

# Political Polarization vs. Fact-checking: Emergence of impudent lies in social media

Samuel Solgon Santos\*

Marcelo de Carvalho Griebeler†

Federal University of Rio Grande do Sul

Federal University of Rio Grande do Sul

19/07

## Resumo

Difícilmente se poderá negar que candidatos(as) à cargos eletivos eventualmente decidem mentir. O que explica a decisão de políticos de mentir de forma descarada? Neste trabalho mostramos que, uma vez que tenham decidido mentir, os candidatos podem optar por um nível de descaramento maior ou menor para a mentira, a depender da polarização do eleitorado e do custo de produção de mentira. Estudamos o problema de um candidato que deve escolher o nível de descaramento de uma mentira que será disseminada por i) *eleitores*, ii) *militantes partidários* e iii) *BOTs* em uma rede social. Estes agentes são incorporados no modelo de difusão proposto por Bass (1969). Em particular, nós propomos uma relação microfundamentada entre as decisões dos eleitores de compartilhamento (ou não compartilhamento) de conteúdo na rede social, com o modelo de disseminação em rede e, por conseguinte, com os incentivos do candidato a produzir mentiras mais (ou menos) descaradas. Concretamente, assumimos que o candidato conhece o nível de polarização do eleitorado e o processo de disseminação e, com base nestas informações, escolherá o grau de descaramento da mentira que maximiza a disseminação até a data da eleição. Este modelo tem a virtude de relacionar os principais elementos que comumente são relacionados à produção e disseminação de mentiras nas redes sociais. O principal resultado do modelo é que o aumento da polarização do eleitorado gera incentivos para a produção de mentiras mais descaradas, e este resultado está de acordo com a observação

---

\*E-mail: samuel.solgon@ufrgs.br. This research was supported in part by Coordination for the Improvement of Higher Education Personnel (CAPES).

†Address: Universidade Federal do Rio Grande do Sul – UFRGS, Faculdade de Ciências Econômicas, Departamento Economia e Relações Internacionais, Avenida João Pessoa 52, Centro Histórico, Porto Alegre – RS, Brazil, ZIP code: 90040-000 – Phone: 55 (51) 3308-3324. E-mail: marcelo.griebeler@ufrgs.br.

de que o recente aumento da polarização do eleitorado foi acompanhado pela disseminação de mentiras descaradas, em particular mentiras sobre temas em política.

**Palavras-chave:** Redes Sociais. Bass model. Polarização do Eleitorado. Militantes Partidários. BOTs.

## Abstract

What explains the increase in the prevalence of impudent lies among politicians? Suppose a candidate to an elective office has decided to lie through a post on a social media website. After having decided to lie, the candidate must choose the lie's level of impudence. We assume the candidate aims to maximize the share of electors who will make contact with the lie until the election date. We also assume that the spread of the lie follows an adaptation of the Bass diffusion model whereupon we incorporate the actions of *i) electors*, *ii) candidate's militants* and *iii) BOTs*. In particular, we incorporate the elector's rational behavior and the aggregate level of polarization of the electorate to the model. In particular, we assume that electors choose rationally whether to share or not share the candidate's message. Furthermore, we aggregate these decisions, in such a way to link elector's individual decisions to the diffusion of the lie, and consequently, to the candidate's decision regarding the level of the impudence of the lie. The main contribution of this model is to link the actors that are often pointed out as the main factors behind the dissemination of lies in social media websites. Additionally, we show that the dissemination of the lie depends fundamentally on the polarization level of the electorate, and that the dissemination is greater when the electorate is more polarized. Furthermore, we show that the share of electors who accepts to share the lie is greater when the polarization is higher, and that the candidate chooses accordingly deciding to send a more impudent lie.

**Keywords:** Social Media. Bass model. Political Polarization. Political Militants. BOTs.

**JEL classification:** D72, D01, F50

**Área da ANPEC:** Microeconomia, Métodos Quantitativos e Finanças.

## 1 Introdução

A produção e disseminação de mentiras com objetivos político-partidários não é um fenômeno recente, entretanto, as mentiras que foram disseminadas durante as eleições americanas de 2016 chamaram atenção pelo seu alto grau de descaramento. Notícias como “*Pope Francis Shocks World, Endorses Donald Trump for President*” e “*ISIS Leader Calls for American Muslim Voters to Support Hillary Clinton*” são histórias inventadas,

descaradamente falsas e que foram compartilhadas por diversos usuários de redes sociais durante os meses anteriores à eleição americana de 2016.

Histórias falsas podem ser publicadas por diferentes agentes e com diferentes objetivos. Adolescentes da Macedônia, por exemplo, perceberam que poderiam atrair milhares de internautas para os seus sites – e assim aumentar sua receita com publicidade – postando histórias inventadas em grupos políticos do Facebook (Subramanian, 2017). Por outro lado, a organização russa Internet Research Agency foi acusada de produzir e disseminar mentiras com o objetivo de interferir nos resultados das eleições americanas de 2016 (of America, 2018).

Não obstante, há evidências de que candidatos à presidência também podem publicar mentiras descaradas. A agência de checagem de notícias Politifact, por exemplo, classificou 34% do total das afirmações de Donald Trump analisadas como “falsas” e, 15% (do total) foram classificadas como imprecisas e ridículas (Politifact, 2019)<sup>1</sup>.

Nosso modelo busca ilustrar os incentivos de candidatos à cargos eletivos (presidência, por exemplo) a mentir de forma descarada. O modelo que propomos não endogeniza a decisão do candidato de mentir, ao contrário, parte do pressuposto que o candidato já decidiu mentir e deve escolher o nível de descaramento da mentira que publicará com o objetivo de maximizar a sua disseminação até a data da eleição.

Assumimos que a disseminação da mentira nas redes sociais depende da atuação de três tipos de agentes que denominamos eleitores, militantes e BOTs. Nós assumimos que, em geral, os eleitores são bem intencionados e evitam compartilhar mentiras, de forma que decidem se compartilham a mensagem do candidato com base na probabilidade com que acreditam que a mensagem é verdadeira. Tendo conhecimento do comportamento “bem-intencionado” dos eleitores, o candidato deve considerar de que forma o grau de descaramento da mentira irá afetar a decisão de compartilhamento dos eleitores.

Note que mentiras descaradas (tais como as citadas no início desta Introdução, por exemplo) poderiam ser facilmente desmascaradas com uma rápida pesquisa *online*, assim, assumimos que quanto maior o nível de descaramento da mentira, menor será o seu custo de checagem e, portanto, mais eleitores estarão dispostos a checa-la. Ainda que mais eleitores desconfiem de mentiras descaradas – e, portanto, menos eleitores aceitem compartilhá-las –, mostramos que a polarização do eleitorado pode induzir o candidato a escolher mentir mais descaradamente. Ressalta-se que no modelo que propomos, os eleitores são penalizados por compartilhar mentiras, de tal forma que parece ser fácil concluir que o nosso resultado se manterá ao incorporarmos a possibilidade de que os eleitores afirmem prazer em compartilhar mentiras.

Observações empíricas exigem que inclua-se BOTs em um modelo de disseminação de mentiras em redes sociais. Varol et al. (2017), por exemplo, estimam que entre 9 e

---

<sup>1</sup>O site checa novas declarações periodicamente, de forma que as porcentagens apresentadas podem variar.

15% das contas ativas do Twitter são BOTs, enquanto o Facebook estima que cerca de 60 milhões de BOTs podem estar infectando a plataforma ([Watts, 2017](#)).

Além disso, consideramos desejável considerar a existência de militantes partidários que, diferente dos eleitores, sempre compartilham as mensagens do candidato. A única especificação que assumiremos em relação ao comportamento dos BOTs e militantes partidários é que ambos recebem as mensagens diretamente dos candidatos<sup>2</sup> e ambos compartilham estas mensagens, independente do nível de descaramento. Estas semelhanças nos levam a tratar BOTs e militantes partidários de forma indistinta e, tendo assim prosseguido, mostramos que a influência de ambos na disseminação da mentira é maior para períodos mais próximos à data da eleição. Adicionalmente, mostramos que, para um dado nível de descaramento da mentira, a atuação de BOTs e militantes pode incentivar eleitores a aceitarem compartilhar a mentira. De forma geral, a atuação de militantes e BOTs acelera o processo de disseminação<sup>3</sup> e é especialmente importante nos instantes iniciais da propagação<sup>4</sup>.

## 2 Revisão de Literatura

Este trabalho relaciona-se com algumas correntes literária-científicas que, em geral, não se comunicam: a teoria microeconômica de firmas produtoras de notícias, a literatura de modelos de difusão de inovações e a literatura de modelos de aprendizagem social e “comportamento de rebanho”.

Parcela importante da literatura sobre firmas produtoras de notícias buscam explicar a produção de notícias viesadas a partir da hipótese de que consumidores sofrem de viés de seleção e que, assim, buscam consumir notícias de fontes alinhadas às suas preferências ideológicas<sup>5</sup>. [Gentzkow and Shapiro \(2006\)](#), por exemplo, consideram que firmas podem enviar consistentemente suas notícias com o objetivo de tornarem-se conhecidas por certos grupos de indivíduos que buscam consumir informações adequadas às suas predisposições ideológicas. Ainda, o viés de seleção dos indivíduos pode levar a resultados eleitorais ineficientes, tais como a eleição de políticos mais corruptos ([Bernhardt et al., 2008](#)), ou mesmo influenciar a escolha de *policies* pelos governantes ([Eisensee and Strömberg, 2007](#)).

Há diversas ferramentas matemáticas que são utilizadas para estudar disseminações em rede. Desde a contribuição seminal de [Kermack and McKendrick \(1991\)](#) para a teoria

---

<sup>2</sup>Pois, por hipótese, BOTs e militantes partidários seguem a página do candidato na rede social.

<sup>3</sup>Tal como observado por [Lazer et al. \(2018\)](#); [Shao et al. \(2018\)](#).

<sup>4</sup>Tal como observado por [Shao et al. \(2018\)](#).

<sup>5</sup>Além dos trabalhos que escolhemos citar no corpo do texto, sugerimos ao leitor interessado na teoria microeconômica por trás da produção de notícias viesadas os trabalhos de [Mullainathan and Shleifer \(2005\)](#), [Suen \(2004\)](#), [Duggan and Martinelli \(2011\)](#) e o livro organizado por [Anderson et al. \(2016\)](#) que conta, inclusive, com capítulos dedicados às mídias sociais e ao mercado de notícias na internet.

matemática das epidemias de doenças contagiosas, a literatura de modelos epidêmicos passou a ser frequentemente citada como um ferramental natural para estudar a disseminação de informações em uma sociedade composta por agentes conectados (Jackson, 2010). Ainda, os trabalhos de Daley and Kendall (1964) e Maki and Thompson (1973) são referências clássicas de modelos que buscam descrever a disseminação de rumores entre agentes interconectados. Entretanto, foi na literatura de modelos de disseminação de inovações entre firmas que encontramos a abordagem mais adequada aos nossos objetivos.

O modelo de disseminação que utilizamos foi desenvolvido por Bass (1969) com o objetivo de prever a dinâmica de disseminação de uma inovação. O modelo proposto pelo autor, juntamente com os trabalhos de Fournier and Woodlock (1960) e Mansfield (1961) são os modelos de difusão de inovações mais conhecidos (Mahajan et al., 1990). Diferente de modelos anteriores que previam o crescimento exponencial do número de usuários da inovação (Fournier and Woodlock, 1960; Haines Jr, 1964), o modelo proposto por Bass prevê que a dinâmica da difusão da inovação segue um padrão tal como ilustrado pela Figura 4. De fato, em Bass (1969) são captadas características essenciais do processo de imitação, onde firmas conectadas por algum laço social/econômico tomam conhecimento de uma inovação por meio de outras firmas que já a adotaram. Segundo o autor, firmas podem obter conhecimento sobre a inovação consultando uma fonte exógena (tal como um jornal ou revista especializada) ou pelo contato com alguma outra firma que já tenha adotado a inovação. A interpretação para a nossa aplicação é análoga. Os partidos agem como a fonte de informação exógena pela qual alguns usuários com viés de seleção (militantes) terão contato com a notícia. A partir do compartilhamento da publicação do partido por estes usuários iniciais, outros indivíduos passam a ter contato com a mentira – ainda que inicialmente não estivessem dispostos a receber informações do partido. O modelo que propomos permite avaliar o efeito de diversos fatores sobre a dinâmica de disseminação da mentira do partido, tal como o nível de descaramento da mentira, o nível de polarização do eleitorado e a força de atuação de BOTs e militantes.

Por fim, nosso modelo se relaciona com a literatura de aprendizado social. Um problema importante nesta literatura é explicar a emergência de “comportamento de rebanho” em contextos onde indivíduos racionais tomam decisões sequencialmente observando as escolhas que foram tomadas anteriormente, tal como no seminal trabalho de Banerjee (1992). A ideia básica por trás destes modelos é que se os agentes não têm certeza sobre qual a melhor ação, então pode ser racional levar em conta decisões tomadas anteriormente com base em conjuntos de informação privada diferentes. Este cenário teórico está intimamente relacionado à propagação de conteúdo em redes sociais onde os usuários observam a quantidade de compartilhamentos anteriores antes de decidir se (re)compartilham um conteúdo ou não. Ainda, o fenômeno da disseminação de mentiras nas redes sociais encontra uma explicação teórica fundamental em Bikhchandani et al. (1992). Em um contexto de decisão sequencial, os autores mostram que, em algum estágio do processo, os agentes

passam a desprezar a informação privada que possuem e agem de acordo com as decisões tomadas anteriormente, gerando o que os autores denominam “*Informational Cascades*”. Ainda, o processo de decisão individual que adotamos neste trabalho foi inicialmente proposto por Papanastasiou (2018) que, assim como o trabalho de Bass (1969), serviu de ponto de partida para o desenvolvimento do modelo que apresentaremos.

### 3 Modelo

Nosso modelo básico é composto por uma massa de eleitores, por um candidato a cargo eletivo (presidência, por exemplo) e por BOTs e militantes do candidato (ou do partido do candidato). O candidato divulga uma mentira em uma rede social durante a campanha para algum cargo eletivo. Tal mensagem pode ser publicada na sua página do Facebook, ou divulgada como uma notícia ou informação no Twitter, por exemplo. Os militantes acompanham as publicações do candidato e, tendo a ela acesso imediato, contribuem com sua disseminação entre os eleitores através do compartilhamento. A função dos BOTs no processo de disseminação é similar à dos militantes e será discutida com mais detalhes na seção 3.3. Os eleitores que tiverem contato com a publicação através do compartilhamento decidirão entre (re)compartilha-la ou não com os eleitores que ainda não tiveram contato. Observe que, enquanto os militantes obtém a informação diretamente da fonte (da página do partido na rede social), os eleitores a recebem via compartilhamento. Além disso, os militantes sempre as compartilham, enquanto os eleitores decidem se o farão.

Nas seções que seguiremos descreveremos formalmente os agentes que compõem o modelo e mostraremos como estes agentes determinam o processo de disseminação da mentira e influenciam o grau de descaramento escolhido pelo candidato.

#### 3.1 Os eleitores

Consideramos uma massa de eleitores que utilizam a rede social (ex.: Facebook) para se informar. Os eleitores são indexados por  $i \in \mathcal{I}$  e não tem certeza quanto a veracidade da publicação do candidato. Cada eleitor  $i$  é inteiramente caracterizado pela probabilidade  $\text{Prob}_i(\cdot)$  que define no espaço amostral  $\{\text{Mentira } (M), \text{Verdade } (V)\}$ . A medida  $\text{Prob}_i(\cdot)$  é definida *ex ante* o contato com a publicação, de forma que  $\text{Prob}_i(V)$  representa a probabilidade com que o indivíduo  $i$  acredita que uma publicação qualquer do candidato seja verdadeira. Denotaremos  $b_i := \text{Prob}_i(V)$  e interpretaremos esta quantidade como a credibilidade que o eleitor  $i$  atribui ao candidato. Os eleitores são heterogêneos na credibilidade que atribuem ao candidato, e denotaremos a distribuição  $b_i$  por  $F(b_i)$  e sua densidade por  $f(b_i)$ . Assumimos que ambas sejam diferenciáveis em todo o domínio e na ordem em que se fizer necessário.

Estamos interessados em modelar o comportamento de eleitores quando estes têm contato com a publicação do candidato por meio do compartilhamento de algum de seus “amigos” da rede social. Ao ter contato com a publicação, os eleitores decidem em dois estágios: no primeiro estágio decidem se checam o conteúdo da publicação e, no segundo estágio, decidem se a compartilham ou não. Pressupomos que os eleitores são “bem-intencionados”, no sentido de que sempre buscam compartilhar informações verdadeiras e nunca compartilhar informações falsas. A ideia é que a recompensa psicológica destes eleitores para cada possível ação depende, exclusivamente, da probabilidade com que acreditam que a ação escolhida é benéfica para a sociedade. Como os eleitores não têm certeza quanto a veracidade da publicação, eles podem decidir checa-la ou não no primeiro estágio. Assumimos que a ação de checar sempre revela o verdadeiro “estado de mundo”, qual seja  $M$ . Além disso, independente da decisão tomada no primeiro estágio, os eleitores são chamados a decidir, no segundo estágio, se irão ou não compartilhar a publicação.

A hipótese de boa intenção é capturada pela seguinte função utilidade, que pressupomos ser compartilhada por todos os eleitores

$$u(s; V) = u(n; M) = 1 \quad \text{e} \quad u(s; M) = u(n; V) = 0 \quad (3.1)$$

onde “ $s$ ” representa a ação de compartilhar e “ $n$ ” representa a ação de não compartilhar. Note que a especificação (3.1) é tal que o indivíduo sempre ganha o *pay-off* máximo de 1 quando compartilha uma mensagem verdadeira ou quando não compartilha uma mensagem mentirosa. Dado que a mensagem do candidato é mentirosa por hipótese, os eleitores que decidirem checar a publicação certamente irão decidir não compartilhá-la. Ainda, como estes eleitores têm certeza de que estão agindo em benefício da sociedade irão auferir a recompensa psicológica máxima. A checagem do conteúdo, entretanto, implica em um custo de tempo e de esforço investigativo que representaremos por  $K \in \mathbb{R}_+$ . Assumimos que o custo de checagem de uma mesma mensagem é homogêneo entre os eleitores, de forma que o *pay-off* daqueles que decidirem checar a notícia no primeiro período será de  $1 - K$ .

Por outro lado, aqueles eleitores que não checarem a publicação não terão certeza sobre a veracidade do conteúdo e deverão escolher entre compartilhar ou não com base na credibilidade que atribuem ao candidato. Segundo a especificação dada em (1), temos que  $u_E(s) = b_i$  e  $u_E(n) = 1 - b_i$ , onde  $u_E(\cdot)$  denota a utilidade esperada do indivíduo  $i$  para todo  $i \in \mathcal{I}$ .

### 3.2 A escolha dos eleitores

Dada a simplicidade do espaço de escolhas e da função utilidade dos eleitores, podemos deduzir condições suficientes que os levam a escolher cada uma de suas ações possíveis sem precisar derivar condições de primeira ordem.

A condição necessária e suficiente para que o indivíduo decida checar a notícia é que  $1 - K > \max\{b_i, 1 - b_i\}$ <sup>6</sup>. Agora, note que se o custo de checagem  $K$  for maior do que  $0,5$ , a recompensa por checar a mensagem e agir com certeza em favor da sociedade,  $1 - K$ , será sempre menor do que  $\max\{b_i, 1 - b_i\}$  para todos os eleitores, de tal forma que nenhum eleitor irá checar a mensagem. Por outro lado, se  $K = 0$ , então  $1 - K > \max\{b_i, 1 - b_i\}$  com probabilidade 1, de tal forma que a massa de eleitores que irão compartilhar a mensagem é igual à 0<sup>7</sup>. Para evitar estes casos, limitaremos  $K$  ao intervalo  $(0, 0,5)$ . Isto implica que para uma mentira com custo de checagem  $K$ , apenas os eleitores mais partidários, com  $b_i \geq 1 - K$  irão compartilhar a mentira sem checa-la<sup>8</sup>.

O eleitor escolherá não checar e compartilhar a publicação se, e somente se,  $b_i \geq \max\{1 - K, 1 - b_i\}$ <sup>9</sup>. Ainda, note que se  $b_i \geq 1 - b_i$  então  $b_i \geq 0,5$ , assim, no que diz respeito a parcela dos eleitores que irão compartilhar a publicação do partido (sem checar), podemos limitar nossa atenção a parcela de indivíduos à direita de  $0,5$  na Figura 1. Não obstante, vale notar que  $b_i \geq 0,5$  não é uma condição suficiente para o compartilhamento, haja visto que, tal como ilustrado pela Figura 2, é possível que  $0,5 \leq b_i < 1 - K$ , caso em que o eleitor decide checar a publicação. Por outro lado, como  $K \in (0, 0,5)$ , temos  $1 - K \in (0,5, 1)$  de tal forma que se  $b_i > 1 - K$ , então automaticamente  $b_i > 1 - b_i$ . Vamos pressupor que se  $b_i = 0,5$  o indivíduo escolhe compartilhar a notícia sem verificar e, assim, concluimos que  $b_i \geq 1 - K$  é uma condição suficiente para o compartilhamento.

A parcela de indivíduos que, caso recebam a publicação a irão compartilhar sem checar é dada por  $q(K; a) := 1 - F(1 - K)$ . Ainda, para qualquer nível de  $K \in (0, 0,5)$ , o tamanho da parcela  $q(K; a)$  dependerá do nível de polarização do eleitorado, tal como ilustrado pela Figura 2. Para introduzir a noção de polarização do eleitorado pressupomos a seguinte especificação para a distribuição de  $b_i$  no intervalo  $[0, 1]$

$$F(b_i; a) = \frac{ab_i^3}{3} - \frac{ab_i^2}{2} + b_i \left(1 + \frac{a}{6}\right), \quad \text{onde } a \in [0, 12] \quad (3.2)$$

a partir da qual derivamos a conveniente função densidade dada por

$$f(b_i, a) = ab_i^2 - ab_i + 1 + a/6 \quad (3.3)$$

e cujos gráficos para diferentes valores do parâmetro  $a$  são ilustrados na Figura (1)

A convexidade da curva  $f(b_i; a)$  é controlada pelo parâmetro  $a$ , ao qual doravante nos

---

<sup>6</sup>Estamos pressupondo que, em caso de empate, a ação de checar é sempre dominada e, portanto, o eleitor só irá checar se a desigualdade for estrita.

<sup>7</sup>Isto será verdade para a função de distribuição  $F(\cdot)$  que iremos adotar.

<sup>8</sup>De fato, há evidências de que se uma mentira é definitivamente boa para um partido, então a mesma tende a ser creditada e compartilhada mais facilmente por apoiadores do partido (Garrett and Weeks, 2013; Shin and Thorson, 2017). Meirick (2013), por exemplo, mostrou que indivíduos que se identificaram como “Republicanos” mostraram-se mais suscetíveis a uma mentira publicada no Facebook pela ex-candidata republicana à vice-presidência Americana, senhora Sarah Palin.

<sup>9</sup>Estamos assumindo que em caso de empate a ação de “não checar e compartilhar” domina qualquer outra ação.



Figure 1: Densidades de  $b_i$  para diferentes valores de  $a$

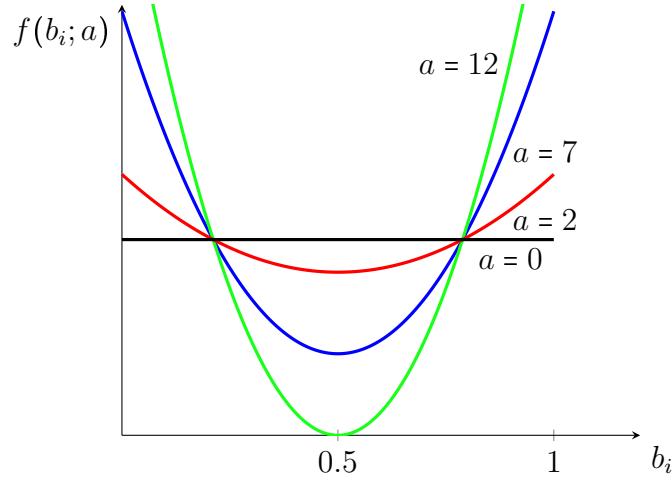
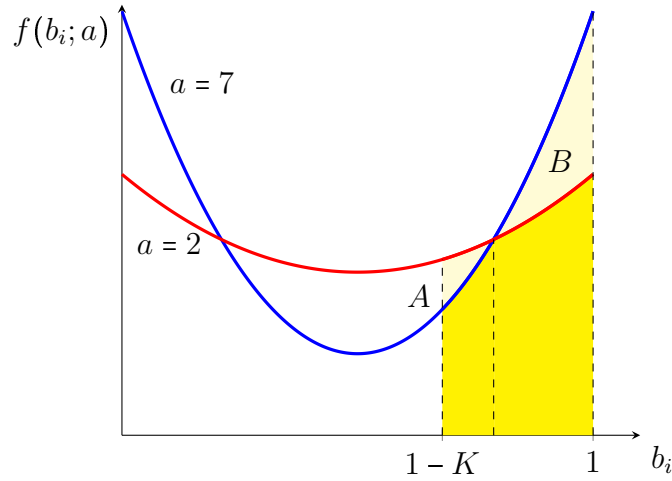


Figure 2: Massa de eleitores que compartilham a publicação



referiremos como “nível de polarização”. Como ilustrado na Figura (1), a convexidade de  $f(b_i; a)$  é crescente no nível de polarização, com  $f(b_i; a = 0)$  correspondendo à distribuição uniforme e o caso  $f(b_i; a = 12)$  representando o nível de polarização máxima do eleitorado.

A estratégia de propor uma forma funcional específica para  $f(b_i; a)$  carrega conveniências importantes, entretanto alguns ressaltos devem ser feitos. A primeira vantagem é que  $f(b_i; a)$  capta a ideia de que um eleitorado mais polarizado é composto por mais indivíduos com crenças extremas e que, eventualmente irão compartilhar a mentira publicada sem verificá-la. Esta ideia é ilustrada pelo fato de que, para todo  $K \in (0, 0,5)$ , a área à direita de  $1-K$  e abaixo da curva  $f(b_i; a)$ , é crescente no parâmetro de polarização  $a$ <sup>10</sup>. Além disso, o gráfico da densidade  $f(b_i; a)$  que propomos facilita a visualização do mecanismo pelo qual diferentes valores de  $K$  - que será a variável de escolha do can-

<sup>10</sup>Este fato pode ser verificado bastando mostrar que a área  $B$  é maior do que a área  $A$  ilustradas na Figura (2).

didato - implicam em diferentes parcelas  $q(K; a)$  de eleitores que, ao serem expostos à publicação, decidem compartilhá-la sem verificar. A ressalva que deve ser levantada é que, dado o nosso atual desconhecimento empírico sobre o formato da distribuição real de  $b_i$ , a estratégia mais segura seria propor uma  $f(b_i, a)$  tão geral quanto possível, de tal forma que qualquer que fosse a distribuição real de  $b_i$ , a mesma estaria contemplada pela teoria. De fato, resultados obtidos sob hipóteses mais fracas são, na verdade, mais fortes e, sendo assim, um possível desenvolvimento posterior deste trabalho é investigar se os resultados aqui obtidos se mantêm sob especificações mais gerais para a distribuição de  $b_i$ .

### 3.3 A disseminação da mentira e os militantes partidários

Queremos representar a disseminação da publicação entre indivíduos conectados em uma rede. Para tanto iremos considerar apenas determinantes “globais” do processo de disseminação, não levando em consideração propriedades estruturais / topológicas da rede. O modelo de disseminação que adotaremos foi inicialmente proposto por Bass (1969) e é baseado na dinâmica intertemporal de  $G(t) :=$  parcela de eleitores que tiveram contato com a publicação do partido até  $t$ .

Segundo a interpretação de Lekvall and Wahlbin (1973), a parcela  $G(t)$  varia por influência de fatores “internos” e “externos”. A influência interna representa a atuação de eleitores que, ao terem contato com a publicação decidem compartilhá-la. A influência externa, por sua vez, representa a atuação de BOTs e militantes partidários que recebem a publicação diretamente do candidato, e recebendo-a, compartilham-na imediatamente atingindo novos eleitores.

A distinção básica entre a influência interna e externa é que a primeira representa a disseminação “boca-a-boca”, enquanto a segunda representa a disseminação via BOTs e militantes que independem da intermediação de outros agentes para terem contato com a publicação.

Formalmente, as influências internas e externas definem a variação de  $G(t)$  da seguinte forma:

$$\frac{dG(t)}{dt} = q(K; a)G(t)(1 - G(t)) + p(1 - G(t)) \quad p > 0 \quad (3.4)$$

O termo  $q(K; a)G(t)(1 - G(t))$  é a influência interna. Representa a atuação de eleitores no processo de disseminação e relaciona o processo de disseminação com o comportamento dos eleitores. A importância de  $q(K; a)$  no processo de disseminação é imediata: quanto maior a parcela de eleitores que aceitam compartilhar a mentira sem checa-la maior será a taxa de disseminação dada por  $dG(t)/dt$ . O termo  $G(t)$  por sua vez, é o principal meio pelo qual captamos a disseminação “boca-a-boca” no modelo. A interpretação é de

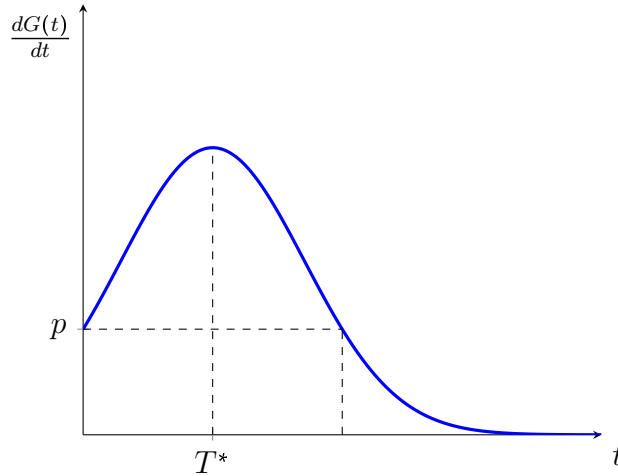
que quanto maior a parcela de eleitores que já tiveram contato com a publicação, mais indivíduos podem disseminá-la e, conseqüentemente, maior a probabilidade de um novo contato em  $t$ . Por último, o termo  $1 - G(t)$  incorpora o fato de que para que  $dG(t)/dt > 0$  é necessário que ainda hajam indivíduos que não tiveram contato com a publicação até  $t$ .

O termo  $p(1 - G(t))$ , por sua vez, é a influência externa e representa a atuação de militantes partidários e BOTs no processo de disseminação. A influência destes agentes ainda depende da existência de indivíduos que não tiveram contato com a notícia  $1 - G(t)$ , entretanto, a hipótese é que BOTs e militantes partidários “seguem” a página do candidato na rede social e, portanto, independem da intermediação de outros indivíduos para terem contato com a publicação<sup>11</sup>.

A equação (3.4) juntamente com a hipótese  $G(0) = 0$  tem solução bem conhecida dada por

$$G(t) = \frac{1 - e^{-(p+q)t}}{1 + (q/p)e^{-(p+q)t}}, \quad \text{onde } q := q(K; a) \quad (3.5)$$

Figure 3: Função  $g(t) = \frac{dG(t)}{dt}$

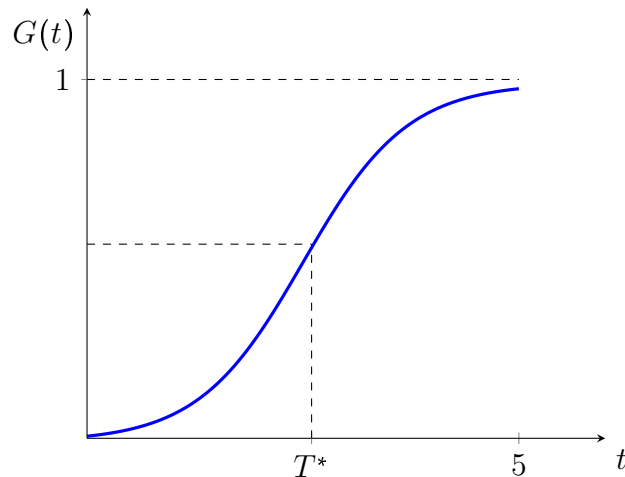


Por (3.4) verificamos que a condição inicial  $G(0) = 0$  implica que a taxa  $dG(0)/dt$  é exclusivamente determinada pela participação dos militantes. Este fato é ilustrado pelo intercepto vertical do gráfico da Figura 3. Na medida em que a publicação se dissemina, a participação dos eleitores aumenta e sobrecompensa a queda da participação dos militantes, de forma que a taxa de disseminação acelera até atingir seu máximo no período  $T^*$ . A partir de então, a disseminação segue à taxas decrescentes até que  $\lim_{t \rightarrow \infty} G(t) = 1$ , tal como ilustrado pela Figura 4<sup>12</sup>.

<sup>11</sup>A atuação de BOTS e militantes nos parece igualmente bem modeladas pelo termo de influência externa, entretanto, nosso modelo não é sofisticado o suficiente para distingui-las. No decorrer do texto interpretaremos o termo de influência externa como captando a atuação de militantes.

<sup>12</sup>Vale reforçar que a previsão do modelo é de que todos os eleitores terão contato com a publicação,

Figure 4: Dinâmica da disseminação



Como detalharemos na próxima subseção, estamos particularmente interessados em estudar a parcela de eleitores que tiveram contato com a publicação até a data da eleição. Para tanto, ressaltamos que a dinâmica da disseminação é inteiramente determinada pelos parâmetros  $p$  e  $q$ , e pressupomos a importante hipótese de que

**Hipótese 1** *Os parâmetros  $p$  e  $q$  são tais que a taxa de disseminação é máxima na data da votação.*

A Hipótese 1 é motivada pela observação de que o interesse dos eleitores em relação aos partidos e candidatos parece atingir seu pico durante o dia da votação. Veja por exemplo os gráficos da Figura 5<sup>13</sup> extraídos do Google Trends que mostram que, na última eleição presidencial do Brasil e dos Estados Unidos houveram picos de interesse pelos principais candidatos durante o dia de votação<sup>14</sup>.

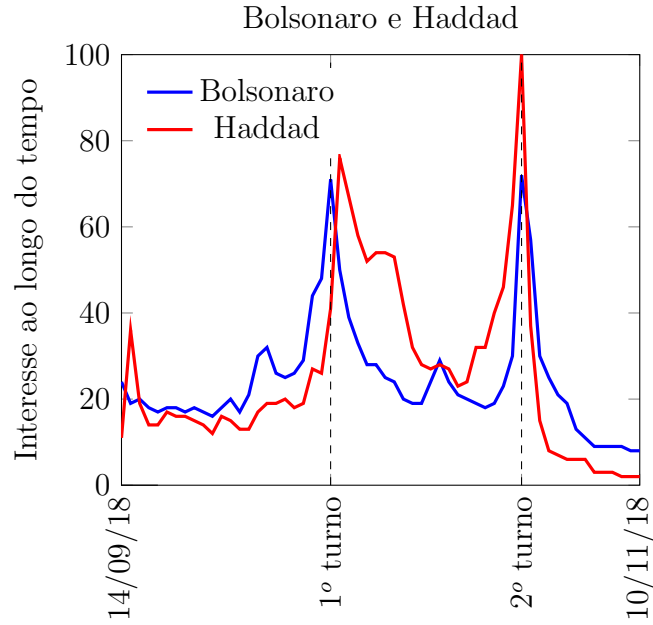
A função da Hipótese 1 é pressupor que a taxa de disseminação da publicação do partido é fortemente correlacionada ao interesse dos eleitores pelos candidatos (ilustrado pela Figura 5), de tal forma que a taxa de disseminação também irá atingir seu máximo na data da votação. Os valores de  $q$  e  $p$  definem a dinâmica de disseminação e, em particular, determinam quanto tempo levará para que a taxa de disseminação da publicação, representada pela Figura 3, atinja seu ponto de máximo. Estamos assumindo que o partido emite a publicação no instante  $t = 0$ , assim, se  $T^* \in \mathbb{R}$  é o intervalo de

e não de que a publicação será compartilhada por todos.

<sup>13</sup>Os eixos verticais dos gráficos da Figura 5 representam o interesse de pesquisa relativo ao ponto mais alto no gráfico em um dado período e em uma região específica (no caso da eleição brasileira, o Brasil e no caso da eleição americana, os Estados Unidos). Um valor de 100 representa o pico de popularidade de um termo. Um valor de 50 significa que o termo teve metade da popularidade. Uma pontuação de 0 significa que não havia dados suficientes sobre o termo.

<sup>14</sup>Ainda que o pico de interesse não tenha ocorrido exatamente no dia da votação no caso da eleição americana, é inegável que a data da votação é o fator por trás do crescimento abrupto do interesse dos internautas sobre os candidatos Donald Trump e Hillary Clinton.

Figure 5: Google trends



tempo entre a publicação e a data de votação, podemos simplesmente dizer que a  $T^*$  é a data da votação. Logo, nossa hipótese implica que  $q$  e  $p$  têm que ser tais que a taxa de disseminação seja máxima no período  $t = T^*$ .

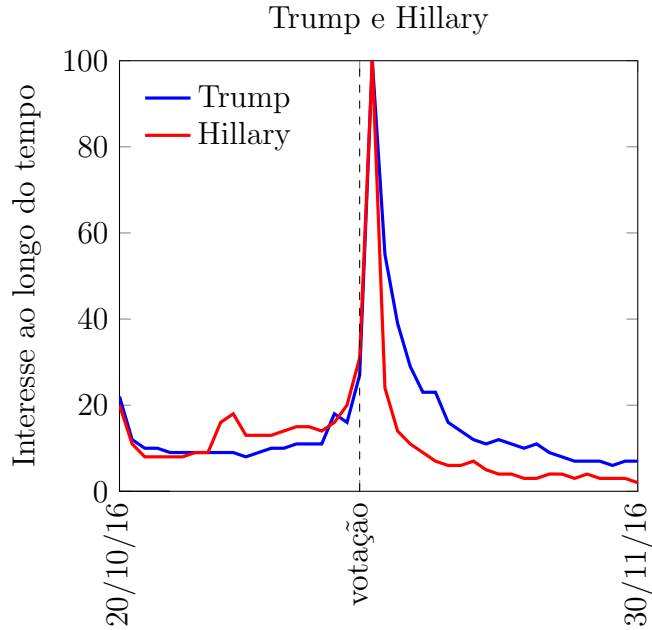
Igualando a derivada da função  $dG/dt$  à zero concluímos que a quantidade de tempo necessário para que, a partir da data emissão, a publicação atinja sua taxa de disseminação máxima é dado por  $\ln(q/p)/(p+q)$ . Assim, nossa hipótese implica que, sendo  $T^*$  o intervalo de tempo entre a publicação do partido e a data da eleição, os parâmetros  $p$  e  $q$  devem ser tais que  $\ln(q/p)/(p+q) = T^*$ .

**Proposição 1** *Com relação à atuação dos militantes, temos:*

a) *Se a distribuição de  $b_i$  for constante em relação ao tempo, então a atuação dos militantes é positivamente relacionada com a proximidade das datas de publicação da mensagem e de votação.*

b) *Se  $q(K; a) < 1/T^*$ , então aumentos da atuação dos militantes implicam no aumento da atuação dos eleitores.*

As proposições acima derivam da restrição de que a taxa de disseminação deve ser máxima na data da votação e, além disso, caracterizam o comportamento de militantes e eleitores em função da proximidade da data da publicação com a data de votação. Ambas foram obtidas a partir da restrição  $\frac{1}{(p+q)} \ln(q/p) = T^*$  e a validade de cada uma depende das hipóteses que aceitarmos pressupor sobre a distribuição de  $b_i$ . Se assumirmos que  $F(b_i)$  é constante, então também o será a parcela  $q(K; a)$ . Neste caso, a Proposição 2-a) mostra que a atuação dos militantes será maior para publicações emitidas mais próximas



à data da votação. Por outro lado, se acreditarmos que a parcela  $q(K; a)$  muda em função da proximidade da data de publicação com a data da votação, então a Proposição 2-b) nos permite caracterizar esta mudança. De fato, a segunda proposição afirma que se a publicação for emitida suficientemente próxima da data da votação, então a parcela de eleitores que aceita compartilhar a publicação do partido responde positivamente a aumentos na atuação dos militantes.

### 3.4 O Candidato

Assumimos que o candidato gostaria de maximizar a parcela  $G(T^*)$  de eleitores que terão contato com a publicação até a data da votação<sup>15</sup>. Buscando maximizar  $G(T^*)$ , o candidato deverá levar em conta o nível de polarização do eleitorado e o parâmetro de influência externa para, então, decidir o nível de descaramento da mentira que publicará. Para incentivar os eleitores a compartilhar sua publicação (desincentivar a checagem), o candidato tem incentivos a publicar uma mentira com alto custo de inspeção  $K \in (0, 0, 5)$ . O custo de inspeção – em que incorrem eleitores que decidam checar a publicação – é proporcional ao nível de detalhamento e complexidade da mentira, bem como seu potencial em se passar por informação verdadeira. A produção de uma mentira com estas características demanda diversos recursos tais como tempo, capital humano e recursos financeiros e assumimos que a demanda por estes insumos é proporcional ao nível de

<sup>15</sup>Argumentamos que esta função objetivo faz sentido na medida em que os eleitores que checam e desmascaram a publicação não compartilham este fato com os demais, assim, o candidato não precisa se preocupar com os efeitos adversos de uma eventual contra-corrente de compartilhamentos de que sua publicação foi mentirosa.

complexidade e detalhamento da mentira. Em suma, é mais custoso (para o candidato) produzir mentiras com maior custo de inspeção. Formalmente, para produzir uma mentira com custo de inspeção  $K \in (0, 0,5)$ , o candidato incorre no custo  $c(K)$ . Ainda, assumimos que além de ser crescente e contínua, a função custo é convexa e

$$\text{H.1) } \lim_{K \rightarrow 0} c'(K) = 0 \quad \text{e} \quad \text{H.2) } \lim_{K \rightarrow 0,5} c'(K) = +\infty \quad (3.6)$$

A hipótese H.1 capta o simples fato de que é muito barato produzir as primeira unidades de disfarce da mentira. Ainda, esta hipótese incentiva o partido a investir algum montante positivo de recursos na produção da mentira e, assim, obter alguma parcela positiva de eleitores que a irão compartilhar, tal como parece ocorrer na realidade. A hipótese H.2, em conjunto com a convexidade de  $c(\cdot)$ , implica que mentiras muito sofisticadas têm seu custo marginal cada vez maior. Dito de outra forma, tornar uma mentira mais plausível é mais difícil quanto mais plausível ela já é. Se imaginarmos que temos uma mentira muito descarada ( $K$  pequeno), torná-la um pouco mais plausível é mais simples – e menos custoso – do que fazer o mesmo para uma mentira menos descarada (com  $K$  alto), que já se parece muito com uma afirmação verdadeira.

Assumimos que o partido conhece a regra de decisão dos eleitores e o nível de polarização, de forma que tem pleno conhecimento do efeito de sua escolha sobre a parcela  $q(K; a)$  e sobre a dinâmica de disseminação dada pela equação (3.5).

### 3.4.1 Análise do Problema do Candidato

Formalmente, nossa hipótese é de que o candidato resolve

$$\max_{K \in (0, 0,5)} G(T^*, K; a) - c(K) \quad (3.7)$$

onde  $G(T^*, K; a) = (1/2)[1 - p/q(K; a)]$  e  $q(K; a) = [2aK^3 - 3aK^2 + K(a + 6)]/6$ . Antes de enunciar o resultado principal da seção, notamos que

**Lema 1** *O problema dado pela equação (3.7) tem solução única, a qual denotaremos por  $K^*$ .*

E então, estamos prontos para afirmar que

**Proposição 2** *Quando o nível de polarização do eleitorado aumenta, o candidato publica mentiras mais descaradas.*

A proposição acima garante que o efeito da polarização sobre  $K^*$  é monótono e vale para qualquer nível de  $a \in [0, 12]$  e dos parâmetros  $p$  e  $q$ . Isto ocorre porque o aumento da polarização o candidato é capaz de “ludibriar” a mesma parcela de eleitores utilizando uma mentira com  $K$  menor.

De fato, mostramos que o benefício marginal de aumentar o custo de inspeção, dado por

$$\frac{3p(6aK^2 - 6aK + a + 6)}{(2aK^3 - 3aK^2 + K(a + 6))^2} \quad (3.8)$$

é decrescente no nível de polarização. Assim, na medida em que a polarização aumenta,  $K^*$  diminui para compensar a queda do benefício marginal. Além disso, o custo marginal  $c'(K)$  é independente de  $a$ , de forma que somente o benefício marginal é afetado pela variação na polarização do eleitorado.

## 4 Prova das Proposições

Por conta da limitação do número de páginas permitidas para esta publicação, iremos fornecer provas bastante curtas e, em alguns casos, apenas indicações.

### Proposição 1-a) e 1-b)

*Prova:* As provas das Proposições 1-a) e 1-b) seguem diretamente do Teorema da Função Implícita aplicado às expressões  $q - p \exp\{T_e(p + q)\} = 0$  e  $(q/e^{T_e q}) - pe^{T_e p} = 0$ , respectivamente (ambas derivadas da restrição  $\ln(q/p)/(p + q) = T^*$  obtida a partir da Hipótese 1).

### Lema 1 (Existência e Unicidade de solução para o problema do partido):

*Prova:* Vamos definir a função  $J(K, a, p)$  baseada na CPO do problema do partido

$$J(K, a, p) = \frac{3p(6aK^2 - 6aK + a + 6)}{(2aK^3 - 3aK^2 + K(a + 6))^2} - c'(K) = 0 \quad (4.1)$$

Para verificar a existência de  $K \in (0, 0,5)$  que satisfaz (4.1), basta notar que  $J(K, a, p)$  é contínua e que  $\lim_{K \rightarrow 0} J(K, a, p) = +\infty$  e  $\lim_{K \rightarrow 0,5} J(K, a, p) = -\infty$ .

Para garantir a unicidade da solução, basta mostrar que, para todo  $a \in [0, 12]$  a função  $J(K, a, p)$  decresce monotonamente no seu domínio  $K \in (0, 0.5)$ .

### Proposição 2

*Prova:* Esta prova é baseada no Teorema da Função Implícita. Tomando o diferencial total da função  $J(K, a, p)$ , obtemos

$$\frac{\partial K}{\partial a} = -\frac{\partial J / \partial a}{\partial J / \partial K} \quad (4.2)$$

O resultado segue da observação (não trivial) de que  $\partial K / \partial a < 0$ .

## References

Anderson, S. P., Waldfogel, J., and Stromberg, D. (2016). *Handbook of Media Economics, vol 1A*. Elsevier.



- Banerjee, A. V. (1992). A simple model of herd behavior. *The quarterly journal of economics*, 107(3):797–817.
- Bass, F. M. (1969). A new product growth for model consumer durables. *Management science*, 15(5):215–227.
- Bernhardt, D., Krasa, S., and Polborn, M. (2008). Political polarization and the electoral effects of media bias. *Journal of Public Economics*, 92(5-6).
- Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, 100(5):992–1026.
- Daley, D. J. and Kendall, D. G. (1964). Epidemics and rumours. *Nature*, 204(4963):1118.
- Duggan, J. and Martinelli, C. (2011). A spatial theory of media slant and voter choice. *The Review of Economic Studies*, 78(2):640–666.
- Eisensee, T. and Strömberg, D. (2007). News droughts, news floods, and us disaster relief. *The Quarterly Journal of Economics*, 122(2):693–728.
- Fourt, L. A. and Woodlock, J. W. (1960). Early prediction of market success for new grocery products. *Journal of marketing*, 25(2):31–38.
- Garrett, R. K. and Weeks, B. E. (2013). The promise and peril of real-time corrections to political misperceptions. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 1047–1058. ACM.
- Gentzkow, M. and Shapiro, J. M. (2006). Media bias and reputation. *Journal of political Economy*, 114(2):280–316.
- Haines Jr, G. H. (1964). A theory of market behavior after innovation. *Management Science*, 10(4):634–658.
- Jackson, M. O. (2010). *Social and economic networks*. Princeton university press.
- Kermack, W. O. and McKendrick, A. G. (1991). Contributions to the mathematical theory of epidemics-i. *Bulletin of mathematical biology*, 53(1-2):33–55.
- Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., et al. (2018). The science of fake news. *Science*, 359(6380):1094–1096.
- Lekvall, P. and Wahlbin, C. (1973). A study of some assumptions underlying innovation diffusion functions. *The Swedish journal of economics*, pages 362–377.

- Mahajan, V., Muller, E., and Bass, F. M. (1990). New product diffusion models in marketing: A review and directions for research. *Journal of marketing*, 54(1):1–26.
- Maki, D. P. and Thompson, M. (1973). Mathematical models and applications: with emphasis on the social life, and management sciences. Technical report.
- Mansfield, E. (1961). Technical change and the rate of imitation. *Econometrica: Journal of the Econometric Society*, pages 741–766.
- Meirick, P. C. (2013). Motivated misperception? party, education, partisan news, and belief in “death panels”. *Journalism & Mass Communication Quarterly*, 90(1):39–57.
- Mullainathan, S. and Shleifer, A. (2005). The market for news. *American Economic Review*, 95(4):1031–1053.
- of America, U. S. (2018). Disctrict of columbia. <https://www.justice.gov/file/1035477/download>. Acessado em: 2019-01-26.
- Papanastasiou, Y. (2018). Fake news propagation and detection: A sequential model. *Available at SSRN 3028354*.
- Politifact (2019). Donald trump’s file. <https://www.politifact.com/personalities/donald-trump/>. Acessado em: 2019-02-10.
- Shao, C., Ciampaglia, G. L., Varol, O., Yang, K.-C., Flammini, A., and Menczer, F. (2018). The spread of low-credibility content by social bots. *Nature communications*, 9(1):4787.
- Shin, J. and Thorson, K. (2017). Partisan selective sharing: The biased diffusion of fact-checking messages on social media. *Journal of Communication*, 67(2):233–255.
- Subramanian, S. (2017). Inside de macedonian fake-news complex. <https://www.wired.com/2017/02/veles-macedonia-fake-news/>. Acessado em: 2019-01-07.
- Suen, W. (2004). The self-perpetuation of biased beliefs. *The Economic Journal*, 114(495):377–396.
- Varol, O., Ferrara, E., Davis, C. A., Menczer, F., and Flammini, A. (2017). Online human-bot interactions: Detection, estimation, and characterization. In *Eleventh international AAAI conference on web and social media*.
- Watts, C. (2017). Extremist content and russian disinformation online: Working with tech to find solutions. *Statement prepared for the Senate Judiciary Committee, Subcommittee on Crime and Terrorism*.