	(1.1393)
Nationality	0.7397***	(0.0740)	0.3397***	(0.0834)
Age	0.1084**	(0.0511)	0.0484	(0.0552)
Homophily (male)	0.8095***	(0.1429)	−0.1298	(0.1324)
Homophily (female)	−0.6711	(0.4190)	0.3789	(0.3353)
Image	2.3788***	(0.4241)	0.4514	(0.5729)
Bio	3.3752***	(0.2128)	−0.2255	(0.2352)
Experience	−0.0698*	(0.0363)	0.1494***	(0.0406)
Wealth	0.0028	(0.0315)	0.0642*	(0.0348)
Income	−0.0270	(0.0380)	0.0276	(0.0411)
Risk	−0.0966**	(0.0471)	−0.0430	(0.0530)
Leader's popularity	0.0052***	(0.0004)	0.0008*	(0.0005)
Leader's activity	−0.0244**	(0.0116)	0.0397***	(0.0123)
Follower's popularity	−0.0572***	(0.0149)	−0.0059	(0.0094)
Follower's activity	−0.0259***	(0.0073)	−0.0853***	(0.0105)
Transitivity	0.0923***	(0.0272)	−0.0385	(0.0654)
Constant	−22.2050***	(6.9228)	2.2508	(1.4399)
Country fixed effects	Yes		Yes	
Log likelihood	-13,239.59		-9,110.82	
Observations	11,000,219		19,744	

Notes: A positive coefficient in the formation process indicates that a link is more likely to form, whereas a positive coefficient in the dissolution process indicates that the link is more likely to be sustained. For better interpretation, the average profit and standard deviation of profit are scaled by a factor of 100, and the average holding time is scaled by a factor of 1/100. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

L. Estimation with investment percentage as the dependent variable

On many social trading platforms (such as ours), when a follower follows a leader (link formation), the follower's account is set up to automatically copy (mirror) the trades executed in the leader's account. As such, the link formed between the follower and leader is directly related to fund flows. Specifically, when a follower follows a leader, the follower's funds flow into the instruments invested by the leader. In contrast, if the follower stops following the leader, then the follower's fund flows out of the instrument invested by the leader. In our main model, we construct a leader-follower network using a binary variable to represent the fund flow—inflows as link formation and outflows as link dissolution. As such our main model identifies whether social ties (communication) affect fund flow.

To obtain a more granular view of network formation and dissolution, i.e., to not only answer the question of whether a link is established or dissolved but also be able to speak to the question of how much a follower invests in a given leader, we additionally study portfolio weights (i.e., the monetary amount allocated to the leader relative to the total portfolio value of the follower) allocated by the follower when following a leader. This information enables us to investigate whether followers allocate a larger or smaller share of their portfolio to a given leader. Hence, we use portfolio weights as an additional measure to examine how social communication affects fund flows. Since the dependent variable is now a percentage, we use a fractional regression model to model the link formation and dissolution processes.

We present the estimation results in Table L-1. The higher the leader's post quantity and quality are, the greater the investment weights (fund flow) allocated to follow and sustain a link with the leader. The higher the positive score of those comments received by the leader is, the higher investment weights (fund flow) allocated to follow and sustain a link with the leader. In contrast, the higher the negative score of those comments received by the leader is, the less investment weights (fund flow) allocated to follow and sustain a link with the leader. Furthermore, the better the leader's financial performance (as measured by higher average profit and lower standard deviation of profit) is, the higher investment weights (fund flow) allocated to follow and sustain a link with the leader.

Table L-1 Estimation results using investment percentage

Variable	Formation		Dissolution	
Leader's post quantity	0.0497***	(0.0040)	0.0061***	(0.0013)
Leader's post quality	0.5540***	(0.0385)	0.0215***	(0.0077)
Leader's number of replies	0.0719**	(0.0329)	−0.0075	(0.0082)
Leader's received positive comment score	1.6176***	(0.2408)	0.1882***	(0.0465)
Leader's received negative comment score	−2.5805*	(1.4593)	−0.3999**	(0.1757)
Leader's average profit	0.1339***	(0.0232)	0.0337***	(0.0080)
Leader's std. dev. profit	−4.7173***	(0.9587)	−1.4769***	(0.3378)
Controls				
Leader's average holding time	0.3285***	(0.0420)	−0.0013	(0.0222)
Leader's lottery preference	0.0329	(0.4909)	0.3495**	(0.1450)
Leader's HHI	−0.8978***	(0.1448)	−0.0405	(0.0366)
Follower's post quantity	−0.0513	(0.0491)	−0.0765***	(0.0293)
Follower's post quality	−0.1761	(0.1735)	0.0179	(0.0495)
Follower's average profit	−0.0201	(0.0272)	0.0148	(0.0179)
Follower's std. dev. profit	−3.2365*	(1.6750)	−3.7496***	(0.7835)
Nationality	0.7827***	(0.1207)	0.1694***	(0.0430)
Age	0.2023***	(0.0749)	0.0270	(0.0184)
Homophily (male)	1.1777***	(0.2027)	−0.1016*	(0.0543)
Homophily (female)	−0.6087	(0.7662)	0.0991	(0.1116)
Image	1.3145***	(0.4522)	−0.0377	(0.1883)
Bio	3.4314***	(0.3002)	−0.0350	(0.0554)
Experience	−0.1709***	(0.0466)	−0.0001	(0.0177)
Wealth	−0.0052	(0.0369)	0.0213	(0.0166)
Income	−0.0820*	(0.0476)	−0.0152	(0.0187)
Risk	−0.0882	(0.0589)	0.0090	(0.0239)
Leader's popularity	0.0076***	(0.0005)	0.0004**	(0.0002)
Leader's activity	0.0183	(0.0145)	0.0392***	(0.0063)
Follower's popularity	−0.0835**	(0.0355)	−0.0024	(0.0015)
Follower's activity	−0.0411*	(0.0225)	−0.0379***	(0.0113)
Transitivity	0.1160**	(0.0586)	−0.0158	(0.0237)
Constant	−12.6529	(8.9395)	−0.2332	(0.4462)
Log likelihood	-3,211.13		-31,694.84	
Observations	11,000,219		19,744	

Notes: A positive coefficient in the formation process indicates that a link is more likely to form, whereas a positive coefficient in the dissolution process indicates that the link is more likely to be sustained. For better interpretation, the average profit and standard deviation of profit are scaled by a factor of 100, and the average holding time is scaled by a factor of 1/100. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

M. Estimation controlling for the leader role

An investor may switch his or her role during our sample period. For example, an investor may start as a follower and then become a leader. Leaders may also follow other leaders while being leaders themselves. In our network analysis, we create an $N \times N$ (N denotes the total number of investors in the sample) adjacency matrix to examine the link formation (dissolution) process, and a link can be established between any two nodes. In other words, for an investor who starts as a follower, the adjacency matrix has already examined his or her probability of becoming a leader. As our data structure in network analysis allows for incorporating the scenario where a leader is also following other leaders, we create a new variable (denoted as $Is\_Follower_{jt}$) to indicate whether leader j is also following other leaders at time t. This time-variant dummy variable captures the dual role dynamic. We add this variable as an additional control variable to our link formation and dissolution models.

Table M-1 presents the estimation results. Our results show that the main variables of interest such as the leader's social communication and financial performance remain qualitatively consistent with those of our main model. Interestingly, we find that the coefficient of $Is\_Follower_{jt}$ is significantly negative in the link formation model, indicating that potential followers are more likely to follow a leader who is not following other leaders (a "pure leader"). However, the coefficient of $Is\_Follower_{jt}$ becomes significantly positive in the link dissolution model, suggesting that existing followers are more likely to dissolve their links with pure leaders. One interpretation of this observation is consistent with disconfirmation theory. According to expectation-disconfirmation theory (EDT) (Oliver 1980), users' satisfaction with a service/product is influenced by their confirmation of expectation. In our study, potential followers may regard pure leaders as more knowledgeable in the trading market since they do not follow others, which results in a higher expectation and a higher probability of forming a link, compared to those leaders who are also followers. However, once a link is formed, the initial higher expectation of pure leaders is more difficult to confirm, which can result in cognitive dissonance among followers and make the link more likely to dissolve (Bhattacherjee 2001).

Table M-1 Estimation results controlling for the leader role

Variable	Formation		Dissolution	
Leader's post quantity	0.0407***	(0.0033)	0.0157***	(0.0042)
Leader's post quality	0.5165***	(0.0256)	0.0652***	(0.0223)
Leader's number of replies	0.1475***	(0.0207)	0.0344	(0.0233)
Leader's received positive comment score	1.3333***	(0.1276)	0.6874***	(0.1422)
Leader's received negative comment score	−3.6830***	(0.7068)	−2.0687***	(0.4738)
Leader's average profit	0.1014***	(0.0140)	0.0783***	(0.0190)
Leader's std. dev. profit	−2.9759***	(0.5770)	−3.0457***	(0.7942)
Is follower	−0.3500***	(0.0687)	0.1407**	(0.0670)
Controls				
Leader's average holding time	0.2940***	(0.0465)	0.0676	(0.0666)
Leader's lottery preference	0.4152	(0.3463)	1.1472***	(0.3976)
Leader's HHI	−0.7234***	(0.0907)	−0.1080	(0.0940)
Follower's post quantity	0.0155	(0.0283)	−0.2770***	(0.0413)
Follower's post quality	−0.2531***	(0.0830)	0.0048	(0.0793)
Follower's average profit	0.0280	(0.0182)	0.0375*	(0.0193)
Follower's std. dev. profit	−3.4627***	(1.1217)	−8.3796***	(1.1210)
Nationality	0.7411***	(0.0706)	0.3401***	(0.0798)
Age	0.1148**	(0.0507)	0.0459	(0.0544)
Homophily (male)	0.9898***	(0.1360)	−0.1145	(0.1172)
Homophily (female)	−0.8626**	(0.4166)	0.2522	(0.3179)
Image	2.4221***	(0.4250)	0.4467	(0.5455)
Bio	2.8952***	(0.1894)	−0.2046	(0.2021)
Experience	−0.0679*	(0.0363)	0.1466***	(0.0407)
Wealth	0.0048	(0.0315)	0.0661*	(0.0348)
Income	−0.0270	(0.0380)	0.0224	(0.0411)
Risk	−0.0953**	(0.0471)	−0.0363	(0.0529)
Leader's popularity	0.0076***	(0.0004)	0.0010***	(0.0004)
Leader's activity	−0.0098	(0.0116)	0.0285**	(0.0126)
Follower's popularity	−0.0627***	(0.0151)	−0.0057	(0.0094)
Follower's activity	−0.0267***	(0.0072)	−0.0847***	(0.0105)
Transitivity	0.1085***	(0.0262)	−0.0323	(0.0653)
Constant	−22.4959***	(6.9132)	0.9761	(0.9505)
Log likelihood	-13,522.56		-9,153.15	
Observations	11,000,219		19,744	

Notes: A positive coefficient in the formation process indicates that a link is more likely to form, whereas a positive coefficient in the dissolution process indicates that the link is more likely to be sustained. For better interpretation, the average profit and standard deviation of profit are scaled by a factor of 100, and the average holding time is scaled by a factor of 1/100. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

N. Bi-weekly network view

In social trading, a follower can follow and unfollow a leader at any time, potentially calling for a network analysis using rather short horizons. At the same time, investment horizons may be not quite as short-lived. What's more, computational intensity increases in more granular analyses. For example, we already have 11,000,219 observations for link formation when organizing the data at the monthly level, and this number increases significantly using shorter horizons. Additionally, the average link duration in the data is around 47 days, and the previous literature (see, e.g., Pelster and Hofmann 2018) makes use of a monthly horizon as well. Therefore, we conduct our main analysis using monthly data. However, given that the median duration of links is 9 days, we additionally organize the data at the bi-weekly level to be able to provide a more detailed reflection of the link dynamics and test the robustness of our findings. We reestimate the link formation and dissolution models using the conditional logit estimator and summarize the estimation results in Table N-1. Our results remain qualitatively unchanged.

Table N-1 Estimation results with a bi-weekly period

Variable	Formation		Dissolution	
Leader's post quantity	0.0481***	(0.0050)	0.0182***	(0.0058)
Leader's post quality	0.3828***	(0.0295)	0.0562***	(0.0189)
Leader's number of replies	0.2107***	(0.0261)	0.0463*	(0.0243)
Leader's received positive comment score	1.9860***	(0.1206)	0.5240***	(0.1206)
Leader's received negative comment score	−1.4545**	(0.5965)	−1.2482***	(0.3411)
Leader's average profit	0.1258***	(0.0131)	0.0614***	(0.0167)
Leader's std. dev. profit	−3.1079***	(0.5862)	−2.2827***	(0.7338)
Controls				
Leader's average holding time	0.2512***	(0.0232)	0.0033	(0.0564)
Leader's lottery preference	0.3242	(0.2564)	0.8334***	(0.3178)
Leader's HHI	−0.4069***	(0.1200)	−0.1854**	(0.0820)
Follower's post quantity	−0.0819	(0.0527)	−0.1830***	(0.0399)
Follower's post quality	0.1494	(0.1280)	−0.1828**	(0.0852)
Follower's average profit	0.0212	(0.0465)	−0.0254	(0.0203)
Follower's std. dev. profit	−1.9800	(2.4241)	−10.3658***	(1.2958)
Nationality	0.8618***	(0.1018)	0.1830**	(0.0738)
Age	0.0463	(0.0565)	−0.0049	(0.0519)
Homophily (male)	1.0123***	(0.1405)	−0.1389	(0.1086)
Homophily (female)	−0.9332**	(0.4261)	0.1182	(0.2772)
Image	3.5562***	(0.6771)	0.1550	(0.5038)
Bio	3.1300***	(0.3001)	−0.1587	(0.1680)
Leader's popularity	0.0087***	(0.0005)	0.0007*	(0.0004)
Leader's activity	−0.0632**	(0.0285)	0.0191*	(0.0112)
Follower's popularity	−0.0995**	(0.0451)	−0.0282	(0.0179)
Follower's activity	−0.0252	(0.0187)	−0.1969***	(0.0170)
Transitivity	0.2539***	(0.0918)	0.2792***	(0.0795)
Log likelihood	-12,061.75		-7,698.92	
Observations	20,120,878		34,042	

Notes: A positive coefficient in the formation process indicates that a link is more likely to form, whereas a positive coefficient in the dissolution process indicates that the link is more likely to be sustained. For better interpretation, the average profit and standard deviation of profit are scaled by a factor of 100, and the average holding time is scaled by a factor of 1/100. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

O. Estimation incorporating market sentiment

One may wonder whether link formation and dissolution vary with varying market conditions. For example, it may be possible for followers to focus on different aspects in times of bull markets (high sentiment) compared to times of bear markets (low sentiment). The sample period between January 2016 and December 2017 covers a rather bullish market period. While markets generally are very bullish during our sample period, sentiment does show more variation within our sample period. To quantify sentiment, we utilize the FEARS measure, which is among the most widely used sentiment measures in the financial literature (Birru and Young 2022). SENT (Baker and Wurgler 2007) is another well-known sentiment index, which—however—is more closely associated with institutional sentiment, while FEARS, based on Google Search Volume (GSV), is more closely linked to retail investor sentiment. As the majority of investors on the platform are retail investors, we choose to use FEARS. Although the FEARS index is negative, on average, during our sample period, it does exhibit some time-series variation and positive periods, indicating negative sentiment. Using the index, we incorporate an interaction term between social communication (leader's post quantity and quality) and market sentiment in the link formation and dissolution models, which allows us to shed light on the varying effects with respect to time-varying sentiment.

We present the estimation results in Table O-1. Both leader's post quantity and quality are less important—have a lower effect size—when the FEARS index is higher (sentiment is lower). One possible interpretation is the following: when sentiment is lower, people are less susceptible to peers' communication.

Table O-1 Estimation results incorporating market sentiment

Variable	Formation		Dissolution	
Leader's post quantity	0.0359***	(0.0040)	0.0112**	(0.0047)
Leader's post quality	0.3741***	(0.0395)	0.0161	(0.0316)
Leader's number of replies	0.1247***	(0.0215)	0.0309	(0.0239)
Leader's received positive comment score	1.4546***	(0.1272)	0.7008***	(0.1432)
Leader's received negative comment score	−3.6285***	(0.7093)	−1.9513***	(0.4747)
Leader's average profit	0.1161***	(0.0135)	0.0838***	(0.0192)
Leader's std. dev. profit	−3.1014***	(0.5598)	−3.2398***	(0.8051)
Leader's post quantity · FEARS	−0.0029***	(0.0009)	−0.0018*	(0.0011)
Leader's post quality · FEARS	−0.0333***	(0.0057)	−0.0102**	(0.0047)
FEARS	−0.0036	(0.0134)	−0.0031	(0.0115)
Controls				
Leader's average holding time	0.3157***	(0.0451)	0.0500	(0.0658)
Leader's lottery preference	0.2990	(0.3611)	1.1992***	(0.3965)
Leader's HHI	−0.6741***	(0.0905)	−0.1554*	(0.0933)
Follower's post quantity	0.0213	(0.0289)	−0.2770***	(0.0413)
Follower's post quality	−0.2631***	(0.0835)	0.0043	(0.0794)
Follower's average profit	0.0272	(0.0184)	0.0386**	(0.0194)
Follower's std. dev. profit	−3.8439***	(1.1561)	−8.5462***	(1.1262)
Nationality	0.7330***	(0.0706)	0.3409***	(0.0799)
Age	0.1169**	(0.0507)	0.0467	(0.0544)
Homophily (male)	0.9716***	(0.1360)	−0.1466	(0.1165)
Homophily (female)	−0.8465**	(0.4167)	0.2797	(0.3173)
Image	2.4701***	(0.4233)	0.4695	(0.5442)
Bio	3.0542***	(0.1891)	−0.2198	(0.2017)
Experience	−0.0669*	(0.0363)	0.1501***	(0.0408)
Wealth	0.0061	(0.0315)	0.0693**	(0.0350)
Income	−0.0274	(0.0379)	0.0195	(0.0412)
Risk	−0.0945**	(0.0471)	−0.0332	(0.0531)
Leader's popularity	0.0074***	(0.0004)	0.0011***	(0.0004)
Leader's activity	−0.0442***	(0.0124)	0.0375***	(0.0120)
Follower's popularity	−0.0636***	(0.0152)	−0.0061	(0.0095)
Follower's activity	−0.0308***	(0.0074)	−0.0847***	(0.0105)
Transitivity	0.0978***	(0.0271)	−0.0313	(0.0652)
Constant	−22.5669***	(6.9076)	1.3462	(0.9567)
Log likelihood	-13,480.12		-9,141.98	
Observations	11,000,219		19,744	

Notes: A positive coefficient in the formation process indicates that a link is more likely to form, whereas a positive coefficient in the dissolution process indicates that the link is more likely to be sustained. For better interpretation, the average profit and standard deviation of profit are scaled by a factor of 100, and the average holding time is scaled by a factor of 1/100.
Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

P. Nonlinear term for post quantity and quality

To further study exactly how the leader's post quantity and quality are related to the likelihood of link formation, we add quadratic terms of the leader's post quantity and quality to our model. We reestimate the link formation and dissolution models and summarize the estimation results in Table P-1. We find a diminishing marginal effect of post quantity and quality on link formation. We observe a diminishing marginal effect of post quantity in the link dissolution model as well. In addition, from descriptive data analysis, we observe that the minimum value of the post quantity and quality of leaders who have followers are both zero, suggesting that there is no minimum amount of posts necessary to become a leader.

Table P-1 Estimation results with squared terms

Variable	Formation		Dissolution	
Leader's post quantity	0.1332***	(0.0105)	0.0402***	(0.0102)
Leader's post quantity^2	−0.0013***	(0.0002)	−0.0004***	(0.0002)
Leader's post quality	1.0291***	(0.0618)	0.0684***	(0.0223)
Leader's post quality^2	−0.1340***	(0.0171)	−0.0098	(0.0145)
Leader's number of replies	−0.0328	(0.0258)	−0.0048	(0.0284)
Leader's received positive comment score	0.9537***	(0.1372)	0.5908***	(0.1447)
Leader's received negative comment score	−4.5708***	(0.7304)	−2.0084***	(0.4754)
Leader's average profit	0.1114***	(0.0135)	0.0799***	(0.0190)
Leader's std. dev. profit	−2.8476***	(0.5548)	−3.1034***	(0.7962)
Controls				
Leader's average holding time	0.3377***	(0.0460)	0.0589	(0.0662)
Leader's lottery preference	0.1611	(0.3577)	1.1595***	(0.3976)
Leader's HHI	−0.5866***	(0.0905)	−0.1344	(0.0944)
Follower's post quantity	0.0145	(0.0280)	−0.2780***	(0.0412)
Follower's post quality	−0.2574***	(0.0827)	0.0085	(0.0794)
Follower's average profit	0.0281	(0.0182)	0.0370*	(0.0193)
Follower's std. dev. profit	−3.4636***	(1.1294)	−8.3729***	(1.1217)
Nationality	0.7374***	(0.0705)	0.3434***	(0.0799)
Age	0.1060**	(0.0508)	0.0472	(0.0544)
Homophily (male)	0.9966***	(0.1359)	−0.1449	(0.1165)
Homophily (female)	−0.8704**	(0.4167)	0.2747	(0.3175)
Image	2.3981***	(0.4246)	0.4653	(0.5458)
Bio	2.9414***	(0.1905)	−0.2276	(0.2022)
Experience	−0.0689*	(0.0362)	0.1478***	(0.0406)
Wealth	0.0041	(0.0315)	0.0656*	(0.0348)
Income	−0.0271	(0.0379)	0.0214	(0.0410)
Risk	−0.0943**	(0.0471)	−0.0348	(0.0528)
Leader's popularity	0.0099***	(0.0005)	0.0014***	(0.0004)
Leader's activity	−0.0373***	(0.0120)	0.0379***	(0.0120)
Follower's popularity	−0.0587***	(0.0150)	−0.0053	(0.0093)
Follower's activity	−0.0273***	(0.0072)	−0.0854***	(0.0105)
Transitivity	0.0969***	(0.0265)	−0.0163	(0.0653)
Constant	−22.2742***	(6.9025)	1.3555	(0.9503)
Log likelihood	-13,463.89		-9,151.40	
Observations	11,000,219		19,744	

Notes: A positive coefficient in the formation process indicates that a link is more likely to form, whereas a positive coefficient in the dissolution process indicates that the link is more likely to be sustained. Post quantity and quality are further mean centered to reduce the correlation between them and their quadratic terms. For better interpretation, average profit and std. dev. profit are scaled by a factor of 100. The average holding time is scaled by a factor of 1/100. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Q. Correlation matrix

Table Q-1 Correlation matrix

		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Transitivity	1	1													
In degree	2	0.537	1												
Out degree	3	0.189	-0.080	1											
Average profit	4	0.018	0.067	-0.074	1										
Std. dev. profit	5	-0.023	0.026	-0.085	0.102	1									
Holding time	6	0.002	0.059	-0.119	0.333	0.151	1								
MDD	7	-0.003	-0.021	0.045	-0.056	0.064	-0.050	1							
Post quantity	8	0.032	0.148	-0.034	0.008	0.020	0.001	-0.005	1						
Post quality	9	0.320	0.661	-0.092	0.064	0.033	0.031	-0.017	0.020	1					
Reply	10	0.112	0.168	-0.065	0.017	0.015	-0.010	-0.007	0.488	0.164	1				
Comment positive	11	0.106	0.224	-0.063	0.034	0.017	0.020	-0.021	0.139	0.323	0.202	1			
Comment negative	12	0.077	0.133	-0.077	0.025	0.030	0.010	-0.004	0.059	0.227	0.163	0.316	1		
Lottery preference	13	0.020	0.010	0.063	-0.008	0.004	0.012	0.046	-0.009	0.016	-0.021	-0.001	-0.003	1	
HHI	14	-0.047	-0.011	-0.254	0.006	0.169	-0.049	0.026	0.013	0.013	0.051	0.046	0.050	-0.161	1