A Hobbesian Solution to Infodemics

In the first volume of *Law, Legislation and Liberty* (1973, Chaps. 1 and 2), F.A. Hayek exposes his famous criticism of the constructivist (or rationalist) approach to human history. As Hayek puts it, the latter approach assumes that humans are fully rational and thus can construct perfect social institutions because reason can advise them on how to impeccably do so. In this regard, Hayek (1973, Chap. 1, p. 12) writes:

“Complete rationality of action in the Cartesian sense demands complete knowledge of all the relevant facts. A designer or engineer needs all the data and full power to control or manipulate them if he is to organize the material objects to produce the intended result. But the success of action in society depends on more particular facts than anyone can possibly know. And our whole civilization in consequences rests, and must rest, on our believing much that we cannot know [Hayek’s italics] to be true in the Cartesian sense.”

Against this view, Hayek (1973) argues that social institutions are rather the result of conscious human activity. Specifically, Hayek (1973) maintains that social institutions are the imperfect result of complex sociopolitical and socioeconomic processes. Moreover, Hayek (1973, Chap. 1) maintains that the constructivist approach to human history suits the purpose of totalitarian regimes. That is why Hayek (1973, Chap. 1, p. 32) writes:

“Reason is merely a discipline, an insight to the limitations of the possibilities of successful action, which often will tell us only what not to do. This discipline is necessary precisely because our intellect is not capable of grasping reality in all its complexity […] Liberalism for this reason restricts deliberate control of the overall order of society to the enforcement of such general rules as are necessary for the formation of a spontaneous order, the details of which we cannot foresee.”
Hence, for Hayek (1973), the existing spontaneous social order results from the decisions and actions of dispersed individuals dealing with different circumstances. That is why, for Hayek (1973), the spontaneous social order emerges as people exercise reason. More precisely, as people learn how to face circumstances by habit and rules of thumb, the latter order emerges—Hayek (1973) argues—because people stipulate common rules that enable them to achieve satisfactory coordination. On this basis, Hayek (1973, Chap. 2) proposes his famous distinction between order and taxis.

By order, Hayek (1973, Chap. 2) means a self-governing entity (or spontaneous social order) whose internal rational structure emerges from people’s boundedly rational actions and decisions. Instead, by taxis, Hayek (1973, Chap. 2) means our overall understanding of the internal rules of such self-governing entity. Most importantly, Hayek (1973, Chap. 2) argues that—in a democratic and economically free society—the order emerges spontaneously because individuals achieve coordinate for the sake their mutual interest. Instead, in a totalitarian regime, it is the dominating political class that imposes some social order upon society in agreement with the dominant ideology of the ruling class.

This difference is essential to understand the role of taxis in a despotistic and in a democratic society. In the first case, according to Hayek (1973, Chap. 2), the dominating political class imposes some rules of behavior upon the citizens to ensure that they behave as per the will of the dominating class. Instead, in the second case, the democratically elected representatives steer the behavior of citizens by enacting forms of legislation whose primary goal is to ensure that free individuals coordinate for the sake their mutual interest in the best possible way. Expressed otherwise, the approach of totalitarian is constructivist, whereas the second approach results from the exercise of reason. Hayek calls the latter approach ecological.

Interestingly, Hayek’s (1973, Chap. 1-2) overall line of reasoning resonates Bentham’s argument presented in Chapter 7 of An Introduction to the Principles of Morals and Legislation (1780). In this work, Bentham (1780) famously argues that legislators should enact laws in agreement with the principle of utility. In other words, legislators ought to promulgate laws rewarding actions that maximize general utility, while they ought also to promulgate laws punishing actions that fail to do so. Like in Hayek’s (1973) case, Bentham’s writing incorporates the spirit of British enlightenment. Namely, Bentham revolts against the rationalist approach to legislation of his predecessors, particularly Hobbes’s.
In the same way, Hayek (1973, Chap. 1-2) revolts against the rationalist approach to political economy. This point is even more explicit in Hayek’s earlier writing “Free Enterprise and Competitive Order” (1947). In the latter work, Hayek (1947) famously maintains that only fair and lean regulation of economic affairs can ensure that a competitive order emerge unhindered. Specifically, Hayek (1947) argues that a fair and lean legal framework regulating trademarks and contracts is the necessary condition for economic competition to be effective. Namely, legislation must be lean and fair because it must guarantee that the market function enhances society’s prosperity. In other words, a lean legislation is the proper taxis to steer the competitive order.

Globalization and the digital revolution

After several decades, facts have proven Hayek (1947; 1973) right. For example, the establishment of The Vienna Convention on the Law of Treaties in 1969 has given firms simple legal guidelines to draft commercial contracts to trade goods globally. Particularly, the Vienna convention gave firms a solid legal umbrella to access unexplored markets. Moreover, the regulation of international trade through the establishment of WTO in 1995 has facilitated the growth of global trade and caused a sharp decrease of tariffs and quotas worldwide. Finally, the promulgation of the International Financial Reporting Standards and the Generally Accepted Accounting Principles provided firms with internationally recognized reporting and accounting guidelines that allow them to communicate financial information globally in a standardized and systematic way.

In Hayek’s terminology, the emergence of the global competitive order results from the individual necessity of firms to trade their goods and services beyond national borders. Instead, its persistence and growth in time results from the existence of solid taxis that steers it, i.e., the introduction of lean international regulations for international trades and fair arbitrage rules for legal disputes on unperformed contracts. Furthermore, the emergence of the global competitive order has provided a spectacular confirmation of the Hayek-Friedman Hypothesis, which—simply put—states that economic freedom is the necessary condition for the emergence of democratic sociopolitical institutions (Lawson and Clark 2010).

Indeed, as remarkably shown by Steven Pinker (2019, Chaps. 11-14), the acceleration of globalization has since the 1990s caused an unprecedented increase in the number of democracies, a sharp decline in violence and conflicts worldwide, and an unparalleled
increase in global prosperity. As shown also by Johan Norberg (2017, chap. 4), the number of poor worldwide decreased to less than 1 billion people for the first time in human history. In this regard, Deaton (2013) has shown that globalization has also caused a rise in longevity that correlates with higher economic standards. Moreover, Bailey and Tupy (2020) have shown that technology facilitated the dispersion of knowledge across the globe to realize such an economic and social success in human history. Finally, Tupy and Pooley (2022) have also shown that the dispersion of knowledge and technology across the globe has given humanity an unprecedented opportunity to contain the horrible side-effects of Malthus Law.

While it is indisputably true that the current digital revolution sustains the global economic order, it is also true that there are reasons that to believe that it might hinder the posited effects of the Hayek-Friedman hypothesis. To see why this line of argument may be justifiable, we shall consider the sociopolitical side-effects of the most potent propeller of the current digital revolution, i.e., social media. On the one hand, social media are trading facilitators for many firms and startups. Namely, social media are pillars of the current global economic order. Moreover, social media give their users new opportunities to gain knowledge about new topics and sociopolitical causes. Yet the algorithms of social media are, on the other hand, dangerous enablers for rapid verifications of Brandolini’s law. That is why social media often become dangerous enablers of political polarization and instability in Western liberal democracies.

In any case, social media are enablers for the emergence of a spontaneous online social order. Interestingly, this fact shifts us back to the Hayek’s distinction between constructivist and ecological approaches to such a social order. Particularly, it forces us to ask ourselves which approach is best for the spontaneous online social order. On the one hand, totalitarian regimes, like China and Iran, pursue a strong constructivist approach to the latter order. Indeed, such regimes have banned citizens’ access to Western social media such as Facebook and Twitter to avoid political instability. On the other hand, in Western liberal democracies, the lack of a clear policy against fake news and hate speech on social media lies at the core of unfortunate events like the Capitol Hill Assault on Jan 6th, 2021. Accordingly, as fact seems to suggest, if free speech in the online world is not subject to

---

1 Brandolini’s Law states that the amount of energy needed to refute bullshit is an order of magnitude bigger than that needed to produce it.
the same limitations as in the offline world, the social function of social media can backfire and violate the general implication of the Hayek-Friedman hypothesis. This line of reasoning leads to a much important research question. Namely, should Western governments enact policies against fake news and the likes even though it might imply running against the free speech principle? Or should Western governments expect that individuals exercise reason as Hayek (1973) proposes? In other words, which approach to the problem of fake news and the likes is the most preferable? A top-down approach or a bottom-up approach? The answer to these questions is not easy to find. That is because the general principles of a Western liberal democracy do not provide a valid justification to introduce some form of taxis to steer the spontaneous online social order. Yet, as we show below, some form of legislation might enhance general social welfare without disrupting the benefits of the current digital revolution.

_Hobbes on seditious doctrines_

The British philosopher Thomas Hobbes figures among those authors Hayek (1973, Chap. 1) criticizes for undertaking a constructivist approach to human history. Particularly, Hayek’s (1973) criticism of Hobbes aims at undermining a common feature of Modern thinkers. Namely, the assumption that knowledge is certain IFF it can be expressed with the same clarity as the statements of geometry and mathematics. In Hobbes’s _Leviathan_, this assumption is openly put forward. Indeed, Hobbes himself (1651, Part 1, Chap. 4, p. 13) writes:

> “Seeing then that truth consists in the right ordering of names in our affirmations, a man who seeks precise truth needs to remember what every name he uses stands for, and to place it accordingly; otherwise he will find himself entangled in words like a bird in lime twigs: the more he struggles the more thoroughly he is belimed. And therefore in geometry, which is virtually the only precise science, men begin by settling the meanings of their words in what they call ‘definitions’, which they place at the start of their calculations.”

These premises lie at the foundations of Hobbes’s (1651) entire political theory. As Slomp (2000) puts it, Hobbes (1651) pursues the ideal of political geometry. Namely, Hobbes (1651)
seeks out to produce a political theory that is as rigorous as geometry. That is why Hobbes (1651) equates logical necessity to causation. In fact, in Chapter 5, Hobbes (1651) argues that the task of philosophy is to discover effects from known causes and vice versa in the same way as geometry discovers the nexuses of logical necessity among definitions. On these grounds, Hobbes (1651) proceeds to define the state of war as the anteceding cause of the organization of political life in a sovereign state (Copleston 1994, Vol 5, p. 4).

In other words, differently than Hayek (1973), Hobbes (1651) openly denies that social institutions emerge from complex social processes. Instead, Hobbes (1651, part II, §15) postulates that the state of war perdures so long as humans do not discover reason because they attain their passions and lack any knowledge of good and evil. As a result, according to Hobbes (1651, part II, §15) perdures so long as humans seek glory through violence and competition. On these premises, Hobbes (1651) endows the human reason with constructive powers to ensure a nexus of causality (or logical necessity) between the primordial state of war and the Hobbesian Commonwealth.

Specifically, Hobbes (1651) maintains that reason allows humans to acknowledge that they would be better off if they avoided war and sought self-preservation instead. In this regard, Hobbes (1651, part I, §13-15) claims that reason allows humans to discover 19 natural laws that are the basis of true moral philosophy. As Hobbes (1651, part I, §13-15) puts it, these laws are moral precepts that dictate humans to avoid what is destructive to them. Particularly, given that Hobbes (1651) considers mathematics and geometry the only forms of perfect knowledge, Hobbes (1651) argues that the 19 natural laws are theorems whose evidence is rationally acknowledgeable by humans (see Hobbes 1651, part I: §15, p. 74). Namely, these laws are theorems whose clear necessity is immediately evident to human reason.

Furthermore, for Hobbes (1651), the discovery of these laws by means of reason is the primary cause for the establishment of the Hobbesian Commonwealth. On one side, according to Hobbes (1651), it enables humans to acknowledge that they would be better off if they avoided war and sought self-preservation instead. On the other, it provides the moral basis for any form of fair legislation within the Hobbesian Commonwealth. That is why Hobbes (1651) maintains that humans can willfully and consciously understand that they ought to submit their wills to the will of a Sovereign who could be either a single man or an assembly of men.
More precisely, for Hobbes (1651, part I, §13-14), reason advises humans on why they ought to stipulate a covenant and let the Sovereign compel them to perform it. Wherefore, in the Hobbesian Commonwealth, liberty coincides with obedience to the rule of law (i.e., the covenant) that also establishes which actions are morally admissible (Hobbes 1651, part II: §21). Thus, in the Hobbesian Commonwealth, liberty cannot be unconditioned. Accordingly, if someone wants to be unconditionally free, they must leave the Commonwealth and return to the state of war.

As Copleston (1994, vol. 5: p. 41) correctly points out, “any fundamental criticism of Hobbes must be directed against his account of human nature rather than against the details of the theory of the covenant.” In this regard, it is imperative to note that Hobbes (1651) saw Modern Monarchy as the only political system that could ensure socio-political stability. That is because Hobbes (1651) identified the Sovereign with the rule of law. On these premises, Hobbes (1651) makes generous assumptions about human rationality to justify his argument that humans can rationally acknowledge that their general welfare would increase if they behaved in accordance with the rule of law (i.e., if they submitted their wills to a Sovereign who compels them to perform the covenant).

However, in Part II, Hobbes (1651) turns around toward his initial assumptions about human nature and explores the consequences of his constructive approach to human institutions. More precisely, Hobbes (1651, Part II) provides a justification for his argument that reason alone allows humans to foresee that they are better off if they never violate the covenant and always respect the Sovereign’s will. In other words, Hobbes (1651) attempts to justify his view that a set of initial binding principles is sufficient to grant that no unexpected action violates the covenant.

In this regard, Hobbes (1651, Part II) maintains that humans should stipulate covenant to grant the Sovereign special rights to ban those seditious doctrines that can cause socio-political instability in the Commonwealth. As Hobbes (1651, part II, §18, p. 82) himself puts it, the Sovereign must be

> “the judge of what opinions and doctrines are threats to peace and what ones tend to support it; and consequently of which men are to be trusted to speak to multitudes of people, on what occasions, and how far they should be allowed to go; and of who shall examine the doctrines of all books before they are published […] For a doctrine that is harmful to peace can’t be true, any more than peace and harmony can be against the law of nature.”
Along these lines, Hobbes (1651, part II, §29) also argues that bad covenants produce commonwealths that stand on a shaky basis. Accordingly, some seditious doctrine can lead to their dissolution. Hence, the covenant should give the Sovereign the right to limit the circulation of any doctrine that can disrupt the Commonwealth. That is why, according to Hobbes (1651), the Sovereign must not be subject to the laws of the Commonwealth. Particularly, for Hobbes (1651, part II: §29, p. 146), the statement “he who has the sovereign power is subject to the civil laws” is hostile to the sociopolitical order of the Commonwealth because the Sovereign cannot be subject to the laws he makes and cannot be subject to his own judgement.

On the one hand, these arguments help Hobbes (1651) avoid logical circularity in his entire argument for the rise of a Commonwealth based on a fair covenant. On the other hand, these arguments lead Hobbes (1651, part II: §29, p. 145) to maintain that “every private man is a judge of good and evil actions” is a seditious doctrine because it entices humans to disavow the rule of law. In this way, Hobbes (1651, part II) attempts to justify that an action is admissible IFF it is coherent with the basic principles of the covenant that bestows the Sovereign with ruling power over the Commonwealth. Expressed otherwise, such a sociopolitical order is for Hobbes (1651) is the endpoint of orderly causal linear processes such that no unexpected action occurs through it.

*Was Hobbes right after all?*

Had Hobbes (1651) been right, the institutions of Western liberal democracies would have emerged from orderly linear social processes like the Hobbesian Commonwealth. Instead, Western democracies have emerged from lengthy and complicated historical processes. That is primarily because *the course of history is determined by what people do with their opportunities*, as Thomas Sowell once said. In this regard, Hayek’s (1973) criticism of Hobbes is undoubtedly justified.

Analogously, if Hobbes (1651) were right about the constructive powers of reason, the constitutions of Western democracies would also have been the products of perfect minds like the Hobbesian covenant. In contrast, the constitutions of Western democracies are the imperfect products of lengthy negotiations amongst opposite political factions within constitutional assemblies. Particularly, had it been otherwise, no Western Constitution
would ever have needed amendments by legislators. Thus, Hayek (1973) certainly scores another point against Hobbes (1973).

Nevertheless, we believe that the Hobbesian lesson nowadays needs some reconsideration for at least two reasons. Firstly, as we shall see below, some of the most recent findings in contemporary psychology reveal that humans are bad at exercising reason when contextual variables and contextual information affect the soundness of the reasoning. In this respect, the algorithms of social media are potent enablers for the propagation of contemporary seditious doctrines (i.e., fake news and the likes) that can potentially disrupt the sociopolitical order of Western democracies.

Secondly, the existence of these flaws in human nature panders to the thesis that lean regulation rather than the exercise of reason by individuals alone can make up for the flaws in human nature. Expressed otherwise, the flaws in human nature require that some form regulation is needed to limit the negative sociopolitical side-effects of the digital revolution on human behavior. In this sense, there seems to be room for a possible reconciliation between Hayek’s (1947; 1973) and Hobbes’s (1651) respective points of views. Indeed, as we argue below, only lean regulation of digital media can enable people to keep reaping the benefits of the digital revolution.

Contemporary Seditious Doctrines and Human Psychology

Though metaphorically only, the Hobbesian covenant can be considered a close substitute of the constitutions of Western liberal democracies. Specifically, the drafting of constitutions in the West has been the precondition for the rise and perdurance of liberal democracies for at least three fundamental reasons. Firstly, Western constitutions are the charters of the rights of Western citizens. Secondly, Western constitutions establish the separation of powers and express the basic principles to produce legislation that does not infringe the principle of common interest in a democratic nation. Thirdly, Western constitutions establish that everyone is equal before the Law, and no one is unboundedly freedom even though Western citizens are generally able to express their thoughts in an unbounded way.

Importantly, the basic legal principle underlying free speech in a democratic country is that people are accountable for their statements in their public and private lives. For this reason, deception, such as fraud and ideological forgery, is punishable by the Law
whenever it has harmful side-effects in the private and public spheres. For the same reason, the Law punishes the propagators of fabrication leading to negative outcomes that harm general wellbeing. That is why, if someone shouts that there is a bomb in a theater full of people under false pretenses, and some people in the theater die or get injured because the crowd attempting to escaping the theater tramples them, the perpetrator of the silly prank is punishable by the Law.

Now, a compelling question is why the same does not apply to people’s behavior online. The answer to this question is far from trivial. Yet the Hobbesian lesson can help us understand why the contemporary seditious doctrines (i.e., fake news and the likes) can potentially undermine the sociopolitical and socioeconomic orders of Western liberal democracies. Particularly, the Hobbesian lesson raises awareness on three aspects of our life online.

Firstly, the contemporary seditious doctrines circulate freely online because the internet often bestows its users with a pernicious illusion of unconditioned freedom (Pagliaro 2017; Ostillio 2018). Secondly, a few internet searches often give internet users an overblown illusion of in-depth knowledge (IoK) about complicated matters (Mazzone 2019). Thirdly, the wide availability of contemporary seditious doctrines on the internet often nourishes the IoK (Butter and Knight 2020) about many subjects with adverse effects on sociopolitical stability (e.g., the case of QAnon and the assault of Capitol Hill).

In light of these remarks, it is also worth noting many Westerners develop IoK on social media, where many contemporary seditious doctrines prosper and become accessible to everyone (Ziccardi 2019). Specifically, as more and more Westerners become active social media users year by year, social media frequently give rise to an unusual social process, which we shall call adverse reversal of pluralistic ignorance.

In simple terms, pluralistic ignorance (PI) occurs when groups of people display uniform behavior while group members privately reject that behavior. For instance, when professors ask students whether they have questions at the end of a confusing lecture and students answer that question with a roaring silence, PI occurs because some students would have questions but comply with the group’s silence (McRaney 2013).

When PI holds, private information disclosure usually reverses PI bringing about dramatic changes in the existing social norms and preferences (O’ Gorman 1975; Sunstein 2019). For example, Bursztyn et al. (2018) run a survey asking young Saudi men whether they favor women’s participation in the labor force or not. Bursztyn et al. (2018) astoundingly
find that most of the surveyed Saudi men support women’s participation in the labor force but believe that other young Saudi men do not do so. Remarkably, Bursztyn et al. (2018) also report that when they disclose the actual survey results to the surveyed Saudi men, the likelihood that their wives apply for jobs increases.

While we do not question that reversals of PI enable dramatic changes in existing social norms and preferences, we also acknowledge that there are reasons to believe that the reversal of PI may sometimes have adverse effects. For example, Burioni (2016; 2017) explains that potential anti-vaxxers usually hold some existing (wrong) beliefs about vaccines but conform to the existing social preferences. In particular, as potential anti-vaxxers initially avoid disclosing their beliefs in their social circles, potential anti-vaxxers initially act within the PI schema. However, this attitude can be reversed once potential anti-vaxxers acknowledge that other people share the same beliefs about vaccines (Burioni 2017, chap. 5). Thus, disclosure reverses PI with adverse effects as it entails that a potential anti-vaxxer becomes an actual anti-vaxxer.

Notably, social media algorithms are frequent enablers of such forms of adverse reversals of PI. That is because social media algorithms obey specific business assumptions allowing social media providers to operate unique double transactions with their users. On one side, users disclose previous information (e.g., personal preferences and inclinations) to providers. On the other, social media algorithms record this information and dispense each user with personalized content. As a result, when social media users uncover too much information about themselves to providers, social media algorithms produce redundant personalized content streams (Quintarelli 2019). In jargon, the latter streams are called filter-bubbles (Spohr 2017; Froelich 2019).

When social media users fall victim of filter-bubbles, social media algorithms generate dangerous availability cascades. In a nutshell, an availability cascade occurs when new and traditional media repeatedly report some piece of information with general adverse effects on collective imagination (Kuran and Sunstein 1998; Sunstein 2005). More concisely, the definition of an availability cascade encloses the famous adage “repeat something long enough, and it will become true.” In filter-bubbles, the latter adage continuously strengthens users’ (wrong) beliefs or feed into their IoK about any subject with personalized content streams.

Furthermore, availability cascades in filter-bubbles enable like-minded social media users to get together. Namely, availability cascades in filter-bubbles increase the likelihood that
like-minded users make contact and disclose aligned (wrong) beliefs with one another, actualizing the adverse reversals of PI. Particularly, availability cascades in filter-bubbles enable like-minded users to flow into online communities featuring intense group polarization (i.e., communities that develop shared dominant narratives and marginalize or reject contrarians), where the reversal of PI is finalized.

It follows that availability cascades in filter-bubbles amplify the abovementioned illusion of unconditioned freedom and overblow IoK about any subject. Indeed, social media users enjoy unconditioned freedom to espouse any thesis without consequences as they mostly interact with like-minded subjects. Consequently, it should not be a surprise that availability cascades in filter-bubbles might initiate destructive behavioral patterns that culminate into subversive actions (e.g., the assault of Capitol Hill).

**The lesson from the Covid19 pandemic**

During the Covid19 pandemic, the online behavioral patterns described above have led some online communities to urge their members to burn 5G antennas (e.g., in Italy and England) or join no-mask public initiatives (e.g., in Italy or Germany). Other communities have urged their members to join covid parties that resulted in massive recoveries in intensive care units. Hence, it should be now clear why freedom cannot be unbounded online and why the *contemporary seditious doctrines* can often become dangerous enablers of the patterns described above.

Furthermore, although we are much aware that Hobbes (1651) did not use the covenant to justify the rise of a perfect democracy, we shall observe that the Hobbesian covenant and the constitutions of Western liberal democracies share another fundamental similarity. Like humans stipulate the covenant to grant the Sovereign special rights to maintain law and order if circumstances require it, the constitutions of Western liberal democracies grant democratically elected governments special rights to take extraordinary actions during times of high distress.

That is why many Western governments initiated emergency protocols known as national lockdowns to curb contagion and contain the circulation of the coronavirus when the Covid19 pandemic erupted worldwide early in 2020. Though aimed entirely at preserving public health, these protocols have encountered fierce opposition in many Western countries because of the harsh socioeconomic implications of national lockdowns. Specifically, many Western political leaders and commentators criticized national
lockdowns typically espousing that national lockdowns do more harm than good due to their disastrous economic consequences, violate fundamental human rights, and pave the way to more state interventionism. Importantly, because lockdown increased the time Westerners spent online, these claims have enjoyed a profound echo on social media since the beginning of the Covid19 pandemic. Indeed, social media have since funneled much of Westerners’ resentment against lockdowns and have since become arenas hosting harsh tiffs between those who favor lockdowns and those who do not. As a result, political polarization increased in many Western democracies.

Worryingly, some evidence suggests criticism of national lockdowns augmented political polarization with adverse effects on governmental response to the Covid19 pandemic. For instance, Milosh et al. (2020) find that political polarization is one of the drivers of America’s failure to provide adequate prompt policy response to the Covid19 pandemic and its devastating consequences. Other evidence suggests criticism of lockdowns has not always been legitimate and well-founded. In this regard, Evanega et al. (2020) find that the eruption of the Covid19 pandemic has brought about an infodemic, i.e., a misinformation epidemic. In their study, Evanega et al. (2020) analyze 38 million online articles in the English language and find that about half (46.6%) contain misinformation or conspiracy theories about the Covid19 pandemic and its origins, whereas only 16.6% include fact-checking. This finding is a spectacular confirmation of Brandolini’s law.

Of particular concern is also the contribution to both the infodemic and political polarization from several Western political outlets. In this respect, some evidence suggests that several Western political outlets intentionally spread alternative narratives about the Covid19 pandemic on social media to enlarge their supporter base and turn their supporters against political opponents. For example, the BBC reported that Matteo Salvini and Donald Trump used these strategies to feather their nests in April 2020. Besides, American Republicans became less prone to see the Covid19 pandemic as a threat to public health than American Democrats due to prolonged exposure to media outlets endorsing Donald Trump (Tyson 2020). For the same reason, several thousands of Europeans joined no-mask demonstrations (Belenky 2020).

Another reason of concern arises from the findings of some opinion polls about the Covid19 vaccine. First and foremost, if public trust in science stays typically high worldwide (Funk et al. 2020), worldwide trends (2015-2019) in people’s confidence in the
importance, safety, and effectiveness of vaccines lean towards mixed patterns (de Figuereido et al. 2020). In particular, the Wellcome Trust (2019) found that Western democracies feature some of the lowest global scores in confidence about vaccine safety. Hence, it is not excludable that the lasting effects of the infodemic may have impacted general willingness to get the Covid19 vaccine in some Western democracies. That is why countries like Portugal and Italy had to prompt their armies to handle the vaccination campaign against the Covid19, whereas France and the US enacted policies that made vaccination against Covid19 mandatory for some period of time.

Interestingly, the findings of the Wellcome Trust (2019) and de Figuereido et al. (2020) find support in two surveys carried out at the beginning of the pandemic highlighted that—indeed—reveal that Westerners were less prone to get vaccinated against Covid19. For example, Lazarus et al. (2021) ask 13,426 respondents from nineteen countries (amounting to 55% of the World’s population) to rate positively or negatively the statement, “If a COVID-19 vaccine is proven safe and effective and is available, I will take it.” Lazarus et al. (2020) find that positive responses are above average (> 71.35%) in ten countries (of which three are Western democracies: USA, Spain, and the UK). Instead, positive responses are below average (< 71.35%) in nine countries (of which six are Western democracies: Italy, Canada, Germany, Sweden, France, and Poland).

On the other hand, a survey by IPSOS for the World Economic Forum reveals similar patterns in 2020. In their study, IPSOS researchers asked 19,519 adults in twenty-seven countries to rate their overall agreement (disagreement) with the statement, “if a vaccine for COVID-19 were available, I would get it.” IPSOS researchers find that 74% of all respondents (37% fully and 37% somewhat) agree with the statement, whereas 26% of all respondents (14% somewhat and 12% fully) disagree with the statement. In particular, IPSOS researchers find that the agreement share is above average (> 74%) in thirteen countries (of which four are Western democracies: Australia, UK, Canada, Japan). Conversely, IPSOS researchers find that the agreement share is below average (< 74%) in fourteen countries (of which ten are Western democracies: Spain, Netherlands, Belgium, Sweden, USA, Germany, Italy, France, Hungary, Poland).

More importantly, IPSOS researchers asked 4,860 adults in 27 countries to state their reasons for not taking a vaccine and found that 56% of all respondents say that they are worried about the side effects, 29% of all respondents do not think it will be effective, 19% of all respondents claim not to be enough at risk from COVID-19, and 17% of all
respondents claim to be against vaccines in general. Two of these findings are particularly worrisome. Firstly, the share of respondents who say they are worried about the side effects is above average (>56%) in twelve of the twenty-seven countries in the sample (of these countries, eight are Western democracies: Spain, Sweden, Poland, Japan, Belgium, France, the USA, Hungary, Germany, the UK). Secondly, the share of respondents who state they are against vaccines in general is above average (>17%) in thirteen of the twenty-seven countries in the sample (of these countries, eight are Western democracies: Italy, France, Belgium, the USA, Germany, Australia, Canada, Poland).

In light of these findings, we might claim that the infodemic and political polarization stand in the same relationship as availability cascades in filter-bubbles and group polarization on social media. Accordingly, like availability cascades in filter-bubbles initiate destructive behavioral patterns that may culminate into subversive actions, it is not excludable that the infodemic and political polarization might have had adverse effects on governmental struggles against Covid19.

**Top-down or bottom-up?**

As we saw above, the bulk of Hayek’s (1973) argument against constructivism is that humans learn to exercise reason to enhance the institutions they share. Particularly, according to Hayek (1973), people learn from their mistakes and enhance their behavior accordingly. In this way, humans enhance institutions so that they prevent them from repeating the same mistakes. Yet, if Hayek (1973) is right, why do humans continuously fall prey of infodemics and their side-effects without any sign of enhancement?

As we saw above, the findings of contemporary psychology tell us that there are cases in which the lack of rational control over behavior makes humans unable to correct their behavior. Thus, if the lack of rational control is most visible when contextual variables and contextual information affect human behavior, is it sound to postulate that lean regulation of social media algorithms can enhance coordination among people in their daily online life? As we saw earlier, any answer to this question depends on which approach to human institutions we choose to pursue.

It is worth noting that the likely answers to the above question are many and multifaceted. If we followed Hayek’s (1973) to the letter, we could—as Hayek (1973) indeed does—resort to John Stuart’s Mill (1859) argument that free speech offers the most convenient
means to reinforce the prevailing view that a community (or a society) shares. In this regard, his *On Liberty*, Mill (1859, pp. 33-34) writes that

“Freedom of opinion, and freedom of the expression, are needed for the well-being of mankind [...] An opinion that is compelled to silence may, for all we can certainly know, be true [...] Even when the silenced opinion is an error, it can and very commonly does contain a portion of the truth; and since the general or prevailing opinion on any topic is rarely if ever the whole truth about it, it is only through the collision of conflicting opinions that the remainder of the truth has any chance of being supplied [...] even if the publicly accepted opinion is not only true but is the whole truth on the subject in question, unless it is vigorously and earnestly disputed most of those who accept it will have it in the manner merely of a prejudice, with little grasp or sense of what its rational grounds are.”

Though undoubtedly agreeable, Mill’s (1859) point holds so long as humans are able to exercise reason, are willing to listen and rationally evaluate others’ opinions, and external manipulations, like in the case of social media algorithms, are out of the picture. Accordingly, we maintain that the answer to our guiding question requires rather abstract reasoning than practical reasoning based on common sense. That is why we attempt to settle the issue of this paper with a simple game-theoretical model.

*The battle of the sexes*

We tackle the research problem of this paper by formalizing it in an intuitive way. To this end, we argue that the battle of the sexes game provides the best itinerary to resolve the dilemma we are facing. In the latter game, a young couple, say John and Annie, quarrels about what they shall do in the evening. While John wants to go see a boxing fight, Annie wants to go see the ballet. If they fail to make a decision, they shall stay home and spend the evening alone. Yet, if Annie gets through the quarrel, they will spend the evening together at the ballet, but she will be happier than John. In contrast, if John gets through the quarrel, they will spend the evening together at the boxing fight, but he will be happier than Annie.

Notably, we assume that the battle of the sexes game may be considered a metaphor that represents a typical tiff between two strongly polarized factions populating some online community. Accordingly, we show that the game has three possible solutions: the first possibility is analogous to the Hobbesian state of War; the second solution is analogous to
the Hayek’s and Mill’s arguments; finally, the third solution is analogous to the Hobbesian covenant. Interestingly, we find that the last solution maximizes welfare better than the previous two. On this basis, we present our final argument in favor of the introduction of some lean regulation for social media algorithms.

**First solution: pure strategy Nash equilibrium**

As the first solution to the battle of the sexes game is analogous to the Hobbesian state of war, we shall assume that John and Annie participate in the game by issuing simultaneous decisions aimed at maximizing their personal interests. Besides, before we proceed to solving the same, we must introduce two fundamental notions in our discussion: the notion of a Nash equilibrium and the notion of a best response strategy.

In plain English, the first notion represents a set of strategies such that none of the players has a reason to change their strategy, given what other players do, whereas the second notion represents the strategy that brings about the most significant pay-off for each of the players. Notably, a pure Nash strategy is nothing but the best response strategy for each of the players. Thus, if two players, like John and Annie, only pursue their highest interest and issue simultaneous decisions that maximize their personal interests, a game like ours shall have at least one Nash equilibrium.

Let us now introduce some notation. The players set $A$ contains only two elements, i.e.,

$$A = \{\text{John, Annie}\}.$$

The set of strategies for each player $S$ contains two elements, i.e.,

$$S = \{\text{Fight, Ballet}\}.$$

The set of all possible outcomes $O$ contains four elements, i.e.,

$$O = S \times S = \{(\text{Fight, Ballet}), (\text{Ballet, Fight}), (\text{Ballet, Ballet}), (\text{Fight, Fight})\}.$$

If $U$ is an ordinal utility function mapping each element of $O$ to the set of rational numbers $\mathbb{R}$, i.e.,

$$U: O \to \mathbb{R},$$

we can define a ranking of preferences for each player using the operators $>$ (strictly preferred) and $\sim$ (indifference).

Given the initial conventions about each players’ preferences, we shall suppose that, for John, each outcome has the following utilities: $\{\text{Fight, Fight}\} = 2$, $\{\text{Ballet, Ballet}\} = 1$, $\{\text{Fight, Ballet}\} = \{\text{Ballet, Fight}\} = 0$. Instead, for Annie, each outcome has the following utilities: $\{\text{Ballet, Ballet}\} = 2$, $\{\text{Fight, Fight}\} = 1$, $\{\text{Fight, Ballet}\} = \{\text{Ballet, Fight}\} = 0$. Therefore, for John, $\{\text{Fight, Fight}\} > \{\text{Ballet, Ballet}\} > \{\text{Fight, Ballet}\} \sim \{\text{Ballet, Fight}\}$. In contrast, for Annie, $\{\text{Ballet, Ballet}\} > \{\text{Fight, Fight}\} > \{\text{Fight, Ballet}\} \sim \{\text{Ballet, Fight}\}$. Thus, if we now organize this
information in a matrix representing the utilities of each player, we shall obtain the following table (Table 1):

<table>
<thead>
<tr>
<th>The battle of the sexes</th>
<th>Annie</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fight</td>
</tr>
<tr>
<td>John</td>
<td></td>
</tr>
<tr>
<td>Fight</td>
<td>2,1</td>
</tr>
<tr>
<td>Ballet</td>
<td>0,0</td>
</tr>
</tbody>
</table>

Table 1 - The battle of the sexes

Given our initial definition of a pure Nash strategy, the outcomes \{Fight, Fight\} and \{Ballet, Ballet\} are John’s and Annie’s pure Nash strategies, respectively. Hence, our game has two Nash equilibrium, i.e., the outcomes marked in red in Table 1. Importantly, this situation is analogous to the Hobbesian state of war because every player only maximizes their personal interest. As a result, the wellbeing of one player is diminished in favor of the wellbeing of the other player. Remarkably, this situation also resembles a typical tiff between polarized factions of users in an online community, where either one faction or the other wins the tiff.

*Second solution: mixed strategy Nash equilibrium*

Let us now complicate the matter at hand by incorporating Hayek’s (1973) and Mill’s (1859) in favor of the view that people can and do exercise reason. Hence, let us assume that each player has some expectation concerning what the other player shall or shall not do. In the jargon of game theory, this assumption leads to the hypothesis that our game can have a mixed strategy solution. Namely, we assign some probability \(p\) to a player’s belief that the other player shall or shall not do something. More precisely, a mixed strategy represents a probability distribution such that each player believes that the other player shall pursue some strategy with probability \(\sigma\) or shall not pursue another strategy with probability \(1 - \sigma\) (i.e., \(\sigma + (1 - \sigma) = 1\)). Using these conventions, we can introduce the notion of rational expectations for each player and derive the mixed strategy Nash equilibria for our game, while leave John’s and Annie’s preferences unchanged. Namely, for John, \{Fight, Fight\} > \{Ballet, Ballet\} >
\{Fight, Ballet\} \sim \{Ballet, Fight\}, and, for Annie, \{Ballet, Ballet\} > \{Fight, Fight\} > \{Fight, Ballet\} \sim \{Ballet, Fight\}.

Let us now add three assumptions to incorporate Hayek’s (1973) and Mill’s (1859) argument into our line of reasoning:

1. Like before, John and Annie issue simultaneous decisions aimed at maximizing their personal interest.
2. Their decisions are consequences of their belief that the other player shall or shall not undertake one of the two strategies of the game.
3. Both John and Annie assign probability \(\sigma\) to their beliefs that their opponent shall undertake one or the other strategy.

Therefore, if players issue simultaneous decisions dependent on their current beliefs, all probabilities \(\sigma\) associated with the belief of each player can be considered independent probabilities (Spaniel 2014; Bonanno 2018).

For the sake of simplicity in notation, we shall now change names to the strategies in Table 1. For Annie, the possible strategies are Left or Right, whereas, for John, the possible strategies are Up or Down. Given that preferences stay unchanged, we can now rewrite the information in Table 1 in the following way:

<table>
<thead>
<tr>
<th>The battle of the sexes</th>
<th>Annie</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Left</td>
</tr>
<tr>
<td>John</td>
<td></td>
</tr>
<tr>
<td>Up</td>
<td>2,1</td>
</tr>
<tr>
<td>Down</td>
<td>0,0</td>
</tr>
</tbody>
</table>

Table 2 - The battle of the sexes with changed strategy names

If players hold expectations based on their current beliefs, their utility is expected. We define the latter as \(EU = \sum_{i=0}^{k} u(x_k)p_k\) for some finite \(k \in \mathbb{N}\), where \(u(x_k)\) is the utility assigned to each outcome and \(p_k\) is its probability to occur. On this basis, we compute the probabilities \(\sigma\) that players individually assign to each pair of strategies in Table 2. Namely, given the definition of a pure Nash equilibrium strategy, we know that if players make simultaneous decisions, they shall choose their best response strategy based on what other players shall do. Nonetheless, if players only have expectations about what the other player shall do or shall not do, their final expected utility depends on all the probabilities \(\sigma\) that each player assigns to each outcome.
On an individual basis, though, Annie’s decision to choose one between \{Left, Right\} depends to the probability \(\sigma\) she assigns to the fact that John may or may not choose the strategy Up. Instead, John’s decision to choose one between \{Up, Down\} depends to the probability \(\sigma\) he assigns to the fact that John may or may not choose the strategy Left. Thus, we obtain two simple systems of linear equations with two identical solutions:

\[
\begin{align*}
\text{Annie} & = \left\{ \begin{array}{l}
U_{\text{left}} = \sigma_{\text{up}}(2) + (1 - \sigma_{\text{up}})(0) \\
U_{\text{right}} = \sigma_{\text{up}}(0) + (1 - \sigma_{\text{up}})(1)
\end{array} \right. \\
\text{John} & = \left\{ \begin{array}{l}
U_{\text{up}} = \sigma_{\text{left}}(1) + (1 - \sigma_{\text{left}})(0) \\
U_{\text{down}} = \sigma_{\text{left}}(0) + (1 - \sigma_{\text{left}})(2)
\end{array} \right.
\]

\[
\begin{align*}
\sigma_{\text{up}} & = \frac{1}{3} \\
\sigma_{\text{down}} & = 1 - \sigma_{\text{up}} = \frac{2}{3} \\
\sigma_{\text{left}} & = \frac{2}{3} \\
\sigma_{\text{right}} & = 1 - \sigma_{\text{up}} = \frac{1}{3}
\end{align*}
\]

Although the existence of identical probabilities assigned to the beliefs each player enacts coordination between them, it also reduces the expected utilities for each player. Importantly, the laws of probability theory tell us that independent probabilities cannot be summed but must be multiplied. If we translate the independence of probabilities into the formula of expected utility, we obtain the following:

\[
EU_{\text{John}} = (\sigma_{\text{up}})(\sigma_{\text{left}})(U_{\text{up}}) + (\sigma_{\text{down}})(\sigma_{\text{left}})(U_{\text{down}}) + (\sigma_{\text{up}})(\sigma_{\text{right}})(U_{\text{up}}) + (\sigma_{\text{down}})(\sigma_{\text{right}})(U_{\text{down}}); \\
EU_{\text{John}} = \left(\frac{2}{3}\right)\left(\frac{1}{3}\right)(2) + \left(\frac{2}{3}\right)\left(\frac{2}{3}\right)(0) + \left(\frac{1}{3}\right)\left(\frac{1}{3}\right)(0) + \left(\frac{2}{3}\right)\left(\frac{1}{3}\right)(2) + \left(\frac{1}{3}\right)\left(\frac{2}{3}\right)(1) = \frac{2}{3}.
\]

\[
EU_{\text{Annie}} = (\sigma_{\text{left}})(\sigma_{\text{up}})(U_{\text{left}}) + (\sigma_{\text{left}})(\sigma_{\text{down}})(U_{\text{left}}) + (\sigma_{\text{right}})(\sigma_{\text{up}})(U_{\text{right}}) + (\sigma_{\text{right}})(\sigma_{\text{down}})(U_{\text{down}}); \\
EU_{\text{Annie}} = \left(\frac{1}{3}\right)\left(\frac{2}{3}\right)(1) + \left(\frac{2}{3}\right)\left(\frac{2}{3}\right)(0) + \left(\frac{1}{3}\right)\left(\frac{1}{3}\right)(0) + \left(\frac{2}{3}\right)\left(\frac{1}{3}\right)(2) = \frac{2}{3}.
\]

It follows that, all things considered, \(EU_{\text{John}} < 1 < 2\) and \(EU_{\text{Annie}} < 1 < 2\). Therefore, this mechanism of coordination through belief coordination does not maximize the personal interest of each player. It rather makes it worse. Thus, if we assume that John and Annie can exercise reason as Hayek (1973) and Mill (1859) suggest, the produced outcome does not make improve their overall situation. Furthermore, all things considered, the solution
described in this section does not exclude that John and Annie shall spend the evening alone and gain utility 0. Namely, it does not resolve conflict between them.

**Third solution: Correlated Equilibrium**

The last solution of the battle of the sexes incorporates the Hobbesian covenant into our game. Thus, we shall now assume that Annie and John want to avoid spending the evening alone. So, they flip a coin and decide what to do depending on whether tail or head turns up. If head turns up, they will go see the fight. Instead, if tail turns up, they will go to the ballet. In this case, our game has a correlated equilibrium (Aumann 1974; Lasaulce and Tembine 2011). Namely, by flipping a coin, John and Annie avoid spending the evening by acting in agreement with the randomized signaling from the coin.

It follows that if our game has a correlated equilibrium, the randomizer assigns *objective* probabilities to the possible outcomes of the game. In contrast, if the game has a mixed strategy Nash equilibrium, players only assign some probability to their beliefs. In other words, in the previous section, the probabilities are *subjectively* assigned to each outcome by players. Thus, we claim that a randomizer resolves conflict because it enhances welfare for the players. Let us now test this hypothesis.

In the last section, players subjectively assigned two identical sets of probabilities to the possible outcomes of the game, i.e., *Annie*: $\begin{pmatrix} 1 & 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}$ and *John*: $\begin{pmatrix} 2 & 1 \\ 3 \\ 4 \end{pmatrix}$. Given this information, we can now compute the expected utility of each outcome in Table 1.

- **{Fight, Fight}**: $EU_{John} = \left( \frac{2}{3} \right) \left( \frac{1}{3} \right) 2 = \frac{4}{9}$, $EU_{Annie} = \left( \frac{2}{3} \right) \left( \frac{1}{3} \right) 1 = \frac{2}{9}$
- **{Ballet, Fight}**: $EU_{John} = 0$; $EU_{Annie} = 0$.
- **{Ballet, Fight}**: $EU_{John} = 0$; $EU_{Annie} = 0$
- **{Ballet, Ballet}**: $EU_{John} = \left( \frac{1}{3} \right) \left( \frac{2}{3} \right) 1 = \frac{2}{9}$; $EU_{Annie} = \left( \frac{2}{3} \right) \left( \frac{1}{3} \right) 2 = \frac{4}{9}$.

<table>
<thead>
<tr>
<th>The battle of the sexes</th>
<th>Annie</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fight</td>
</tr>
<tr>
<td>John</td>
<td></td>
</tr>
<tr>
<td>Fight</td>
<td>$\frac{4}{9}$</td>
</tr>
<tr>
<td>Ballet</td>
<td>0,0</td>
</tr>
</tbody>
</table>

Table 3 - The expected utilities per outcome
If John and Annie decide to flip a coin, the set of probability is unique for both. That is because tail and head are equiprobable with probability \( \frac{1}{2} \) to turn up. On one side, flipping a coin eliminates the situation in which both players obtain 0 expected utility. On the other side, it eliminates the problem associated with independent probabilities and increases the expected utility of each outcome for each player. In other words, by flipping a coin, John and Annie obtain the following utilities:

- If head turns up, the outcome is \( \{\text{Fight, Fight}\} \): John’s utility is \( \left( \frac{1}{2} \right) (2) = 1 \), and Annie’s utility is \( \left( \frac{1}{2} \right) (1) = \frac{1}{2} \).
- If tail turns up, the outcome is \( \{\text{Ballet, Ballet}\} \): John’s utility is \( \left( \frac{1}{2} \right) (1) = \frac{1}{2} \), and Annie’s utility is \( \left( \frac{1}{2} \right) (1) = 1 \).

<table>
<thead>
<tr>
<th>Coin</th>
<th>Outcome</th>
<th>John</th>
<th>Annie</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>( {\text{Fight, Fight}} )</td>
<td>1</td>
<td>( \frac{1}{2} )</td>
</tr>
<tr>
<td>Tail</td>
<td>( {\text{Ballet, Ballet}} )</td>
<td>( \frac{1}{2} )</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4 - Correlated equilibrium

John’s and Annie’s utilities in Table 4 are lower than their utilities in Table 1. Nonetheless, Table 4 excludes that John and Annie shall spend the evening alone. On the other hand, John’s and Annie’s utilities in Table 4 are higher than their expected utilities in Table 3. Notably, \( \frac{2}{9} < \frac{4}{9} < \frac{1}{2} < 1 \). Furthermore, if we now compute the expected utilities resulting from flipping a coin, we again obtain identical (but higher) expected utilities: \( EU_{John} = 2 \left( \frac{1}{2} \right) + 1 \left( \frac{1}{2} \right) = \frac{3}{2} \); \( EU_{Annie} = 1 \left( \frac{1}{2} \right) + 2 \left( \frac{1}{2} \right) = \frac{3}{2} \). Importantly, \( \frac{3}{2} > \frac{2}{3} \).

**Concluding remarks**

Though not free of simplifications, the above model provides a strong case in favor a top-down approach to the regulation of social media algorithms. Importantly, the above model shows that players generally enjoy greater welfare when they interact following fair rules. Remarkably, this conclusion does not only provide a confirmation of the fact that some
form of Hobbesian covenant is sometimes the only way to escape conflict. It also scores a point in favor of Hayek’s (1947) argument in favor of the view that only lean and fair regulation can enhance the perdurance of a competitive order. In the same way, the existence of lean and fair regulation of the social media algorithms could enhance the perdurance of the online spontaneous order. Thus, there is a possibility for arguing that Hobbes’s and Hayek’s views can find some common ground.

Indeed, as we argued above, the Vienna Convention, the WTO, the GAAP, and the IRS offer lean legal frameworks that sustain the perdurance of the global competitive order. Accordingly, it is arguably true that the introduction of international standards for the regulation of social media algorithms could sustain and enhance the perdurance of the online spontaneous social orders. Expressed otherwise, the existence of such standards could enable people to keep reaping the benefits of the digital revolution, while it could also reduce the sociopolitical risks associated with infodemics. Hence, the Hobbesian lesson is perhaps worth listening to as we live in times when the contemporary seditious doctrines become direct threat to sociopolitical stability in Western liberal democracies.

Bibliography


https://www.earlymoderntexts.com/assets/pdfs/bentham1780.pdf


McRaney, David. You are now less dumb: How to conquer mob mentality, how to buy happiness, and all the other ways to outsmart yourself. Penguin, 2013.


Acknowledgement: I thank dr. Giulio Sciacca for his fundamental remarks and advice on the drafting of this paper.