# Curation Algorithms and Filter Bubbles in Social Networks*

Ron Berman† Zsolt Katona‡

**Abstract**

Social platforms often use curation algorithms to match content to user tastes. Although designed to improve user experience, these algorithms have been blamed for increased polarization of consumed content. We analyze how curation algorithms impact the number of friends users follow and the quality of content generated on the network, taking into account horizontal and vertical differentiation. Although algorithms increase polarization for fixed networks, when they indirectly influence network connectivity and content quality their impact on polarization and segregation is less clear.

We find that network connectivity and content quality are strategic complements, and that introducing curation makes consumers less selective and increases connectivity. In equilibrium, content creators receive lower payoffs because the competition leads to a prisoner's dilemma.

Filter bubbles are not always a consequence of curation algorithms. A perfect filtering algorithm increases content polarization and creates a filter bubble when the marginal cost of quality is low, while an algorithm focused on vertical content quality increases connectivity as well as lowers polarization and does not create a filter bubble. Consequently, although user surplus can increase through curating and encouraging high quality content, the type of algorithm used matters for the unintended consequence of creating a filter bubble.

Keywords: Social Media, Polarization, Filter Bubbles, Echo Chambers, Fake News, Algorithmic Curation, Game Theory.

---

*We thank the NET Institute (www.netinst.org) for its generous support. Feedback from seminar participants at UTD Bass Conference, HUJI, SICS Conference, New York City Media Seminar, FCC, HBS, MIT Media Lab, IDC and Kellogg is acknowledged.

†Assistant Professor of Marketing, The Wharton School, email: ronber@wharton.upenn.edu

‡Associate Professor of Marketing, Haas School of Business, email: zskatona@berkeley.edu

# 1 Introduction

When social media platforms emerged, one of their major appeals was the promise of personalized content. By telling the platform who their friends are, users could have more social interactions and consume content that was tailored to their tastes. Because they could connect to homophilous friends, consumers were assumed to be exposed to content which felt more personal and within their domain of interest. If in the past, for example, national newspapers did not delve into local matters while local newspapers did, social platforms were meant to be the ultimate source for relevant and personalized content. A second appeal of social platforms was that users could also become content creators, while sharing it easily with their friends and followers. Creators of content could thus gain utility (through fame, reputation or intrinsic enjoyment) by having readers interact with the content without having to compete with other creators over limited publishing space.

The adoption of these platforms and accessibility of their content creation technology has led to a dramatic increase in the frequency and volume of available content. Statistics published by leading platforms[1] indicate that over $48,000$ photos were published on Instagram every minute in 2015, while Twitter users posted over $347,000$ tweets per minute in the same time frame. This large volume of content has often led to information overload and to desensitization to new ideas (Rodriguez et al. 2014).

To help users manage this overload, social platforms introduced curation algorithms whose goal is to select or rank the best content for users to consume. In the case of Facebook's newsfeed, for example, the platform selects stories that "are influenced by your connections and activity on Facebook. This helps you to see more stories that interest you from friends you interact with the most."[2] Instagram introduced a ranking algorithm "so you'll see the moments you care about first."[3] The problem the platform was trying to solve was that "On average, people miss 70 percent of their feeds. It's become harder to keep up with all the photos and videos people share as Instagram has grown."[4]

[1] See, *e.g.* `https://www.networkworld.com/article/2915538/7-staggering-social-media-use-by-the-minute-stats.html`, accessed September 12, 2019.

[2] `https://www.facebook.com/help/327131014036297/`, accessed July 9, 2019.

[3] `https://instagram-press.com/blog/2016/03/15/see-the-moments-you-care-about-first/`, accessed July 9, 2019.

[4] Ibid.

These curation algorithms constitute a major product design choice for social media platforms. Although the introduction of these algorithms was aimed to increase user activity and satisfaction, this had not always been the outcome. In April 2016 it emerged that users of Facebook have been sharing less original personal content over time and opted to post and share more professionally produced content such as news and other information.[5] Consequently, in January 2018 Facebook announced that they are changing their newsfeed algorithm because "public content – posts from businesses, brands and media – is crowding out the personal moments that lead us to connect more with each other".[6]

A second issue often attributed to the adoption of curation algorithms is the creation of "filter bubbles"[7]—a state in which social media consumption exhibits increased polarization and segregation of consumed content, and where ideas and concepts, limited in their diversity, match the consumer's beliefs. Early work by Van Alstyne and Brynjolfsson (2005) has shown that increased connectivity between online users can either fragment or better connect communities depending on the content and community preferences. As a result, "cyber-balkanization" may arise online. Recently, the issue has become so prevalent that The Onion, a satirical website, published an article with the sarcastic title "Horrible Facebook Algorithm Accident Results In Exposure To New Ideas".[8]

Despite the belief that algorithms are the root cause of increased online polarization, there is not much evidence to support this claim (Gentzkow 2016, Boxell et al. 2017). On one hand, because users can now choose to connect only with friends with similar tastes, we would expect decreased exposure to diverse ideas and increased polarization in content consumption. On the other hand, because an algorithm now assists the users in filtering for the best content, they may choose to change who they consume content from, and this may result in altered consumption patterns. We therefore make the distinction between *polarization* of content consumption, which may exogenously stem from individual self-selection over horizontally differentiated content, and a *filter bubble*, which entails increased polarization of the content consumed because of a curation

[5] `http://www.bloomberg.com/news/articles/2016-04-07/facebook-said-to-face-decline-in-people-posting-personal-content`, accessed July 9, 2019.

[6] `https://www.facebook.com/zuck/posts/10104413015393571`, accessed July 9, 2019.

[7] The term was coined in Pariser (2011).

[8] `http://www.theonion.com/article/horrible-facebook-algorithm-accident-results-expos-53841`, accessed July 9, 2019.

algorithm. In other words, although polarization may exist on social media (e.g., because of taste biases), we ask about the contribution of the platform to this polarization by catering to user preferences.

In this paper we analyze the effects that selection and filtering algorithms have on social network structure, the quality of content produced by network members and the polarization of content consumed. This process, which we call *algorithmic curation* may have several possible effects on consumers and creators of content which we call *receivers* and *senders*, respectively. For receivers, if algorithmic curation is able to improve their experience, they may become less selective in their friend selection resulting in longer friends lists. For senders, if algorithmic curation increases network connectivity, an individual sender may now face stronger competition in receiving attention for her content. Moreover, the sender may have a harder time passing the filtering bar of the algorithm, and a harder time passing the bar of quality that elicits a positive interaction from a receiver. All in all, algorithmic curation may cause an increase or a decrease in the quality of content on the platform, which may result in more selective or less selective friending by receivers.

To analyze these effects we set up an analytical model with three types of players; a platform, receivers and senders. The platform's goal is to design an algorithm that maximizes receiver utility, which would cause receivers to keep visiting the platform in the long run and allow the platform to maximize its income from advertising. Receivers derive utility from consuming content items that were produced by senders they follow. These items have two stochastic components to their utility which depend on the distance of a sender from a receiver in terms of their tastes, as well as the quality of the content item which may fluctuate among items produced by the same sender. To control and maximize the utility they receive from items, receivers decide whom to friend as well as how many content items to consume. We also model senders who set their quality level strategically and derive utility when their content is liked by receivers. Sender payoffs are thus dependent on the quality they provide to receivers and on the competition they face from other senders.

We compare three types of algorithms available to platforms with a benchmark case when no curation is in effect, while later we show that these algorithms encompass the range of options a platform will pick in equilibrium. The Perfect Algorithm (PA) filters out the content which is below a threshold for a given receiver and only shows content to the receiver which is above the threshold.

This algorithm is “perfect” in that it can observe the true utility the receiver will obtain from content. The Quality Algorithm (QA) selects the content that passes a minimum quality bar, but cannot tell for each receiver if the item is close to her taste or far from it. In this way, the platform encourages high quality content, and focuses less on the individual match to each consumer. The third algorithm, which we call the Distance Algorithm (DA), selects for each consumer the content from the closest friends to her, but cannot tell whether each item is of high or low quality. An extension in the online appendix further analyzes filtering based on the popularity of content and shows that the results are similar to the quality algorithm.

The analysis focuses on the equilibrium quality of content and number of friends receivers follow. Our first results, focusing on the receivers, show that when the platform uses a curation algorithm, receivers are incentivized to connect to more senders because the platform provides assurances that the receiver will not experience too low utility. The perfect algorithm provides more assurance than the quality algorithm and elicits higher connectivity. Interestingly the distance algorithm does not increase connectivity, because filtering on horizontally differentiated senders does not incentivize receivers to increase connectivity which remains unchanged compared to the no algorithm case.

Examining the filtering outcome on network connectivity allows us to compare the average social distances of consumed messages between a sender/receiver pair as a measure of content polarization. Intriguingly, when the average quality of messages is low compared to consumption costs, both the perfect algorithm and the quality algorithm increase the average distance, resulting in *decreased* polarization. This stems from the increased connectivity achieved by these algorithms when compared to the no algorithm case, leading to connections to individuals at a higher distance. Even though some messages may come from individuals far away in the social space, they will pass through the filters if their quality is high enough. Thus, a filter bubble does not form in this case, but on the contrary, filtering helps *reduce* polarization through encouraging more distant friendships.

Only when the average quality of messages is high does the perfect algorithm lead to a lower average distance and increased polarization. The intuition is that above a certain level of quality, receivers have connected to many senders, including some at high distances. The algorithm, however, still filters content with decently high quality, trying to compensate for the increased distances. This yields lower average distances observed by the receiver, a characteristic of a filter bubble.

As apparent, the popular argument that filtering always causes these bubbles may not hold in general. Content polarization may be high online because users are allowed to pick the "friends" they consume content from. If these connections were not impacted by curation algorithms, the filtering would have indeed increased polarization. However, because algorithms can indirectly influence the network structure, only the perfect algorithm increases polarization and only when messages are of high quality in a highly connected network. In a typical setting of relatively low quality messages, filtering algorithms tend to increase the average distance, increase content diversity and reduce polarization.

Analyzing the full model with endogenized sender behavior reveals that the equilibrium content quality under the quality and perfect algorithms is higher than the benchmark case when the cost of quality is high. When quality costs are low, senders produce high enough quality that no filtering is needed and the algorithms are not effective. When the cost of quality is high, however, the algorithms effectively threaten the senders to increase their quality or face filtering. This allows the platform to provide receivers with content that maintains a higher level of quality compared to the no algorithm case.

An interesting outcome of this analysis demonstrates that the user surplus (welfare) under a curation algorithms is higher than without an algorithm only when quality is relatively cheap to produce. Although the algorithms may reduce the payoff for senders, they more than compensate with increased user surplus, as long as the platform makes it cheap for senders to invest in quality. However, when quality costs are high, the social welfare declines when using an algorithm by placing too much pressure on content creators. While we do not directly study the business model and revenues of the platform, the user surplus analysis provides insights on product design as the platform arguably profits most when the sum of payoffs in the ecosystem is the highest. In essence, platforms face a conflicting choice; either increase welfare and create a bubble, or increase diversity but lower overall surplus.

The implications of our research apply to a growing trend of firms to employ machine learning algorithms to improve user experience, sometimes resulting in surprising reduction in user satisfaction, increased polarization in communication patterns, and dissatisfaction of advertisers. Taking into account the equilibrium effects of a curation algorithm in these two-sided platforms should help platforms better design their algorithmic products.

The remainder of the paper is structured as follows. Section 2 overviews related work on social algorithms, recommendation systems, and the impact of filtering and ranking algorithms. Section 3 introduces our model as well as the different types of algorithmic curation. Section 4 analyzes the receivers' decision problem and resulting connectivity with exogenous senders, while Section 5 explores the distance consumed messages arrive from to address the issue of filter bubbles. Next, Section 6 analyzes and endogenizes the senders' choice and receivers "like" activities, and Section 7 solves for the full equilibrium choice of filtering algorithm and analyzes the full equilibrium outcomes. Finally, Section 8 discusses limitations to our model and concludes. The proofs are relegated to the Appendix throughout the paper, while the online appendix analyzes other extensions to our model as well as the impact on user surplus.

## 2 Related Literature

The literature analyzing the impact of selection and filtering algorithms has initially focused on how to design better algorithms, famous examples of which are Shardanand and Maes (1995) and Linden et al. (2003). Recently, more focus has been put on the economics and impact of such algorithms on consumers and the diversity of consumed products (See, *e.g.*, Latzer et al. (2016) for a survey).

One stream of this recent research has focused on the impact of recommendation systems on product consumption trying to determine whether the diversity and volume of products consumed increases after a recommendation system is introduced. Oestreicher-Singer and Sundararajan (2012) and Hosanagar et al. (2013) find that the introduction of recommendation systems tends to increase the diversity of products bought by customers, as well as to increase their consumption volume.

A second stream of research focuses on the impact of news and media aggregators on the network structure and quality of content created online. Athey et al. (2017) and Chiou and Tucker (2017) find that news aggregators serve as a complement to news websites, and allowing news aggregators to include excerpts of content from news sites benefits news sites in the long run as the number of their visitors increase. In addition visitors are able to visit and explore more niche content. Dellarocas et al. (2013) and Roos et al. (2015) find that on one hand, websites that create content can benefit from aggregators helping visitors find their content, but on the other hand the same aggregators appropriate advertising profits as well as cause more competition for the content creator websites. In this sense, the findings are similar to ours that a filtering platform that helps consumers find better content causes more linking within the network which results in stronger competition among content creators.

Works analyzing polarization of online content consumption provide mixed evidence about whether algorithms drive an increase or decrease in online polarization and segregation. Bakshy et al. (2015) have shown that algorithmic curation does limit somewhat the cross transfer of political ideas between the virtual isles on social media, but the filtering effect of the algorithm has a smaller impact than the preferences of consumers on whether to consume or not consume the content. Nguyen et al. (2014) and Flaxman et al. (2016) found some evidence for polarization in the form of lowered average distance between the preferences of consumers, but also found an increase in the dispersion of content consumed. Lee et al. (2014) and Barberá (2014) have found that frequent use of social media is associated with more heterogeneous online connections (in terms of political ideology), and that increased social media use does not increase polarization, but rather the opposite. Finally, Boxell et al. (2017) provide empirical evidence against the argument that increased internet usage is associated with rising political polarization. We contribute to this research by considering how the endogenous formation of social media links and content production are impacted by curation algorithms, and how these affect the polarization of content consumption.

Su et al. (2016) analyze Twitter's "Who to follow" suggestions to users. The paper shows that popular users on the platform will "get richer" and will gain more followers, but that in general the entire network will have more connections created. Their analysis does not look at the strategic impact on content quality, polarization or impact of different algorithms. The complementary work in Iyer and Katona (2015) analyzes the incentives of social media members to become senders based on the structure of the network they face and the resulting amount of competition. The authors find that increased connectivity increases competition among senders, a finding which is similar to ours. This results in higher message intensity, but in equilibrium results in a lower amount of senders creating content. In our model we endogenize the network structure as well as analyze different possible algorithms for curation and their impact. Our findings on the impact on distance and quality of content are also an addition to this literature.

# 3 Model

We have three types of players in the market. The first one is the social networking platform that provides the infrastructure for consumers to befriend each other and share content. The second and third types are $N$ consumers of whom some are senders who produce content and some are receivers who consume content. Let $\alpha$ denote the proportion of consumers who take on the role of a sender. Thus, we have $\alpha N$ senders and $(1-\alpha)N$ receivers in the market. Throughout the paper we assume that a user can be either a sender or a receiver. The online appendix analyzes a model which endogenizes $\alpha$ and allows users to mix between sending and receiving.

## 3.1 Receivers

Receivers are identical, and are evenly distributed in the social network. Let $R$ be a prototypical receiver. When $R$ consumes an item posted by sender $j$, the item yields utility $U_j = Q_j - D_j$. The quality of the item $Q_j$ is distributed uniformly in the $[q_j, q_j + 1]$ interval, where $q_j$ is the baseline quality level set by sender $j$. The distance $D_j$ between the sender $j$ and receiver $R$ is also a random variable and is distributed uniformly between $[0, 2\varphi]$, where $\varphi$ measures the connectivity of a receiver with $0 < \varphi < 1$. Receivers set $\varphi$ by deciding how many senders to connect to, where the extreme case of $\varphi = 1$ means that the receiver is connected to all of the senders in the network. When connecting to senders, receivers will take into account the expected utility of the content generated by each sender. If the receiver will be indifferent between connecting and not connecting with a sender, we assume they decide to not connect with a sender. This is equivalent to the assumption that connecting to senders has a small cost associated with it.[9] We assume that receivers and senders are uniformly distributed throughout the social network, therefore the number of senders a receiver is connected to is $\varphi\alpha N$.

The setup has two main features. First, the receiver does not exactly observe the distance of content from a given friend,[10] but has an idea about the distribution of distances. This feature captures the notion that social distance is hard to observe, might also depend on a certain topic, and can fluctuate over time. Second, the expected distance becomes higher as an individual selects more friends. This is a natural assumption given that people tend to become friends with homophilous

[9] We elected not to model this cost to maintain the model's parsimony.

[10] We adopt the term "friend", officially used by Twitter to denote a user who is followed on Social media.

individuals. The specific distribution we use was selected to facilitate parsimony but is not required for the results to hold. The main assumption is that there are cases when receivers will receive negative expected utility if they connect to too many senders and if the quality of content produced is not high enough to compensate for that. The topology of a Salop circular model exhibits these same features. We chose this symmetric setup for the sake of parsimony; thus it may not capture the skewed nature of connections and the presence of "celebrity" nodes in a real social network. In the online appendix we show how introducing heterogeneity and generalizing our model can lead to a more realistic model and outcome.

When a receiver consumes content, she decides to read $J$ content items, each item from one of $J$ random friends, where $J < \varphi\alpha N$, the number of friends the receiver has among the senders. The cost of reading each item is $c$ and the utility of content from a random friend $j$ is $U_j = Q_j - D_j$ as defined before, with $j = 1 \dots J$.[11]

Consequently, receiver $R$'s payoff is:

$$\pi_R = \sum_{j=1}^{J}(U_j - c) = \sum_{j=1}^{J}(U_j) - Jc \tag{1}$$

Receivers also express their liking of content which gives utilities to senders. When a piece of content consumed by a receiver is above a certain threshold, the receiver likes it. The details of this process are described in the section that details the utilities of the senders.

### 3.2 The platform and curation

Consumers cannot be sure about the utility they get from a piece of content until they consume it. Due to the variation in content quality and fit, they often decide to consume content that turns out to yield low or potentially negative utility ex-post. We model algorithmic curation as a process that filters content in order to maximize receiver payoff.

An algorithm is described by a scoring formula that assigns a score $S_j$ to each content item and a threshold $t$ that filters all items with a score below the threshold. The score is based on the item's sender, receiver and content quality, and is defined as $S_j = \beta_Q \cdot Q - \beta_D \cdot D$. The parameters $\beta_Q$ and $\beta_D$ capture the weights that the social network platform assigns to the quality and distance of each content item. The threshold $t$ is also set by the social network platform.

[11]The online appendix shows the results hold when costs of consumption are convex.

In reality, platforms use algorithms that are imperfect and depend on the investment the platforms make in measuring $Q$ and $D$. We assume that the platform needs to invest to observe $Q$ or $D$, each at a cost of $\gamma_Q$ or $\gamma_D$ respectively.

Our main analysis compares the following algorithms, while Section 7.1 shows the conditions in which each of these algorithms is optimal:

- The perfect algorithm (PA) observes all relevant information and can perfectly determine receiver utility, setting the score equal to the receiver utility, that is $S_j = U_j$. The cost of implementation is $\gamma_Q + \gamma_D$.
- The quality algorithm (QA) does not observe social distance nor takes it into account. It observes quality perfectly and sets the score equal to the quality of a piece of content, that is $\beta_Q = 1$, $\beta_D = 0$ and $S_j = Q_j$, at a cost of $\gamma_Q$.
- The distance algorithm (DA) ignores quality and sets the score equal to the distance between the sender of a specific item and the receiver. In this case, the algorithm filters above a certain distance threshold. Thus, here we set $S_j = -D_j$, because $-D_j > t$ is equivalent to $D_j < -t$. The platform therefore sets $\beta_D = 1$ and $\beta_Q = 0$ at a cost of $\gamma_D$.

We assume that the objective of the platform is to maximize receiver utility net of algorithm implementation costs. This assumption is consistent with a platform that profits from advertising or from charging subscription fees. If the platform charges subscription fees, maximizing the utility of the receiver will allow it to extract the most surplus in fees. If the platform uses advertising as an income source, it can impose a higher disutility of advertising on receivers if the receivers have more surplus from reading non-advertising items (Godes et al. 2009). In the long-run, because platforms' income from advertising depends on the number of content readers, if the platform provides users with low utility items, they may also choose to abandon the platform and consume content on a competing platform that provides higher surplus.[12] Hence, the platform's payoff is:

$$\pi_R(\beta_Q, \beta_D, t) - \mathbb{I}(\beta_Q \neq 0)\gamma_Q - \mathbb{I}(\beta_D \neq 0)\gamma_D \tag{2}$$

where $\mathbb{I}(\cdot)$ is the indicator function.

[12] See, e.g., `http://www.emarketer.com/newsroom/index.php/facebook-losing-younger-users-faster-pace/` on young consumers abandoning Facebook for competing platforms. Accessed July 10, 2019.

### 3.3 Senders

Following evidence that users may change their behavior in order to generate more liking, commenting and sharing of the content they post,[13] we assume senders derive utility from receivers liking their content.[14] In order to receive more likes sender $j$ can increase her content quality by setting a higher $q_j$, resulting in a higher random quality $Q_j \approx U[q_j, q_j + 1]$. Higher quality comes at a cost of $kq_j$, making the sender's utility

$$\pi_S^j = \#(likes) - kq_j \tag{3}$$

An item can be liked by a receiver if it was observed by a receiver (was not filtered by the algorithm), as well as provided enough utility to merit a like. We assume that a receiver picks one item from the $J$ items she reads and likes it if it passes a reservation utility $r \geq c$. The restriction $r \geq c$ means a sender can receive a like only if they generated net positive utility for the receiver. Since the receiver picks an item randomly from her set of friends, having a larger set of friends means that the sender will face stronger competition for receiving likes from receivers with more friends. In addition, if receivers have a high bar for liking content, senders will have a tougher time generating likes.

Summing up the likes from all receivers and adding the cost, we can write the sender's payoff:

$$\pi_S^j = \varphi(1-\alpha)N\frac{\Pr(U_j > r \text{ and } S_j > t)}{J} - kq_j \tag{4}$$

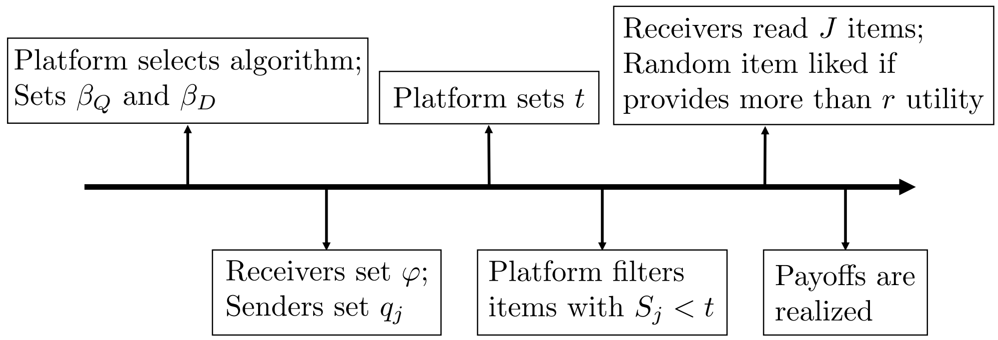
### 3.4 Timing

Figure 1 illustrates the timing of the game. First, the platform decides to invest in measuring $Q$ and $D$, and sets $\beta_Q$ and $\beta_D$ accordingly. Second, receivers pick $\varphi$ to determine the number of friends they have and senders choose their quality levels $q_j$ simultaneously. Third, the platform chooses the threshold $t$ for their algorithm. When the content is produced by the senders, the platform filters some of the items based on the algorithm chosen. Finally, receivers decide how many content items to read and like some of them, at which point payoffs are realized.

[13]See, e.g., "Society's New Addiction: Getting a "Like" over Having a Life", `https://www.vitalsmarts.com/press/2015/03/societys-new-addiction-getting-a-like-over-having-a-life/`, accessed July 10, 2019.

[14]The online appendix shows that the results are robust to assuming senders receive utility from impressions instead of "likes".

Figure 1: Timing of Actions




Investing in curation is a costly long-term effort of the platform, which is why we assume it is made before receivers connect to friends and senders generate their content. The analysis in the online appendix shows that allowing the platform to set the threshold $t$ before receivers make their connection decisions does not change the results. Our assumption that senders and receivers make their decisions simultaneously stems from the fact that it is unclear if senders respond in their content creation to how many followers they have and how much competition they have, or alternatively, that receivers elect to connect to senders that have a-priori set higher quality.

# 4 Analysis of Receivers and Platform Filtering

We first analyze the receivers' decision making process and how the platform's curation algorithm affects these decisions. Therefore, in this section we assume that senders are passive and they all set a baseline quality of $q_j = q$. After an analysis of a benchmark case where the platform does not apply algorithmic filtering, we focus on analyzing the three cases mentioned above (PA, QA and DA).

## 4.1 Benchmark

We begin with the case where the platform does not use any filtering algorithm. A prototypical receiver's expected payoff is

$$\mathbb{E}(\pi_R) = J \cdot (\mathbb{E}(U_j|\varphi) - c) \tag{5}$$

Note that the receiver cannot distinguish between senders and the expected utility is the same for all messages before reading, hence the receiver either reads content from no friend or from all friends. Given that $\mathbb{E}(U_j|\varphi) = q + \frac{1}{2} - \varphi$, the receiver will not read any content if $c - q > 1/2 - \varphi$, resulting in $J = 0$. Otherwise, the receiver will set $J = \varphi\alpha N$ and read content from all friends.

In the first stage, when the receiver decides on $\varphi$, the receiver solves:

$$\max_{\varphi} \mathbb{E}(\pi_R) = \max_{\varphi} \varphi\alpha N \left(q - c + \frac{1}{2} - \varphi\right) \tag{6}$$

The interior solution is $\varphi^*_{NA} = \frac{q-c}{2} + \frac{1}{4}$. Furthermore, if $c - q > 1/2$ receivers will not connect to anyone and set $\varphi^*_{NA} = 0$, whereas if $c - q < -3/2$, receivers will connect to everyone and set $\varphi^*_{NA} = 1$. The following proposition summarizes the main implications for the benchmark case: content quality and content consumption costs have opposite effects on the number of friends a receiver chooses. When quality is high, receivers are more open to friendships expecting higher content. However, a high consumption cost has the opposite effect, reducing the number of connections in the network.

**Proposition 1.** *The number of friends receivers choose is increasing in the content quality, but is decreasing in the cost of content consumption.*

A particular example that we will use for comparison between the cases is that of $c = q$, when we get $\varphi^*_{NA} = 1/4$. In this case the majority of friends the receiver will connect to will be her closer friends, with distances lower than $1/2$.

### 4.2 Perfect Algorithm

We now analyze the case of the perfect algorithm (PA) which fully observes and filters based on both message quality and sender distance. Starting at the last stage, recall that the platform intends to maximize receiver payoff. With a given $t$ threshold, the payoff is

$$\mathbb{E}(\pi_R) = J \cdot \Pr(U_j > t)\mathbb{E}(U_j - c|\varphi, U_j > t) \tag{7}$$

As in the benchmark case, the receiver cannot distinguish between different content pieces before reading so she either reads none or all. Thus the platform simply maximizes $\Pr(U_j > t)\mathbb{E}(U_j - c|\varphi, U_j > t)$, which reaches its maximum at $t^* = c$. The intuition is fairly straightforward: given that the platform can perfectly observe the utility that the receiver will get, it filters out all content that would decrease the receiver's total payoff. Solving for the receiver's optimal friendship choice, we derive the following proposition:

**Proposition 2.** *Under the perfect algorithm:*

- *When $-1 < c - q < 1$, receivers set $\varphi^* = \frac{1+q-c}{2}$. When $c - q \leq -1$ receivers connect to all senders in the network. When $c - q \geq 1$ receivers do not connect to any sender.*

- *The proportion of content filtered out by the algorithm increases in content consumption cost and decreases in content quality.*

The proposition tells us that the algorithm increases consumer utility by filtering out content that could have a negative impact on total consumer utilities. Interestingly, the algorithm also incentivizes receivers to connect to more friends than without an algorithm. In particular, consumers may have an incentive to connect to all senders when the cost is low enough compared to quality. The reason is that consumers can now afford to risk getting content from friends who are at a larger social distance, because the algorithm offers insurance against mismatching content.[15]

To illustrate our results better, we present the solution in the case of $c = q$ (the solution of the general case is presented in the Appendix). For $c = q$, we get

$$\Pr(U_j > c)\mathbb{E}(U_j - c|\varphi, U_j > c) = \begin{cases} \frac{1}{6}\left(4\varphi^2 - 6\varphi + 3\right) & \varphi < \frac{1}{2} \\ \frac{1}{12\varphi} & \varphi \geq \frac{1}{2} \end{cases} \tag{8}$$

Thus, in this case, the receiver maximizes

$$\max_{\varphi} \mathbb{E}(u_R) = \max_{\varphi} \varphi\alpha N \Pr(U_j > c)\mathbb{E}(U_j - c|\varphi, U_j > c)\big|_{c=q} \tag{9}$$

yielding the solution of $\varphi^*_{PA} = 1/2$ which is clearly higher than the benchmark case of $\varphi^*_{NA} = 1/4$. As opposed to (6), the benefit from increasing the number of friends has a positive quadratic term in the top line of (8), making it worthwhile to set a higher $\varphi$ than in the NA case.

Another good way to measure the algorithm's behavior and compare the different cases is to calculate the level of filtering, which is the probability that a piece of content from a random friend is hidden from the receiver. Naturally, with no filtering, this probability is 0. In case of the perfect algorithm, we get $p^*_{PA} = \Pr(U_j \leq c) = 1/2$ in the focal case of $c = q$.

It is interesting to contrast this result to the no algorithm result. Although the receivers double their number of friends with the perfect algorithm, half the content is filtered in this focal case.

[15] The tie-breaking rule where receivers do not connect if they are indifferent leads to the interior solution. The results are robust with the milder assumption that receivers connect with probability lower than 1 when indifferent.

The ratio of the amount of content consumed under the no-algorithm case to the perfect algorithm is 1, which means that the perfect algorithm does not effectively increase content consumption when $c = q$. However, the algorithm does make a difference in the average distance of a consumed message as we derive later in Section 5.

### 4.3 Quality Algorithm

In case of the quality algorithm (QA), the platform maximizes

$$\mathbb{E}(\pi_R) = J \cdot \Pr(Q_j > t)\mathbb{E}(U_j - c|\varphi, Q_j > t) \tag{10}$$

As we derive in the proof, $\Pr(Q_j > t)\mathbb{E}(U_j - c|\varphi, Q_j > t)$ reaches its maximum at $t^* = c + \varphi$. As long as $q < t^* < q + 1$, we obtain

$$\Pr(Q_j > t^*)\mathbb{E}(U_j - c|\varphi, Q_j > t^*) = \frac{1}{2}(c - q + \varphi - 1)^2 \tag{11}$$

Solving the receiver's maximization problem, we get

$$\varphi^*_{QA} = \frac{q - c + 1}{3}, \ \ p^*_{QA} = \frac{2(c - q) + 1}{3} \tag{12}$$

for $-1/2 < c - q < 1$, yielding the following result.

**Proposition 3.** *Under the quality algorithm:*

- *Receivers connect to more senders than without a filtering algorithm if $c - q > -1/2$. When $c - q \leq -1/2$ the quality algorithm does not filter any items and the receivers connect to the same number of friends as with the no algorithm case.*

- *Receivers connect to fewer friends and receiver utility is lower than under the perfect algorithm.*

The proposition demonstrates that the quality algorithm is generally a less effective solution than the perfect algorithm in increasing receiver utility and connectivity. However, consumers receive enough assurance from the algorithm to connect to more friends than without an algorithm as long as $c - q$ is not too low. The perfect algorithm is more efficient than the quality algorithm in both increasing connectedness as well as increasing receiver utility. The level of filtering also increases as the cost grows, offering a more selective experience to consumers, but due to the imperfect filtering it cannot keep up with the reduction in expected utility, thereby decreasing consumer incentives to connect to friends.

### 4.4 Distance Algorithm

In case of the distance algorithm (DA), the platform maximizes

$$\mathbb{E}(\pi_R) = J \cdot \Pr(-D_j > t)\mathbb{E}(U_j - c|\varphi, -D_j > t) = J \cdot \Pr(D_j < -t)\mathbb{E}(U_j - c|\varphi, D_j < -t) \qquad (13)$$

As we derive in the proof, $\Pr(D_j < -t)\mathbb{E}(U_j - c|\varphi, D_j < -t)$ reaches its maximum at $t^* = c - q - \frac{1}{2}$. Solving the receiver's maximization problem, we get the following result.

**Proposition 4.** *Under the distance algorithm:*

- *Receivers set* $\varphi^*_{DA} = \frac{q-c}{2} + \frac{1}{4}$, *when* $-3/2 < c - q < 1/2$. *When* $c - q \leq -3/2$ *receivers connect to all senders in the network. When* $c - q \geq 1/2$ *receivers do not connect to any sender.*
- *No content is filtered out by the algorithm in equilibrium.*

Interestingly, the distance algorithm does not encourage more connectedness in the network than in the benchmark NA case. The reason is that since the algorithm is set to filter out all content with some distance above a threshold, there is no value of adding friends with higher distances, which results in no actual content being filtered out in equilibrium.

## 5 Polarization and Filter Bubbles

By employing curation algorithms social networks found themselves in a controversy regarding the diversity of content each consumer reads. The claim is that by catering to the preferences of their users, social networks effectively limit the exposure to new ideas, decrease heterogeneity of content consumed and increase online polarization.

Since we make a distinction between exogenous polarization caused by the preferences of receivers for content that matches their tastes and increased polarization because of a filtering algorithm (a filter bubble), we can compare the equilibrium average distances of consumed content in our model to understand the impact of curation algorithms on polarization. On one hand, curation increases connectivity, which may lead to potentially more diverse content being consumed in terms of distance from the receiver. On the other hand, curation filters more content, which may lower the actual distance of consumed content. If in equilibrium receivers consume content which is more distant, then we say that content polarization has decreased in the market.

Let $\overline{d}$ denote the expected distance of an item that a receiver reads in equilibrium. We obtain the following results:

**Corollary 5.** *When $c - q$ is low, $\overline{d}_{QA} = \overline{d}_{NA} = \overline{d}_{DA} > \overline{d}_{PA}$. For intermediate values of $c - q$, $\overline{d}_{PA} > \overline{d}_{QA} \geq \overline{d}_{NA} = \overline{d}_{DA}$, and for high values of $c - q$, $\overline{d}_{QA} = \overline{d}_{PA} > \overline{d}_{NA} = \overline{d}_{DA}$.*

The results clearly show the different ways the three algorithms affect content available to readers. An immediate implication of the corollary is that when the quality is high ($c - q$ is low), the average distance of content seen by the receiver under all algorithms is higher than that of the perfect algorithm, resulting in more diverse content being displayed and lowered polarization. The intuition is that above a certain level of quality, receivers have connected to many senders under all algorithms and except for the perfect algorithm, none of them filter any content, resulting in higher average distances. The perfect algorithm, however, still filters content with quite high quality, trying to compensate for the increased distances, yielding lower average distance observed by the receiver.[16]

When quality is low, it is interesting that focusing on improving the user's experience and utility increases average distances and content diversity compared to the no algorithm case. This is a result of the increased connectivity achieved by these algorithms compared to the no algorithm case, leading to a shift in the distribution of distances further and allowing the receiver to experience more diverse content.

Figure 2 compares the impact of changes in $c - q$ on the connectivity of the receivers under the different algorithms (top panels) and the average distance of content consumed by the receiver with the different algorithms (bottom panels). To emphasize the differences between the algorithms we graph the absolute connectivity and average distances (left panels) and the difference of the PA and QA cases from the NA cases (right panels).

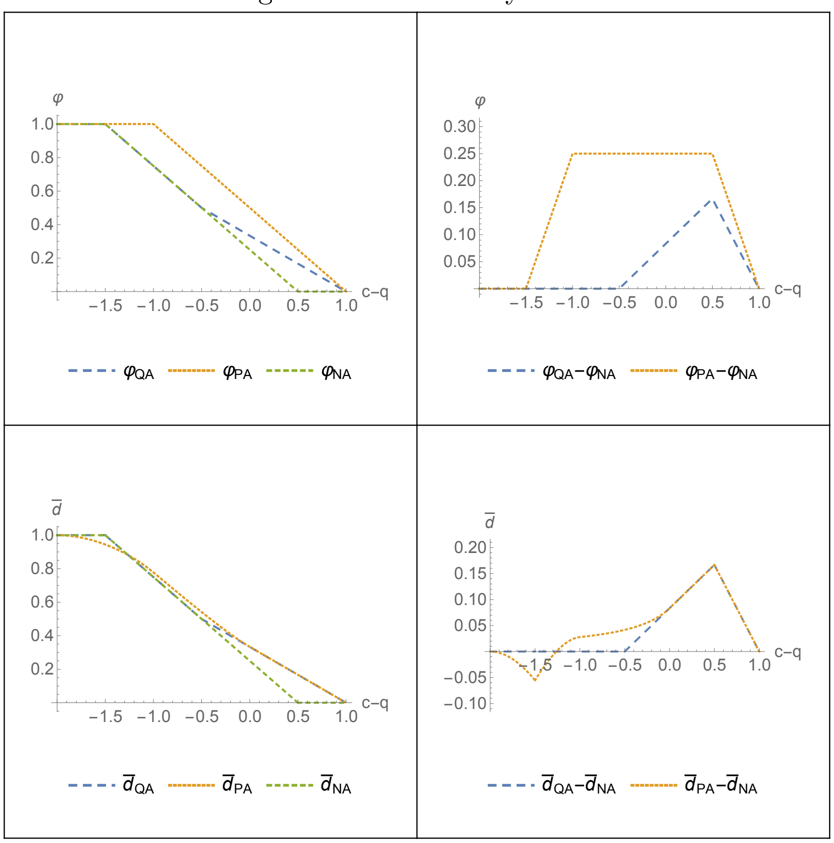
In the top panels we can see that both the perfect and quality algorithms generally increase connectivity compared to the new algorithm, and that this effect is stronger when the quality of content $q$ is low, or the cost of consuming content $c$ is high. Despite the increased connectivity, the bottom panels show that the filtering effect of the algorithms decreases the difference in average distances compared to the differences in connectivity, and in the case of the perfect algorithm, this

[16]The online appendix extends this analysis to heterogeneous levels of connectivity and lower levels of connectivity. The results remain unchanged.

may result in shorter average distances to senders, a characteristic of increased polarization because of a filter bubble.

The pattern we find sheds lights on which algorithms may be responsible for filter bubbles. The popular argument that filtering always causes these bubbles may not hold generally. Our results show that only the perfect algorithm decreases the average distance and only under special circumstances: when quality is high compared to consumption costs and when consumers connect to many senders to start with. In the typical situations—with relatively low quality senders—filtering algorithms tend to increase the average distance between a sender and receiver of a consumed message, and thus decrease polarizarion.

Figure 2: Connectivity and Polarization




Top: Receiver connectivity $\varphi$ as a function of $c - q$; Bottom: average distance $\overline{d}$.

Left: function values; right: difference from the no algorithm case.

## 6 Analysis of Senders

In this section we analyze the senders' behavior. The decision they make is the quality level that they set to shift the distribution of message quality available to receivers. As all senders are identical we are looking for a symmetric equilibrium where each sender sets the same level of quality. Throughout the section, we assume that $\alpha = 1/2$ for the sake of parsimony, implying that the number of senders equals the number of receivers. The online appendix extends the analysis to users who can endogenously choose $\alpha$.

### 6.1 No curation

We start with the benchmark case of no curation. When examining sender $j$'s decision, we denote the quality level set by her as $q_j$, while all other senders set $q_{-j}$. Let $G(\cdot)$ denote the CDF of the utility distribution $U_j = Q_j - D_j$ when $q_j = 0$ and let $g(\cdot)$ denote its PDF.[17] In the absence of curation, sender $j$ maximizes her expected payoff as follows:

$$\max_{q_j} \mathbb{E}(\pi_S^i) = \max_{q_j} \frac{\varphi N}{2} \cdot \frac{1 - G(r - q_j)}{J} - kq_j. \tag{14}$$

When receivers find consuming messages worthwhile they set $J = \varphi N/2$, hence the first order condition becomes:

$$g(r - q_j) = k \tag{15}$$

Since the second order conditions of the maximization problem hold only on the increasing parts of $g(\cdot)$, we find that the best response quality of a sender ($q^*$) to the share of friends each receiver has ($\varphi$) is

$$q^*(\varphi) = r + 2\varphi(1 - k) \tag{16}$$

We also assume $0 \leq k < \overline{k}_{NA}$ to make sure that the local maximum obtained above is also a global maximum.[18] When $k \geq \overline{k}_{NA}$ the equilibrium quality level is $q^* = 0$.

It is apparent that the best-response quality increases in $r$, the threshold set by the receiver, and it naturally decreases in $k$, the cost of quality.

[17] Exact formulas for the CDF and PDF are given at the beginning of the Appendix.

[18] The value is fully derived in the Appendix.

To derive the equilibrium, we use the formula from the previous section, where we determined that the receiver sets $\varphi = \frac{2(q-c)+1}{4}$. Substituting this into (16), we get a quadratic equation for $\varphi$ with the following solutions:

$$\varphi^* = \begin{cases} 1 & 0 < k \leq \frac{2(r-c)+1}{4} \\ \frac{1+2(r-c)}{4k} & \frac{2(r-c)+1}{4} < k < \overline{k}_{NA} \end{cases} \tag{17}$$

The strategic interaction between quality and connectivity drives up the number of friends receivers pick to the maximum level if costs are below a certain threshold. As it becomes costlier to produce quality the connectivity declines. A comparative statics analysis reveals the following:

**Proposition 6.** *Both $q^*$ and $\varphi^*$ increase in $r$ and decrease in $c$ and $k$.*

Proposition 6 shows that increasing the minimum utility $r$ required for a like will increase the number of friends a receiver chooses to connect to, as well as will increase the quality of content produced by the senders. In this sense, being more selective helps the receivers and improves the quality of content on the platform. This results in the receivers being less careful about selecting their friends, since giving likes more carefully helps align the senders to produce higher quality content.

When the cost of consuming content $c$ is increasing, the receiver decreases the amount of friends she connects to, because this will yield a lower distance from the senders on average, and will give a higher probability of consuming content whose utility passes the minimal cost facing the receiver. As a result the sender will face less competition in the market, and the quality will decrease. Finally, when $k$ increases, senders face a higher cost of producing quality and will lower the quality they choose to provide.

It is interesting to note that the positive interaction between quality and connections hurts senders in general. While more connections provide senders more opportunities to reach an audience, the limited rewards from a single receiver cancels this benefit out. At the same time, increased connectivity results in potentially larger social distances forcing senders to work harder and to invest more in quality in order to overcome the distances.

## 6.2 Perfect Algorithm

As we have shown in Section 4, the perfect algorithm filters out any messages that yield negative payoff to a receiver. Hence if the difference of the realized quality and the realized distance is less than the cost of consuming a message, the receiver does not see or read the message. Each message is thus potentially being filtered out with a probability depending on the quality level chosen by the sender. The choice set of messages the receiver can like changes depending on the realized qualities and distances yielding an a-priori random number of messages to choose from. The probability that a given sender who sets quality $q$ is filtered out is $G(c-q)$. We thus obtain the payoff for a given sender $j$, when all other senders set quality $q_{-j}$ as follows:

$$\max_{q_j} \frac{\varphi N}{2} \cdot (1 - G(r - q_j)) \sum_{L=0}^{J-1} \left[ \frac{1}{L+1} \binom{J-1}{L} (1 - G(c - q_{-j}))^L G(c - q_{-j})^{J-1-L} \right] - kq_j \tag{18}$$

The first order condition is thus

$$\frac{\varphi N}{2} \cdot g(r - q_j)) \sum_{L=0}^{J-1} \left[ \frac{1}{L+1} \binom{J-1}{L} (1 - G(c - q_{-j}))^L G(c - q_{-j})^{J-1-L} \right] = k \tag{19}$$

To obtain a tractable formula we assume that $N \to \infty$, and conduct all further analysis involving senders in the limit. The first order condition becomes

$$\frac{g(r - q_j)}{(1 - G(c - q_{-j}))} = k. \tag{20}$$

Compared to the benchmark case, senders have a higher incentive to invest in quality for a fixed like threshold, $r$. This is driven by the reduced competition due to filtering. On the other hand, the filtering threshold increases in the cost $c$ and potentially lowers incentives to invest in quality. To fully understand the combination of these forces, we determine the equilibrium quality level. As we show in the Appendix, senders will always set $q_j^* \geq r$. Substituting $q_j = q_{-j} = q^*$ into (20), and solving for the equilibrium, we obtain the equilibrium quality:[19]

$$q^* = \begin{cases} 2(1-k) + r & 0 < k \leq \frac{r-c}{2} \\ \frac{1+(2+c)k - \sqrt{2k(c+2k-r)+1}}{k} & \frac{r-c}{2} < k < \overline{k}_{PA} \end{cases} \tag{21}$$

Analyzing the result above we obtain the following results:

**Proposition 7.**

[19] The value of $\overline{k}_{PA}$ is derived in the Appendix.

- *$\varphi^* = 1$ and the receivers connect to all senders.*

- *$q^*$ increases in $r$ and in $c$ and decreases in $k$.*

- *For low values of $k$ senders set the same quality and obtain the same payoff as without filtering.*

- *For intermediate values of $k$, senders set a higher quality and obtain a higher payoff than without filtering.*

- *For high values of $k$, senders set a higher quality than without an algorithm, but obtain a lower payoff when $r - c$ is low.*

The first part of the proposition states that receivers connect in equilibrium to all senders. Given the full insurance the platform provides against low utility items, there is no risk for the receivers now and they might as well connect to all senders. That is, the result we obtained by just examining receivers is reinforced by the presence of active senders. The perfect algorithm induces receivers to connect to more senders and this incentivizes senders to invest in content quality, further increasing the willingness of senders to establish connections. In reality, users on social networks do not connect to all other users, of course. This is an artifact of our assumption that the platform can perfectly observe utilities and filter out negative utility items. A weaker assumption where the platform can only observe the utility with some noise would lead to an outcome where the equilibrium connectivity is not full. Moreover, the online appendix analyzes a model with connectivity heterogeneity where not all receivers connect to all senders. The results of the analysis on the impact of filtering on connectivity and filter bubbles are unaffected.

The second part of the proposition shows that $r$ and $k$ affect the amount of effort spent on quality in a similar way to the case of no curation. More demanding readers generally entice senders to produce higher quality as long as the cost is not high.

Interestingly, the second part also reveals that increasing receivers' cost has the opposite effect compared to without an algorithm. This parameter is an important component of our model as the basic idea behind curation is to prevent readers from spending valuable time on content yielding low utility. As our results show, curation has a double positive effect on receivers. In addition to the baseline effect of filtering out low quality content, it also incentivizes senders to create higher quality content. Higher costs increase the pressure on senders, since the probability of being filtered out increases with $c$. Therefore, senders have to spend more effort on quality to avoid being hidden from receivers.

The remaining results establish a comparison between the case of no filtering and perfect filtering. While qualities and sender payoffs are equal between the two cases when the cost of quality is low, filtering increases quality efforts when $k$ is high enough as discussed above. This threat of being filtered out makes senders increase quality over and above what they would set when the only incentive is to pass a reader's like threshold. The result also shows a potentially unintended consequence: together with the increased effort comes a reduced payoff for senders for high values of $k$ when $r - c$ is low enough. While a higher quality results in more likes in total, the effort is not worth it. But interestingly, filtering makes it more appealing for senders to spend more on quality because they expect fewer competitors for a given receiver due to curation. That is, senders are not driven to spend extra by the direct effect of filtering, but by the indirect, competitive effect akin to a prisoner's dilemma.

### 6.3 Quality Algorithm

As the quality algorithm filters out any message with quality less than $t^*$, senders can increase their base quality $q_j$ to make sure their items are not being filtered by the platform. Likewise, because receivers only like content above the threshold $r$, a sender can increase their quality $q_j$ to generate items with utility that passes this threshold.

Using the result from Section 4, suppose the platform sets $t^* = c + \varphi$ as the threshold for filtering quality, then the probability that an item from any senders passes filtering is $Pr(Q > t^*) = \min(q + 1 - t^*, 1)$. In order to receive a like, sender $j$'s item needs to pass the quality filter as well as the utility threshold $r$ set by the receiver. We denote as $P(q_j) = Pr(Q_j > t^*, U_j > r)$ the probability of item $j$ not being filtered for quality and being eligible for receiving a like.[20] The sender then maximizes

$$\max_{q_j} \frac{P(q_j)}{\min(q_{-j} + 1 - t^*, 1)} - kq_j \tag{22}$$

Examining the profit function and its first and second order conditions, the profit is locally maximized only when $q_j \geq t^*$. As we show in the Appendix, when $r$ is high, the qualities coincide

[20] The full derivation is available in the Appendix.

with the case of the no algorithm. For intermediate values of $r$, however, solving for the first order condition when assuming that in equilibrium $q_j = q^*$, we obtain[21]

$$q^* = \begin{cases} r + 2\varphi(1-k) & 0 < k \leq \frac{1}{2} + r - c \\ c + \varphi & \frac{1}{2} + r - c < k \leq \overline{k}_{QA} \end{cases} \tag{23}$$

Analyzing the equilibrium quality level and connectivity, we find that

$$\varphi^*_{QA} = \begin{cases} 1 & 0 < k \leq \frac{2(r-c)+1}{4} \\ \frac{1+2(r-c)}{4k} & \frac{2(r-c)+1}{4} < k \leq \frac{1}{2} + r - c \\ \frac{1}{2} & \frac{1}{2} + r - c < k \leq \overline{k}_{QA} \end{cases} \tag{24}$$

The following proposition summarizes the results.

**Proposition 8.** *Under the quality algorithm:*

- *No filtering takes place in equilibrium.*
- *When $k$ is low, the equilibrium behaves as in the no algorithm case. $q^*$ increases in $r$ and decreases in $k$ and $c$.*
- *When $k$ is high, senders set a higher quality in equilibrium, yet obtain a lower payoff compared to without filtering.*

The first part of the proposition finds that quality filetring is effective in the sense that senders always set a high enough quality to make sure their content is not filtered. This effectively increases the cost of quality facing the senders.

The second part finds similar effects for $r$, $k$ and $c$ as in the no algorithm case, and generally shows that when quality is cheap, quality filtering is not necessary.

The third part shows that quality filtering results in higher equilibrium quality levels for high quality cost $k$ but a lower profit for the senders. In such cases, quality filtering acts as a threat for the senders to maintain a higher quality even if the cost is too high, and as a result, senders maintain a fixed quality level above the level they would have set without filtering. This enhanced quality causes higher connectivity in equilibrium, but results in lower payoff for senders because of

[21] The value of $\overline{k}_{QA}$ is derived in the Appendix.

additional costs faced by the senders. The intuition is that the quality filter puts a sharp bound below which the senders face no competition, and as such, moving from setting a quality level below $t^*$ to above $t^*$ has a larger increase compared to the case of no curation. This results in a bigger incentive for each sender to pass the threshold as compared to before.

### 6.4 Distance Algorithm

As the results in Section 4 have shown, using a distance algorithm does not cause receivers to change their connectivity, which results in no filtering taking place for any value of $q$ that may be set by the senders.

Consequently, since senders can only affect the quality of the content they create, when facing a distance filtering algorithm senders have no incentive to invest in quality differently than in the case of no filtering. The conclusion is that the results of Proposition 6 apply to this case as well.

## 7 Platform Choice and Filter Bubbles in Equilibrium

After analyzing the Receiver and Sender's choices for connectivity and quality, we now analyze the platform's equilibrium algorithm choice, and verify that our previous results regarding filter bubbles hold in the full equilibrium with endgenous actions by the receviers, senders and the platform.

### 7.1 Equilibrium Algorithm Choice by the Platform

First, we show that the perfect algorithm maximizes receiver utility when qualities are exogenous. Second, we derive the equilibrium choice by the platform given the implementation costs.

Finding the platform's equilibrium choice with endogenous Senders turns out to be intractable. Consequently, we prove that the algorithm choice results do hold in a full equilibrium when the platform can pick between the perfect ($\beta_Q = \beta_D = 1$), quality ($\beta_Q = 1$, $\beta_D = 0$) and distance ($\beta_Q = 0$, $\beta_D = 1$) algorithms. We further provide numerical evidence that these three algorithms are the only choices the platform will consider when $r = c = 0$, implying the results will follow through when $\beta_Q$ and $\beta_D$ are not constrained.

When $q$ is exogenous and $\beta_Q$ and $\beta_D$ can be set arbitrarily, the platform sets

$$t^* = \arg\max_t J \cdot \Pr(S_j > t)\mathbb{E}(U_j - c|\varphi, S_j > t)$$

while the receiver sets $\varphi^* = \arg\max_\varphi J \cdot \Pr(S_j > t^*)\mathbb{E}(U_j - c|\varphi, S_j > t^*)$.

The derivation in the Appendix shows how the platform will set a different filtering threshold $t^*$ depending on the values of $c$, $q$, $\varphi$, $\beta_Q$ and $\beta_D$. Taking into the account the utility maximizing behavior of the receiver and picking $\beta_Q$ and $\beta_D$ that will maximize the equilibrium receiver's utility, we obtain the following results:

**Proposition 9.**

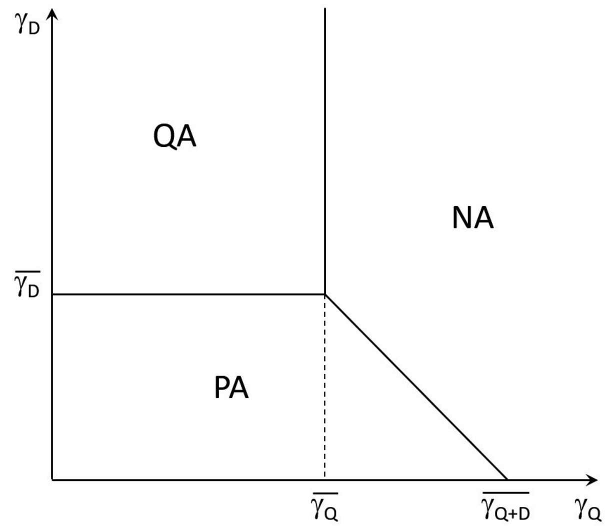
- *The algorithm that maximizes receiver's utility has $\beta_Q = \beta_D$. Specifically, the PA algorithm with $\beta_Q = 1$ and $\beta_D = 1$ maximizes the receiver's utility.*
- *There are thresholds such that the platform will curate using the PA algorithm when both $\gamma_Q + \gamma_D \leq \overline{\gamma_{Q+D}}$ and $\gamma_D \leq \overline{\gamma_D}$, will use the QA algorithm when $\gamma_Q \leq \overline{\gamma_Q}$ and $\gamma_D > \overline{\gamma_D}$, and will forgo any algorithm, when both $\gamma_Q + \gamma_D > \overline{\gamma_{Q+D}}$ and $\gamma_Q > \overline{\gamma_D}$.*
- *Specifically, when $q = c$, the platform will choose the PA algorithm iff $\gamma_Q + \gamma_D \leq \frac{\alpha N}{48}$ and $\gamma_D \leq \frac{\alpha N}{108}$, it will pick the QA algorithm if $\gamma_Q \leq \frac{5\alpha N}{432}$ and $\gamma_D > \frac{\alpha N}{108}$, and no algorithm otherwise.*

Because the receiver and the platform aim to maximize the same value (receiver's utility), once an algorithm has been implemented, they do not have misaligned incentives. Thus, when $\beta_Q$ and $\beta_D$ are set, the receiver adjusts the connectivity $\varphi$ which in turn changes the equilibrium filtering threshold $t$. Intuitively, filtering based on better information results in higher receiver utility hence the perfect algorithm maximizes the total receiver utility.

Taking into account the costs of implementing the different algorithms, the second and third items describe the platform's choice given these costs as illustrated by Figure 3. The perfect algorithm is preferred when the sum of both costs is sufficiently low. However, if the cost of implementing the distance is relatively high within the sum, the platform prefers to switch to the quality algorithm. This latter algorithm is chosen as long as the cost of implementing quality filtering is not too high, but that of distance filtering is high. If both costs are high, no algorithm is chosen.

Finally, note that the distance algorithm is never picked because it does not improve customer satisfaction, but comes at a cost. It is interesting to compare the magnitudes of the thresholds when $q = c$. In this case, the threshold for the sum of costs is $\overline{\gamma_{Q+D}} = \frac{\alpha N}{48}$ with the threshold

Figure 3: Algorithm choice of platform depending on implementation costs




separating the QA and PA choices is $\overline{\gamma_D} = \frac{4}{9}\overline{\gamma_{Q+D}}$, while the threshold that separates the QA choice from NA is $\overline{\gamma_Q} = \frac{5}{9}\overline{\gamma_{Q+D}}$. This implies that if we were to restrict the two costs per sender to be the same, i.e. $\gamma = \gamma_D = \gamma_Q$, the platform's choice is PA for a low cost, QA for intermediate, and NA for high.

The following corollary establishes that the results from Proposition 9 hold in the full equilibrium (that endogenizes the Sender's quality decision) when $r = c$:

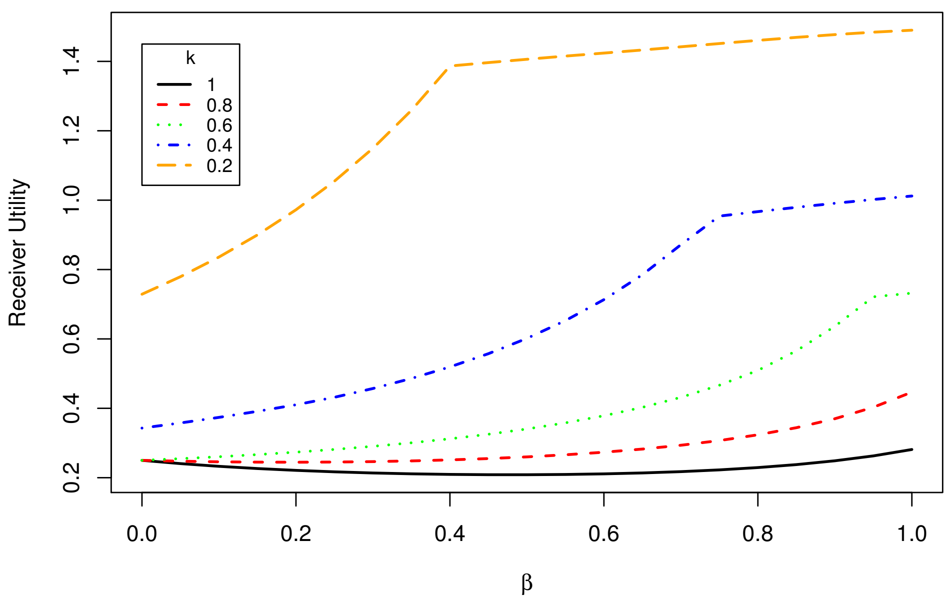
**Corollary 10.** *When $r = c$, in equilibrium with endogenous actions by senders and receivers, $\pi_R^{PA} \geq \pi_R^{QA} \geq \pi_R^{NA}$.*

The corollary shows that in equilibrium the perfect algorithm achieves the highest receiver utility, followed by the quality algorithm and then followed by not using an algorithm. As this is the same condition which is necessary and sufficient to prove Proposition 9, the results of the proposition also apply in the full equilibrium, although the values of the cost thresholds $\overline{\gamma_{Q+D}}$,$\overline{\gamma_Q}$ and $\overline{\gamma_D}$, will be different.

To provide further evidence about the equilibrium choice of algorithm by the platform, we turn to numerical analysis. The online appendix details the numerical approach and the limitations which allow us to find a solution for the cases of $r = c = 0$. As part of the analysis, we show that any scoring algorithm with parameters $(\beta_Q, \beta_D)$ can be represented by the one-dimensional ratio $\beta = \frac{\beta_D}{\beta_Q}$. Thus, $\beta = 1$ is the perfect algorithm while $\beta = 0$ is the quality algorithm.

Figure 4 illustrates the utility of the receiver for different values of $k$ and $\beta$. The figure shows that the results from Proposition 9 hold such that the receiver's utility is maximized when $\beta = 1$ and the platform uses the perfect algorithm. Interestingly, there are cases (when k is high) where the quality algorithm yields higher utility than an algorithm that uses partial weight ($\beta < 1$) for the distance $D$.

Figure 4: Receiver utility in full equilibrium when $r = c = 0$ for different values of $\beta$ and quality cost $k$




## 7.2 Equilibrium Effect on Filter Bubbles and Sender Payoffs

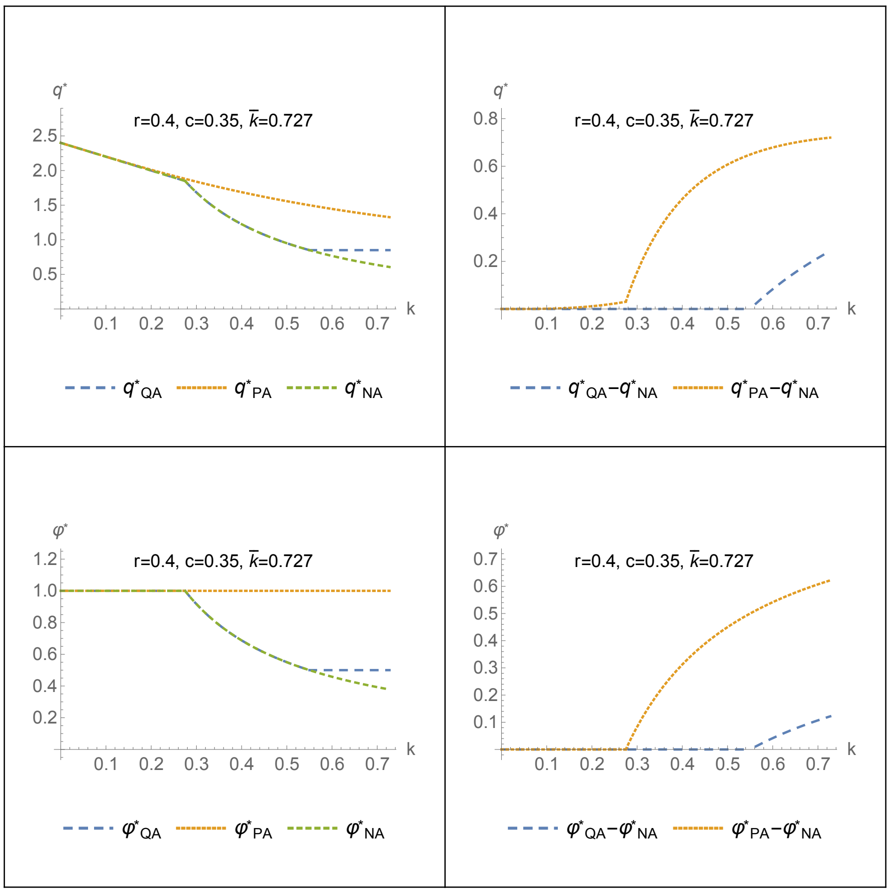
In this section we compare the equilibrium outcomes on connectivity, quality, payoffs and filtering, to illustrate the main results from our analysis in equilibrium.

Figure 5 illustrates the equilibrium values of $q^*$ and $\varphi^*$ as functions of $k$ when fixing $c = 0.35$ and $r = 0.4$.[22] Comparing the top and bottom left panels, we notice the correlation between the quality and connectivity, which is a result of the strategic complementarity between quality and

[22]These values were chosen because they allow a wide range of $k$ values where the equilibrium is not a corner solution, but the same general pattern holds for all parameter values.

connectivity. We also see that when $k$ is low, the filtering algorithms do not make a difference as senders elect to produce high quality content which in turn entices receivers to connect to all senders. For high values of $k$, the quality algorithm leads to somewhat higher quality and connectivity levels compared to the case of no algorithm or the distance algorithm, but the perfect algorithm generally achieves higher qualities and full connectivity. The right panels illustrate how increases in $k$ cause an increase in the difference of both quality and connectivity as compared to the no algorithm case.

Figure 5: Equilibrium Quality and Connectivity




Top: Equilibrium sender quality $q^*$; Bottom: receiver connectivity $\varphi^*$.

Left: function values; right: difference from the no algorithm case.

$k < \bar{k} \approx 0.727$, $c = 0.35$ and $r = 0.4$.

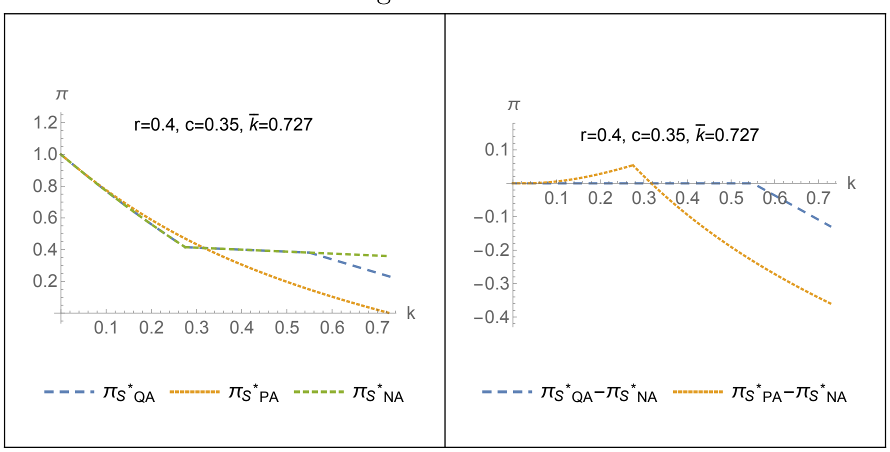
Focusing on filter bubbles, we can also compare the equilibrium average distance $\bar{d}$ of content receivers consume. For the NA and QA cases, the average distance is equal to the equilibrium connectivity $\varphi^*$. When comparing the NA and PA cases, we arrive at the following result:

**Corollary 11.** *There exists a* $\frac{r-c}{2} + \frac{1}{4} < \underline{k} < min(\overline{k}_{PA}, \overline{k}_{NA})$ *s.t. if* $0 < k < \underline{k}$ *then* $\overline{d}_{NA} > \overline{d}_{PA}$*. When* $k > \underline{k}$*,* $\overline{d}_{NA} < \overline{d}_{PA}$*.*

The corollary shows that the findings of corollary 5 hold in equilibrium, and that for low values of $k$, the perfect algorithm will increase polarization and create a filter bubble where receivers are exposed to less diverse content compared to without an algorithm.

Finally, in Figure 6 we illustrate the payoffs achieved by the senders in the conditions of Figure 5. We see that the PA profit declines continuously in the cost of quality $k$, while under the no algorithm case and quality algorithm the profit is initially lower than the PA profit, but as it declines slower with $k$, for high values of $k$ it is essentially higher. This is a result of the higher quality created by the perfect algorithm, which eventually costs too much and yields lower profit. In summary, these figures illustrate our main results well. Filtering algorithms force senders to create higher quality content especially in the case of the perfect algorithm. However, unless content creation costs are low, this is a burden for senders, lowering their profits and hindering their potential willingness to enter the market.

Figure 6: Sender Profit




Left: profit values; Right: difference from the no algorithm case.

$k < \overline{k} \approx 0.727$, $c = 0.35$ and $r = 0.4$.

## 8 Discussion and Conclusion

The implementation of curation algorithms by social platforms raises questions about their impact on content diversity experienced by users. If filter bubbles are indeed a result of using curation algorithms, platforms should carefully consider which algorithms best serve their users. As an example, strong curation may allow "fake news" to spread since they match the preconceived notion of some readers, but they never reach other readers who are able to flag them as fake.

Despite the important implications, arguments surrounding these algorithms tend to be simplistic and ignore many facets of this complicated ecosystem. The research about which specific details of algorithms have the most impact is limited. One contributor is the fact that platforms are tight lipped about the details of their curation. Media interviews with the platforms[23] commonly documented the following features as impacting the prominence of a content item being displayed to readers: (1) media richness - whether the item is an image, or has a video attached to it, (2) recency of item, (3) engagement of other users with this item, (4) engagement of a user with similar past items and, (5) the strength of connection between sender and receiver in terms of past engagement with past items. Our model advances this field of research by including creators of content (senders), consumers of content (receivers) and a curating platform. The platform can measure the quality of content, similarly to the richness of the media aspect, and the relationship between sender and receiver, as measured by their social distance. The analysis in the online appendix also considers the engagement of other users with each item. As a result, our analysis provides clear theoretical predictions about the impact of curation algorithms on network structure and user behavior. These predictions can form a basis for further empirical work that examines these issues.

The results show that product design choices behind curation algorithms have non-trivial implications. Algorithmic curation can alter the structure of the network as well as the quality and diversity of content on a social network. Curation may cause receivers to become less selective in choosing whom to follow, while senders in general will have to increase the quality of their content when the cost of producing quality is high. Furthermore, we find that algorithms such as

[23] E.g., `http://www.slate.com/articles/technology/cover_story/2017/03/twitter_s_timeline_algorithm_and_its_effect_on_us_explained.html` for Twitter, `https://techcrunch.com/2018/01/11/facebook-time-well-spent/` for Facebook, and `https://techcrunch.com/2018/06/01/how-instagram-feed-works/` for Instagram, Accessed July 14, 2019.

the perfect algorithm may increase polarization of matched content between receivers and senders and create filter bubbles, while other algorithms (such as the QA) may decrease said polarization. Consequently the popular arguments that algorithms by design increase polarization because they over-match content to user preferences seem to be inflated.

This non-monotonic impact of the different algorithms may explain why in some cases it is observed that introducing algorithms encourages higher quality items on networks and more pluralistic content consumed by different receivers, while in other cases the effects are the opposite. Taking these potential effects into account should help platforms in the process of making modifications to the user-experience they offer as part of their product design process.

Our analysis of the two-sided market required simplifying assumptions and is naturally not without limitations. These simplifications facilitated both parsimony of exposition and the ability to perform more specific analyses of different aspects of the resulting content distribution. Because the algorithms employed by the platforms are confidential and often complex, we did not model all aspects of the network that may impact the level of curation. Among these, further exploration of the possibility of a bias in an algorithm that incorrectly learns from observed data about user preferences is warranted. For example, Microsoft recently had to shutdown a Twitter Bot after it shifted to promoting hate speech,[24] while Facebook has had trouble with implementing algorithms that correctly distinguish hate speech from legitimate opinion.[25]

Another assumption we have made is that the market has homogeneous receiver and sender marginal costs. Consequently, some notable aspects of social networks, such as the empirically observed heterogeneous distribution of degree of connectivity among users, are not modeled by our setup. The extensions in the online appendix show that our results are robust to introducing different forms of heterogeneity to the model, while featuring outcomes which are more realistic (such as heterogeneity in degree of connectivity in the network). Having said that, we note that our results might not apply to specific cases which have potential for future work, such as social media behavior and followersip of celebrities. The online appendix provides an example of how our model can be extended to considering the popularity of items while filtering, and shows that such cases are similar to the analysis of the quality filtering algorithm.

[24] `https://www.techrepublic.com/article/why-microsofts-tay-ai-bot-went-wrong/`, accessed July 14, 2019.

[25] `https://www.propublica.org/article/facebook-hate-speech-censorship-internal-documents-algorithms`, accessed July 14, 2019.

From an incentive angle, we have abstracted away from the explicit goal of the platform beyond focusing on specific aspects of consumer utility. Moreover, there are other possibilities to describe the receiver consumption decision and the "like" process, including more sophisticated search or satisficing processes which are left for future work.

An interesting implication of our finding is that by making it easier for senders to create high quality content, platforms can alter the incentives of receivers to connect to senders as well as the resulting equilibrium diversity of content. The resulting higher quality may increase connectivity, but may potentially reduce sender payoffs and diversity of content consumed by receivers. The analysis in the online appendix shows that while curation algorithms may increase user surplus as a whole, this may come at the cost of increasing content polarization and creating a filter bubble. One way platforms can approach this problem is by supplying tools that allow producers easier creation and editing of content, while focusing their algorithms on encouraging overall higher quality content without trying to match consumer tastes.

## References


Susan Athey, Markus M Mobius, and Jenő Pál. The impact of aggregators on internet news consumption. Stanford University Graduate School of Business Research Paper, 2017.

Eytan Bakshy, Solomon Messing, and Lada A Adamic. Exposure to ideologically diverse news and opinion on facebook. *Science*, 348(6239):1130–1132, 2015.

Pablo Barberá. How social media reduces mass political polarization. evidence from germany, spain, and the us. Working Paper, 2014.

Levi Boxell, Matthew Gentzkow, and Jesse M Shapiro. Is the internet causing political polarization? evidence from demographics. Technical report, National Bureau of Economic Research, 2017.

Lesley Chiou and Catherine Tucker. Content aggregation by platforms: The case of the news media. *Journal of Economics & Management Strategy*, 26(4):782–805, 2017.

Chrysanthos Dellarocas, Zsolt Katona, and William Rand. Media, aggregators, and the link economy: Strategic hyperlink formation in content networks. *Management Science*, 59(10):2360–2379, 2013.

Seth Flaxman, Sharad Goel, and Justin M Rao. Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly*, 80(S1):298–320, 2016.

Matthew Gentzkow. Polarization in 2016. Technical report, Stanford University, 2016. URL `https://web.stanford.edu/~gentzkow/research/PolarizationIn2016.pdf`.

David Godes, Elie Ofek, and Miklos Sarvary. Content vs. advertising: The impact of competition on media firm strategy. *Marketing Science*, 28(1):20–35, 2009.

Kartik Hosanagar, Daniel Fleder, Dokyun Lee, and Andreas Buja. Will the global village fracture into tribes? recommender systems and their effects on consumer fragmentation. *Management Science*, 60(4):805–823, 2013.

Ganesh Iyer and Zsolt Katona. Competing for attention in social communication markets. *Management Science*, 62(8):2304–2320, 2015.

Michael Latzer, Katharina Hollnbuchner, Natascha Just, and Florian Saurwein. The economics of algorithmic selection on the internet. In *Handbook on the Economics of the Internet*, chapter 19, pages 395–425. Edward Elgar Publishing, 2016.

Jae Kook Lee, Jihyang Choi, Cheonsoo Kim, and Yonghwan Kim. Social media, network heterogeneity, and opinion polarization. *Journal of communication*, 64(4):702–722, 2014.

Greg Linden, Brent Smith, and Jeremy York. Amazon. com recommendations: Item-to-item collaborative filtering. *IEEE Internet computing*, 7(1):76–80, 2003.

Tien T Nguyen, Pik-Mai Hui, F Maxwell Harper, Loren Terveen, and Joseph A Konstan. Exploring the filter bubble: the effect of using recommender systems on content diversity. In *Proceedings of the 23rd international conference on World wide web*, pages 677–686. ACM, 2014.

Gal Oestreicher-Singer and Arun Sundararajan. Recommendation networks and the long tail of electronic commerce. *MIS Quarterly*, 36(1):65–84, 2012.

Eli Pariser. *The filter bubble: What the Internet is hiding from you.* Penguin UK, 2011.

Manuel Gomez Rodriguez, Krishna Gummadi, and Bernhard Schoelkopf. Quantifying information overload in social media and its impact on social contagions. In *Eighth International AAAI Conference on Weblogs and Social Media*, 2014.

Jason MT Roos, Carl F Mela, and Ron Shachar. The effect of links and excerpts on internet news consumption. Available at SSRN 2678938, 2015.

Upendra Shardanand and Pattie Maes. Social information filtering: algorithms for automating "word of mouth". In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 210–217. ACM Press/Addison-Wesley Publishing Co., 1995.

Jessica Su, Aneesh Sharma, and Sharad Goel. The effect of recommendations on network structure. In *Proceedings of the 25th International Conference on World Wide Web*, pages 1157–1167. International World Wide Web Conferences Steering Committee, 2016.

Marshall Van Alstyne and Erik Brynjolfsson. Global village or cyber-balkans? modeling and measuring the integration of electronic communities. *Management Science*, 51(6):851–868, 2005.

# Appendix

Preliminaries: We first derive a number of properties of the utility distribution we use throughout the proofs. Let $Q \sim U[0,1]$, $D \sim U[0,2\varphi]$ and $g(), G()$ be the pdf and cdf of $U = Q - D$. We have:

$$\mathbb{E}[U] = \frac{1}{2} - \varphi, \tag{25}$$

$$g(u) = \begin{cases} 1 + \frac{u}{2\varphi} & -2\varphi \leq u < \min(1-2\varphi, 0) \\ \frac{1}{2\varphi} & \min(1-2\varphi, 0) \leq u < 0 \\ 1 & 0 \leq u < \max(0, 1-2\varphi) \\ \frac{1-u}{2\varphi} & \max(0, 1-2\varphi) \leq u \leq 1 \end{cases} \tag{26}$$

$$G(u) = \begin{cases} \frac{(u+2\varphi)^2}{4\varphi} & -2\varphi \leq u < \min(1-2\varphi, 0) \\ 1 + \frac{2u-1}{4\varphi} & \min(1-2\varphi, 0) \leq u < 0 \\ u + \varphi & 0 \leq u < \max(0, 1-2\varphi) \\ 1 - \frac{(1-u)^2}{4\varphi} & \max(0, 1-2\varphi) \leq u \leq 1 \end{cases} \tag{27}$$

$$\Pr(U>t)\mathbb{E}(U|U>t)=\begin{cases}\frac{1}{6}\left(-\frac{t^3}{\varphi}-3t^2+4\varphi^2-6\varphi+3\right) & -2\varphi\le t<\min(1-2\varphi,0)\\ \frac{1-3t^2}{12\varphi} & \min(1-2\varphi,0)\le t<0\\ \frac{1}{6}\left(-3t^2+4\varphi^2-6\varphi+3\right) & 0\le t<\max(0,1-2\varphi)\\ \frac{(t-1)^2(2t+1)}{12\varphi} & \max(0,1-2\varphi)\le t\le 1\end{cases} \tag{28}$$

PROOF OF PROPOSITION 1: The payoff function $\varphi\alpha N\left(q-c+\frac{1-2\varphi}{2}\right)$ is quadratic in $\varphi$ and reaches its maximum at $\varphi^*=\frac{q-c}{2}+\frac{1}{4}$. The latter is clearly increasing in $q$ and decreasing in $c$. □

PROOF OF PROPOSITION 2: The expected payoff $\mathbb{E}(\pi_R)=J\cdot\Pr(U_j>t)\mathbb{E}(U_j-c|\varphi,U_j>t)$ can be written as

$$\mathbb{E}(\pi_R)=J\int_t^{q+1}(x-c)g(x-q)dx \tag{29}$$

By increasing $t$ the integration interval decreases, but the integrand does not change. Hence, the integral is maximized when the integrand is all non-negative, which occurs at $t^*=c$.

To calculate the optimal $\varphi$, we first need to calculate $J\cdot\Pr(U_j>c)\mathbb{E}(U_j-c|\varphi,U_j>c)$ as a function of $\varphi$. To use the results obtained in (28), we transform

$$\begin{aligned}E(\varphi):=\Pr(U_j>c)\mathbb{E}(U_j-c|\varphi,U_j>c)=\Pr(U_j>c)\mathbb{E}(U_j|\varphi,U_j>c)-c\Pr(U_j>c)=\\ \Pr(U>c-q)\mathbb{E}(U|U>c-q)-(c-q)(1-G(c-q))\end{aligned} \tag{30}$$

since the $U_j$ used here is shifted by $q$ to the right compared to the $U$ in (28). The resulting function is increasing for every $\varphi<\frac{1-\tilde{t}}{2}$ when $\tilde{t}=c-q$. Hence, the receiver sets $\varphi^*_{PA}=\frac{1+q-c}{2}$.

To prove the second part of the proposition, the probability of content being filtered by the algorithm is $Pr(U_j\le c|\varphi^*)=Pr(U\le\tilde{t}|\varphi^*)=G(\tilde{t}|\varphi^*)$ when $\tilde{t}=c-q$. We then get

$$\frac{dG(\tilde{t},\varphi^*)}{d\tilde{t}}=g(\tilde{t}|\varphi^*)+\frac{\partial G(\tilde{t},\varphi^*)}{\partial\varphi^*}\frac{\partial\varphi^*}{\partial\tilde{t}} \tag{31}$$

This expression is always positive for $-1\le\tilde{t}\le 1$, and since $\tilde{t}$ is increasing in $c$ and decreasing in $q$, this proves the result. □

PROOF OF PROPOSITION 3: In case of the QA, the expected payoff can be written as

$$\begin{aligned}\mathbb{E}(\pi_R)=J\cdot\Pr(Q_j>t)\mathbb{E}(Q_j-D_j-c|\varphi,Q_j>t)=J\cdot\Pr(Q_j>t)(\mathbb{E}(Q_j-\mathbb{E}(D_j+c)|\varphi,Q_j>t))=\\ J\cdot\Pr(Q_j>t)\left(\mathbb{E}\left(Q_j-\varphi-c|\varphi,Q_j>t\right)\right)=J\int_t^{q+1}(x-\varphi-c)\cdot 1dx\end{aligned} \tag{32}$$

As in the case of the perfect algorithm, the integral is maximized when the integrand is non-negative, that is at $t^* = \varphi + c$. To obtain the optimal $\varphi$, we calculate the integral

$$\int_{\varphi+c}^{q+1} (x - \varphi - c) \cdot 1dx = \frac{1}{2}(c - q + \varphi - 1)^2 \tag{33}$$

Plugging the integral into the receiver's utility function, we find that it has a local maximum at $\varphi^* = \frac{1}{3}(1 - c + q)$, and a local minimum at $1 - c + q$. We require that $c + \varphi^* < q + 1$ to make sure the algorithm does not filter all content, which means the local maximum is also a global maximum. In addition, since the algorithm can filter for qualities between $q$ and $q + 1$, when $c - q < -\frac{1}{2}$, the algorithm does not filter at all, and the receiver experiences the no algorithm case.

Comparing this value to the value found for the PA algorithm in proposition 2, we find that $\varphi^*_{PA} > \varphi^*_{QA}$ when $1 > c - q > -\frac{1}{2}$. The other items are straightforward comparisons to the values found for the other algorithms. $\square$

Proof of Proposition 4: Similarly to the previous case, the expected payoff can be written as

$$\mathbb{E}(\pi_R) = J \Pr(D_j < -t)\mathbb{E}(U_j - c|\varphi, D_j < -t) \tag{34}$$

$$= J \int_0^{-t} \left(q + \frac{1}{2} - x - c\right) \cdot g_D(x)dx \tag{35}$$

When $g_D()$ is the PDF of $D_j$. The expected payoff is maximized at $t = c - q - \frac{1}{2}$. At this threshold the expected payoff is

$$\mathbb{E}(\pi_R) = \begin{cases} \alpha N \varphi \left(q - c - \varphi + \frac{1}{2}\right) & c - q \leq \frac{1}{2} - 2\varphi \\ \frac{\alpha N}{16}(2q - 2c + 1)^2 & \frac{1}{2} - 2\varphi < c - q \leq \frac{1}{2} \end{cases} \tag{36}$$

The function is maximized at $\varphi^* = \frac{2(q-c)+1}{4}$ for $-\frac{3}{2} < c - q < \frac{1}{2}$. When $c - q \leq -\frac{3}{2}$, $\varphi^* = 1$ and when $c - q \geq \frac{1}{2}$, $\varphi^* = 0$.

Using this result, the amount of content filtered by the algorithm amounts to zero, because the receiver has no benefit of adding senders with distance that will be filtered later. $\square$

Proof of Corollary 5: When there is no curation, the average distance from which a receiver reads a message is simply the expected distance which is $\varphi$ with $\varphi^*_{NA} = \frac{q-c}{2} + \frac{1}{4}$, that is

$$\overline{d}_{NA} = \begin{cases} 1 & c - q \leq -\frac{3}{2} \\ \frac{q-c}{2} + \frac{1}{4} & -\frac{3}{2} < c - q \leq \frac{1}{2} \\ 0 & c - q > \frac{1}{2} \end{cases} \tag{37}$$

In case of the perfect algorithm messages with utility below $c$ are filtered out. Hence, the average distance encountered by a customer is $\mathbb{E}_D[D|U_j > c]$ with $\varphi = \varphi^*_{PA}$. Calculation results in:

$$\overline{d}_{PA} = \begin{cases} 1 & c - q \leq -2 \\ \frac{1}{3}\left(\frac{1}{c-q} - \frac{5}{c-q+4} + q - c + 4\right) & -2 < c - q \leq -1 \\ \frac{1+3(q-c)+3(q-c)^2}{3(1+2(q-c))} & -1 < c - q \leq 0 \\ \frac{q-c+1}{3} & 0 < c - q \leq 1 \\ 0 & c - q > 1 \end{cases} \tag{38}$$

Under the quality algorithm, the distance is independent from the filtering, thus the expected distance is $\varphi$ with $\varphi^*_{QA} = \frac{q-c+1}{3}$ when $1 > c - q > -1/2$ and $\varphi^*_{NA}$ otherwise, that is:

$$\overline{d}_{QA} = \begin{cases} 1 & c - q \leq -\frac{3}{2} \\ \frac{q-c}{2} + \frac{1}{4} & -\frac{3}{2} < c - q \leq -\frac{1}{2} \\ \frac{q-c+1}{3} & -\frac{1}{2} < c - q \leq 1 \\ 0 & c - q > 1 \end{cases} \tag{39}$$

In case of the distance algorithm, the expected distance is the same as under the no algorithm case.

Comparison of these values shows that $\overline{d}_{QA} \geq \overline{d}_{NA} = \overline{d}_{DA}$ for for every $1 \geq c - q \geq -2$ with strict inequality when $1 > c - q > -\frac{1}{2}$ in the applicable range of values. In addition, $\overline{d}_{PA} > \overline{d}_{NA}$ if and only if $A < c - q < 1$ where $A \approx -1.263$. □

Proof of Proposition 6: Before examining comparative statics, we provide details on how we obtained the solution in equation (17). A sketch is already presented in the main text. To ensure that the first order condition results in a local maximum, we need to check that the second order condition associated with the first order condition presented in (15) is negative. This is satisfied when $q_j \geq r$ and $r + 2\varphi \geq q_j \geq r + 2\varphi - 1$. We also have to make sure that the solution of (15) yields not only a local, but also a global maximum. The marginal cost crosses the marginal revenue in at most two points, one being identified by the FOC. As the SOC reveals, the marginal cost exceeds the marginal revenue after this point, so the payoff declines above this point. However, it is possible that the marginal cost function crosses the marginal revenue once more, from above for small values of $q_j$, creating a potential for $q_j = 0$ corner solution. To avoid this, we assume that $k < \overline{k}_{NA}$, where in equilibrium the profit of the sender is positive when $k < \overline{k}_{NA}$. Solving for the positive profit results in $\overline{k}_{NA} = \min(\frac{8c}{2(r+c)-1} - 2, 1)$ ensuring that we obtain a non-trivial solution in (17) as long as $r < 1$.

For the comparative statics results we differentiate the $\varphi^*$ function given in (17) with respect to $r, c$ and $k$ and obtain the stated signs. To obtain the results on equilibrium quality, we differentiate $q^* = r + 2\varphi^*(1 - k)$. □

Proof of Proposition 7: Before proving comparative statics, we detail how we obtained $q^*(\varphi)$. From the point of view of receiver $j$, since the marginal cost of quality is fixed at $k$, her best response will fall in the range where the marginal revenue $\frac{g(r-q_j)}{1-G(c-q_{-j})}$ is decreasing in $q_j$. Examination of $g(r - q_j)$ shows this happens only when $q_j \geq r$ and $r + 2\varphi \geq q_j \geq r + 2\varphi - 1$. Adding the condition $k < \overline{k}_{PA}$ with

$$\overline{k}_{PA} = \min\left(2\left(\sqrt{\frac{r(r+4)\left((c-r)^2-8\right)^2}{\left(c^2-r(r+4)\right)^2\left(c(c+8)-r(r+4)+16\right)^2}} - \frac{(c-r)(c(r+2)-(r-2)(r+4))}{\left(c^2-r(r+4)\right)\left(c(c+8)-r(r+4)+16\right)}\right), 1\right) \tag{40}$$

similarly to the previous section is sufficient to ensure a global maximum. Solving the FOC $g(r - q) - k(1 - G(c - q)) = 0$ along with the solution $\varphi^* = \frac{1+q-c}{2}$ from Proposition 2 and the above conditions, shows the only solution is the one we report. The comparative statics in part two follows in a straightforward manner from $q^*$.

For parts three to five, the equilibrium qualities are:

$$q^*_{NA} = \begin{cases} r + 2(1-k) & 0 < k \leq \frac{r-c}{2} + \frac{1}{4} \\ \frac{(2c-1)(k-1)+2r}{2k} & \frac{r-c}{2} + \frac{1}{4} \leq k < \overline{k}_{NA} \end{cases} \tag{41}$$

$$q^*_{PA} = \begin{cases} r + 2(1-k) & 0 < k \leq \frac{r-c}{2} \\ \frac{k(2+c)+1-\sqrt{2k(c+2k-r)+1}}{k} & \frac{r-c}{2} \leq k < \overline{k}_{PA} \end{cases} \tag{42}$$

The senders' equilibrium profits are:

$$\pi_S^{NA} = \begin{cases} k^2 - k(r+2) + 1 & 0 < k \leq \frac{r-c}{2} + \frac{1}{4} \\ \frac{1}{4}(-2c(k-2) - 2(k+2)r + k + 2) & \frac{r-c}{2} + \frac{1}{4} \leq k < \overline{k}_{NA} \end{cases} \tag{43}$$

$$\pi_S^{PA} = \begin{cases} k^2 - k(r+2) + 1 & 0 < k \leq \frac{r-c}{2} \\ \frac{k^2(c(c+4)-r(r+4)+8) - 2k(r-2c-2) - 2(1+k(c+2))\sqrt{2k(c+2k-r)+1}+2}{2\left(\sqrt{2k(c+2k-r)+1} - ck + kr - 1\right)} & \frac{r-c}{2} \leq k < \overline{k}_{PA} \end{cases} \tag{44}$$

For $0 < k \leq \frac{r-c}{2}$ and $\frac{r-c}{2} < k \leq \frac{r-c}{2} + \frac{1}{4}$ the results follow from a direct comparison. When $\frac{r-c}{2} + \frac{1}{4} < k \leq \min(\overline{k}_{NA}, \overline{k}_{PA})$, we find that $\pi_S^{NA} > \pi_S^{PA}$ when:

$$k > \tilde{k}(r,c) = \tag{45}$$

$$-2\frac{16c^3 + c^2(34-48r) + \sqrt{-\left((c-r)^2 - 8\right)^2 \left(4c^2 - 8c(r-1) + 4(r-2)r - 33\right)}}{(13c - 13r + 34)(5c - 5r - 2)} \tag{46}$$

$$-2\frac{4c(r(12r-17)-8) + 2r(r(17-8r)+16) - 36}{(13c-13r+34)(5c-5r-2)} \tag{47}$$

Solving for $r$ and $c$ in the condition $\frac{r-c}{2} + \frac{1}{4} < \tilde{k}(r,c) \leq \min(\overline{k}_{NA}, \overline{k}_{PA})$, yields that for $r < \frac{1}{17}\left(53 - 8\sqrt{33}\right)$ there exists a range of values of $0 < c < r$ s.t. the profits under the NA case are higher than under the PA case. For example, when $r < 0.21$, this condition holds for any $c < r$.

□

Proof of Proposition 8: The probability of receiving a like, $P(Q_j)$ depends on the values of $r$ and $t$ and can be broken up into three cases:

When $r \geq t$:

$$P(Q_j) = \begin{cases} 0 & q_j \leq r - 1 \\ \frac{(q_j - r + 1)^2}{4\varphi} & r - 1 < q_j \leq \min(r, r + 2\varphi - 1) \\ \frac{2q_j - 2r + 1}{4\varphi} & r < q_j \leq r + 2\varphi - 1 \\ q_j - r - \varphi + 1 & r + 2\varphi - 1 < q_j \leq r \\ 1 - \frac{(q_j - r - 2\varphi)^2}{4\varphi} & \max(r, r + 2\varphi - 1) < q_j \leq r + 2\varphi \\ 1 & q_j > r + 2\varphi \end{cases} \tag{48}$$

When $r + 2\varphi \geq t > r$:

$$P(Q_j) = \begin{cases} 0 & q_j \leq t-1 \\ \frac{(q_j-t+1)(q_j-2r+t+1)}{4\varphi} & t-1 < q_j \leq \min(r+2\varphi-1, t) \\ \frac{2q_j-2r+1}{4\varphi} & t < q_j \leq r+2\varphi-1 \\ q_j - \frac{(r-t)^2}{4\varphi} - r - \varphi + 1 & r+2\varphi-1 < q_j \leq t \\ 1 - \frac{(q_j-r-2\varphi)^2}{4\varphi} & \max(t, r+2\varphi-1) < q_j \leq r+2\varphi \\ 1 & q_j > r+2\varphi \end{cases} \tag{49}$$

The case $t > r + 2\varphi$ is not possible since $t^* = c + \varphi$ and $r \geq c$, and thus we omit it.

When using each one of these functions in the profit function, we notice (similar to the NA case) that the marginal revenue may cross the marginal cost up to two times, in which case one point will be a local maximum, and the other a local minimum. In addition, since both functions have linear parts, it is possible to have a discontinuity in the first order from a positive marginal profit to a negative one, depending on the value of $k$.

The second order condition of Equation (22) is only negative on the concave parts of the functions, which always happens in the range $q_j \geq t$. In addition, on the linear parts the maximum is achieved at the right edge of the range, which is always the beginning of a concave part of the function. The best response of the receiver is therefore in the range $q_j \geq t$, which means that in equilibrium $P(Q_{-j} > t) = 1$, and effectively the content of senders is not filtered because of low quality.

Using this simplification, we find that when $r \geq t$, the solution to the first order condition coincides with the NA result, but when $r + 2\varphi > t \geq r$, the solution has a discontinuity, yielding the one appearing in (23):

$$q(\varphi) = \begin{cases} r + 2\varphi(1-k) & 0 < k \leq \frac{1}{2} + r - c \\ c + \varphi & \frac{1}{2} + r - c < k \leq \overline{k}_{QA} \end{cases} \tag{50}$$

with $\overline{k}_{QA} = \min(1, \frac{-4c^2+c(8r+4)-4r(r+1)+7}{8c+4})$.

Using this result and the result from section 4.3 that $\varphi^*_{QA} = \frac{q-c+1}{3}$ when $c - q > -\frac{1}{2}$ and $\frac{2(q-c)+1}{4}$ when $c - q \leq -\frac{1}{2}$ we find the feasible equilibrium values $q^*_{QA}$ and $\varphi^*_{QA}$ which adhere to the constraint that the profit is positive for some positive value of $k$:

$$\varphi_{QA}^* = \begin{cases} 1 & 0 < k \leq \frac{2(r-c)+1}{4} \\ \frac{1+2(r-c)}{4k} & \frac{2(r-c)+1}{4} < k \leq \frac{1}{2} + r - c \\ \frac{1}{2} & \frac{1}{2} + r - c < k \leq \overline{k}_{QA} \end{cases} \tag{51}$$

The comparative statics follow directly from the equilibrium values of $q^*$. $\square$

Proof of Proposition 9: Setting $\beta_Q = 0$ implements a form of the DA algorithm (with $\beta_D$ allowed to vary) which we have shown is equivalent to the NA case. Since implementing algorithms has a cost, we can assume there will be no case with $\beta_D \neq 0$ and $\beta_Q = 0$. Hence, when we assume that $\beta_Q \neq 0$, and allow $\beta_Q$ and $\beta_D$ to vary freely, the utility of the receiver is $J \cdot E\varphi$ when:

$$E\varphi = \begin{cases} -c + q - \varphi + \frac{1}{2} & t \leq \beta_Q q - 2\beta_D \varphi \\ \frac{-8\beta_D^4\varphi^3 + 4\beta_D^3\varphi^2(3\beta_Q c + 4\beta_Q\varphi - 3t) + 6\beta_D^2\varphi(\beta_Q q + \beta_Q - t)(\beta_Q(-2c+q-2\varphi+1)+t)}{12\beta_D^2\beta_Q^2\varphi} & \\ + \frac{\beta_D(t-\beta_Q q)^2(3\beta_Q c - 2\beta_Q q - t) + \beta_Q(\beta_Q q - t)^3}{12\beta_D^2\beta_Q^2\varphi} & \beta_Q q - 2\beta_D\varphi < t \leq \beta_Q q \\ \frac{-4\beta_D^2\varphi^2 + \beta_Q\left(6\varphi(\beta_D c + t) + 8\beta_D\varphi^2 + 6ct\right) - 6\beta_D t\varphi + 3\beta_Q^2(q+1)(-2c+q-2\varphi+1) - 3t^2}{6\beta_Q^2} & \beta_Q q < t \leq \beta_Q(q+1) \\ 0 & \beta_Q(q+1) < t \end{cases} \tag{52}$$

The platform picks $t^*$ to maximize $E\varphi$. The values of $t^*$ where both the first order and second order conditions hold are:

$$t^* = \begin{cases} \beta_Q(c+\varphi) - \beta_D\varphi & \varphi \geq \frac{\beta_Q(q-c)}{\beta_Q - \beta_D} \\ \frac{\beta_Q(2\beta_D c + (\beta_Q - \beta_D)q) + 2\beta_D\varphi(\beta_Q - \beta_D)}{\beta_D + \beta_Q} & \frac{\beta_Q(q-c)}{\beta_Q - \beta_D} > \varphi > \frac{q-c}{2} \end{cases} \tag{53}$$

Maximizing the receiver's utility $J \cdot E\varphi$ results in:

$$\varphi^* = \begin{cases} \frac{\beta_D(1+2q-2c)+\beta_Q}{4\beta_D} & \beta_D \geq \beta_Q \\ \frac{\beta_Q(c-q-1)}{\beta_D - 3\beta_Q} & \beta_D < \beta_Q \end{cases} \tag{54}$$

Plugging into the receiver's profit function yields $\alpha N \frac{3\mathrm{bD}^2(-2c+2q+1)^2 + 2\mathrm{bDbQ} - \mathrm{bQ}^2}{48\mathrm{bD}^2}$ when $\beta_D \geq \beta_Q$ and $\alpha N \frac{\mathrm{bQ}(\mathrm{bD} - 2\mathrm{bQ})(c-q-1)^3}{3(\mathrm{bD} - 3\mathrm{bQ})^2}$ otherwise. Both functions are maximized at $\beta_Q = \beta_D$, and specifically with $\beta_Q = \beta_D = 1$.

Using the notation of $t = c - q$, the maximum value is:

$$\pi^*_{R,PA}(t) = \pi^*_R(\beta_D = \beta_Q) = \begin{cases} \alpha N \frac{2+6t^2+t^3}{12} & -\frac{3}{2} < t \leq -1 \\ \alpha N \frac{1-3t+3t^2}{12} & -1 < t \leq 0 \\ \alpha N \frac{(1-t)^3}{12} & 0 \leq t \leq 1 \end{cases} \tag{55}$$

If implementing a distance measurement is too costly (high $\gamma_Q$), we can perform the same analysis to find the optimal algorithm that uses only quality measurement $\beta_Q \neq 0$ with $\beta_D = 0$.

The resulting receiver profit does not depend on $\beta_Q$:

$$\pi^*_{R,QA}(t) = \pi^*_R(\beta_D = 0, \beta_Q \neq 0) = \begin{cases} \alpha N \frac{(1-2t)^2}{16} & -\frac{3}{2} < t \leq -\frac{1}{2} \\ \alpha N \frac{2(1-t)^3}{27} & -\frac{1}{2} < t \leq 1 \end{cases} \tag{56}$$

hence it is also maximized at $\beta_Q = 1$.

For completeness, the value of the receiver's utility without an algorithm is:

$$\pi^*_{R,NA}(t) = \begin{cases} \alpha N \frac{(1-2t)^2}{16} & -\frac{3}{2} < t \leq \frac{1}{2} \\ 0 & \frac{1}{2} < t \leq 1 \end{cases} \tag{57}$$

It is easy to verify that $\pi^*_{R,PA}(t) \geq \pi^*_{R,QA}(t) \geq \pi^*_{R,NA}(t)$ for every $-3/2 < t < 1$. Hence, when $\gamma_Q + \gamma_D$ is high, the platform will not apply any curation. When $\gamma_Q + \gamma_D$ is low, the platform will apply the perfect algorithm. For intermediate values of $\gamma_D + \gamma_Q$ when:

$$\gamma_Q < \pi^*_{R,QA}(t) - \pi^*_{R,NA}(t) = \begin{cases} \alpha N \frac{(5-8t)(2t+1)^2}{432} & -\frac{1}{2} < t \leq \frac{1}{2} \\ \alpha N \frac{2(1-t)^3}{27} & \frac{1}{2} < t < 1 \end{cases} \tag{58}$$

$$\gamma_D > \pi^*_{R,PA}(t) - \pi^*_{R,PA}(t) = \alpha N \frac{(1-t)^3 + 9t^3}{108} \tag{59}$$

the platform will implement the quality algorithm. □

Proof of Corollary 10: We first verify that when $r = c$ and $k \leq \min(\bar{k}_{PA}, \bar{k}_{QA}, 1)$ in the full equilibrium $\pi_R^{PA} \geq \pi_R^{QA} \geq \pi_R^{NA}$. We plug-in the equilibrium levels of connectivity and qualities found in the previous proofs to derive the receiver utilities:

$$\pi_R^{NA} = \begin{cases} \frac{3}{2} - 2k & k \leq 1/4 \\ \frac{1}{16k^2} & k > 1/4 \end{cases} \tag{60}$$

$$\pi_R^{QA} = \begin{cases} \frac{1}{4} & k \geq 1/2 \\ \frac{3}{2} - 2k & k < 1/4 \\ \frac{1}{16k^2} & 1/4 < k \leq 1/2 \end{cases} \tag{61}$$

$$\pi_R^{PA} = \begin{cases} \frac{31k^2+(6+15k)(1-\sqrt{4k^2+1})}{12k^2} & k \geq 2/3 \\ \frac{9k^3+(2-4k^2)\sqrt{4k^2+1}-2}{6k^3} & k < 2/3 \end{cases} \tag{62}$$

A simple comparison shows that $\pi_R^{PA} \geq \pi_R^{QA} \geq \pi_R^{NA}$. This means that the conclusion of Proposition 9 also hold in our case with different cost boundaries on $\gamma_Q$ and $\gamma_D$. $\square$

Proof of Corollary 11: As mentioned in the text, $\overline{d}_{NA} = \varphi^*_{NA}$ and $\overline{d}_{QA} = \varphi^*_{QA}$. For the perfect algorithm case:

$$\overline{d}_{PA} = \begin{cases} \frac{1}{3}\left(\frac{1}{c+2k-r-2} - \frac{5}{c+2k-r+2} - c - 2k + r + 6\right) & 0 < k \leq \frac{r-c}{2} \\ \frac{1}{3}\left(\frac{1-\sqrt{2ck+4k^2-2kr+1}}{k} - \frac{2k\left(-\sqrt{2ck+4k^2-2kr+1}+3k+1\right)}{\sqrt{2ck+4k^2-2kr+1}-ck+kr-1} + 6\right) & \frac{r-c}{2} < k < \overline{k}_{PA} \end{cases} \tag{63}$$

The result follows from direct comparison of the values when $r \leq 1$. $\square$