Homophily in Social Media and News Polarization

Luis Abreu
Toulouse School of Economics, University of Toulouse Capitole

Doh-Shin Jeon
Toulouse School of Economics, University of Toulouse Capitole, CEPR

* The Networks, Electronic Commerce, and Telecommunications (“NET”) Institute, http://www.NETinst.org, is a non-profit institution devoted to research on network industries, electronic commerce, telecommunications, the Internet, “virtual networks” comprised of computers that share the same technical standard or operating system, and on network issues in general.
Homophily in Social Media and News Polarization∗

Luis Abreu† Doh-Shin Jeon‡

September 30, 2019

Abstract

We study how media bias is affected by the structure of social networks on social media. We consider an ad-financed media firm which chooses the ideological location of its news and targets consumers who can share the news with their followers on an online social media. After studying how a targeted consumer’s incentive to share the news is shaped by the network structure of her followers, we study the firm’s strategy to maximize the breadth of news sharing and find that when the mean (respectively, the variance) of followers’ ideological locations is a convex (respectively, concave) function of a direct consumer’s location, the media firm is likely to produce polarized news. The analysis of the case in which consumers are uniformly distributed reveals that news polarization is more likely to occur as the degree of homophily increases. We also find that media competition makes polarization more likely.

Keywords: Media Bias, Online Social Networks, Homophily, Sharing, Polarization

∗We thank Heski Bar-Issac, Francis Bloch, Giacomo Calzolari, Alexandre de Cornière, Oliver Falck, Renato Gomes, Grace Gu, Emeric Henry, Leslie Marx, Alessandro Pavan, Andrew Rhodes, Marcella Scrimitore, Bruno Strulovici, Frek Van Gils, Jidong Zhou and those who attended our presentation at the 46th EARIE, Hitotsubashi University, Osaka University, TSE Seminar, the 17th CSIO/TSE Conference on Industrial Organization, European University Institute, the 3rd Doctoral Workshop on the Economics of Digitization, the 17th IIOC, the 4th Economics of Media Bias Workshop, the 12th Digital Economics Conference and the 16th Media Economics Workshop. Abreu acknowledges financial support from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement No 670494). Authors also gratefully acknowledge financial support from the NET Institute (www.NETinst.org).
†Toulouse School of Economics, University of Toulouse Capitole, Toulouse, France, luis.martins-abreu@ut-capitole.fr
‡Toulouse School of Economics, University of Toulouse Capitole, CEPR, dohshin.jeon@gmail.com
“What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention...”
(Simon, 1971)

1 Introduction

Social media has become a very important source for news consumption. According to Matsa and Shearer (2018), as of August 2018, two-thirds of Americans report that they get at least some of their news on social media – with two-in-ten doing so often. Among different social media, Facebook remains the most relevant as 45% of Americans get news on Facebook.

This success of social media, however, is viewed under suspicion as some fear that the tendency that users of social networks end up mainly consuming like-minded news would prevent them from digesting diverse viewpoints about important issues. Although people express such concerns by pointing out problems such as filter bubbles (Pariser, 2011) or echo chambers (Sunstein, 2017), one might think that homophily or social network structure on online social media is a more fundamental source of problem. Indeed, by analyzing data from U.S. Facebook users, Bakshy, Messing and Adamic (2015) found that homophily is the most important factor limiting their exposure to attitude-challenging content.\footnote{Similarly, Halberstam and Knight (2016) analyzed information from 2.2 million Twitter users the day before the 2012 U.S. general elections and found that, due to homophily, people are disproportionately exposed to tweets from like-minded others.}

Even if there is a consensus that homophily on social media leads to more consumption of like-minded news, there exists mixed evidence about whether social media leads to political polarization; Barberá (2015) and Boxell, Gentzkow and Shapiro (2017) find evidence inconsistent with polarization while Allcott et al. (forthcoming) and Yanagizawa-Drott, Petrova and Enikolopov (2019) find evidence consistent with polarization. For instance, Barberá (2015) argues that social media reduces political extremism by facilitating exposure to messages from those with weak ties, who are more politically heterogeneous than citizens’ immediate personal networks and provides empirical support.

This paper addresses the following questions. How does the structure of social networks on online media affect news sharing and thereby shape the incentive for a media firm to choose the ideological bias of its news? Under which conditions is the media firm incentivized...
to provide polarized news? To answer these questions, we consider a simple three-tiered hierarchy; a media firm, direct consumers and followers on social media. The media firm sends its news to (targeted) direct consumers, who in turn decide to share or not the news with their followers. We show that, the incentive for a direct consumer to share a given news depends crucially on the ideological location of the news as well as the mean and the variance of the locations of her followers, which in turn depend on her own ideological position. In other words, to understand news polarization it is crucial to understand how extremists’ followers are different from those of moderates in terms of the mean and the variance of their locations. Indeed, Kümpel, Karnowski and Keyling (2015), who review 109 papers on news sharing in social media published by communication or computer scientists, call for theory emphasizing the role of social networks of followers on the decision to share news. As the main result, we find that when the mean (respectively, the variance) of followers’ ideological locations is a convex (respectively, concave) function of a direct consumer’s location, the media firm is likely to produce polarized news. In contrast, when the mean is concave or the variance is convex, it is likely to produce unbiased news.

Our questions are partly motivated by some anecdotal evidence from the last US presidential election. Over a hundred of websites with made-up content were created by teenagers from a small town in Macedonia, seeking for advertising revenues propelled by the large sharing of its news on Facebook (Silverman and Alexander, 2016). As Craig Silverman, one of the reporters that revealed the story, explained in an interview to NPR, the teenagers were using Facebook to drive the traffic to the website where they had ads from Google. Their sites produced misleading partisan content and got more engagement than op-eds and commentary pieces from major media (NPR, 2016). Even though our model does not make distinction between fake news and true news, our results help to understand when polarized news is more likely to be shared than less polarized news.

We build a baseline model where consumers are distributed over an interval $[0, 1]$ of ideological space and a media firm chooses the ideological location of its news in order to maximize its advertising revenue, which is proportionate to the measure of readers. In addition, the media firm chooses an interval of given length to target direct consumers. We assume that the utility from reading the news has a quadratic loss. We adopt an altruistic motive for news sharing (Boyd, Golder and Lotan, 2010; Holton et al., 2014; Small, 2011) and assume that a targeted consumer shares the news if it is relevant enough to her followers,
i.e., if the average benefit of her followers from reading the news is larger than the attention tax she imposes on them by sharing the news. The attention tax captures opportunity cost of attention as sharing a news with a friends induces the latter to pay attention to the news. Using Simon (1971)’s expression, we assume that sharing news “consumes the attention of its recipients”. In a world of information overload, the attention tax can be high. Importantly, the average benefit of followers can be decomposed into a constant utility from reading the news, minus the variance of the followers’ locations, minus a square of the difference between the news location and the mean of the followers’ locations.

In the beginning we consider no targeting in the sense that the target interval is equal to $[0,1]$. For simplicity, we assume that the constant utility from reading the news is large enough such that all consumers who have access to the news read it. Under the assumption, the profit maximization is equivalent to the maximization of the measure of the direct consumers who share the news, which we define as the breadth of news sharing. If the attention tax is low enough, by locating the news at the middle, the media firm can induce all direct consumers to share the news. Hence, we focus on the case in which the attention tax is not small. As a main result, we find that news polarization is likely to occur if the mean of the followers’ locations is a convex function of a direct consumer’s location and/or the variance is a concave function. For instance, suppose that the mean is increasing and convex, which implies that the mean followers of two extremists are more closely located than the mean followers of two moderates. This can induce news polarization as the benefit from news sharing decreases slowly when it is located on an extreme point than when it is located around the mean. The polarization does not mean choosing news location equal to 0 or 1 as the polarization stops once it is extreme enough to induce the most extreme consumer to be indifferent between sharing the news or not, which we call the limit polarization strategy.

After studying news polarization for a given distribution of the mean and the variance of followers’ locations, we study how the degree of homophily affects the distribution. We propose two definitions of homophily, a weak definition and a strong definition. For this purpose, we introduce a parameter of maximum dissimilarity $d$ between a direct consumer and her followers as an inverse measure of the homophily in the social network. In the case of the strong definition, the followers of a consumer located at $x$ are distributed over an interval given by $[0, 1] \cap [x-d, x+d]$ according to the truncated distribution of the users of the social media. In the case of weak definition, we allow some followers to be distributed outside the
interval \([0, 1] \cap [x-d, x+d]\) as well. We show how homophily affects the distribution of the mean and the variance of followers’ locations by considering three different distribution of users; uniform, beta and normal.

Then, we consider the case of the uniform distribution with the strong definition of homophily and find that the breadth maximization leads to the limit polarization for high enough homophily and high enough attention tax and no polarization otherwise. To provide an intuition, we call those consumers located in \([d, 1-d]\) moderate, those located in \([0, d]\) left-wing extremists and those located in \((1-d, 1]\) right-wing extremists. Given the location of the news, we have two marginal consumers (indifferent between sharing the news and not), one on the left and the other on the right side of the news location. Suppose that when the news is located at the middle, both marginal consumers are moderate. Suppose now that the media firm locates the news a bit to the left from the middle. As long as both marginal consumers are moderate, change in the location of the news has no impact on the breadth of the news sharing. This is because, for all moderates, the variance is the same constant and the mean increases with a constant slope. A further change in news location to the left induces the media firm to reach a region in which the left-side marginal consumer is a left-wing extremist and the right-side marginal consumer is a moderate. We find that within this region, any further polarization increases the breadth up to the point it reaches the limit polarization as both the mean effect and the variance effect makes the gain from extremists larger than the loss from moderates.\(^2\)

In Section 4, we consider two competing media firms. All direct consumers receive both news but each of them can share at most one news. We find that under very mild conditions, as long as polarization occurs under monopoly, polarization occurs also under duopoly. In particular, in the case of the uniform distribution previously studied, we find that under duopoly, the limit polarization equilibrium always exists whereas no bias equilibrium never exists.

In Section 5, we provide two extensions of the baseline model. The first allows the firm to target an interval of consumers to show the news. When the direct consumers are

\(^2\)First, the mean effect is such that the mean increases with a smaller slope among extremists than among moderates. This induces polarization for the reason explained above. Second, the variance effect is such that the variance is increasing among extremists while it is constant among moderates. This means that a marginal polarization of the news further reduces the variance for the extremists and thereby induces extra extremists to share the news while such positive effect is absent among moderates.
distributed uniformly, we find that it is optimal for the firm to choose the same location as in the baseline model and to make the left end of the target interval coincide with the left marginal (direct) consumer. In the second extension, we consider re-sharing the shared news along the hierarchical layers of communication and characterize the strategy that maximizes the depth of sharing, i.e. the number of times the news is shared following down the layers of communication. Without loss of generality, we limit the media firm to target only one location of consumer: the media firm chooses the location of the news and the location of the targeted consumer. Then, we find that as long as the attention tax is not small, depth-maximization requires targeting the consumer with the lowest variance and locating the news at the mean location of the targeted consumer’s followers. Therefore, as long as the variance is not minimized at the middle, depth maximization leads to some bias.

Our work relates to the literature on demand-driven media bias (Gentzkow, Shapiro and Stone, 2015), which means that consumers enjoy reading news that confirm their biased beliefs. Mullainathan and Shleifer (2005) and Gabszewicz, Laussel and Sonnac (2001) use a Hotelling kind of model to explain a possible ideological bias of news. Mullainathan and Shleifer (2005) consider a duopoly price competition, and media bias emerges as each firm adopts a maximal differentiation strategy in order to soften price competition. By contrast, Gabszewicz, Laussel and Sonnac (2001) obtain minimal differentiation as the firms’ main revenue sources are advertising and prices are constrained to be non-negative, eliminating the force of differentiation to relax price competition. In addition, in both papers, a monopoly has no gain from choosing a location different from the middle. To the best of our knowledge, we are the first to embed social network of followers into a model of demand-driven media bias in order to study how news sharing and news bias are shaped by the main characteristics of the social network. In our model, the news is free and the media firm maximizes its advertising revenue. Empirical papers on the demand-driven media bias include Gentzkow and Shapiro (2010) and Larcinese, Puglisi and Snyder (2011). Gentzkow and Shapiro (2010) find that readers have an economically significant preference for like-minded news and that firms respond strongly to consumer preferences. Larcinese, Puglisi and Snyder (2011) find evidence that newspapers cater to readers’ partisan tastes on news about unemployment, trade deficit and budget deficit.

The literature on social media is mainly empirical and studies how social media affects voting, protests, xenophobia, polarization and consumption of fake news (see Zhuravskaya,
Petrova and Enikolopov, forthcoming, for a survey). Our paper is more related to the papers studying polarization, which we reviewed in the beginning of the introduction. There are a few theoretical papers on social media which are related to our paper. In terms of the categories of players, our paper is similar to Fainmesser and Galeotti (2018), which studies the interaction among marketers, influencers and followers as we study the interaction among media firms, direct consumers and followers; direct consumers in our model play the role of influencers in their model. Apart from this common point, the two papers are very different as they study competitive equilibria where each influencer chooses her amount of sponsored and organic content and marketers compete to buy sponsored content from the influencers. De Cornière and Sarvary (2018) study how content bundling by social media, i.e., social media shows news content together with user-generated content (UGC), affects the profit of newspapers and their incentive to invest in quality. However, they study neither news sharing nor social network. Campbell, Leister and Zenou (2019) build a dynamic model with two types of news (mass market and niche market) and two types of individuals who have preference for recommending one or other type of news. Each individual receives news from randomly sampled friends and shares one news; as long as she has received a news of her preferred type, she shares that type. They study the steady-state-equilibrium and find that greater connectivity and homophily concurrently increase the prevalence of the niche market content and polarization. Bloch, Demange and Kranton (2018) study transmission of rumors on social networks in a model in which there are two possible states of nature (0 or 1) and two types of agents (unbiased or biased). With probability less than 1, a perfect signal about the state is generated and each agent has an equal chance to become the only agent who receives the signal. Only this agent can send a message, which can be 0 or 1. An agent who receives a message from a neighbor can transmit it or not to her neighbors. After the communication stage is over, all agents vote between two alternatives 0 and 1; unbiased agents want the alternative to match the state of nature whereas biased agents prefer alternative 1. They find that a social network can serve as a filter: unbiased agents block messages from parts of the network that contain many biased agents, so that sufficiently credible information circulates. Our paper is similar to Bloch, Demange and Kranton (2018) and Campbell, Leister and Zenou (2019) as a direct consumer’ incentive to share a news is affected by the type of news, her type and her social network structure. The main difference with respect to the two papers is that we study endogenous choice of media
bias.

In Section 2, we present the baseline model. In Section 3 we study the breadth-maximization strategy. In Section 4, we study two competing media firms. Section 5 provides two extensions: targeting and depth-maximization. Section 6 concludes. All proofs are placed in the Appendix.

2 The Baseline Model

Our model has three categories of players: media firms, direct consumers and indirect consumers (i.e., followers). In the baseline model, we consider one media firm and we limit attention to a single layer of indirect consumers. We study competition between two media firms in Section 4 and study resharing of shared news in Section 5.

The media firm and consumers The media firm has to choose the ideological position of its news content, \( y \in [0, 1] \). We consider only free news, which is financed by advertising revenue. There is a continuum of consumers, who are also users of the social media. Each consumer is located on the interval \([0, 1]\). A consumer’s location, say \( x \), represents the ideal news she would like to read. By reading the news located at \( y \), a consumer gets a surplus equal to \( u \) and incurs a disutility \((x - y)^2\) from the mismatch between the news opinion location, \( y \), and her ideal one, \( x \). That is, the utility that a consumer located at \( x \) obtains from reading a news located at \( y \) is

\[
U(x, y) = u - (x - y)^2.
\]

Direct consumers and followers We assume that the media firm can target an interval of consumers of length \( l \in (0, 1] \), where \( l \) is a parameter depending on the financial resource of the media firm. The media firm can choose \( a \in [0, 1 - l] \) such that it targets consumers belonging to the interval \([a, a + l] \subset [0, 1]\). Each targeted consumer receives the news from the media firm. Those consumers who can be targeted by the media firm are called direct consumers. The particular case of \([a, a + l] = [0, 1]\) is called no targeting. In the case of no targeting, direct consumers are distributed according to the distribution function \( F \) with density \( f > 0 \), which is assumed to be symmetric around 1/2.
Each direct consumer has followers, who are also called indirect consumers. Those who follow a direct consumer and have access to the news only if it is \textit{shared} by the direct consumer. We assume that every direct consumer has a \textit{distinct} group of followers of equal measure, which is normalized to one. Given a direct consumer located at \( x \), her followers are distributed according to \( \tilde{g}(\cdot; x) \). We assume that the distribution of followers \( z \in [0,1] \) of a direct consumer located at \( x \) is symmetric to the distribution of followers of a consumer located at \( 1-x \) such that \( \tilde{g}(z; x) = \tilde{g}(1-z; 1-x) \).

\textbf{News sharing} \hspace{1em} Consumers read the news whenever they have access to it and get a non-negative utility from reading it. The existing literature from communication science and computer science identifies two motives for news sharing: altruistic motive (Boyd, Golder and Lotan, 2010; Holton et al., 2014; Small, 2011) and self-serving motive (Boyd, Golder and Lotan, 2010; Lee and Ma, 2012; Ma, Lee and Goh, 2011). We adopt the altruistic approach and assume that a direct consumer shares it if she reads it \textit{and} if she finds the news relevant enough for her followers. More precisely, we assume that conditional on reading it, she shares the news if and only if the average benefit that her followers obtain from reading the news is larger than the attention tax \( \tau \) she imposes on them by sharing the news. The attention tax represents the opportunity cost of attention as news sharing by a direct consumer induces each follower to pay attention to it, which does not necessarily mean reading it.

Let \( B(x, y) \) denote the average benefit that the followers of a direct consumer located at \( x \) obtain when the latter shares the news located at \( y \), that is,

\[ B(x, y) = \int_0^1 \max\{U(z, y), 0\} \tilde{g}(z; x) \, dz. \]  

(1)

Our altruistic approach regarding news sharing can be reconciled with the standard utility maximization based on self-serving motive such as social status if the sharer attaches positive (negative) utilities to the positive (negative) reactions from her followers as \textit{likes} or dislikes on Facebook or \textit{tweethearts} on Twitter, etc.\(^3\)

\(^3\)Suppose a direct consumer receives a \textit{like} (resp., \textit{dislike}) from a follower located at \( z \) whenever \( U(z, y) \geq \tau \) (resp., \( U(z, y) < \tau \)) and be the difference \( U(z, y) - \tau \) (resp., \( \tau - \max\{U(z, y), 0\} \)) a measure of the intensity of the feedback received. Then, we have:

\[ \mathbb{E}(\text{number of likes}) = \int U(z, y) \geq \tau (U(z, y) - \tau) \tilde{g}(z; x) \, dz, \]

\[ z \in [0,1] \]
**Profit** The media firm maximizes the advertising revenue, which is assumed to proportionate to the traffic to the firm’s news site (Gentzkow, Shapiro and Stone, 2015). The traffic is equivalent to the measure of readers. Given the location of its news at $y$, the profit of the firm is given by

$$\pi(y) = (1 - \alpha) D_0(y) + \alpha D_1(y),$$

where $D_0(y)$ and $D_1(y)$ are, respectively, the demand from the direct readers and the one from the indirect readers, and $\alpha$ captures the importance the firm gives to indirect demand relative to the indirect demand.

We assume that $u$ is large enough such that every consumer has an incentive to read the news. This assumption is made to simplify our analysis as under the assumption, every targeted consumer reads the news and every follower also reads it as long as the news is shared. Therefore, under the assumption, the profit maximization is equivalent to maximization of the measure of targeted consumers who share the news.

**Assumption 1.** $u$ is large enough that $U(x, y) > 0$ for any $(x, y) \in [0, 1]^2$.

Due to the symmetry of the problem, if $y^*$ maximizes the overall demand (i.e. the weighted sum of the direct demand and the indirect demand), so does $1 - y^*$. Therefore, it is sufficient to analyze only over a half of the interval. From now on we restrict attention to $y \in [0, 1/2]$, w.l.o.g.

**Definition 1 (Breadth).** The breadth of news sharing is the measure of direct consumers who share the news.

Suppose that $F$ is uniformly distributed and the direct consumers who share the news belongs to an interval. Then, the breadth of news sharing is just equal to the length of this interval. As we previously wrote, under Assumption 1, the profit maximization is equivalent to maximization of the breadth. In the next section, we study the breadth-maximizing location in the absence of targeting.

\[
E(\text{number of dislikes}) = \int_{z \in [0, 1]} U(z, y) < \tau (\tau - \max\{U(z, y), 0\}) \tilde{g}(z; x) dz.
\]

In that case, if a direct consumer shares the news if and only if $E(\text{number of likes}) - E(\text{number of dislikes}) \geq 0$ then she shares the news if and only if $B(x, y) - \tau \geq 0$. 


3 Breadth-maximizing location: No targeting

In this section, we study the breadth-maximizing strategy of the firm in terms of the location of the news in the absence of targeting, which means that the target interval is equal to \([0, 1]\).

Let \(\mu(x)\) and \(\sigma^2(x)\) represent, respectively, the mean and the variance of the locations of the followers of a consumer located at \(x\). Precisely,

\[
\mu(x) = \int_{0}^{1} z \tilde{g}(z; x) \, dz \quad \text{and} \quad \sigma^2(x) = \int_{0}^{1} (z - \mu(x))^2 \tilde{g}(z; x) \, dz.
\]

We first consider a general model in which the main primitives are the distribution of \(\mu(x)\) and \(\sigma^2(x)\) over \(x \in [0, 1]\) and study the properties of the distribution which leads to news polarization. Second, we introduce some definitions of homophily and generate the distribution of \(\mu(x)\) and \(\sigma^2(x)\) for some well-known distribution functions of population \(G\). Third, we consider the case in which \(G\) is uniform distribution, analyze the location choice of the media firm and perform comparative statics with respect to the degree of homophily and the level of attention tax.

3.1 General analysis

We here consider a general model in which the main primitives are the distribution of \(\mu(x)\) and \(\sigma^2(x)\) over \(x \in [0, 1]\). Recall that \(f\) represents the density of the direct consumers.

The next lemma shows that the benefit from sharing \(B(x, y)\) is a function of \(\mu(x)\) and \(\sigma^2(x)\) regardless of the distribution of the followers \(\tilde{g}(z; x)\). This is due to the fact that we assume a quadratic loss function when we define the utility from reading.

**Lemma 1.** Under Assumption 1, the average benefit that followers of a direct consumer located at \(x\) obtain when the latter shares the news located at \(y\) is \(B(x, y) = u - \sigma^2(x) - (y - \mu(x))^2\).

Therefore, to each direct consumer, the benefit of sharing decreases with an increase in the ideological dispersion of her followers, \(\sigma^2(x)\), and with the mismatch between the location of the news and the ideological location of her average follower, \((y - \mu(x))^2\).

We introduce a technical assumption to simplify our analysis: under the assumption, for any given attention tax level \(\tau\) and location of the news \(y\), the set of locations of direct consumers that share the news is an interval.
**Assumption 2.** For any \( y \), \( B(\cdot, y) \) is strictly quasiconcave and differentiable.

Let \( x_l(y) \) and \( x_r(y) \), with \( x_l(y) \leq x_r(y) \), denote the limits of such interval. From Assumption 1, every follower reads the news if she gets access to it. Thus, the indirect demand reached by the media firm when locating its news at \( y \) is equal to the breadth of the news sharing, that is,

\[
D_1(y) = \int_{x_l(y)}^{x_r(y)} f(x) \, dx = F(x_r(y)) - F(x_l(y)).
\]

We first identify a straightforward case of no polarization.

**Proposition 1.** Under Assumptions 1 and 2, if the attention tax is low enough such that even the most extremist consumer is willing to share an ideologically neutral news, then no polarization, i.e. \( y^* = 1/2 \), maximizes the breadth of news sharing as \( D_1(1/2) = 1 \).

Therefore, from now on, we consider \( \tau > B(x = 0, y = 1/2) \). In this case, it is always possible to have a location \( y \) (around 1/2) such that both \( x_l(y) > 0 \) and \( x_r(y) < 1 \). The locations where either \( x_l(y) = 0 \) or \( x_r(y) = 1 \) are of special interest and will be called the limit polarization locations.

**Definition 2** (Limit polarization locations). Let \( y^\text{L} \) (resp., \( y^\text{R} \)) denote the closest location from the center to which \( x_l(y) = 0 \) (resp., \( x_r(y) = 1 \)). We will refer to \( y^\text{L} \) as the left limit-polarization location and to \( y^\text{R} \) as the right limit-polarization location.

The maximization of the indirect demand leads, respectively, to the following first and second order conditions

\[
\frac{\partial D_1}{\partial y}(y) = f(x_r) \frac{\partial x_r}{\partial y} - f(x_l) \frac{\partial x_l}{\partial y}
\]

\[
\frac{\partial^2 D_1}{\partial y^2}(y) = f'(x_r) \left( \frac{\partial x_r}{\partial y} \right)^2 - f'(x_l) \left( \frac{\partial x_l}{\partial y} \right)^2 + f(x_r) \frac{\partial^2 x_r}{\partial y^2} - f(x_l) \frac{\partial^2 x_l}{\partial y^2}.
\]

Since a consumer located at \( \hat{x}(y) \in \{x_l(y), x_r(y)\} \) is just indifferent between sharing the news and not, we have \( B(\hat{x}, y) = \tau \) or, equivalently,

\[
u - \sigma^2(\hat{x}) - (y - \mu(\hat{x}))^2 = \tau.
\]

Then, by differentiating \( \hat{x} \) with respect to \( y \) on the above equation we have

\[
\frac{\partial \hat{x}}{\partial y} = \frac{2 (y - \mu)}{2 \mu_{\hat{x}} (y - \mu) - \sigma_{\hat{x}}^2}.
\]

We introduce our last assumption.
Assumption 3. We assume that there is some degree of homophily such that for any y, \( \mu(x_l(y)) < y < \mu(x_r(y)) \).

If there is no homophily at all, \( \mu(x) \) is a constant which does not depend on \( x \). In this case, clearly the above assumption cannot be satisfied. If the assumption is not satisfied such that \( \mu(x_l(y)) < \mu(x_r(y)) < y \) holds (respectively, \( y < \mu(x_l(y)) < \mu(x_r(y)) \) holds), from Lemma 1 and Assumption 2, moving \( y \) to the left (respectively, to the right) increases breadth. Therefore, the assumption makes the breadth maximization more interesting.

Lemma 2. From Assumptions 1-3, \( x_l \) and \( x_r \) are strictly increasing in \( y \in [y, \bar{y}] \).

Lemma 3. From the symmetry of the distribution of followers we have that, for any \( x \) and \( y \),

1. \( \mu(x) = 1 - \mu(1 - x) \);
2. \( \sigma^2(x) = \sigma^2(1 - x) \); and
3. \( x_r(y) = 1 - x_l(1 - y) \).

By differentiating equation (2) with respect to \( y \) we have the following second order derivative

\[
\frac{\partial^2 \hat{x}}{\partial y^2} = \frac{2(\sigma_x^2)^2 - (2\mu_x(y - \mu) - \sigma_x^2)^2}{(2\mu_x(y - \mu) - \sigma_x^2)^3}.
\] (3)

We have the following result:

Proposition 2. Assume the density of direct readers is such that \( f'(x) \geq 0 \) for \( x \in [0, 1/2) \). Then, the breadth of news sharing is maximized as follows:

For all \( y \in [y, \bar{y}] \),

1. \( x_l(y) \) strictly convex is a sufficient condition for no polarization (i.e. \( y^* = 1/2 \)).
2. \( x_l(y) \) concave enough is a sufficient condition for the optimality of the limit polarization strategy (i.e., \( y^* = y \) or \( y^* = \bar{y} \)).

Even if the proposition is general, it is stated in terms of convexity or concavity of \( x_l(y) \), of which the meaning is not obvious to grasp. Therefore, in what follows, we consider that \( f \) is uniformly distributed and try to understand which distribution of \( \mu(x) \) and \( \sigma^2(x) \) generates (no) polarization.
Figure 1: On the left, convex $x_l(y)$ implies that indirect demand is maximized with no polarization, $y^* = 1/2$. On the right, concave $x_l(y)$ implies that indirect demand is maximized with limit polarization, $y^* = \frac{y}{2}$ or $y^* = \overline{y}$.

Suppose that $f$ is uniformly distributed. Consider $x_l \in (0, 1/2)$, then from equation (3) follows that

- **Case 1:** $\sigma^2$ constant. (No variance effect)
  \[
  \frac{\partial^2 x_l}{\partial y^2} = -\frac{\mu_{xx}}{\mu_x^2}
  \]
  Polarization occurs if $\mu$ is convex on $[0, 1/2]$.

- **Case 2:** Both $\mu$ and $\sigma^2$ linear. (No second derivative effect)
  \[
  \frac{\partial^2 x_l}{\partial y^2} = \frac{2(\sigma_x^2)^2}{(2\mu_x (y - \mu) - \sigma_x^2)^3} > 0
  \]
  Polarization never occurs.

- **Case 3:** $\mu$ convex on $[0, 1/2]$, and $\sigma^2$ linear.
  \[
  \frac{\partial^2 x_l}{\partial y^2} = \frac{2(\sigma_x^2)^2 - 8\mu_{xx}(y - \mu)^3}{(2\mu_x (y - \mu) - \sigma_x^2)^3}
  \]
  Polarization occurs if $\mu$ is convex enough on $[0, 1/2]$.

- **Case 4:** $\mu$ linear and $\sigma^2$ concave on $[0, 1/2]$.
  \[
  \frac{\partial^2 x_l}{\partial y^2} = \frac{2(\sigma_x^2)^2 + 4\sigma_x^2(y - \mu)^2}{(2\mu_x (y - \mu) - \sigma_x^2)^3}
  \]
  Polarization occurs if $\sigma^2$ is concave enough on $[0, 1/2]$. 

13
The table provides the results for various cases. Summarizing, we obtain our central result.

<table>
<thead>
<tr>
<th>Constant variance</th>
<th>Concave variance</th>
<th>Linear variance</th>
<th>Convex variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concave mean</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Linear mean</td>
<td>No</td>
<td>Yes, if variance concave enough</td>
<td>No</td>
</tr>
<tr>
<td>Convex mean</td>
<td>Yes, if mean convex enough or variance concave enough</td>
<td>Yes, if mean convex enough</td>
<td>Yes, if mean convex enough</td>
</tr>
</tbody>
</table>

Table 1: Cases where polarization occurs when \( f \) is uniform.

**Proposition 3.** Assume that the density of direct readers \( f(x) \) is uniform.

1. No polarization (i.e. \( y^* = 1/2 \)) occurs if (i) mean \( \mu(x) \) is concave enough or variance \( \sigma^2(x) \) is convex enough for \( x \in [0, 1/2] \); or (ii) both mean \( \mu(x) \) and variance \( \sigma^2(x) \) are linear for \( x \in [0, 1/2] \).

2. The limit polarization (i.e, \( y^* = \underline{y} \) or \( y^* = \overline{y} \)) occurs if mean \( \mu(x) \) is convex enough or variance \( \sigma^2(x) \) is concave enough for \( x \in [0, 1/2] \).

We here try to provide an intuition for why convex mean \( \mu(x) \) for \( x \in [0, 1/2] \) can lead to polarization. Note first that \( \mu(x) \) will typically increase with \( x \). Hence, convex \( \mu(x) \) implies that it increases in a increasing way for \( x \in [0, 1/2] \); the mean followers of two left-wing extremists are more closely located than the mean followers of two moderates (see the Figure 2), which makes the benefit from news sharing decrease slowly when the news is located close to the extremists than when it is located close to the moderates, leading to the limit polarization.

In order to provide an intuition for why concave variance \( \sigma^2(x) \) for \( x \in [0, 1/2] \) can lead to polarization, let us assume that \( \sigma^2(x) \) is increasing for \( x \in [0, 1/2] \). The symmetry around the middle implies that \( \sigma^2(x) \) is concave for \( x \in [0, 1] \), increases for \( x \in [0, 1/2] \) and decreases for \( x \in [1/2, 1] \). Now consider \( y < 1/2 \) and examine what happens when \( y \) further moves to the left. Then, this movement will induce some left-wing extremists to start to share
Figure 2: Convexity of $\mu(x)$ on $[0, 1/2]$ means that mean followers of extremists are more closely located to each other than followers of moderates.

the news while inducing some moderates to stop sharing it. However, when $\sigma^2(x)$ is very concave, the decrease in $\sigma^2$ experienced by the former is much larger than any change in $\sigma^2$ experienced by the latter, which leads to an overall expansion of the breadth. This logic implies that limit polarization is optimal.

Let us now examine the effect of the distribution function of direct readers. As long as the function

$$\phi(y) = f(x_l(y)) \frac{\partial x_l(y)}{\partial y}$$

is increasing, it leads to $\frac{\partial D_1}{\partial y} > 0$ for $y < 1/2$ and $\frac{\partial D_1}{\partial y} < 0$ for $y > 1/2$, resulting in no polarization as the breadth-maximizing strategy. Analogously, $\phi$ decreasing leads to limit polarization as the optimal strategy.

The derivative of $\phi$ is given by

$$f'(x_l) \left( \frac{\partial x_l}{\partial y} \right)^2 + f(x_l) \frac{\partial^2 x_l}{\partial y^2}.$$ 

In our previous analysis, we fixed the sign of $f'$ to be non-negative on $[0, 1/2]$ and discussed how no polarization (resp., limit polarization) results from $x_l$ being convex (resp., concave enough). Now, we want to call attention to the fact that $f'(x) \geq 0$ (resp., $\leq 0$) generates a force towards no polarization (resp., limit polarization). If, for example, one consider target consumers randomly selected from the population, such that $F = G$, that suggests that the dramatic increase in the polarization of partisan preferences over the past 40 years in the US (see Lazer et al., 2018) makes polarization of online news media more likely.
3.2 Homophily and social networks

Up to now, the primitives of the model in terms of the social network have been $\mu(x)$ and $\sigma^2(x)$. But these are a function of the degree of homophily. In order to show how the degree of homophily affects these primitives, we here introduce a parameter $d \in [0, 1]$, which is an inverse measure of homophily, together with two definitions of homophily: a strong definition and a weak definition. Let $G$ represent the original distribution of the users of social media. Then, by applying the definitions to $G$, we can generate a distribution of $\mu(x)$ and $\sigma^2(x)$, which is a function of $d$. This offers a framework which allows us to study how an increase in the degree of homophily affects news polarization.

Definition 3 (Strong definition of homophily). The followers of a consumer located at $x$ are distributed over $[0, 1] \cap [x - d, x + d]$ according to the truncated distribution of the original distribution $G$.

In the strong definition of homophily, $d$ represents the maximum distance between a direct reader’s location and her follower’s location. Hence, if $d$ is relatively small and $x$ is not so close to an extreme, then the followers of a consumer located at $x$ are distributed over $[x - d, x + d]$ according to the truncated distribution of the original distribution $G$. However, as $x$ approaches an extreme, for instance, say the extreme left, then $x - d$ becomes smaller than 0, and in this case, we assume that the followers of a consumer located at $x$ are distributed over $[0, x + d]$ according to the truncated distribution of the original distribution $G$. Figure 3 illustrates this point.

In one extreme case of $d = 0$ all followers have the same preference as the direct consumer they follow, hence, the degree of homophily is extremely high; at the opposite extreme case of $d = 1$, all direct consumers have their followers distributed in an identical way over the unit interval no matter their location, in which case we have zero homophily.

We will write the truncation of a distribution $G$ to the interval $[0, 1] \cap [x - d, x + d]$ as $\tilde{G}(\cdot; x)$ and its density as $\tilde{g}(\cdot; x)$. Therefore,

$$\tilde{G}(z; x) = \begin{cases} 0 & \text{if } z < \tilde{z} \\ \frac{G(z) - G(z)}{G(\tilde{z}) - G(z)} & \text{if } z \in [\tilde{z}, \bar{z}] \\ 1 & \text{if } z > \bar{z} \end{cases}$$

and

$$\tilde{g}(z; x) = \begin{cases} \frac{g(z)}{G(\tilde{z}) - G(z)} & \text{if } z \in [\tilde{z}, \bar{z}] \\ 0 & \text{otherwise}, \end{cases}$$

with $\tilde{z} = \max\{0, x - d\}$ and $\bar{z} = \min\{1, x + d\}$. 

16
\[ \tilde{g}(z; x) = \frac{g(z)}{G(x + d) - G(x - d)} \]

Figure 3: Density of mass of followers under strong homophily.

In the weak definition of homophily, we relax the strong definition by assuming that the followers of any consumer located at \( x \in [0, 1] \) are distributed over the full support \([0, 1]\) but the mass of followers within the interval \([0, 1] \cap [x - d, x + d]\) is larger than the mass of the original population within the same interval:

**Definition 4 (Weak definition of homophily).** The distribution of the followers of a consumer located at \( x \) is obtained by redistributing some positive mass of population located outside of the interval \([0, 1] \cap [x - d, x + d]\) into the interval such that

1. the followers within the interval and outside the interval are distributed according to the truncated distribution of the original distribution \( G \); and
2. the total mass of followers within the interval is the same for all \( x \in [0, 1] \).

In other words, consider independent random variables \( Z \) and \( \tilde{Z} \), where \( Z \) is distributed according to \( G \) and \( \tilde{Z} \) is distributed according to \( \tilde{G}(\cdot; x) \), then the distribution of the followers of a consumer located at \( x \) in the weak definition is the same as the distribution of the random variable \( W = (1 - \beta) Z + \beta \tilde{Z} \), with \( \beta \in (0, 1) \) denoting the strength of homophily. In the limit case of \( \beta = 0 \) the distribution of followers of a consumer located at \( x \) is equal the population distribution, hence leading to zero homophily, and at the other extreme, as \( \beta \) go to 1 we move back toward the strong definition of homophily.

The following are some examples of the effect of the degree of homophily over the mean and variance functions to the Uniform(0,1), Beta(6,6) and Normal(0,1) distributions under the strong definition and weak definition.
Examples: Strong homophily

µ(x) for Uniform(0,1)

σ²(x) for Uniform(0,1)

µ(x) for Beta(6,6)

σ²(x) for Beta(6,6)

µ(x) for Normal(0,1)

σ²(x) for Normal(0,1)
Examples: Weak homophily

\[
\begin{align*}
\text{mean} & \quad 0.0 & 0.2 & 0.4 & 0.6 & 0.8 & 1.0 \\
\text{variance} & \quad 0.016 & 0.020 & 0.024 \\
\end{align*}
\]

\(\mu(x)\) for Uniform(0,1)

\[
\begin{align*}
\text{mean} & \quad 0.3 & 0.4 & 0.5 & 0.6 & 0.7 \\
\text{variance} & \quad 0.006 & 0.008 & 0.010 & 0.012 & 0.014 \\
\end{align*}
\]

\(\mu(x)\) for Beta(6,6)

\[
\begin{align*}
\text{mean} & \quad -4 & -2 & 0 & 2 & 4 \\
\text{variance} & \quad 0.4 & 0.5 & 0.6 & 0.7 & 0.8 \\
\end{align*}
\]

\(\mu(x)\) for Normal(0,1)
In the case of beta and normal distributions with strong homophily, we find that \( \mu(x) \) is concave for \( x \in [0, 1/2] \) and that \( \sigma^2(x) \) is initially convex and then concave for \( x \in [0, 1/2] \). If we apply the insight from Proposition 3, we find forces against polarization around \( x = 0 \) as the mean is convex and the variance is convex. In contrast, around \( x = 1/2 \), there are forces for polarization as long as the variance is convex enough relative to the concavity of mean. For instance, this is clearly the case when \( d = 1.0 \) since then the mean is almost linear. Therefore, we expect some intermediate level of polarization to occur in this case.

### 3.3 The case of uniform distribution

We here analyze the case in which both \( F \) and \( G \) are the uniform distribution. We apply the strong definition of homophily with the inverse measure of homophily \( d \) smaller than or equal to \( 1/2 \) (which means that only the direct consumer located at \( 1/2 \) can have her followers over the whole interval \([0, 1]\)). Therefore, consider

\[
\tilde{g}(z; x) = \begin{cases} 
\frac{1}{z-x} & \text{if } z \in [\bar{x}, \overline{z}] \\
0 & \text{otherwise}, 
\end{cases}
\]

\( z = \max\{0, x-d\} \) and \( \overline{z} = \min\{1, x+d\} \). We have that

\[
\mu(x) = \int_{\tilde{z}}^{\bar{z}} z \tilde{g}(z; x) \, dz = \frac{\bar{x} + \tilde{z}}{2} = \begin{cases} 
x + \frac{d}{2} & \text{if } x < d \\
\frac{x + d}{2} & \text{if } d \leq x \leq 1 - d \\
\frac{1 + x - d}{2} & \text{if } x > 1 - d 
\end{cases}
\]

and

\[
\sigma^2(x) = \int_{\tilde{z}}^{\bar{z}} (z - \mu(x))^2 \tilde{g}(z; x) \, dz = \frac{(\bar{x} - \tilde{z})^2}{12} = \begin{cases} 
\frac{(x + d)^2}{12} & \text{if } x < d \\
\frac{d^2}{3} & \text{if } d \leq x \leq 1 - d \\
\frac{(1 - x + d)^2}{12} & \text{if } x > 1 - d. 
\end{cases}
\]

Denoting \( L = [0, d], M = [d, 1 - d], \) and \( R = (1 - d, 1] \), we will say a consumer \( x \) is an left-wing extremist (resp., moderate or right-wing extremist) if \( x \in L \) (resp., \( M \) or \( R \)).

When the left-side marginal consumer is a moderate, \( x_l \in [d, 1 - d], \) we have \( \mu_x = 1, \sigma^2_x = 0 \) and

\[
\left. \frac{\partial x_l}{\partial y} \right|_{x_l \in M} = \frac{2(y - \mu)}{\frac{2(y - \mu)}{3}} = 1.
\]
When the left-side marginal consumer is a left-wing extremist, \( x_l \in [0,d) \), we have \( \mu_x = 1/2, \sigma_x^2 > 0 \) and \( \frac{\partial x_l}{\partial y} \bigg|_{x_l \in L} > \frac{2(y - \mu)}{(y - \mu)} = 2 \).

When the left-side marginal consumer is a right-wing extremist, \( x_l \in (1-d,1] \), we have \( \mu_x = 1/2, \sigma_x^2 < 0 \) and \( \frac{\partial x_l}{\partial y} \bigg|_{x_l \in R} < \frac{2(y - \mu)}{(y - \mu)} = 2 \).

In the case of uniform distribution some of our regularity assumptions cannot be satisfied. First, Assumption 2 ensures that for any given attention tax level \( \tau \) and location of the news \( y \), the set \( \{x : B(x,y) \geq \tau\} \) is an interval. In the case of uniform distribution, the benefit function is sometimes bimodal but some restriction in the parameters as follows is enough to ensure that the set of sharers will be an interval:

**Assumption 4.** \( d \in (0,3/8) \) and \( \tau < u - d^2/3 \).

Second, Assumption 3 can be violated when the benefit function has some skewness, that is, when the location of the direct consumer with the highest benefit of sharing is somewhat far from the news location. However, we can show that whenever \( \tau < u - d^2/3 \), we have that \( x_r(y) > y \) for all \( y \in [0,1/2] \).

Therefore, under Assumption 4 our previous analysis applies to the uniform case. In what follows, we replace Assumptions 2-3 with Assumption 4 and study the breadth-maximizing strategy. After describing a useful lemma, we provide the main result.

**Lemma 4.** Under Assumptions 1 and 4, in the case of uniform distribution, the maximum breadth is always reached either with no polarization, \( y^* = 1/2 \), or with limit polarization.

**Proposition 4** (Polarization in the case of uniform distribution). Under Assumptions 1 and 4, for any \( d \), there exists \( \tau^*(d) \) such that whenever \( \tau > \tau^*(d) \), the breadth of news sharing is maximized by the limit polarization location; otherwise, there is no polarization (\( y^* = 1/2 \)). In the case of the limit polarization, the firm chooses \( y^* = \overline{y} \) where

\[
\overline{y} = \mu(0) + \sqrt{u - \tau - \sigma^2(0)} = d/2 + \sqrt{u - \tau - d^2/12}.
\]

We below provide an intuition of the above result on limit polarization. When the attention tax is large enough, conditional on \( y = 1/2 \), both marginal consumers are moderate,
Figure 4: Summary of the cases of polarization: Polarization occurs when the degree of homophily is high and the attention tax is not negligible.

i.e., \( x_l \in [d, 1-d] \) and \( x_R \in [d, 1-d] \). As long as both marginal consumers are moderate, any small change in \( y \) has no impact on the breadth as
\[
\frac{\partial x_l}{\partial y} \bigg|_{x_l \in M} = \frac{\partial x_l}{\partial y} \bigg|_{x_r \in M} = 1.
\]

As \( y \) decreases, we reach a region in which the left marginal consumer is left-wing extremist and the right marginal consumer is moderate. In this region, any reduction in \( y \) increases breadth until the left marginal consumer’s location becomes equal to zero. This is because we have
\[
\frac{\partial x_l}{\partial y} \bigg|_{x_l \in L} > \frac{\partial x_l}{\partial y} \bigg|_{x_r \in M} = 1.
\]

Once the left marginal consumer’s location becomes equal to zero, any further reduction in \( y \) results in a loss in the breadth.

Therefore, the force that leads to the limit polarization is captured by (6). This force can be decomposed into the mean effect and the variance effect, both of which go to the same direction. First, the convexity of the mean function implies that two extremists have their mean followers more closely located that the mean followers of two moderates. This creates an incentive for polarization as the benefit of sharing decreases less fast relative to the case in which news is located at the center. Second, left-wing extremists have an increasing variance while moderates have a constant variance. This implies that, the more extreme is a
consumer, the lower is the variance of her followers and, consequently, the higher is the gross benefit of sharing, $u - \sigma^2$. Thus, such effect makes it easier for the firm to induce sharing from extremists than from moderates.

![Diagram](image)

Figure 5: When $y = 1/2$, both marginal consumers are moderate (i.e., $x_l \in [d, 1-d]$ and $x_r \in [d, 1-d]$) and any small change in $y$ has no impact on the breadth. As $y$ decreases, we reach a region in which the left marginal consumer is left-wing extremist and the right marginal consumer is moderate. In this region, any reduction in $y$ increases breadth until $y$ reaches the left-limit polarization location and the left marginal consumer’s location becomes equal to zero.

We have two comparative statics results. The first is about the regime change:

**Proposition 5.** *The threshold level of attention tax under which the media firm is indifferent between no polarization and limit polarization is such that $\partial \tau^*/\partial d > 0$. Hence, as the attention tax increases or the degree of homophily increases, the limit polarization is more likely.*

The above proposition is well explained in Figure 4. The fact that $\tau^*$ increases with $d$ implies that as the attention tax increases or the degree of homophily increases, the limit polarization is more likely.

The second result is about the comparative statics within the regime of limit polarization.
Proposition 6 (Comparative statics for the case of uniform distribution). The limit polarization location gets more polarized as the attention tax increases but the effect of an increase in the degree of homophily (i.e. a decrease in $d$) is ambiguous.

Let us first examine the effect of an increase in $\tau$ on polarization. If the firm is at the left-limit polarization location $y$, then a slight increase in the attention tax from $\tau$ to $\tilde{\tau}$ leads to $B(0, y) < \tilde{\tau}$, causing the most extremist consumer to stop sharing. To recover the extremist, the media firm has to move towards the extreme, at the cost of losing some moderate sharers. Consequently, the new optimal location, $\tilde{y}$, would be such that $\tilde{y} < y$ and $B(0, \tilde{y}) = \tilde{\tau}$.

We now examine the effect of an increase in $d$ (i.e., a reduction in homophily) on polarization. As the most extreme direct consumer has her followers located in the interval of $[0, d]$, a marginal increase in $d$ changes the consumer’s net benefit of sharing by $(U(d, y^*) - \tau)$. If this is positive (negative), the media firm finds it optimal to increase $y^*$ (decrease), leading to less (more) polarization of news.

4 Duopoly

How does competition affect news polarization? In this section, we consider competition between two media firms, $L$ and $R$. Each firm chooses the ideological position of its news. Let $y_L$ be the ideological location of the news produced by firm $L$ and $y_R$ the location of the news produced by firm $R$. Without loss of generality we can consider $y_L \leq y_R$ in equilibrium. We consider no targeting and assume that all direct consumers receive both news but each of them shares at most one news. A direct consumer shares the news that generates larger average benefit conditional on it being larger than the attention tax $\tau$. We first consider the general model and then the case of the uniform distribution.

Let $[x_l(y_L), x_r(y_L)]$ (resp. $[x_l(y_R), x_r(y_R)]$) be the set of locations of direct consumers that prefer to share the news from firm $L$ (resp. firm $R$) than do not share any news. The functions $x_l(\cdot)$ and $x_r(\cdot)$ are determined the same way as in the monopolistic model. Suppose that $x_l(y_R) < x_r(y_L)$. A direct consumer is indifferent between sharing $L$ news and sharing $R$ news if she is located at $x_m = x_m(y_L, y_R)$ such that $B(x_m, y_L) = B(x_m, y_R)$, which is equivalent to $\mu(x_m) = (y_L + y_R)/2$. Therefore, the firms’ respective indirect demands are

$$D_L(y_L, y_R) = F(x_m(y_L, y_R)) - F(x_l(y_L))$$
and
\[ D_R(y_L, y_R) = F(x_r(y_R)) - F(x_m(y_L, y_R)). \]

When attention tax is low enough, we have \( x_l(1/2) = 0 \) and \( x_r(1/2) = 1 \). In that case, as in the monopolistic model, we have no polarization.

**Proposition 7.** Under duopoly, if the attention tax is low enough such that even the most extremist consumer is willing to share an ideologically neutral news, the only equilibrium is the one where both firms adopt the no polarization strategy, i.e., \( y^*_L = y^*_R = 1/2 \).

The above proposition is similar to the result obtained by Gabszewicz, Laussel and Sonnac (2001) in a Hotelling model in which the non-negative pricing constraint binds due to high advertising revenue.

Now, consider the case where \( \tau > B(x = 0, y = 1/2) \) and let’s check when it is still optimal that both firms choose to be located at 1/2. Suppose \( y_R = 1/2 \). Then the maximization of the indirect demand of firm \( L \) leads to
\[
\frac{\partial D_L}{\partial y_L}(y_L, 1/2) = f(x_m) \frac{\partial x_m}{\partial y_L} - f(x_l) \frac{\partial x_l}{\partial y_L}.
\]

Therefore, \( \partial D_L/\partial y_L > 0 \) if and only if \( \partial x_l/\partial y_L < (f(x_m)/f(x_l)) \partial x_m/\partial y_L \).

**Proposition 8.** Under duopoly, when the attention tax is such that the most extremist consumer is not willing to share an ideologically neutral news, then

1. \( \mu(x) \) concave \([0, 1/2]\),
2. \( \sigma^2(x) \) decreasing on \([0, 1/2]\), and
3. \( f(x) \) increasing on \([0, 1/2]\)

are characteristics of the network that contribute to \( y^*_L = y^*_R = 1/2 \) to be an equilibrium.

In the above proposition, the second result is different from the result obtained in the monopoly case; in the monopoly case, convex variance contributed to no polarization. This is because in the case of duopoly, \( x_m \) does not depend on the variance \( \sigma^2(x) \) while \( x_l \) depends on it. A decreasing variance for \( x \in [0, 1/2] \) decreases the loss from moving \( y_L \) to the right (i.e., \( \partial x_l/\partial y_L \)) but does not affect the gain (i.e. \( \partial x_m/\partial y_L \)).

The next proposition shows that under mild conditions, limit polarization in monopoly implies limit polarization in duopoly.
Proposition 9. Limit polarization in monopoly implies limit polarization in duopoly whenever, for \( x \in [0, 1/2] \),

1. \( \mu \) is convex and \( \sigma^2_x \leq 0 \) or \( \sigma^2_x > 0 \) but small; or

2. \( \mu \) linear and \( \sigma^2 \) concave but with slope not very negative.

Note first that the conditions in the proposition such as convex mean and concave variance are consistent with the limit polarization in the monopoly (See Proposition 3). The reason for which the limit polarization occurs more often under duopoly is that the marginal consumer who is indifferent between sharing L news and sharing R news is less sensitive to a change in \( y_L \) (for instance) relative to the marginal consumer who is indifferent between sharing L news and not sharing. This is because the former has an outside option of sharing R news and the benefit from sharing R news is an increasing function of his/her location \( x_m \) whereas the latter has no outside option.

Suppose that the L firm chooses some \( y_L < \bar{y} \) given that \( y_R = \bar{y} \). Then, the interval of direct consumers who share L news is given by \([x_l(y_L), x_m(y_L)]\). Consider a monopoly media firm choosing the same \( y = y_L \). Then, the interval of direct consumers who share the news is given by \([x_l(y_L), x_r(y_L)]\). Now consider moving \( y_L \) a bit. First, suppose that \( y_L \leq 1/2 \) and consider moving \( y_L \) to the left side. Then the gain from such movement is \(|\partial x_l/\partial y_L|\) for both the duopolist and the monopolist. However, the loss is \(|\partial x_m/\partial y_L|\) for the duopolist while it is \(|\partial x_r/\partial y_L|\) for the monopolist, which is larger than \(|\partial x_m/\partial y_L|\) for the reason explained earlier. Hence, in this case, the marginal net benefit from the same deviation is larger for the duopolist than for the monopolist. Second, suppose that \( y_L \geq 1/2 \) and consider moving \( y_L \) to the right side. Then the loss from such movement is \(|\partial x_l/\partial y_L|\) for both the duopolist and the monopolist. However, the gain is \(|\partial x_m/\partial y_L|\) for the duopolist while it is \(|\partial x_r/\partial y_L|\) for the monopolist, which is larger than \(|\partial x_m/\partial y_L|\). Hence, in this case, the marginal net benefit from the same deviation is smaller for the duopolist than for the monopolist. Finally, combining these two results together with the symmetry of the monopolist’s profit around the mean implies that it is optimal for the L firm to choose \( y_L = y \) if it is optimal for the monopolist to choose \( y_L = y \) (see the Figure 6).

Finally, we consider the model of Section 3.3 in which \( F \) and \( G \) are the uniform distribution and study the duopoly competition. We have the following result.
Figure 6: Indirect demand functions of the monopolist and the duopolist. The monopolist has a larger loss (respectively, gain) than the duopolist for $y < 1/2$ (respectively, $y > 1/2$).

**Proposition 10.** Consider duopoly competition in the case in which $F$ and $G$ are the uniform distribution. Under Assumptions 1 and 4,

(i) There is no equilibrium in which both firms choose $y_L = y_R = 1/2$

(ii) There always exists an equilibrium in which each firm chooses the limit polarization $y_L = \underline{y}$ and $y_R = \overline{y}$.

The results can be easily explained as we have; on the one hand,

$$|\partial x_m/\partial y_L| = \frac{1}{2};$$

on the other hand,

$$|\partial x_l/\partial y_L| = 1 \quad \text{for moderate}$$

$$|\partial x_l/\partial y_L| > 2 \quad \text{for left-wing extremist}.$$ 

Therefore, given $y_R = 1/2$ or $\overline{y}$, L firm finds it optimal to choose $y_L = \underline{y}$.

5 Extensions

We here provide two extensions of the monopoly case.

5.1 Targeting

We here extend the baseline model by enabling the media firm to target an interval of direct readers. For instance, the Macedonian teenagers, mentioned in the introduction, purchased
bogus Facebook accounts and used them to target certain profiles of users to spread their fake news. In the baseline model, all direct consumers on the Hotelling line receive the news, which can be interpreted as the case of no targeting since it is conceptually similar to the case in which the direct consumers who receive the news is randomly selected from the Hotelling line.

Suppose that the media firm can target direct readers belonging to an interval of length \( l \in [0, 1) \) to send its news. Suppose also that the targeted consumers are uniformly distributed: \( F \) is uniformly distributed. Hence, the media firm now should choose not only the location of its news \( y \) but also the target interval \([a, a + l] \subset [0, 1]\). We show that when the media firm maximizes the breadth, it is a weakly dominant strategy to mimic the optimal location without targeting by choosing \( y = y^* \) and to choose the target interval by choosing \( a = x_l(y^*) \).

**Proposition 11** (Targeting Strategy). Suppose that the media firm can choose \( a \) to target direct consumers belonging to \([a, a + l] \subset [0, 1]\) with \( l \in [0, 1) \) in addition to choosing the news location \( y \). Suppose that \( F \) is uniformly distributed.

1. Whenever limit polarization is optimal without targeting, it is a weakly dominant strategy to choose \( y \) equal to the left limit polarization location and \( a = 0 \);

2. Whenever no polarization is optimal without targeting, it is a weakly dominant strategy to choose \( y = 1/2 \) and \( a = (1 - l)/2 \).

The proposition shows that the limit polarization strategy is still optimal even if we allow for targeting. This result is very consistent with the practice of the Macedonian teenagers and suggests that it would be optimal for them to use their bogus Facebook accounts to target the most extreme segment of consumers.

### 5.2 Resharing and the depth-maximization strategy

We can consider the possibility for indirect readers to have their own followers and to share the news. We can go down different layers of sharing starting from sharing of the news by direct consumers, resharing of the news shared, resharing of the reshared news and so on. Here, we provide an extension which, focuses on the depth of sharing. We maintain the assumption that each group of followers is distinct at any layer.
For instance, suppose that the constant $u$ depreciates over time such that it also depreciates when the layer of sharing increases. That is, the utility that a consumer located at $x$ obtains from reading a news located at $y$ after it has been shared $n = 1, 2, \ldots$ times would be

$$U_n(x, y) = \delta^{n-1}u - (x - y)^2,$$

with $\delta \in (0, 1)$, and the benefit generated by sharing the news one extra time given by

$$B_{n+1}(x, y) = \int_0^1 \max\{U_{n+1}(z, y), 0\} \, \tilde{g}(z; x) \, dz.$$

But for the sake of computational simplicity, we adopt a continuous version of the problem where sharing and consumption of the news occur instantaneously, in other words, lets $t \in [0, +\infty)$ be the time at which a news located at $y$ reaches an indirect consumer located at $x$, then a consumer’s utility of reading the news and her benefit of resharing it will be given, respectively, by

$$U_t(x, y) = \delta^t u - (x - y)^2 \quad \text{and} \quad B_t(x, y) = \int_0^1 \max\{U_t(z, y), 0\} \, \tilde{g}(z; x) \, dz.$$

Then, we can check how long the communication keeps going and take it as a measure of the number of reshares a news will receive.

**Definition 5 (Depth).** We define the depth of news sharing as the maximum number of times the news is shared following down the hierarchical layers of communications.

We allow the firm to target only one direct consumer, as this simplifies our exposition. Hence, the firm chooses the location of the targeted consumer and the location of the news.

The next result characterizes the depth-maximization strategy when the attention tax is not small.

**Proposition 12 (Depth Maximizing Strategy).** Suppose that the attention tax is not small. The media firm’s optimal strategy to maximize the depth of news sharing is characterized as follows:

1. It targets the consumer located at $x^*_{\text{depth}}$ whose $\sigma^2(x)$ is equal to $\min_{x \in [0, 1/2]} \sigma^2(x)$.

2. It chooses the location of the news $y^*_{\text{depth}} = \mu(x^*_{\text{depth}})$.

**Corollary 1.** Suppose that the attention tax is not small.
1. The depth-maximization leads to some polarization as long as \( \sigma^2(1/2) > \min_{x \in [0,1/2]} \sigma^2(x) \);

2. If \( \sigma^2(x) \) is increasing, it leads to \( x_{depth}^* = 0 \).

The assumption that the attention tax is not small implies that for consumers at the target location, as the news depreciates, the net benefit of sharing becomes zero before the utility of reading becomes zero. Given a consumer, her benefit from sharing is maximized when the news has the same location as that of her mean follower, and from (1) this leads to a benefit equal to \( u - \sigma^2(x) \). Therefore, it is optimal to target the location of the direct consumer with the lowest variance \( \sigma^2(x) \).

6 Conclusion

This paper studies how the social network of followers shapes a user’s incentive to share news on social media and how this in turn creates or not incentives for a profit-maximizing media firm to provide partisan content. In particular, we focus on how the distribution of the mean and the variance of followers’ ideological locations affects the ideological location of news. We find that both convex mean and concave variance contribute to the polarization when the media firm maximizes the breadth of news sharing. The analysis of the uniform distribution case reveals that polarization is more likely to occur as the degree of homophily increases and the attention tax imposed on followers by news sharing increases. We also found that increasing variance leads to limit polarization when the media firm maximizes the depth of news sharing.

If one considers that false political news tend to be hyperpartisan (Silverman et al., 2016), our results provide an explanation for the findings of Vosoughi, Roy and Aral (2018) that false news diffused deeper and more broadly than the true news, specially for false political news. The authors suggest that the degree of novelty and the extent to which the news is emotionally charged may explain their findings. Our results suggest that their findings can be also explained by the structure of social networks.

We think that the most interesting avenue for future research is to examine the role of news feed algorithm. In the case of Facebook, after a user decides to share news with her friends, the algorithm determines which subset of the friends would have exposure to the news. In our model, a social media platform has a clear incentive to employ an algorithm to
maximize the amount of time users spend on social media. Will such engagement-maximizing algorithm lead to more or less polarization of news?

References


Silverman, Craig, Lauren Strapagiel, Hamza Shaban, Ellie Hall, and Jeremy Singer-Vine. 2016. “Hyperpartisan Facebook Pages Are Publishing False And Misleading Information At An Alarming Rate.” *BuzzFeed*.


Zhuravskaya, Ekaterina, Maria Petrova, and Ruben Enikolopov. forthcoming. “Political Effects of the Internet and Social Media.” *Annual Review of Economics*.
A Proofs

Lemma 1. Under Assumption 1, the average benefit that followers of a direct consumer located at $x$ obtain when the latter shares the news located at $y$ is $B(x, y) = u - \sigma^2(x) - (y - \mu(x))^2$.

Proof. 

\[
B(x, y) = \int_0^1 \max\{U(z, y), 0\} \tilde{f}(z; x) dz \\
= \int_0^1 (u - (z - y)^2) \tilde{f}(z; x) dz \\
= u \int_0^1 \tilde{f}(z; x) dz - \int_0^1 ((z - \mu(x)) - (y - \mu(x)))^2 \tilde{f}(z; x) dz \\
= u - \int_0^1 (z - \mu(x))^2 \tilde{f}(z; x) dz + 2(y - \mu(x)) \int_0^1 (z - \mu(x)) \tilde{f}(z; x) dz \\
- (y - \mu(x))^2 \int_0^1 \tilde{f}(z; x) dz \\
= u - \sigma^2(x) - (y - \mu(x))^2 .
\]

\[\square\]

Proposition 1. Under Assumptions 1 and 2, if the attention tax is low enough such that even the most extremist consumer is willing to share an ideologically neutral news, then no polarization, i.e. $y^* = 1/2$, maximizes the breadth of news sharing as $D_1(1/2) = 1$.

Proof. The quasiconcavity of the benefit function implies that, for any $y$, $B(x, y) \geq \min\{B(0, y), B(1, y)\}$, for all $x$. Additionally, from the symmetry of the problem, $\max_{y \in [0,1]} \min\{B(0, y), B(1, y)\} = B(0, 1/2) (= B(1, 1/2))$.

Therefore, if $\tau \leq B(0, 1/2)$, by locating its news at $y^* = 1/2$ the media firm can induce sharing from all direct consumers, thus maximizing the breadth of news sharing. \[\square\]

Lemma 2. From Assumptions 1-3, $x_l$ and $x_r$ are strictly increasing in $y$.

Proof. Note that

\[
\frac{\partial B(x, y)}{\partial x} = 2\mu_x(x)(y - \mu(x)) - \sigma^2_y(x).
\]

The strict concavity and symmetry of $B$ implies that

\[
\frac{\partial B(x_l(y), y)}{\partial x} > 0 \quad \text{and} \quad \frac{\partial B(x_r(y), y)}{\partial x} < 0.
\]
Taking that together with the Assumption 3 give us that both \( \partial x_l / \partial y \) and \( \partial x_r / \partial y \) are positive.

**Lemma 3.** From the symmetry of the distribution of followers we have that, for any \( x \) and \( y \),

1. \( \mu(x) = 1 - \mu(1 - x) \);

2. \( \sigma^2(x) = \sigma^2(1 - x) \); and

3. \( x_r(y) = 1 - x_l(1 - y) \).

**Proof.**

1. 

\[
\mu(x) = \int_0^1 z \hat{f}(z; x) \, dz \\
= \int_0^1 z \hat{f}(1 - z; 1 - x) \, dz \\
= \int_0^1 (1 - w) \hat{f}(w; 1 - x) \, dw \\
= 1 - \mu(1 - x)
\]

2. 

\[
\sigma^2(x) = \int_0^1 (z - \mu(x))^2 \hat{f}(z; x) \, dz \\
= \int_0^1 (z - 1 + \mu(1 - x))^2 \hat{f}(1 - z; 1 - x) \, dz \\
= \int_0^1 (w - \mu(1 - x))^2 \hat{f}(w; 1 - x) \, dw \\
= \sigma^2(1 - x)
\]

3. 

\[
B(\hat{x}, y) = u - \sigma^2(\hat{x}) - (y - \mu(\hat{x}))^2 \\
= u - \sigma(1 - \hat{x}) - (y - 1 + \mu(1 - \hat{x}))^2 \\
= B(1 - \hat{x}, 1 - y)
\]
\(x_l(1 - y)\) is the smallest \(x\) that solves \(B(x, 1 - y) = \tau\), but from the above if \(x_l(y)\) and \(x_r(y)\) are solutions for \(B(x, y) = \tau\), then \(1 - x_l(y)\) and \(1 - x_r(y)\) are solutions for \(B(x, 1 - y) = \tau\), and since \(x_l(y) < x_r(y)\) implies \(1 - x_l(y) > 1 - x_r(y)\), we have that 
\(x_l(1 - y) = 1 - x_r(y)\).

\[\square\]

**Proposition 2.** Assume the density of direct readers is such that \(f'(x) \geq 0\) for \(x \in [0, 1/2)\). Then, the breadth of news sharing is maximized as follows:

For all \(y \in [\underline{y}, \overline{y}]\),

1. \(x_l(y)\) strictly convex is a sufficient condition for no polarization (i.e. \(y^* = 1/2\)).

2. \(x_l(y)\) concave enough is a sufficient condition for the optimality of the limit polarization strategy (i.e., \(y^* = \underline{y}\) or \(y^* = \overline{y}\)).

**Proof.**

1. Take \(y < 1/2\).

(a) \(x_l\) strictly increasing (Lemma 2), \(f' \geq 0\) for \(x \in [0, 1/2)\) and symmetric imply that \(f(x_l(y)) \leq f(x_l(1 - y))\); 
(b) \(x_l\) strictly convex implies \(\frac{\partial x_l(y)}{\partial y} < \frac{\partial x_l(1 - y)}{\partial y}\); and 
(c) \(x_r(y) = 1 - x_l(1 - y)\) (Lemma 3) implies \(\frac{\partial x_r(y)}{\partial y} = \frac{\partial x_l(1 - y)}{\partial y}\).

From (a)-(c) and from the symmetry of \(f\) follows that
\[
\frac{\partial D_1}{\partial y}(y) = f(x_r(y))\frac{\partial x_r(y)}{\partial y} - f(x_l(y))\frac{\partial x_l(y)}{\partial y} \\
= f(x_l(1 - y))\frac{\partial x_l(y)}{\partial y}(1 - y) - f(x_l(y))\frac{\partial x_l(y)}{\partial y}(y) > 0.
\]

It is straightforward to verify that \(y > 1/2\) implies \(\frac{\partial D_1}{\partial y}(y) < 0\). Therefore, the breadth is maximized at \(y^* = 1/2\).

2. For any \(y < \underline{y}\) we have that \(x_l(y) = 0\) and
\[
\frac{\partial D_1}{\partial y}(y) = f(x_r(y))\frac{\partial x_r(y)}{\partial y} \geq 0.
\]
On the other hand, for any \( y < y < 1/2 \) we have that if \( x_i(y) \) is concave enough such that
\[
\frac{\partial^2 x_i}{\partial y^2} < -\frac{f'(x_i)}{f(x_i)} \left( \frac{\partial x_i}{\partial y} \right)^2,
\]
then the product \( f(x_i) \frac{\partial x_i}{\partial y} \) is decreasing in \( y \). Resulting to
\[
\frac{\partial D_1}{\partial y}(y) = f(x_i(1 - y)) \frac{\partial x_i}{\partial y}(1 - y) - f(x_i(y)) \frac{\partial x_i}{\partial y}(y) < 0.
\]
Hence, the limit polarization strategy is optimal.

\[\square\]

**Proposition 3.** Assume that the density of direct readers \( f(x) \) is uniform.

1. No polarization (i.e. \( y^* = 1/2 \)) occurs if (i) mean \( \mu(x) \) is concave enough or variance \( \sigma^2(x) \) is convex enough for \( x \in [0, 1/2] \); or (ii) both mean \( \mu(x) \) and variance \( \sigma^2(x) \) are linear for \( x \in [0, 1/2] \).

2. The limit polarization (i.e, \( y^* = \underline{y} \) or \( y^* = \overline{y} \)) occurs if mean \( \mu(x) \) is convex enough or variance \( \sigma^2(x) \) is concave enough for \( x \in [0, 1/2] \).

**Proof.** When direct readers are uniformly distributed it follows that
\[
\frac{\partial D_1}{\partial y}(y) = \frac{\partial x_r}{\partial y}(y) - \frac{\partial x_l}{\partial y}(y)
= \frac{\partial x_l}{\partial y}(1 - y) - \frac{\partial x_l}{\partial y}(y).
\]
Therefore, no polarization (resp., limit polarization) occurs if \( x_i(y) \) is convex (resp., concave) on \([y, \overline{y}]\). From Equation (3) we have that
\[
\frac{\partial^2 x_i}{\partial y^2} = \frac{2 \left( \sigma_x^2(x_i) \right)^2 - (2\mu_{xx}(x_i)(y - \mu(x_i)) - \sigma_{xx}^2(x_i)) \left( y - \mu(x_i) \right)^2}{\left( 2\mu_{xx}(x_i)(y - \mu(x_i)) - \sigma_{xx}^2(x_i) \right)^2}.
\]
Hence,
\[
\frac{\partial^2 x_i}{\partial y^2} \geq 0 \quad \text{if and only if} \quad 2\mu_{xx}(x_i)(y - \mu(x_i)) - \sigma_{xx}^2(x_i) \leq \frac{1}{2} \left( \frac{\sigma^2_x(x_i)}{y - \mu(x_i)} \right)^2,
\]
and the proposition follows.

\[\square\]

**Lemma 4.** Under Assumptions 1 and 4, in the case of uniform distribution, the maximum breadth is always reached either with no polarization, \( y^* = 1/2 \), or with limit polarization.
Proof.

- Suppose \( x_l \in R \) and \( x_r \in R \), then from (2), (4) and (5) follows that

\[
\left. \frac{\partial x_l}{\partial y} \right|_{x_l \in R} < 2 \quad \text{and} \quad \left. \frac{\partial x_r}{\partial y} \right|_{x_r \in R} > 2.
\]

Hence, \( D_1(y) \) is increasing for all \( y \in [1/2, \bar{y}] \) such that \((x_l(y), x_r(y)) \in (R, R)\). For \( y > \bar{y} \), \( D_1(y) \) is decreasing (as \( \partial x_r / \partial y = 0 \)). Therefore, the maximum is reached at the limit polarization location.

- Suppose \( x_l \in M \), then

\[
\left. \frac{\partial x_l}{\partial y} \right|_{x_l \in M} = 1.
\]

- Case 1: if \( x_r \in M \), then

\[
\left. \frac{\partial x_r}{\partial y} \right|_{x_r \in M} = 1
\]

and \( \frac{\partial D_1}{\partial y}(y) = 1 - 1 = 0 \). Hence, \( D_1(y) \) is constant for all \( y \) such that \((x_l(y), x_r(y)) \in (M, M)\), and \( y = 1/2 \) is a local maximum.

- Case 2: if \( x_r \in R \), then

\[
\left. \frac{\partial x_r}{\partial y} \right|_{x_r \in R} > 2.
\]

Hence, \( D_1(y) \) is increasing for all \( y \) such that \((x_l(y), x_r(y)) \in (M, R)\). So, as \( y \) increases, either there is a regime change and \( x_l \) becomes an extreme-rightist, or the maximum is reached at the limit polarization location.

- Suppose \( x_l \in L \), then from (2), (4) and (5) follows that

\[
\left. \frac{\partial x_l}{\partial y} \right|_{x_l \in L} > 2.
\]

- Case 1: \( x_r \in L \),

\[
\left. \frac{\partial x_r}{\partial y} \right|_{x_r \in L} < 2.
\]

Hence, as long as \( x_l(y) > 0 \) we have that \( \frac{\partial D_1}{\partial y}(y) < 0 \). That is, \( D_1(y) \) is decreasing for all \( y \in [y, 1/2] \) such that \((x_l(y), x_r(y)) \in (L, L)\). Therefore, the maximum is reached at the limit polarization location.
- Case 2: \( x_r \in M \)
\[
\left. \frac{\partial x_r}{\partial y} \right|_{x_r \in M} = 1.
\]

Hence, \( \frac{\partial D_1}{\partial y}(y) < 0 \) and \( D_1(y) \) is decreasing for all \( y \in [y, 1/2] \) such that \( (x_l(y), x_r(y)) \in (L, M) \). Therefore, as \( y \) decreases, either there is a regime change and \( x_r \) becomes an extreme-leftist, or the maximum is reached at the limit polarization location.

- Case 3: \( x_r \in R \)
\[
\left. \frac{\partial x_r}{\partial y} \right|_{x_r \in R} > 2.
\]

In this case, to determine the signal of \( \frac{\partial D_1}{\partial y}(y) \) we need to recur to the second order derivative. From (3), (4) and (5) follows that
\[
\left. \frac{\partial^2 x_l}{\partial y^2} \right|_{x_l \in L} = \frac{2 \left( \sigma^2(x_l) \right)^2 + \frac{2}{3} (y - \mu(x_l))^2}{(y - \mu(x_l) - \sigma^2(x_l))^3} > 0
\]

and
\[
\left. \frac{\partial^2 x_r}{\partial y^2} \right|_{x_r \in R} = \frac{2 \left( \sigma^2(x_r) \right)^2 + \frac{2}{3} (y - \mu(x_r))^2}{(y - \mu(x_r) + \sigma^2(x_r))^3} < 0.
\]

The signs of the above expressions result from the fact that their numerators are both positive and their denominators are equal to \( B_x(x_l, y) > 0 \) and \( B_x(x_r, y) < 0 \).

Hence, \( \frac{\partial^2 D_1}{\partial y^2}(y) < 0 \) for all \( y \in [y, \bar{y}] \).

And, since by Lemma 3, we must have
\[
\left. \frac{\partial x_l}{\partial y} \right|_{y=1/2} = \left. \frac{\partial x_r}{\partial y} \right|_{y=1/2},
\]

i.e., \( \frac{\partial D_1}{\partial y}(1/2) = 0 \) and \( y = 1/2 \) is a local maximum.

Therefore, in any case the breadth is maximized either with no polarization or with limit polarization. \( \square \)

**Proposition 4** (Polarization in the case of uniform distribution). *Under Assumptions 1 and 4, for any \( d \), there exists \( \tau^*(d) \) such that whenever \( \tau > \tau^*(d) \), the breadth of news sharing is maximized by the limit polarization location; otherwise, there is no polarization (\( y^* = 1/2 \)). In the case of the limit polarization, the firm chooses \( y^* = y \) where
\[
y = \mu(0) + \sqrt{u - \tau - \sigma^2(0)}
= d/2 + \sqrt{u - \tau - d^2/12},
\]

39
Figure 7: Regions of regime changes.

Proof. When the attention tax is such that

- $\tau \leq B(0, 1/2)$ (Region I in Figure 4) we have that $D_1(1/2) = 1$. That is, $y^* = 1/2$.

- $B(0, 1/2) < \tau < B(1 - d, \tilde{y})$ (Region II in Figure 4) we have that $(x_l, x_r) \in (L, R)$ for all $y \in (\tilde{y}, 1/2]$. And since for $(x_l, x_r) \in (L, R)$ we have that $D_1(y)$ is increasing on $[y, 1/2]$ follows that $y^* = 1/2$.

- $B(1 - d, \tilde{y}) \leq \tau < B(d, 1/2)$ (Region III in Figure 4) we have that $(x_l, x_r) \in (0, M)$ if $y \in [0, \tilde{y}]$
  $(x_l, x_r) \in (L, M)$ if $y \in (\tilde{y}, y)$
  $(L, R)$ if $y \in (\tilde{y}, 1/2]$, where $\tilde{y} < 1/2$ is such that $x_r(\tilde{y}) = 1 - d$.

Therefore,

$D_1(y)$ is increasing for $y \in [0, \tilde{y}]$
$D_1(y)$ is decreasing for $y \in (\tilde{y}, y]$ 
$D_1(y)$ is increasing for $y \in (\tilde{y}, 1/2]$.

Hence, $y^* = y$ or $y^* = 1/2$.  

40
• $B(d, 1/2) \leq \tau < B(1/2, 1/2) = u - d^2/3$ (Region IV in Figure 4) we have that

\[
(0, M) \quad \text{if } y \in [0, \hat{y}]
\]
\[
(x_l, x_r) \in (L, M) \quad \text{if } y \in (\hat{y}, \hat{y})
\]
\[
(M, M) \quad \text{if } y \in [\hat{y}, 1/2],
\]

where $\hat{y} < 1/2$ is such that $x_l(\hat{y}) = d$.

Therefore,

increasing for $y \in [0, \hat{y}]$

$D_1(y)$ is decreasing for $y \in (\hat{y}, \hat{y})$

constant for $y \in [\hat{y}, 1/2]$.

Hence, $y^* = \hat{y}$.

Therefore, the continuity of the problem assures that for any $d$, there exists $\tau^*(d)$, with $B(1 - d, \hat{y}) < \tau^*(d) < B(d, 1/2)$, such that the breadth of news sharing is maximized with no polarization if $\tau < \tau^*(d)$, and with limit polarization if $\tau > \tau^*(d)$.

\[\square\]

**Proposition 5.** The threshold level of attention tax under which the media firm is indifferent between no polarization and limit polarization is such that $\partial \tau^*/\partial d > 0$. Hence, as the attention tax increases or the degree of homophily increases, the limit polarization is more likely.

**Proof.** Consider $B(1 - d, \hat{y}) \leq \tau < B(d, 1/2)$ (Region III in Figure 4). Then,

1. $x_l(1/2) \in L$ and from $B(x_l(1/2), 1/2) = \tau$ follows that

\[
u - \frac{(x_l(1/2) + d)^2}{12} - \left(\frac{1}{2} - \frac{x_l(1/2) + d}{2}\right)^2 = \tau \Rightarrow x_l(1/2) = \frac{3}{4} - d - \sqrt{3 \left(u - \tau - \frac{1}{16}\right)}.
\]

Hence,

$D_1(1/2) = x_r(1/2) - x_l(1/2)$

$= 1 - 2 x_l(1/2)$

$= 2d - \frac{1}{2} + 2 \sqrt{3 \left(u - \tau - \frac{1}{16}\right)}.$
2. $x_r(y) \in M$ and from $B(x_r(y), y) = \tau$ follows that

$$u - \frac{d^2}{3} - (y - x_r(y))^2 = \tau \Rightarrow x_r(y) = y + \sqrt{u - \tau - \frac{d^2}{3}}.$$ 

Therefore,

$$D_1(y) = x_r(y) - x_l(y)
= y + \sqrt{u - \tau - \frac{d^2}{3}} - 0
= \frac{d}{2} + \sqrt{u - \tau - \frac{d^2}{12}} + \sqrt{u - \tau - \frac{d^2}{3}}.$$ 

Hence, $\tau^*(d)$ is such that

$$\frac{d}{2} + \sqrt{u - \tau^*(d) - \frac{d^2}{12}} + \sqrt{u - \tau^*(d) - \frac{d^2}{3}} = 2d - \frac{1}{2} + 2\sqrt{3\left(u - \tau^*(d) - \frac{1}{16}\right)}.$$ 

Implicitly differentiating the equation above with respect to $d$, follows that

$$\frac{1}{2} - \frac{d}{12b} - \frac{d}{3c} - \left(\frac{1}{2b} + \frac{1}{2c}\right) \frac{\partial \tau^*}{\partial d} = 2 - \frac{\sqrt{3}}{a} \frac{\partial \tau^*}{\partial d}$$
$$\left(\frac{\sqrt{3}}{a} - \frac{1}{2b} - \frac{1}{2c}\right) \frac{\partial \tau^*}{\partial d} = \frac{3}{2} + \frac{d}{12b} + \frac{d}{3c}$$
$$\sqrt{3bc - a(b + c)/2} \frac{\partial \tau^*}{\partial d} = \frac{18bc + (4b + c)d}{12bc} \quad (7)$$

with $a = \sqrt{u - \tau^*(d) - 1/16}$, $b = \sqrt{u - \tau^*(d) - d^2/12}$ and $c = \sqrt{u - \tau^*(d) - d^2/3}$.

Since the RHS of equation (7) is positive, $\partial \tau^*/\partial d$ has the same sign of $\sqrt{3bc - a(b + c)/2}$.

But for any $d < 3/8$ we have that $a > c > b$, thus

$$\sqrt{3bc - a(b + c)/2} > \sqrt{3b^2 - a^2} > 1/16 - \sqrt{3}d^2/12 > 0.$$ 

Therefore, $\tau^*(d)$ is strictly increasing in $d$. 

**Proposition 6** (Comparative statics for the case of uniform distribution). The limit polarization location gets more polarized as the attention tax increases but the effect of an increase in the degree of homophily (i.e. $a$ decreases in $d$) is ambiguous.

*Proof.* Whenever $y = d/2 + \sqrt{u - \tau - d^2/12}$ we have that

$$\frac{\partial y}{\partial \tau} = -\frac{1}{2\sqrt{1/4 - d^2/12 - \tau}} < 0.$$
Thus, an increase in the attention tax causes an increase in polarization.

On the other hand,
\[
\frac{\partial y}{\partial d} = \frac{1}{2} \left( 1 - \frac{d/6}{\sqrt{u - \tau - d^2/12}} \right).
\]

Consequently,
\[
\frac{\partial y}{\partial d} > 0 \iff \sqrt{u - \tau - d^2/12} > \frac{d}{6} \iff \tau < u - \frac{d^2}{9}.
\]

Therefore, an increase in homophily (decrease in \(d\)) leads to further polarization only in case that \(\tau < u - \frac{d^2}{9}\).

**Proposition 7.** Under duopoly, if the attention tax is low enough such that even the most extremist consumer is willing to share an ideologically neutral news, the only equilibrium is the one where both firms adopt the no polarization strategy, i.e., \(y^*_L = y^*_R = 1/2\).

**Proof.** Clearly, \(y^*_L = y^*_R = 1/2\) is an equilibrium since by deviating from the center a firm does not increases the number of sharing from extremists (since all they were already sharing) while it loses sharings from some moderates that now prefers to share the news of its competitor.

In any other tentative equilibrium where one firm, say firm \(L\), plays \(y_L \neq 1/2\), the best reply for firm \(R\) is to chose a location \(y_R\) between \(y_L\) and \(1/2\), the closest possible of \(y_L\). But since whenever \(y_R \neq 1/2\) then the best reply of firm \(L\) is to choose location \(y_L\) between \(y_R\) and \(1/2\), the closest possible of \(y_R\), the only possible equilibrium is \(y^*_L = y^*_R = 1/2\).

**Proposition 8.** Under duopoly, when the attention tax is such that the most extremist consumer is not willing to share an ideologically neutral news, then

1. \(\mu(x)\) concave \([0, 1/2]\),

2. \(\sigma^2(x)\) decreasing on \([0, 1/2]\), and

3. \(f(x)\) increasing on \([0, 1/2]\)

are characteristics of the network that contribute to \(y^*_L = y^*_R = 1/2\) be an equilibrium.

**Proof.** Let’s start by considering the case where \(f\) uniform. Then, we need to have
\[
\frac{\partial D_L}{\partial y_L}(y_L, 1/2) = \frac{\partial x_m}{\partial y_L} - \frac{\partial x_l}{\partial y_L} \geq 0.
\]

In case we have
• $\sigma^2$ constant:

\[
\frac{\partial D_L}{\partial y_L}(y_L, 1/2) = \frac{1}{2\mu_x(x_m)} - \frac{1}{\mu_x(x_l)}.
\]

Therefore, $\partial D_L/\partial y_L \geq 0$ requires $\mu(x)$ to be sufficiently concave on $[0, 1/2]$.

• both $\mu$ and $\sigma^2$ linear on $[0, 1/2]$:

For $x \in [0, 1/2]$, let $\mu(x) = mx$ and $\sigma^2(x) = vx$. Hence,

\[
\frac{\partial D_L}{\partial y_L}(y_L, 1/2) = \frac{1}{2m} - \frac{2(y - mx)}{2m(y - mx) - v}.
\]

Therefore, $\partial D_L/\partial y_L \geq 0$ requires $v \leq -2m(y - mx)$. Since, from Assumption 3, we must have $m > 0$ and $(y - mx) > 0$, it results that $v$ needs to be sufficiently negative.

• $\mu$ concave on $[0, 1/2]$ and $\sigma^2$ linear on $[0, 1/2]$:

Let $\sigma^2(x) = vx$, for $x \in [0, 1/2]$. Hence,

\[
\frac{\partial D_L}{\partial y_L}(y_L, 1/2) = \frac{1}{2\mu_x(x_m)} - \frac{2(y - \mu(x_l))}{2\mu_x(x_l)(y - \mu(x_l)) - v}.
\]

Therefore, $\partial D_L/\partial y_L \geq 0$ requires $\mu_x(x_l) - 2\mu_x(x_m) \geq v |2(y - \mu(x_l))|^{-1}$. Hence, the more negative (positive) is the slope of $\sigma^2(x)$, the less (more) concave $\mu(x)$ needs to be.

• $\mu$ linear and $\sigma^2$ decreasing on $[0, 1/2]$:

Let $\mu(x) = mx$, for $x \in [0, 1/2]$. Hence,

\[
\frac{\partial D_L}{\partial y_L}(y_L, 1/2) = \frac{1}{2m} - \frac{2(y - mx)}{2m(y - mx) - \sigma^2_x(x_l)}.
\]

Therefore, $\partial D_L/\partial y_L \geq 0$ requires $\sigma^2(x_l) \leq -2m(y - mx_l)$, i.e., $\sigma^2$ needs to be sufficiently negative.

Consequently, when $f$ is uniform, $\mu(x)$ concave on $[0, 1/2]$ and $\sigma^2(x)$ decreasing on $[0, 1/2]$ contribute $y_L^* = y_R^* = 1/2$ be an equilibrium.

It remains to be shown that $f'(x) \geq 0$ on $[0, 1/2]$ contributes to $y_L^* = y_R^* = 1/2$ be an equilibrium.

When $f'(x) \geq 0$ for $x \in [0, 1/2]$, we have that

\[
\frac{\partial D_L}{\partial y_L}(y_L, 1/2) = f(x_m) \frac{\partial x_m}{\partial y_L} - f(x_l) \frac{\partial x_l}{\partial y_L} \geq f(x_l) \left( \frac{\partial x_m}{\partial y_L} - \frac{\partial x_l}{\partial y_L} \right).
\]
Therefore, whenever $y_L^* = y_R^* = 1/2$ is an equilibrium for $f$ is uniform, it is also an equilibrium when $f' \geq 0$ on $[0, 1/2]$. Additionally, $f(x_m) > f(x_l)$ makes it possible to have $\partial D_L / \partial y_L \geq 0$ even in cases where $\partial x_m / \partial y_L < \partial x_l / \partial y_L$.

\[ \frac{\partial D_1(y)}{\partial y} = \frac{\partial x_r}{\partial y}(y) - \frac{\partial x_l}{\partial y}(y) = \begin{cases} < 0 & \text{for } y \in (y, 1/2) \\ > 0 & \text{for } y \in (1/2, \bar{y}) \end{cases} \]

In the case of the duopoly, we consider the deviation of the left side media firm given that the right side firm chooses the right-limit polarization location $\bar{y}$. That is,

\[ \frac{\partial D_L(y, \bar{y})}{\partial y_L} = \frac{\partial x_m}{\partial y_L}(y, \bar{y}) - \frac{\partial x_l}{\partial y_L}(y). \]

Given that the rival is located at $\bar{y}$, we have $1/2 \leq x_m(y, \bar{y}) \leq x_r(y)$ for any $y \in [y, \bar{y}]$.

\begin{itemize}
  \item Claim 1: $\frac{\partial D_L}{\partial y} (y, \bar{y}) \leq \frac{\partial D_1}{\partial y} (y)$ for any $y \in (y, 1/2)$.
  
  This claim implies $|\frac{\partial D_L}{\partial y} (y, \bar{y})| \geq |\frac{\partial D_1}{\partial y} (y)| > 0$ for $y \in (y, 1/2)$. Hence, if the left side media firm moves to the left in $(y, 1/2)$, its indirect demand increases more than the increase in indirect demand experienced by the monopolist from the same change in $y$. Hence, it is optimal for the former to choose $\bar{y}$ when we consider $y \leq 1/2$.

  \item Claim 2: $\frac{\partial D_L}{\partial y} (y) \geq \frac{\partial D_L}{\partial y} (y, \bar{y})$ for any $y \in (1/2, \bar{y})$.

  This claim implies that if the left side media firm moves to the right in $(1/2, \bar{y})$, its indirect demand increases less than the increase in indirect demand experienced by the monopolist from the same change in $y$.

As the monopolist’s indirect demand is symmetric around $1/2$ and is maximized when the media firm chooses $y = \bar{y}$ or $y = y$, Claim 1 and 2 imply that it is optimal for the left-side media form to choose $\bar{y}$. (See Figure 6)
Claim 1 is equivalent to \( \frac{\partial x_m}{\partial y_L}(y, \overline{y}) \leq \frac{\partial x_r}{\partial y_L}(y) \) for \( y \in (y, 1/2) \) and Claim 2 is equivalent to \( \frac{\partial x_r}{\partial y_L}(y) \geq \frac{\partial x_m}{\partial y_L}(y, \overline{y}) \) for \( y \in (1/2, \overline{y}) \).

Hence, to prove both claims, it is enough to show \( \frac{\partial x_m}{\partial y_L}(y, \overline{y}) \leq \frac{\partial x_r}{\partial y_L}(y) \), for any \( y \in [y, \overline{y}] \).

Now, we have

\[
\frac{\partial x_m}{\partial y_L}(y, \overline{y}) = \frac{1}{2\mu_x(x_m)}
\]

and

\[
\frac{\partial x_r}{\partial y}(y) = \frac{2(\mu(x_r) - y)}{2\mu_x(x_r)(\mu(x_r) - y) + \sigma_x^2(x_r)},
\]

therefore, \( \frac{\partial x_m}{\partial y_L} \leq \frac{\partial x_r}{\partial y} \) if, and only if,

\[
2 \left( 2\mu_x(x_m) - \mu_x(x_r) \right) (\mu(x_r) - y) \geq \sigma_x^2(x_r).
\]

1. In case we have \( \mu \) convex on \([0, 1/2]\) and \( \sigma^2 \) constant on \([0, 1/2]\):

From \( \mu(x) \) convex on \([0, 1/2]\), and from symmetry, follows that \( \mu(x) \) is concave on \([1/2, 1]\). Therefore, for any \( y \in [y, \overline{y}] \)

\[
\mu_x(x_m(y, \overline{y})) > \mu_x(x_r(y))
\]

which is equivalent to

\[
\frac{1}{\mu_x(x_m(y, \overline{y}))} < \frac{1}{\mu_x(x_r(y))}.
\]

Therefore,

\[
\frac{\partial x_m}{\partial y_L}(y, \overline{y}) = \frac{1}{2\mu_x(x_m(y, \overline{y}))} < \frac{1}{\mu_x(x_m(y, \overline{y}))} < \frac{1}{\mu_x(x_r(y))} = \frac{\partial x_r}{\partial y}(y).
\]

From the above argument, it is clear that \( \partial x_m/\partial y_L \leq \partial x_r/\partial y \) holds when \( \sigma_x^2 < 0 \) on \([0, 1/2]\) (or, equivalently, \( \sigma_x^2 > 0 \) on \([1/2, 1]\)) as well.

Furthermore, as we have some strict gap created by comparison between \( \frac{1}{2\mu_x(x_m)} \) and \( \frac{1}{2\mu_x(x_r)} \), the result carries out even if \( \sigma_x^2 > 0 \) but small on \([0, 1/2]\).

2. In case we have \( \mu \) linear and \( \sigma^2 \) concave on \([0, 1/2]\):

Let \( \mu(x) = mx \), for \( x \in [0, 1/2] \). Hence, \( \mu(x) = 1 - m(1 - x) \), for \( x \in (1/2, 1] \).

In that case, \( \frac{\partial x_m}{\partial y_L}(y, \overline{y}) \leq \frac{\partial x_r}{\partial y}(y) \) requires \( \sigma_x^2(x_r) \leq 2m (\mu(x_r) - y) \), that is, the slope of \( \sigma^2 \) cannot be very negative on \([0, 1/2]\) (equivalently, not very positive on \([1/2, 1]\)).
The cases where (i) $\sigma^2$ is linear and $\mu$ is either linear or concave; and (ii) $\mu$ is linear and $\sigma^2$ is convex, imply that limit polarization is not optimal in the monopoly case. Therefore, those cases are ruled out from our analysis.

**Proposition 10.** Consider duopoly competition in the case in which $F$ and $G$ are the uniform distribution. Under Assumptions 1 and 4,

(i) There is no equilibrium in which both firms choose $y_L = y_R = 1/2$

(ii) There always exists an equilibrium in which each firm chooses the limit polarization $y_L = \underline{y}$ and $y_R = \overline{y}$.

**Proof.** We will prove that for any $y_R \in [1/2, \overline{y}]$ we have that $D_L(y_L, y_R)$ is only maximized when $y_L = \underline{y}$.

Let $x_m(y_L, y_R)$ satisfy $\mu(x_m) = (y_L + y_R)/2$. Therefore,

$$\frac{\partial x_m}{\partial y_L} = \frac{1}{2\mu(x_m)} = \begin{cases} 1 & \text{if } x_m < d \text{ or } x_m > 1 - d \\ 1/2 & \text{if } d \leq x_m \leq 1 - d. \end{cases}$$

1. Consider $y < \underline{y}$. Then, $x_l(y) = 0$ and

$$\frac{\partial D_L}{\partial y_L}(y, y_R) = \frac{\partial x_m}{\partial y_L}(y, y_R) > 0.$$

2. Consider $\underline{y} < y \leq y_R$. Then, $x_l(y) > 0$.

From $\tau < u - d^2/3$ (Assumption 4), follows that $\overline{y} > d$. Therefore, $x_m(y, y_R) \in [y, y_R] \subset [d, 1 - d]$ and

$$\frac{\partial D_L}{\partial y_L}(y, y_R) = \frac{\partial x_m}{\partial y_L}(y, y_R) - \frac{\partial x_l}{\partial y_L}(y) = \begin{cases} < 1/2 - 2 & \text{if } 0 < x_l(y) \leq d \\ 1/2 - 1 & \text{if } d \leq x_l(y) \leq 1 - d, \end{cases}$$

thus $\frac{\partial D_L}{\partial y_L}(y, y_R) < 0$ for any $y \in (y, y_R]$.

Therefore, $y_L = \underline{y}$ is firm’s $L$ the best response to any $y_R \in [1/2, \overline{y}]$. By the symmetry of the problem follows that the unique equilibrium is the one where each firm chooses a limit polarization strategy, that is, $y_L = \underline{y}$ and $y_R = \overline{y}$.

**Proposition 11** (Targeting Strategy). Suppose that the media firm can choose $a$ to target direct consumers belonging to $[a, a + l] \subset [0, 1]$ with $l \in [0, 1)$ in addition to choosing the news location $y$. 

47
1. Whenever limit polarization is optimal to the no targeting version of the problem, it is a weakly dominant strategy to choose \( y \) equal to the left limit polarization location and \( a = 0 \);

2. Whenever no polarization is optimal to the no targeting version of the problem, it is a weakly dominant strategy to choose \( y = 1/2 \) and \( a = (1 - l)/2 \).

Proof.

1. By choosing \( y = y \) and \( a = 0 \), if \( l < x_r(y) \) then all targeted consumers share the news and, obviously, one cannot do better than that; if, otherwise, \( l \geq x_r(y) \) then all direct consumers on \([0, x_r(y)]\) share the news. Hence, we have as much news sharing as in the no-targeting case and again we reach an upper bound.

2. By choosing \( y = 1/2 \) and \( a = (1 - l)/2 \), if \( a > x_l(1/2) \) then \( a + l < x_r(1/2) \) and it is optimum as all targeted consumers share the news; if, otherwise, \( a < x_l(1/2) \) then \( a + l > x_r(1/2) \) and all direct consumers on \([x_l(1/2), x_r(1/2)]\) share the news. Once more it is optimal because in this case we have as much news sharing as in the no-targeting case.

\[ \square \]

**Proposition 12** (Depth Maximizing Strategy). Suppose that the attention tax is not small. The media firm’s optimal strategy to maximize the depth of news sharing is characterized as follows:

1. It targets the consumer located at \( x_{depth}^* \) whose \( \sigma^2(x) \) is equal to \( \min_{x \in [0, 1/2]} \sigma^2(x) \).

2. It chooses the location of the news \( y_{depth}^* = \mu(x_{depth}^*) \).

Proof. Let \( t^* \) denote the depth of the news sharing of the news located at \( y \). Since, by assumption, a consumer located at \( x \) reshares the news located at \( y \) whenever both \( U_{t(y)}(x, y) \geq 0 \) and \( B_{t^*}(x, y) \geq \tau \). If the attention tax is not small the latter inequality is the first to bind. Therefore,

\[
B_{t^*}(x, y) = \int_0^1 U_{t^*}(z, y) \tilde{g}(z; x) \, dz \\
= \delta^* u - \sigma^2(x) - (y - \mu(x))^2.
\]
And from $B_{t(y)}(x, y) = \tau$ follows that

$$
\delta^{t^*} u - \tau = \sigma^2(x) + (y - \mu(x))^2.
$$

As the LHS of the equation above is decreasing in $t^*$, the depth is maximized when the RHS is minimized. The result follows.

**Corollary 1.** Suppose that the attention tax is not small.

1. The depth-maximization leads to some polarization as long as $\sigma^2(1/2) > \min_{x \in [0, 1/2]} \sigma^2(x)$;
2. If $\sigma^2(x)$ is increasing, it leads to $x_{\text{depth}}^* = 0$.

**Proof.** It follows immediately from from the application of the results of Proposition 12.