Influence Trading:  
A Theory of Social Media*

Apostolos Filippas    John J. Horton
Fordham          MIT & NBER

December 4, 2020

Abstract
Social media users obtain consumption utility by consuming information produced by users they follow, and influence utility by producing information consumed by their followers. Influence can be obtained by producing highly-valued information that attracts followers, but also by reciprocation: users agree to follow each other, essentially trading influence. Influence trading profoundly changes the production and consumption of information on social media, compared to traditional media: “clubs” of low-quality users emerge, whose “members” reciprocally follow each other and actively produce information. If low-quality information is less truthful, influence trading creates pools of disinformation that other users can pick up and spread.

JEL Codes: L82, L14, B14

*Thanks to Bo Cowgill, Andrey Fradkin, and Evan Sadler for their helpful comments. Author contact information and code are currently or will be available at http://www.john-joseph-horton.com/.
1 Introduction

Social media matter. They are used by millions, and can influence their beliefs and behaviors (Alatas et al., 2019).\(^1\) Although many people clearly enjoy using social media—or at least prefer them to their next best alternative—the full social and political consequences of social media may be less benign. For example, a common critique is that social media serve as a vector for misinformation, giving voice to conspiracy theorists, white supremacists, state security forces, and propagandists (Allcott and Gentzkow, 2017; Vosoughi et al., 2018). Without the vetting of traditional media and editorial judgement, rumors and deliberate falsehoods can reach millions of users within hours.\(^2\) Progress in designing and operating social media platforms, as well as in advancing the corresponding policy debate, requires a better understanding of the economic forces that shape what information is produced on social media, who produces it, and who consumes it.

In this paper, we develop an economic model of social media platforms, where users are potentially both producers and consumers of information. It is an economic model in the sense that users make rational production and consumption choices, and these choices depend on the choices of other users. We then use this model to examine the economic effects that emerge on social media due to coexistence of consumption and influence.

Consider the case of Twitter, a large online social media platform. Twitter serves as the main motivating example throughout the rest of the paper; for brevity, we henceforth use “Twitter” in lieu of “social media,” although our analysis applies to social media in general. On Twitter, each user consumes the tweets of users she follows, and can produce tweets consumed by her followers. The consumption decision is conceptually similar to these made by traditional media users: people consume tweets for the same reason they listen to the radio, watch TV, read newspapers, and so on.

The production decision—whether to tweet—is less straightforward. Why do some users tweet, when almost no one is paid to tweet? The reason users tweet is because they derive influence utility from having followers, that is, other users who consume their tweets (Toubia and Stephen, 2013). The number of a user’s followers depends in part on how good she is at producing tweets other users value—similar to traditional media that differ in quality, and attract audiences accordingly. However, a user can also gain followers by engaging in reciprocal following relationships, that is, following users that follow her back (Hopcroft et al., 2011). We call this phenomenon influence trading.

The possibility for influence trading is a key distinction between social media and traditional media, and is what necessitates an equilibrium perspective. We first examine the case

\(^1\)http://www.pewinternet.org/2018/03/01/social-media-use-in-2018.

\(^2\)“Information” that would have been confined to poorly photocopied newsletters and email chains is now readily shared by the 45th President of the United States to his more than 70 million Twitter followers.
where users are vertically differentiated with respect to their tweeting ability, and otherwise identical. Because there is an opportunity cost to consumption, a user first decides on the lowest tweet quality that she is willing to consume, or, equivalently, on the lowest-ability user she is willing to follow “organically.” Two users that follow each other organically both receive some consumption utility from consuming each other’s tweets, and some influence utility from gaining one follower.

Users can also follow other users “reciprocally.” A reciprocal follower is a users who follows another user only if that user “follows back.” By definition, a reciprocal following relationship always decreases consumption utility. In addition to the consumption utility decrease, users also incur monitoring costs to sustain a reciprocal following relationship; these costs include the effort a user exerts in ensuring that her reciprocal followers hold their end of the bargain. When deciding whether to engage in a reciprocal following relationship, users weigh the benefits against the costs—the influence utility increase against the consumption utility decrease, and the monitoring costs. Users prefer, ceteris paribus, to engage in reciprocal following relationships with users of higher ability, as following them results in smaller consumption utility decrease.

We derive the platform equilibria when reciprocation—and hence influence trading—is not possible, and when it is. Without influence trading, every user simply follows the highest-ability users, down to some marginal user. The marginal user who actively tweets generates tweets with quality equal to the opportunity cost of consuming a tweet. This maximizes users’ consumption utility, but only a small number of high-quality users actively tweet.

With influence trading, the quantity of information produced increases, but the average quality of information produced and consumed decreases. Every user still follows the best users, but reciprocal following now takes place: “clubs” comprising users of similar abilities follow each other reciprocally, but do not reciprocally follow any user outside their club. Club “members” actively tweet, but would have otherwise not tweeted: without reciprocal followers these users would have no followers, and without followers they would not obtain any influence utility by tweeting. The consumption utility of club members decreases, as they engage in reciprocal following relationships. The tweet quality of the marginal user who actively tweets is equal to the marginal utility of gaining a single follower: this is exactly the first user whose ability is so low that she would not reciprocally follow herself.3 Because club members are of low ability, reciprocal following clubs become natural “reservoirs” of low-quality tweets.4

The baseline model assumes that users have identical preferences for other users’ tweets.

---

3 On a somewhat more general context than social media, Groucho Marx famously made a similar observation: “I don’t want to belong to any club that will accept me as one of its members.”

4 Reservoirs of disinfection can be thought of as the social media analogue to the epidemiological concept of reservoirs of pathogens (Haydon et al., 2002).
This approximation can be reasonable for Twitter communities such as #EconTwitter or #StatsTwitter, which consist of users with similar preferences, but some degree of horizontal differentiation is certainly present in real-life. Allowing for horizontal preferences dampens reciprocal following, because all users gain a number of organic followers in accordance to their abilities. However, because high-ability users may follow low-ability users, the edges of the reciprocal following clubs are “smoothed out,” and hence tweets of low-ability users can reach the entire platform user population (Bakshy et al., 2011). Introducing other sources of heterogeneity, such as idiosyncratic preferences for influence, has similar implications.

Compared to a platform where only the highest-ability users attract followers, influence trading results in lower-quality individuals generating and spreading more information. While user ability is defined solely in terms of consumption utility in our model, one could readily interpret a high-ability user as “truthful” (rather than “entertaining” or “educational.”) Indeed, users within some communities may want biased, or even untruthful information (Gentzkow and Shapiro, 2010). However, even if we assume that all users prefer truthful tweets, influence trading implies that, in equilibrium, some users will be willing to consume less truthful information in exchange for being listened to. As such, our model provides a supply-side explanation for untruthful tweets; this contrasts demand-side explanations where people want to be lied to.

Our model also differs from other supply-side explanations. In other models, the marginal user is made worse off because the barriers to entry are lower, and lower information dissemination costs lead to higher consumption utility (perhaps at the expense of variety, and a few dominant producers (Rosen, 1981). Instead, our model implicates influence trading as a key determinant of who produces, and the quality of what is being produced.

Several empirical findings can be explained by our model. Kwak et al. (2010) find a relatively small number of Twitter users with a large number of followers, but who only follow few users—corresponding to our users that receive followers without reciprocating. The remaining users follow a number similar to how many users follow them, creating a follower-to-following ratio close to 1, similar to clubs in our model. Kwak et al. also find that a substantial number of Twitter users purely consume: these users have few if any followers and do not tweet, as in our model. As tweeting ability is monotonic in both the number of followers and the follower-to-following ratio, our model explains why people pay for followers—they are exploiting the equilibrium relationship to make themselves appear better without actually being better.\(^5\) Our results also explain Twitter’s policy against “aggressive follow churn” behavior: Twitter prohibits following large numbers of users (some of whom will reciprocally follow back) and later unfollowing them, assuming some will not notice the defection.\(^6\)

The rest of the paper is organized as follows. A brief introduction to Twitter is presented in Section 2. The model is developed in Section 3. Results and extensions are discussed in Section 4. Empirical tests for model predictions are carried out in Section ?? . Section 5 concludes.

2 How Twitter works

Twitter is a micro-blogging platform that allows users to generate tweets—snippets of text, URLs, images and videos, which do not exceed 280 characters—and share them with other users. Tweets are publicly available, unless a user has elected to “protect” her tweets by making her account private.

A user chooses to see the tweets of other users by “following” them. Upon logging in Twitter, a user’s “home timeline” displays a stream of tweets from users she has chosen to follow. Twitter allows unilateral following (as do Instagram and Youtube), with any user free to “follow” any other user, in contrast to other online social media, such as Facebook and LinkedIn, where relationships are reciprocal.\(^7\) In addition to tweets of users followed, other content may be displayed on a user’s home timeline. This content includes suggestions of users to follow, tweets from users not followed, advertisements, and so on. To determine what additional content is displayed, Twitter uses proprietary algorithms that utilize a variety of signals, including how popular content is, and how the user’s network—the users she follows—are interacting with it.\(^8\)

The number of users a user follows on Twitter—as well as the number of people that follow her—is prominently displayed. Figure 1a shows a Groucho Marx Twitter parody account. It reports the number of user being followed, the number of users following the user, along with additional information the user may opt to provide, such as the user’s location. When a user views other users’ profile page, she can see whether she is following them, and whether they are following her back. Figure 1b shows an example of how other users’ profiles are viewed when logged in as a user. In this example, the two users follow one another, which is conveyed by the “Following” and “Follows you” tags, which are prominently displayed.

A user can interact with any tweet that is visible to her by “liking” it, commenting on it by generating another tweet, and re-tweeting it, which makes the original tweet appear on the home timeline of her followers. The user who generated the original tweet receives notifications about the interaction other users have with her tweets. Figure 1c shows one such tweet.

\(^7\)A user can “block” another user from seeing her tweets.

Figure 1: Screenshots from Twitter

(a) The profile page of a twitter user.

(b) Features of the relationship of two Twitter users.

(c) Features of a tweet.

Notes: This figure plots some screenshots from twitter. The top panel is a screenshot of a Groucho Marx parody Twitter profile, taken on December 4, 2020. Note the prominent display of the number of users followed by this user, and the number of users following the user. The middle panel is a screenshot of the Twitter account profile page for The Friars Club taken December 4, 2020, a user that follows this profile. The page indicates the The Friars Club follows the user back, and that the user follows The Friars Club, creating a bidirectional following relationship. Note that this relationship may either be organic or reciprocal. The bottom panel is a screenshot of a tweet. The number of comments, retweets, and likes associated with that tweet, appear below the contents of the tweet.
3 A model of influence trading

3.1 Setup

Consider a unit mass of Twitter users. Each user has ability $\alpha \in [0, 1]$, and generates tweets of quality $q(\alpha)$, where $q$ is strictly increasing. Abilities are publicly known, and independently and identically drawn from a given distribution. For ease of exposition, we assume throughout the rest of the paper that $q(\alpha) = \alpha$, and that abilities are uniformly distributed on the unit interval.

Users obtain consumption utility by following other users and reading their tweets. The consumption utility of a user is the sum of the tweet quality of the users she follows, minus an opportunity cost $q_0 \in [0, 1]$ for every user followed. We assume that all users face the same opportunity cost, that is, users have homogeneous preferences for consumption.

Users also obtain influence utility when followed by other users who read their tweets. Influence utility is increasing in the number of followers, exhibits diminishing returns, and does not depend on the identity of followers—only the number of followers matters. Formally, influence utility is a function $I : [0, 1] \to [0, 1]$ such that $I' > 0$ and $I'' < 0$. Users generate tweets only if they can obtain positive influence utility. Users without followers read but do not generate tweets.

The coexistence of consumption and influence allows two distinct follower types to emerge: organic followers and reciprocal followers. An organic follower is a user who follows some other user that increases her consumption utility. Because preferences are homogeneous, only users of ability $\alpha > q_0$ have organic followers. Every other user does not have organic followers, but may obtain followers by engaging in reciprocal following relationships.

A reciprocal follower is a user of ability $\alpha < q_0$ that follows a user of ability $\alpha' < q_0$, only if that user follows her back. In reciprocal following relationships, users trade decreases in consumption utility for increases in influence utility. In addition to the consumption utility decrease, we assume that users also incur a cost $c$ for every reciprocal following relationship they engage in. This cost $c$ captures the effort of maintaining reciprocal relationships, such as the cost a user incurs to periodically monitor that her reciprocal followers hold their end of the bargain, and have not stopped following her.\footnote{In Section 4, we examine how various extensions to this model change its predictions.}

We examine the equilibrium outcomes of two platform configurations: (i) when influence trading is impossible, which happens when follower information is not public, and hence not contractible, and (ii) when influence trading is possible, which happens when follower information is public.
3.2 Equilibria

Because users have homogeneous preferences, we can refer to them by their ability. Throughout the rest of this paper, user \( \alpha \) refers to a user of ability \( \alpha \), whenever convenient.

Let \( f_\alpha : [0, 1] \rightarrow \{0, 1\} \) denote the follow strategy of user \( \alpha \), such that \( f_\alpha(x) = 1 \) if she follows user \( x \), and \( f_\alpha(x) = 0 \) otherwise. The set of follow strategies of a user is \( F \), and the set of follow strategy profiles is \( \mathcal{F} = \times_{\alpha \in [0, 1]} F \).

Consider a follow profile \( f = (f_\alpha, f_{-\alpha}) \in \mathcal{F} \), where \( f_{-\alpha} \) denotes the follow profile of all users other than users of ability \( \alpha \). Under this follow profile, user \( \alpha \) (i) follows \( n_\alpha(f) = \int_0^1 f_\alpha(x)dx \) users, and obtains consumption utility \( U_\alpha(f) = \int_0^1 f_\alpha(x)(x - q_0)dx \), (ii) is followed by \( m_\alpha(f) = \int_0^1 f_x(\alpha)dx \) users, and obtains influence utility \( I_\alpha(f) = I(m_\alpha(f)) \), and (iii) engages in \( r_\alpha(f) = \int_0^{q_0} f_\alpha(x)dx \) reciprocal following relationships, and incurs monitoring cost \( C_\alpha(f) = r_\alpha(f)c \). User \( \alpha \) obtains total utility equal to

\[
V_\alpha(f) = U_\alpha(f) + I_\alpha(f) - C_\alpha(f).
\] (1)

We will restrict attention to follow profiles where any two users \( \alpha \leq q_0 \) and \( \alpha' \leq q_0 \) will either engage in a reciprocal following relationship, or will not follow each another. Formally, we will examine follow profiles in the set

\[
\mathcal{F}^+ = \{ f \in \mathcal{F} : f_\alpha(\alpha') = f_{\alpha'}(\alpha) \ \forall \alpha < q_0 \ \forall \alpha' < q_0 \}.
\] (2)

In words, if user \( \alpha \) unfollows a reciprocal follower \( \alpha' \), then \( \alpha' \) will also unfollow user \( \alpha \).

Consider now a follow profile \( f \in \mathcal{F}^+ \) where users \( \alpha \) and \( \alpha' \) reciprocally follow each other, and a follow profile \( f' \in \mathcal{F}^+ \) where they do not, but is otherwise identical to \( f \). The change in the utility of user \( \alpha \) from unfollowing reciprocal follower \( \alpha' \) is:

\[
V_\alpha(f') - V_\alpha(f) = (\underbrace{U_\alpha(f') - U_\alpha(f)}_{\Delta U_\alpha > 0} + \underbrace{I_\alpha(f') - I_\alpha(f)}_{\Delta I_\alpha < 0} + \underbrace{C_\alpha(f') - C_\alpha(f)}_{\Delta C_\alpha > 0}).
\] (3)

Disengaging from a reciprocal relationship leads to an increase in consumption utility, a decrease in influence utility, and a decrease in monitoring costs; engaging in a reciprocal relationship has the opposite effect. It is this fundamental tradeoff that our paper examines.

A follow profile \( f^* \) is an equilibrium follow profile if

\[
V_\alpha(f^*) \geq V_\alpha(f) \ \forall \alpha \ \forall f \in \mathcal{F}^+.
\] (4)

Note that our solution concept is stronger than that of a Nash equilibrium. First, users take into account that disengaging from a reciprocal relationship will result in being unfollowed back, capturing the fact that the underlying game is not a one-shot game. Second, equilibria
are required to be robust to coalitional deviations of arbitrary size.

3.3 Influence trading is impossible

We begin by examining the equilibrium outcomes in a platform where influence trading is impossible. This can be thought of as the case where follower information is hidden, and hence non-contractible. In terms of the parameters of our model, no influence trading takes place when the monitoring cost $c$ is large enough that engaging in reciprocal following always decreases a user’s total utility.

When influence trading is impossible, it is straightforward to show that every user follows only those users that increase her consumption utility, i.e., users of ability $\alpha > q_0$.

**Proposition 1.** In a platform without influence trading, there exists a unique equilibrium $f^*$ such that for all $\alpha$, $f^*_\alpha(x) = 1$ for $x > q_0$, and $f^*_\alpha(x) = 0$ otherwise.

In Figure 2a, we illustrate the equilibrium of a platform where influence trading is impossible, for the case of $I(x) = \sqrt{x}/2$, $q_0 = 0.8$, and $c = 0.2$. All users obtain the same consumption utility

$$U_\alpha(f^*) = \int_{q_0}^{1} x - q_0 \, dx = \frac{1}{2}(1 - q_0)^2.$$  \hspace{1cm} (5)

Users of ability $\alpha \leq q_0$ obtain total utility equal to their consumption utility, that is, $U_\alpha(f^*) = V_\alpha(f^*)$. These users do not generate tweets because they have no followers; we call such users “lurkers.” Users of ability $\alpha > q_0$ are followed by every user, generate tweets, and obtain influence utility $I_\alpha(f^*) = I(1)$.

3.4 Influence trading is possible

We next examine the equilibrium outcomes in a platform where influence trading is possible. This is the status-quo configuration of several social media platforms, including Twitter and Instagram: follower information is publicly accessible to users, thus enabling reciprocal following. In terms of our model parameters, the monitoring cost $c$ is small enough that some users are incentivized to engage in reciprocal following relationships.

To derive the equilibrium in this case, we need to make a number of observations. First, for any equilibrium profile $f^*$, $f^*_\alpha(x) = 1$ for all $x > q_0$, that is, organic following is invariant to whether influence trading takes place. Every user continues to follow users of ability $\alpha > q_0$ in equilibrium—it is straightforward to show that every follow profile where this does not hold can be improved upon. As such, users of ability $\alpha > q_0$ are followed by every user, and hence do not engage in reciprocal following. These users obtain the same consumption and influence utilities as in the case where influence trading is impossible. It is only users of ability $\alpha \leq q_0$ will have incentive to engage in reciprocal following relationships.
We next characterize the structure of optimal reciprocal follow strategies.

**Proposition 2.** In any equilibrium $f^*$, users of ability $\alpha \leq q_0$ will reciprocally follow users of abilities in a continuum $[q(\alpha), \overline{q}(\alpha)]$, that is, $f^*_\alpha(x) = 1$ if $x \in (q(\alpha), \overline{q}(\alpha)]$, and $f^*_\alpha(x) = 0$ otherwise. Furthermore, for any users $\alpha_1$ and $\alpha_2$ such that $\alpha_1 > \alpha_2$, $q(\alpha_1) \geq q(\alpha_2)$ and $\overline{q}(\alpha_1) \geq \overline{q}(\alpha_2)$.

**Proof.** Consider any equilibrium profile where user $\alpha_1$ engages in reciprocal following, but does not employ a continuum follow strategy. Then user $\alpha_1$ does not reciprocally follow some user $\alpha_2$, but reciprocally follows some user $\alpha_3 < \alpha_2$. User $\alpha_1$ would then be better off reciprocally following user $\alpha_2$, and incur a smaller consumption decrease. For the same reason, all users of ability higher than $\alpha_1$ that reciprocally follow user $\alpha_2$ will also reciprocally follow user $\alpha_1$; this proves the second part of the proposition. \(\square\)

Users of ability $\alpha$ who reciprocally follow will then increase their total utility by

$$R\left(\frac{q(\alpha)}{2}, \overline{q}(\alpha)\right) = \int_{\frac{q(\alpha)}{2}}^{\overline{q}(\alpha)} x - q_0 dx - \int_{\frac{q(\alpha)}{2}}^{\overline{q}(\alpha)} c dx + \frac{\overline{q}(\alpha) - q(\alpha)}{2}$$

We now provide a constructive proof of the unique follow equilibrium.

**Proposition 3.** In a platform with influence trading, there exists a unique equilibrium $f^*$, that can be computed as follows

1. For any $\alpha$, $f^*_\alpha(x) = 1$ for all $x > q_0$.
2. For users of ability $\alpha > q_0$, $f^*_\alpha(x) = 0$ for all $x \leq 0$.
3. Set $\alpha_1 = q_0$. For all users of ability $\alpha_1$, $f^*_\alpha(x) = 1$ iff $x \in (q(\alpha_1), \overline{q}(\alpha_1)]$, where $\overline{q}(\alpha_1) = \alpha_1$ and $q(\alpha_1) = \arg\max_{x \in [0, \overline{q}(\alpha_1)]} \left\{ R \left( x, \frac{q(\alpha_1)}{2}, \overline{q}(\alpha_1) \right) \right\}$. Users of ability $\alpha \in (q(\alpha_1), \overline{q}(\alpha_1)]$ reciprocally follow each other, and only each other.
4. Iterate by setting $\alpha_{i+1} = \overline{q}(\alpha_i)$, and repeating the previous step.
5. The process stops when either $\overline{q}(\alpha_i) = \overline{q}(\alpha_i)$, or $\overline{q}(\alpha_i) = 0$.

**Proof.** Amongst all users that choose to reciprocate, users of ability $\alpha = q_0$ are the most “valuable” ones, as they decrease their reciprocal followers’ consumption utility by the least amount. In any equilibrium follow profile, users that are reciprocally followed by those users will reciprocate—reciprocally following a higher ability user always is utility-improving for a lower-ability user, so long as it is utility-improving for a higher-ability user. As such, any profile where these users do not follow the highest-ability reciprocal followers can be improved.
upon, and is not an equilibrium profile. Because users are homogeneous, if a user belongs to the continuum of a higher-ability user, then that user’s continuum is identical. Any continuum that stops before marginal utility starts decreasing can be improved upon, and is not an equilibrium. Evoking the same arguments, uniqueness is obtained by contradiction.

In Figure 2a, we illustrate the equilibrium of a platform where influence trading is possible, for the case of \( I(x) = \sqrt{x}/2, \ q_0 = 0.8, \) and \( c = 0.2. \) Users of ability \( \alpha > q_0, \) are unaffected by the presence of influence trading; these users are followed by every other user, do not engage in reciprocal following, and do not experience any change in utility. The set of lurkers shrinks, as lower-ability users can improve their total utilities by forming reciprocal follower clubs.

Five reciprocal follower clubs form in our example, comprising users of abilities in the continua \([q_1, q_0], \ [q_2, q_1], \ [q_3, q_2], \ [q_4, q_3], \) and \([q_5, q_4]. \) Note that the extent of reciprocal following—the size of these clubs—gradually becomes smaller, as reciprocally following lower-quality users is costlier. Reciprocation stops when users no longer find it profitable to follow their own type: users of ability \( \alpha \leq q_5 \) maintain their lurker status, and continue to only consume.\(^\text{10} \) Club members are better off in terms of total utility but their consumption utility decreases, as they are forced to follow users below the opportunity cost \( q_0. \) Higher-ability users belong to more populous clubs and incur both larger monitoring costs, and higher consumption utility decreases.

### 3.5 Network structure

Many economic and social networks exhibit a core–periphery structure: a small number of users—the core—gather a disproportionate amount of connections, while other users—the periphery—maintain only few relationships (Krugman, 1991; Borgatti and Everett, 2000; Hojman and Szedl, 2008; Csermely et al., 2013).

In our model, the equilibrium network topology in a platform without influence trading resembles this core–periphery structure: the core comprises users of ability \( \alpha > q_0 \) who every platform user follows, and the periphery comprises users of ability \( \alpha \leq q_0 \) who are not followed by any user. The reciprocal following clubs that form when influence trading becomes possible can be thought of as peripheral “self-facing cores.” These clubs are cores in the sense that the influence of their members increases, but remain peripheral because only no user outside the club follows members belonging to the club. In Section 4.1, we show how introducing heterogeneity affects this structure.

\(^\text{10}\)This is, again, the (Groucho) Marx condition: a user does not want to be a member of any club that would have her as a member.
Figure 2: Following, consumption, and influence with homogeneous user preferences.

(a) Equilibrium when influence trading is impossible.

(b) Equilibrium when influence trading is possible.

Notes: This figure plots users’ followers, consumption utilities, and influence utilities, with and without follower information. It illustrates the case of uniformly distributed valuations on the unit interval, \( q_0 = 0.8 \), and \( I(x) = \sqrt{x}/2 \). In all panels, the yellow-shaded area depicts the number of organic followers, the blue-shaded area depicts the number of reciprocal followers, the green-shaded area depicts consumption utility minus monitoring costs, and the red-shaded area depicts influence utility. Panel 2a plots the users’ followers (left panel) and utilities (right panel) in a market where influence trading is impossible. Only high-ability users are followed, and those users are followed by every user. Every user obtains the same consumption utility, but only high-ability users obtain influence utilities. Panel 2b plots the users’ followers (left panel) and utilities (right panel) in a market with influence trading. The monitoring cost parameter is set to \( c = 0.2 \). Reciprocal following clubs emerge, in which users trade off decreases in consumption utility for increases in influence utility. The number of lurkers decreases, and both lurkers and high-ability users experience no utility change.
4 Discussion

4.1 Heterogeneity, and information propagation

The model of Section 3 assumes that users are endowed with different tweeting abilities, but are otherwise identical. To be sure, real-life users are heterogeneous in more ways than one, including in their consumption and influence preferences. We next examine how introducing additional sources of heterogeneity affects the predictions of our model.

Users may derive different utilities from consuming the content that a user produces. One way to model consumption preference heterogeneity in our framework is to assume that the opportunity cost \( q_0 \) a user faces is idiosyncratic and non-symmetric. Introducing horizontal consumption preferences does not change the results of the model significantly. One interesting change is that a reciprocal follower no longer follows users of similar abilities exclusively, but may also follow users of lower abilities for whom she faces low enough opportunity costs. However, reciprocal following relationships are still more likely to form between users of similar abilities. As a result, “smooth” reciprocal following clubs emerge, and no user lurks in equilibrium.\(^\text{11}\)

Users may also differ in how much utility they derive from exerting influence. A user \( \alpha \) that obtains higher influence utility from attracting followers sets a lower reciprocal follow bound \( q(\alpha) \), and reciprocally follows users outside her club. These users will follow her back, as her ability is higher than that of their own reciprocal followers—the members of the club they belong to. Having gained a more valuable reciprocal follower, these users may then choose to increase their respective reciprocal follow bound, unfollowing their lowest-quality reciprocal followers. This effect ripples through the population of reciprocal followers, and smooth reciprocal follower clubs form in equilibrium. The effect of introducing heterogeneous monitoring costs is identical.\(^\text{12}\)

Reciprocal follower clubs can be usefully thought of as echo chambers: they comprise users who are similar in terms of abilities, and who would not tweet had influence trading been impossible. With homogeneous preferences, the tweets of those users only reach other members of the same club. Introducing heterogeneity allows content bred within these clubs to enter the broader platform user population. Note that heterogeneity in consumption preferences implies connections between users of very different abilities, whereas heterogeneity in influence preferences implies connections between users of “adjacent” clubs, and hence similar abilities. Of course, there are other channels through which this information can flow, such as recommender systems that do not take into account the influence trading that takes

\(^{11}\)All derivations for this extension can be found in Appendix A.

\(^{12}\)In the extreme, a user whose preference for influence far outweighs her preference for consumption will engage in reciprocal following relationships with any willing user. Those are the social media users who openly advertise their “follow-for-follow” policy.
place in equilibrium.

4.2 Followers’ influence

The model of Section 3 assumes that a user’s influence utility does not depend on the identity of her followers, but only on the number of her followers. It is quite plausible, however, that becoming the 83rd user followed by @elonmusk closely resembles an endorsement, but becoming the 2,177th user followed by @boredelonmusk is far less consequential.\textsuperscript{13}

Consider the case where a user obtains influence that is a function of her follower’s organic followers. The reasoning behind this assumption is that a user’s organic followers are more likely to enjoy consuming her tweets, as well as interact with it by liking it and re-tweeting it. The resulting equilibrium is similar to this of Appendix A, with the only difference being that this “second-order influence” is factored in each user’s preference ranking.

It is worth noting that if a user obtains influence that is analogous to her reciprocal followers’ total followers, then multiple equilibria exist. In some of these equilibria, low-ability users attract higher numbers of reciprocal followers than high-ability users.

4.3 Production, consumption, and quality of information

When influence trading is possible, lower-quality users start generating tweets, decreasing the average quality of the information produced. These low-quality tweets result in a decrease in the average quality of information consumed only for users who belong to the same club when preferences are homogeneous, and for users not belonging to these clubs when preferences are heterogeneous. To the extent that low-quality users generate less truthful tweets, influence trading may increase the spread of disinformation. These negative effects are lesser in magnitude if tweeting intensity is analogous to a user’s followers—instead of the stepwise function assumed in the model of Section 3.

It is possible that influence trading may have positive effects. Consider the case where users face an opportunity cost for being an active user—either a consumer or both a producer and consumer—of a social media platform. Influence trading may incentivize users whose consumption utility does not outweigh this opportunity cost to be an active user of the platform. These users may in turn increase the influence utility of high-quality users they follow, incentivizing high-quality users to produce more highly-valued tweets. Another example can be found with the case where users may produce a a “viral” tweet—an extremely valuable tweet—with low probability every time they generate a tweet. Influence trading may then increase both the amount of viral tweets produced, and the probability that these tweets will

\textsuperscript{13}Follower data as of December 4, 2020. See https://twitter.com/elonmusk, and https://twitter.com/boredelonmusk for up-to-date data.
spread to the entire platform population.\footnote{Influence trading presumably gave us what is arguably the best meme of 2019—see \url{https://twitter.com/indiemoms/status/1189587264654974980}.}

### 4.4 Follower-to-following ratio

An oft-discussed proxy for user ability on social media is the follower-to-following ratio. The follower-to-following ratio of a user is simply defined as the number of followers a user has over the number of following she follows. In the model of Section 3, the follower-to-following ratio is greater than 1 for high-ability users, and less than 1 for low-quality users. A user’s ratio is (weakly) increasing in her ability, with members of the same reciprocal following club having the exact same ratio. The same intuition carries through in the case of heterogeneous preferences, with one important difference being that follower-to-following ratios are strictly increasing in a user’s ability (see Appendix A). This result helps partially explain why social media users attempt to emulate this equilibrium outcome, by keeping their follower-to-following ratio close to 1. To illustrate pictorially, in Figure 3 we plot users’ followers, following, and follower-to-following ratio when influence trading is possible. Panel 3a depicts the case of homogeneous preferences, and Panel 3b depicts the case of heterogeneous preferences.

**Figure 3:** Follower statistics when influence trading is possible

![Graph showing follower statistics](image)

*Notes:* This figure plots equilibrium quantities for the case of heterogeneous user preferences and public follower information. Panel 3a shows the range of reciprocation for each user ability $\alpha$, depicted as the range between the corresponding user’s lower bound $q(\alpha)$ (solid line) and the upper bound $\overline{q}(\alpha)$ (dashed line). Panel 3b shows the number users followed (solid line), the number of followers (dotted line), and the ratio of followers over users followed for each user ability $\alpha$.

It is worth noting that follower-to-following ratios may no longer be increasing in user
ability if users have heterogeneous monitoring costs, or heterogeneous preferences for influence utility (see Section 4.1). For example, a user with a strong preference for influence utility will engage in more reciprocal following relationships, making her follower-to-following ratio potentially greater than that of a higher-ability user with a weaker preference for influence.

5 Conclusion

This paper examines influence trading on social media, and elucidates its root cause by focusing on users’ rational decisions about consuming and producing information. We find that influence trading profoundly affects the production and consumption of information.

For would-be social media designers, our paper illustrates a core market design problem. Social media platforms regularly employ mechanisms aimed at increasing the consumption utility of platform users, such as lists that allow users to consume content generated by a subset of the users they follow, the option to block content generated by a user without unfollowing her, and content rankings that display tweets in non-chronological order. These mechanisms alleviate the consumption utility penalty associated with influence trading, but have no effect on influence utility, thereby making influence trading more compelling.

Whether there are effective platform design responses to this phenomenon is an open question. Platforms could disincentivize influence trading by increasing the monitoring costs. For example, platforms could disallow users to see how many followers they have, or make it harder to see whether a user “follows back.” Platforms could also impose a time limit before being able to unfollow a user. As influence trading may have positive effects, an interesting future direction would be to measure its costs and benefits for platforms.
References


A Heterogeneous consumption preferences

We extend the model of Section 3 to allow for horizontal consumption preferences. Our modeling approach is to no longer assume that opportunity costs are fixed, but instead that a user’s opportunity cost for following another user is independently and identically drawn from a given distribution. This is equivalent to assuming that users do not share one universal ordering of abilities, but rather that each user has an individual preference ordering for the other platform users. As such, there is now both a vertical component (user ability) and a horizontal component (idiosyncratic opportunity costs) to consumption preferences, but high-ability users are more likely to rank higher in a user’s ordering. For ease of exposition, we fix the opportunity cost distribution to be the uniform distribution on the unit interval.

Because there is no universal ordering of user abilities, each user now decides on the maximum consumption utility decrease that she is willing to incur in order to engage in a reciprocal following relationship. We adjust the follow strategy notation, and denote by $f_\alpha \in [0,1]$ the maximum consumption utility decrease that user $\alpha$ is willing to incur to reciprocally follow another user. If $f_\alpha = 0$, user $\alpha$ follows only organically, whereas if $f_\alpha > 0$, user $\alpha$ is willing to engage in reciprocal relationships.

A.1 Influence trading is impossible

When influence trading is not possible, reciprocal following relationships cannot be sustained. There exists a unique equilibrium $f^*$, where $f_\alpha^* = 0$ for all users. In this equilibrium, only organic followers exist. The probability that a user $\alpha$ will organically follow a user $x$ is $\Pr(q_0 \leq x) = x$. User $\alpha$ follows

$$n_\alpha(f^*) = \int_0^1 x \, dx = \frac{1}{2} \quad (A1)$$

users, and obtains consumption utility

$$U_\alpha(f^*) = \int_0^1 E[x - q_0 \mid q_0 \leq x] \Pr(q_0 \leq x) dx = \int_0^1 \frac{x^2}{2} da = \frac{1}{6}, \quad (A2)$$

where $E[x - q_0 \mid q_0 \leq x]$ is the expected consumption utility obtained conditional on a user of ability $x$ being followed organically. Similarly, user $\alpha$ attracts

$$m_\alpha(f^*) = \int_0^1 \alpha \, dx = \alpha \quad (A3)$$

followers, from whom she obtains influence utility $I_\alpha(f^*) = I(\alpha)$.

In Figure 4a, we illustrate the equilibrium of a platform where influence trading is impossible, for the case of $I(x) = \log(1 + x)$ and $c = 0.2$. Allowing for heterogeneous consumption
preferences smooths out following relationships, as each user attracts a followers in accordance to her ability. This is in contrast to the homogeneous preferences case, where the number of followers is a step function (see Figure 2a).

A.2 Influence trading is possible

When influence trading is possible, user $\alpha$ may increase her threshold $f_\alpha$, thereby trading off a decrease in her consumption utility for an increase in her influence utility.

By symmetry, users of the same ability will choose the same follow strategy in any equilibrium. Consider a user $\alpha$, and a follow strategy profile $f = (f_\alpha, f_{-\alpha})$. Under this profile, user $\alpha$ has a number of reciprocal followers equal to

$$r_\alpha(f) = \int_0^1 (1 - \alpha)(1 - x) \Pr(x + f_\alpha > q_0|x < q_0) \Pr(\alpha + f_x > q_0|\alpha < q_0) dx$$

$$= \int_0^1 (1 - \alpha)(1 - x) \min \left\{ \frac{f_\alpha}{1 - x}, 1 \right\} \min \left\{ \frac{f_x}{1 - \alpha}, 1 \right\} dx. \quad (A4)$$

It is worth talking through the terms of Equation (A4). Consider some ability level $x$, and observe that user $\alpha$ is followed organically by an $\alpha$ fraction of users of any ability level, including users of ability $x$. Amongst the remaining $1 - \alpha$ users that do not follow her, user $\alpha$ follows a fraction $x$ organically herself, leaving a fraction $(1 - \alpha)(1 - x)$ to be candidates for reciprocal following. Amongst these potential relationships, a fraction $\min \left\{ \frac{f_\alpha}{1 - x}, 1 \right\} \min \left\{ \frac{f_x}{1 - \alpha}, 1 \right\}$ materializes for follow strategies $f_\alpha, f_x$, giving a total of $r_{a,x}(f)$.

Under the follow profile $f$, user $\alpha$ follows

$$n_\alpha(f) = \frac{1}{2} + r_\alpha(f) \quad (A5)$$

users, and is followed by

$$m_\alpha(f) = \alpha + r_\alpha(f) \quad (A6)$$
users. User $\alpha$ obtains total utility

$$V_\alpha(f) = \frac{1}{6} + \int_0^1 r_{\alpha,x}(f)E[x - q_0|q_0 - f_\alpha \leq x \leq q_0]dx - c \int_0^1 r_{\alpha,x}(f)dx + I(\alpha + r_\alpha(f))$$

(A7)

$$= \frac{1}{6} - \int_0^1 r_{\alpha,x}(f)\frac{f_\alpha}{2}dx - c \int_0^1 r_{\alpha,x}(f)dx + I(\alpha + r_\alpha(f))$$

(A8)

$$= \frac{1}{6} - \left(\frac{f_\alpha}{2} + c\right)r_\alpha(f) + I(\alpha + r_\alpha(f))$$

(A9)

Because $V_\alpha(f_\alpha, f_{-\alpha})$ is concave in $f_\alpha$ for any $f_{-\alpha}$, a unique equilibrium exists (Rosen, 1965). Because influence is a concave increasing function, lower-ability users have stronger incentives to increase their influence utility, and will tolerate larger consumption utility decreases in equilibrium. As such, $\alpha > \alpha'$ implies $f^*_\alpha < f^*_{\alpha'}$.

The equilibrium of the market with hidden follower information is depicted in Figure 4b. In this simulation, we set the influence utility to $I(x) = \log(1 + x)$, and the monitoring cost to $c = 0.2$. In the left panel, we observe that users find it utility-improving to engage in reciprocal following—except users of the highest ability, $\alpha = 1$, who are already followed by every user of the platform. In the right panel, we plot the corresponding utilities. All users obtain lower consumption utility, because they follow users of abilities less than their opportunity cost, and they incur monitoring costs to maintain the reciprocal relationships. Lower-ability users engage in reciprocation most intensely: these users reciprocally follow higher-ability users, and hence obtain less consumption utility, but also the highest influence utility increase.
Figure 4: Following, consumption, and influence with heterogeneous user preferences.

(a) Hidden follower information.

(b) Public follower information.

Notes: This figure plots users’ followers, consumption utilities, and influence utilities, with and without follower information. It illustrates the case of uniformly distributed valuations on the unit interval and $I(x) = \log(1 + x)$ and $c = 0.2$. In all panels, the yellow-shaded area depicts the number of organic followers, the blue-shaded area depicts the number of reciprocal followers, the green shaded-area depicts consumption utility, and the red shaded area depicts influence utility. Panel 4a plots the users’ followers (left) and utilities (right) in a market with hidden follower information. Only organic followers exist, and users are followed by a number of followers commensurate to their ability. Every user obtains both consumption and influence utility: all users obtain the same consumption utility, but influence utility is commensurate to their abilities. Panel 4b plots the users’ followers (left) and utilities (right) in a market with public follower information. Reciprocal following now emerges, and users trade off decreases in consumption utility for increases in influence utility.