1 Foreword

This Element is structured as an extended essay intended to be read sequentially but divided into sections that can be read separately. The first three sections clear the ground for the analysis that follows, and are necessary to justify my approach. The next four sections introduce my Supply and Demand Framework and draw connections between relevant literatures and the distinct systems which are part of what I call the YouTube Apparatus. The following section presents my empirical evidence, the final section my suggestions for future research.

My substantive thesis is that demand creates its own supply: The automatically measured and quantified audience causes YouTube creators to adapt their offerings. My empirical strategy is to classify a large sample of political YouTube channels as either affiliated with some larger media entity or as independent – the latter are the classic "YouTube creators," whose personalistic style and intense audience cultivation engenders para-social relationships with their viewers. With over fifteen years of data, I demonstrate that this distinction between affiliated channels and independent YouTubers results in significant system-level differences in a variety of platform-defined audience metrics.

My meta-scientific thesis, however, is that dry science-y prose like the preceding paragraph is the final but perhaps least important step in the study of social media. Setting the academic agenda is paramount; our answers can only be as good as our questions. So I begin by applying theories of media agenda setting reflexively, in order to understand why such disproportionate effort has been applied to the relatively narrow question of the effect of the YouTube recommendation algorithm on viewers' political beliefs.

It is not yet possible to bring comprehensive data to bear on this meta-scientific question. My hope is that beginning to theorize about setting the academic agenda will inspire more systematic investigation. But this area of inquiry is distinct from our normal science work: When we study ourselves, we immediately change ourselves. So for this process to yield benefits in our study of others, we need some sense of our goals, of what we are trying to achieve. It is both untenable and irresponsible to maintain a distinction between positive and normative meta-science.

My goal, then, is to move the academic agenda for studying social media in the direction of my Supply and Demand Framework. Compared to the status quo, this means more attention to the production of content, to the forces that shape the way in which people decide to post. More broadly, my goal is to align the study of social media with *what social media is*. This is why I study

YouTube, the platform for which the ratio of academic attention to social and political importance seems to me to be the most out of whack. Methodologically, I eschew causal inference in favor of pure quantitative description, the most important tool for understanding *what* and *how much* social media is.

As I argue in the final section, another underappreciated method for setting the academic agenda is poetry. The creation and circulation of resonant metaphors carves up the conceptual space in ways that imply certain research questions (and therefore also the answers generated by rigorous normal science). So while the reader may or may not find compelling the poetic flourishes throughout, I hope you understand why I feel they are necessary.

To restate my substantive conclusion in the words of media theorist Vilém Flusser: "Those who participate actively in the production of information … are themselves being programmed by the mass-media meat choppers for information production."

Or more simply, YouTubers are not "Creators" but *Creations*. Audiences, rationalized by the platform, and the vloggers who upload the videos those audiences consumer are not separable either theoretically or empirically. Together, they make up the YouTube Apparatus.

2 Setting the Academic Agenda

To date, research on social media and politics has been somewhat haphazard. Because of the speed of technological change, the sheer scale of our object of inquiry, the necessity of disciplinary and methodological pluralism, and a cascade of policy-relevant crises, our epistemic community is not yet mature. Perhaps this is for the best; some "mature" areas of social science are facing a variety of crises grappling with methodological revolutions, and the creation of new subfields is an integral part of the scientific process.

This generally congenial anarchy has some drawbacks, however. The chain of academic knowledge production is only as strong as its weakest link, and our first link is the weakest. There is a lack of rigor allocated to the selection of *questions* we ask.

There is a meta-irony here. In this Element, I will argue in favor of a Supply and Demand Framework for studying social media, building on the literature in media economics and how these models have understood previous media-technological regimes. I will advocate for the use of these models in the study of both the contemporary media environment and the "media systems" at the heart of a variety of social media platforms.

And I will also, in passing, use these models *reflexively* to explain how our sociotechnical environment has affected the process by which the academic agenda is set.

For example, the theory of "media storms" developed in Boydstun (2013) falls under what I call the "industrial organization" of news media. This means centering the *physical* aspect of the logistical supply chain required for news production, rather than abstracting away from temporality and distance. This approach yields the insight that news is unexpectedly "sticky": The fixed costs involved in paying for reporters' and producers' travel reduce the speed at which news organizations can switch from one story to another. Furthermore, the physical proximity of news teams to a story makes it more likely that they uncover unexpected aspects or follow-ups to that story. Boydstun, Hardy, and Walgrave (2014) argues that "imitation plays a big role in determining the duration of a storm; until another hot item hits, news outlets are loathe to be the first to drop coverage of an ongoing storm, even if the event or issue itself has run its course" (p513).

Like all models, this model should not be trivially applied beyond the context it was developed to explain. To understand "academic storms," we must use our substantive (qualitative and quantitative) knowledge to adjust the parameters Boydstun identifies as important. Two such parameters strike me. The audience for academic journal articles, and the physical and temporal reality confronted by the authors of journal articles, is different than that of general news media and reporters, respectively. The psychological aspect of the drive for imitation, a key part of human nature, remains the same across contexts.

A related example comes from Usher (2014), an ethnographic investigation of the status games within the *New York Times*. I call this work "economic sociology," a crucial refinement of the lived experience of a class of news producers that provides a more realistic picture of the incentives and pressures *they perceive* as most important. She describes the status competition among veteran journalists not to maximize online news sharing or some other immediately relevant economic metric, but to get their bylines on the front page of the print edition. While the baseline approach from Economics would be to proceed with standard assumptions about the agents' utility functions, this more sociological approach begins by actually understanding the day-to-day experience of those actors – undeniably a more accurate micro-foundation, but one which is necessarily more difficult to generalize from.

Without putting too fine a point on it, we likely all appreciate the distance between the nominal goals of academics to generate knowledge or to inform

policy debates and the daily experience of academics embedded in our own kind of status competition.

Matt Hindman's work provides a final example, as well as quantitative models that directly inspire some of the analysis in this Element. Hindman (2008) and Hindman (2018) both argue against two of the fundamental myths of contemporary political communication: that the lack of physical frictions from previous media regimes means that the online media landscape is more *democratic* and that it can *change overnight*.

Rather than analyzing individual actors' preferences, incentives, and information flows, Hindman's macro-level approach makes claims about the system as a whole. In a memorable example, Hindman (2018) argues that Trump-era Federal Communications Commissioner Ajit Pai is incorrect that the contemporary digital news environment requires fewer regulations because "we get news and analysis throughout the day from countless national and local websites, podcasts and social media outlets": "it is flatly not true that there are 'countless' local digital news outlets. We know, *because we counted them*" (p131). Hindman's analysis also demonstrates that the forces of concentration in digital media outweigh the forces of democratization, and that advantages to scale tend to *compound* over time, further reinforcing the advantage of established incumbents. Although the technical capacity to broadcast information has indeed been spread more evenly by digital media, the reality of *attention* as the scarce resource has tended to dominate.

Within academic research, everyone is now technically able to share their research freely and to anyone in the world. Consonant with Hindman's theory, however, the use of Twitter to share published research in Political Science tends to reinforce rather than undercut existing status hierarchies (Bisbee, Larson, and Munger, 2022). And despite the revolutionary potential of online models for academic publishing, peer review at top journals is grinding to a halt as their gatekeeping function has become even more entrenched.

Although I have laid out the applications of these models and the parameters I believe have shifted between the news media and scholarship, I expect that my readers bring a wealth of their own knowledge to this area of meta-scientific inquiry. I have ideas about how the academic agenda is set now and how it might be set differently, some of which I lay out in the following section.

My primary goal, however, is to *put the question of academic agenda setting on the academic agenda*, to argue that this is an important link in the chain of knowledge creation, and to suggest the Supply and Demand Framework as a useful first step.

I am far from the first to make such a claim. Indeed, I draw heavily on Bennett and Iyengar (2008), a high-profile and influential effort to reset the academic

agenda in light of "the changing foundations of political communication." They argue that paradigms developed by the pioneers of the discipline may no longer be relevant, specifically citing "agenda setting" as a broadcast-era theory that has outlived its relevance.

Much of the ensuing debate rested on their provocative title claim of a return to the era of "minimal effects," but this is not my focus. Ironically, debating the validity of old paradigms merely serves to reinforce their relevance: Agenda setting, after all, tells us that the media (here, academic journal articles) is effective at telling us *what to think about.*

So I agree with Bennett and Iyengar (2008) that the shadow of Theory in a dynamic world can obscure social science practice. They predicted "another time of unsettled findings accompanied by the risk of undertheorized sociotechnological conditions" (p708), a prediction which I believe has been borne out. I will pick up on several trends they identify as newly important in the current media environment, an environment that has changed faster and more dramatically than most scholars (and indeed, most humans) thought possible.

In the broadest possible strokes, the "broadcast paradigm" is primarily focused on the following relationship:

Producer → Media → Consumer

There are a fantastic array of refinements that have made this a fruitful area of research, but this relationship is the core of the previous era.

Bennett and Iyengar (2008) argue that the coming era, our era, will be primarily characterized by a rise in the importance of the *exact inverse* of this relationship:

Consumer → Feedback → Producer

and the following prediction by Chaffee and Metzger (2001) seems even more prescient than it did in 2008: "The key problem for agenda setting theory will change from what issues the media tell people to think about to what issues people tell the media they want to think about" (p375).

Of course, both of these phenomena exist; the relationship between audience and creators is bidirectional, reciprocally causal, and embedded in overlapping sociotechnical contexts. I focus on how this relationship develops on each of the social media platforms that have come to dominate online communication; that is, this Element centers the "technology that shapes consumption, distribution, and content production" (p712). Or, as I call it, the Supply and Demand Framework.

The central dynamic of my framework is circular, so it is difficult to fully capture in the linear medium of text or with the unidirectional causality

approaches currently in vogue in much of social science. As much as I love writing and causal inference, our objects of study do not always contort themselves into ideal forms for our preferred models and paradigms. And I believe that both the general media ecosystem and the local environments on a given platform are best understood through my Supply and Demand Framework. This Element provides a detailed application of the Framework to the study of YouTube Politics, one of the most important yet under-studied components of the contemporary US media environment.

As a political scientist, my primary (but far from exclusive) interest is in the effect of YouTube on the beliefs of citizens who consume YouTube videos and who might vote based on the information they receive in doing so. Taking this larger question of "What does YouTube do to American Politics?" as largely fixed, both by disciplinary convention and self-evident importance to a citizen of the United States, my framework can be used to divide up this daunting task into more manageable, empirical study-sized chunks. This is what I mean by "setting the academic agenda."

My hope is that this framework will allow scholars to situate and synthesize existing approaches to studying social media. That is, I do not see this framework as invalidating any existing research, but rather enhancing its value by suggesting pathways of knowledge circulation. In particular, I argue that mapping out all of the potential inputs to the YouTube system is helpful in keeping track of what scholars do and do not yet know, and therefore to decide what they should study next.

This cannot be a deterministic process, and different scholars can of course contribute in different ways based on their preexisting interests and competencies. I will, however, propose a heuristic for researchers deciding on a specific research question:

Heuristic: Study what is under-studied.

This sounds tautological, but it requires keeping two parameters in mind. First, we should spend more of our time studying what is more important; I'm not arguing for an equal distribution of scholarly attention across every conceivable area of inquiry. Note also that this parameter is intrinsically normative: While I will propose some baseline metrics for importance, this parameter is ultimately determined by our collective or respective *goals*.

Second, we need to invest in some empirical measures of how we are *currently* distributing our scholarly attention. Even if we know how much a given topic should be studied, in order to know whether we should study it more or less, we need to know what we're doing now.

Like all human endeavor, this process is not perfectible: Some kind of overreaction to thermostatic correction is inevitable, and the social phenomena

we study are far from static. Worse, these two processes are not independent: Part of the knowledge generated by research is precisely about the importance of a given topic. The edge case of this problem are the "unknown unknowns": When the world produces some novel phenomenon, we start out knowing nothing about it and thus cannot deliberate over its importance or decide to allocate our time to studying it.

As a scholar of digital media for nearly a decade, this very first link in the chain of knowledge production has frequently posed a challenge. My long-standing habits of trawling obscure corners of the internet (now fully rationalized as "exploratory research") reveal some new platform, trend, or ideological current that might be important. How can I check my intuitions – that is, how can I make this case to myself?

More importantly for the academic environment, how can I convince funders and peer reviewers that this new phenomenon is important? In 2015, my first research on Twitter faced significant headwinds from my peers, who did not see why online behavior was important for understanding contemporary politics. Sure, Twitter was important as a tool for protesters and activists spreading tactical information; the Arab Spring and the wave of anti-capitalists protests in the early 2010s had put that topic firmly on the academic agenda. But why, as the old saw goes, should we care about what anyone had for lunch?

3 There Is Something Wrong on the Internet

Despite an explosion of academic interest in digital media over the past decade, the topic is still *dramatically* understudied. If you believe, like I do, that we are living through a media-technological upheaval rivaled only by that occasioned by the printing press, it is difficult to imagine studying anything else.

But even a skeptical, hard-nosed social scientist, beginning with the Aristotelian premise that humans *are* what we *do*, is forced to recognize that digital media production and consumption makes up an increasingly large percentage of the average human's waking hours.

Allen et al. (2020) provides one of the most comprehensive pictures of total media consumption in the United States to date; they find that as of 2018, the average American spends 460 minutes a day (7.5 hours) consuming media. Among the youngest age group, eighteen to twenty-four years old, this number is just under six hours a day (351 minutes); 207 of those minutes are spent on a mobile device, and another fifty-four minutes on a desktop computer.

Unfortunately for traditional methods for measuring media consumption, the diversity of options makes survey enumeration useless for all but the broadest categories. Someone might be able to tell you that they spent fifteen minutes on TikTok, but asking them "which videos" or even "which creators" they watched is not a robust measurement strategy.

The situation on the supply side of digital media is if anything more dire: The number of people empowered to *produce* media has exploded, rendering quantitative comparison absurd. More hours of video media are recorded and uploaded to YouTube and TikTok in a day than were created in the world in a decade during heyday of the postwar broadcast media.

This deluge of content makes the task of setting the academic agenda difficult. Where do we even begin!?

One entry point for many scholars is simply to use our intuitions, developed from our personal experience with digital media. This was a reasonable strategy with broadcast media; Walter Cronkite was in fact broadly representative of the political media diet of the country as a whole.

The rise of the internet changed everything. It allows production and consumption at a scale far beyond mass media, rendering it impossible for an individual to know what everyone else is up to simply by being an avid and broad consumer of political media. This in turn makes the nonrepresentativeness of social scientists' media diets a problem for our understanding of important trends in the media sphere.

We know that social scientists are intensely nonrepresentative of the populations that we study. Despite considerable progress over the previous decades, this nonrepresentativeness persists for demographic groups other than white American men. This legacy of discrimination persists in the inherited boundaries of the canonical areas of study within media and communication, particularly when it comes to the boundary of "political" media (Freelon, Malmer, and Pruden, 2023).

Even if these biases were fully rectified, we would still face a fundamental problem: *By construction*, political science professors are much better educated and more interested in politics, and we have the capacity to consume massive amounts of political media. There is widespread recognition that we need to remind ourselves of these baseline facts – the central insight of perhaps the most important book about American public opinion (Converse, 1964) is that most people simply don't care about politics–but it is all too easy to rely on the intuitions of our colleagues, peer reviewers, and funders.

We are deeply, professionally invested in the technologies of writing, deliberation, and scientific debate. As the sociologist of science Bruno Latour

pointed out, if we simply observe the actions that scientists take rather than accepting our stated higher purpose, it becomes clear that our job is to *read and write papers* (Latour and Woolgar, 2013). Boundary-pushing weirdos in media studies departments aside, we are not primarily invested in watching and recording videos.

In fact, through both introspection and discussion with colleagues, it is clear that many academics actively can't stand to consume information in video format; we'd much rather read it. The average person, unfortunately for the academic agenda, is the exact opposite: They dislike reading, and have abandoned text-based digital media like blogs as soon as possible. Twitter, despite being perhaps the best-studied social media platform, has never been used by more than 25 percent of US adults. Many professors have noticed a steep decline in the writing ability of incoming undergrads. People who have been teaching for forty years noticed a slight degradation over the decades, but the past ten and especially five years have seen a collapse.

The academic agenda has yet to adapt to the reality that the majority of time spent on social media is spent on non-textual platforms. For a variety of reasons, some legitimate others merely omphalocentric, we spend far too much time studying Facebook and especially Twitter.

The media is how we learn about the world outside of our own experience (Mutz, 1998). We rely on the media to tell us about important trends in other realms of social life. The problem is that academics and media professionals inhabit a mutually constituted and reinforcing echo chamber.

The term "echo chamber" here is not used casually. This *phrase* has been disproportionately effective in setting the academic agenda for studying social media, especially in the first decade after its birth. "Echo chamber" simply rings true, a fundamentally *poetic* achievement that is upstream of quantitative research. One of the ways the academic agenda is set, then, is through ***resonant metaphors***; pun on "echo chamber" intended and not at all incidental. For example, Simon and Camargo (2023) investigate the usage and implications of the "infodemic" metaphor that emerged to describe misinformation around COVID-19.

Considerable energy has been spent investigating the phenomenon of echo chambers. My summary of the literature is that it finds that they don't exist – except among users in specialized (partisan or professional) networks.

What threshold of diversity of media diet would be sufficient to falsify the existence of "echo chambers?" No one denies that online media diets are highly skewed. But the threshold can't just be greater than 50 percent congruent information; that's trivial given any amount of media choice. For now, let's consider a definition based on the offline baseline: Online echo chambers

exist if the online news consumers get a higher percentage of their news from congruent sources than do offline news consumers.

Using this conception, the most comprehensive peer-reviewed paper on the topic, Guess (2021) concludes that "if 'echo chambers' exist, they are a reality for relatively few people who may nonetheless wield disproportionate influence and visibility in society" (p1007). Writing in a Knight Foundation White paper, Guess and co-authors did not feel the need to be quite so circumspect: "public debate about news consumption has become trapped in an echo chamber about echo chambers that resists corrections from more rigorous evidence" (p15).

This realization is unpleasant for diehard Popperians who still believe that science progresses through falsification, that there could exist a "critical experiment" that would convince social scientists to stop saying the words "echo chamber." The meta-scientific approach I advocate here requires that we understand science as to a significant degree a *social* process. We started saying "echo chamber" for nonquantitative reasons; it is therefore implausible to expect quantitative evidence to convince us to stop saying "echo chamber." Falsification simply does not describe the process of social science (Feyerabend, 1975).

In addition to resonant metaphors, another important way in which the academic agenda is set is the sudden eruption of some social media phenomenon from outside of academics' direct experience into the public consciousness; in short, through ***media panics***.

The most dramatic example is of course the avalanche of academic research on online mis/disinformation or "Fake News" in the wake of the 2016 US Presidential Election. This makes sense; democratic societies around the world continue to grapple with this novel epistemic problem, one that threatens the foundations of the liberal reason at the heart of our political system.

We now recognize the magnitude and scope of the problem posed by online misinformation, tragically reinforced by the immediate physical harm caused by COVID-19 vaccine denialism. But we might productively ask why it took the shock electoral victory of Donald Trump to get the topic pride of place on the academic agenda. Though Trump was undoubtedly an innovator in the form, it is difficult to imagine that US-based academics would be spending quite as much time studying misinformation in the Global South today if Hillary Clinton had won a few hundred thousand more votes in Michigan, Pennsylvania, and Wisconsin.

Online misinformation, and even anti-vaccine narratives, were certainly topics of research before 2016 (Bode and Vraga, 2015; Cook, Ecker, and Lewandowsky, 2015; Garrett and Weeks, 2013; Kata, 2010, 2012). They were