



This is a repository copy of *Algorithmic governance : a modes of governance approach*.

White Rose Research Online URL for this paper:
<https://eprints.whiterose.ac.uk/183926/>

Version: Published Version

Article:

Gritsenko, D. and Wood, M. orcid.org/0000-0002-0433-2972 (2022) Algorithmic governance : a modes of governance approach. *Regulation & Governance*, 16 (1). pp. 45-62. ISSN 1748-5983

<https://doi.org/10.1111/rego.12367>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:
<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Algorithmic governance: A modes of governance approach

Daria Gritsenko 

Aleksanteri Institute and Helsinki Center for Digital Humanities (HELDIG), University of Helsinki, Helsinki, Finland

Matthew Wood 

The Department of Politics, University of Sheffield, Sheffield, UK

Abstract

This article examines how modes of governance are reconfigured as a result of using algorithms in the governance process. We argue that deploying algorithmic systems creates a shift toward a special form of design-based governance, with power exercised *ex ante* via choice architectures defined through protocols, requiring lower levels of commitment from governing actors. We use governance of three policy problems – speeding, disinformation, and social sharing – to illustrate what happens when algorithms are deployed to enable coordination in modes of hierarchical governance, self-governance, and co-governance. Our analysis shows that algorithms increase efficiency while decreasing the space for governing actors' discretion. Furthermore, we compare the effects of algorithms in each of these cases and explore sources of convergence and divergence between the governance modes. We suggest design-based governance modes that rely on algorithmic systems might be reconceptualized as algorithmic governance to account for the prevalence of algorithms and the significance of their effects.

Keywords: algorithm, automation, design-based governance, modes of governance.

1. Introduction

In 2010, a New York-based entrepreneur tried to improve office life for corporate clients in the city by designing a computer program – WeWork – that would optimize work conditions by re-designing office space (Anderson *et al.* 2018). Because human architects' recommendations are sometimes poorly conceived, they thought a computer program might be able to do architectural design work with better results. WeWork is an algorithm – a set of computational decision rules into which you can put all the data you have about the office space you need designing – that comes out with office table designs to optimally maximize all the goods (sunlight, space, access) you want from your open plan office in downtown Manhattan. The idea that a computer algorithm can be used to create conditions for improved social processes – in the case of WeWork, improved office environment for more productive work – has gained increased policy attention in the past decade. Such an attempt to design social and economic processes through the use of algorithmic systems is the focus of this article.

Today, algorithms are a part of everyday life, but their impacts are only beginning to be discussed. From social media feeds and tailored search engine recommendations to airport flight scheduling, credit checks, and financial trading decisions, algorithms have become virtually indispensable for analysis and decisionmaking in our data-saturated environments (Gillespie 2014). Algorithms in public governance have been advocated as a powerful tool to increase efficiency and effectiveness of government through a blend of automation and new modes of coproducing public services (O'Reilly 2011; Margetts & Dunleavy 2013; Williamson 2014). Proponents have noted tangible benefits for collective action enabled by algorithm-driven technologies, for example, in making quick decisions that can improve economic productivity and facilitate social interaction (Athey 2017) and even solving pressing social issues, such as low-carbon energy transition (Zhou & Yang 2016) or preventing physical checks on goods crossing the land border on the Island of Ireland following 'Brexit' (Botton 2018). Other

Correspondence: Matthew Wood, The Department of Politics, University of Sheffield, Elmfield, Northumberland Road, Sheffield, S10 2TU, UK. Email: m.wood@sheffield.ac.uk

Declaration of conflict of interest: There are no conflicts of interest to be reported.

Accepted for publication 5 October 2020.

examples, such as the use of bots in elections, the spread of fake news, and “citizen scoring” systems, have raised critical questions, including who is accountable for their operation, whether they can create injustices and erode civic norms, and how we should resolve their (un)intended consequences (Pasquale 2015; Yeung 2018). The implications of accelerated deployment of algorithmic systems within public life informed a renewed interest toward the role of new technologies across disciplines, leading to the emergence of an interdisciplinary body of literature on *algorithmic governance* (AG, for an overview see Danaher *et al.* 2017).

Although there is no common definition, most authors situate AG at the intersection of digitalization, datafication, and governance through technology (Wilsdon 2001; Danaher *et al.* 2017). The literature highlights how digitalization creates information assets beyond the processing capacity of human intelligence (often referred to as Big Data), while offering an opportunity to “get things done” through computer-programmed tools (software) powered by a set of logical rules and commands (algorithms). In these complex computational systems, algorithms define the ways in which software mobilizes information assets to reach particular ends (Gillespie 2014).

Algorithmic forms of governance have been scrutinized in various social science disciplines, including sociology and Science and Technology Studies (STS) (Aneesh 2009; Beer 2017; Cheney-Lippold 2017), media and communication (Bucher 2012; Gillespie 2014; Just & Latzer 2017), and socio-legal studies (Pasquale 2015; Yeung 2018). However, political science, including the regulation and governance scholarship, provides only a limited contribution to theorizing AG (Campbell-Verduyn *et al.* 2017; Yeung & Lodge 2019). We argue that to enable this latter project, conceptual links are required between well-documented modes of governance in political science, including hierarchical, self-governance, and co-governance (Kooiman 2003) and “design-based” governance identified in behavioral economics concerned with constructing “choice architectures” that improve decisionmaking (Madrian 2014; Thaler 2015) and in the legal studies literature on the regulatory power of the internet (Lessig 1996; Yeung 2011). To be clear, we do not argue that these modes of governance are distinctly new. Rather, we argue that analyzing them through a common framework can help situate the power and impacts of algorithms. Within this framework, “algorithmic governance” does not exist yet, and we do not claim it does. We wish to take a step back from “treating the phenomenon monolithically” (Eyert *et al.* this issue, 4) to specifying more carefully how algorithms impact on modes of governance, and specifically to suggest the dimensions along which their effects are particularly significant.

We analyze three cases documenting prominent governance problems – car speeding, media disinformation, and the sharing economy – and show that when governing actors use algorithms, they contribute to an evolution of the modes of governance. The three examples we choose traditionally have been governed through “paradigmatic” examples of hierarchical (speeding), self-governance (disinformation), and co-governance (sharing economy) processes. These processes are increasingly incorporating algorithms. Our analysis shows that by introducing “design-based” elements, algorithmic systems alter the hybridity of governance modes already present within these cases. While this impact of algorithms is consistent, the effects vary. Strikingly, our analysis suggests that algorithms have most significant effects on co-governance even though algorithms are often seen as enabling highly complex network-like interactions. Since the impact of increasingly complex and ubiquitous software on governance will be growing in the next decade, we suggest that an analytically separate mode of “algorithmic governance” would serve as a useful conceptual tool to analyze the evolution of governance in the digital age.

The article proceeds in four sections. First, we critically review existing literature showing the increasing importance of algorithms, highlighting the added value of a political science approach using modes of governance. Second, we deploy three cases illustrating the emergence of hybrid modes of hierarchical, self-governance, and co-governance. Third, we compare the effects of algorithms on these evolving governance modes and explore sources of convergence and divergence. We conclude by highlighting the implications of algorithms in governance and invite empirical and methodological development.

2. Algorithms in governance

2.1. Setting the scene: The “social power of algorithms” debate

Strictly speaking, algorithms are “encoded procedures for solving a problem by transforming input data into a desired output” (Yeung 2018, p. 506), but scholars in humanities and social sciences often use the term as a

shorthand for forms of computational and cybernetic systems, as “multiples” (Seaver 2017) enacted through our interactions with these systems. While the “social power of algorithms” has been long recognized, especially in STS, the ways of analyzing how algorithms become intertwined with society, politics, and culture draw on (at least) two distinct analytical traditions (Neyland & Möllers 2017).

The first line of theorizing dates back to Langdon Winner (1980), who suggested that technological artifacts are “materialized laws” (Lee & Björklund Larsen 2019), hence, they have politics and capacity to structure the social world. A more nuanced version of this argument suggests that algorithms are instruments used to exert power, shape reality, and, ultimately, (re)construct social order (Just & Latzer 2017). This perspective, applied to algorithms (e.g. Lessig, 2009; Bucher 2012; Pasquale 2015), highlights the power and normativities designed into them (Lee & Björklund Larsen 2019). An alternative conceptualization draws on sociologist Latour’s (2005) insight that power does not precede human interaction with technology, but rather results from “assemblages” of humans and technological artifacts. The power of algorithms is achieved through associations and realized situationally (Neyland & Möllers 2017). This perspective regards socio-technical systems as a product shaped by many “actants,” so that conceiving of algorithms apart from the governance system as a whole or judging their effects upon the governing actors is not plausible.

Our approach comes from political science, which privileges the generation of analytically distinct categories to examine the outcomes of power relations between different socio-economic forces within governance processes (*who gets what, when, and how*). This approach, which has not, to date, been prominent in studies of algorithmic governance, takes algorithms as *analytically distinct* socio-economic forces, conceptualized as tools in the governance process, which can be used by governing actors to achieve desired governance outcomes (Howlett 2018). Algorithms can be deployed as tools in steering governance processes, which may substantively shift the power relations within those processes and alter distributive outcomes (see Johns and Compton on “data jurisdictions,” this issue, for an alternative approach). As described by Latzer and Festic (2019, p. 2), algorithmic governance comprises “the intentional and unintentional steering effects of algorithmic-selection systems.” The key question, for us, is how algorithmic systems affect the main features characteristic of different governance processes, leading to new configurations and, arguably, new governance modes. We therefore focus on *algorithms as tools of governance, rather than the regulation/governance of algorithms* (for the latter, see Bellanova and de Goede; Gellert, this issue).

2.2. Studying algorithms in public governance

At base, we contend, governance refers to the coordination of social interactions aimed at addressing collective action problems. The ‘governance’ we are talking about is public, meaning, related to the production and implementation of ideas, plans, regulations, and policies that concern the general public, as opposed to corporate governance that happens within private organizations. In addition, our perspective on governance is interactive, rather than state-centered, and sees governance as a *process* of patterned/orderly – not chaotic – interactions between individuals and collectives structured by the frameworks within which they operate (Kooiman 2003). Diverse structures – defined as enduring institutional arrangements, such as bureaucracies, networks, and markets – emerge from and are sustained through ongoing governance processes.

The literature has pinpointed three distinct challenges that algorithmic systems pose to public governance. The first challenge arises from automated decisionmaking. Discrimination, information distortion, and “echo chambers” are among the examples of the so-called automation bias, including errors of omission and commission (Skitka *et al.* 2000), stemming from both technical imperfections, such as biased algorithms and data (Kay *et al.* 2015; DuPont 2017), and from depriving human agents of judgment, reason, and intervention (Zarsky 2016; Festic, this issue). In the instance of echo chambers, individuals may have wrong assumptions about the distribution of public opinion on an online discussion, because social media feeds only provide information tailored to what the individual (conceived as a consumer of information) “wants,” while an individual may assume that all relevant data are being presented to them. This is a clear example of automation bias because the result is the use of “decision aids” (in this case, the use of news feeds to make decisions about public opinion, political campaigns, and the like) in biased ways (Skitka *et al.* 2000, p. 703).

The second challenge relates to invisibility of algorithmic systems and opacity of their operations. Pasquale (2015) claimed that we live in a “Black Box” society where invisible algorithms trump the public interest due to corporate secrecy. Burrell (2016) portrayed opacity as a persistent characteristic of algorithmic systems that manifests at different levels and requires both technical and non-technical solutions to prevent harm. The adverse effects of algorithmic systems are increasingly problematized in relation to acute social (Eubanks 2018) and political (Benkler *et al.* 2018) problems of contemporary societies. Finally, issues related to ethics and values have been raised, challenging assumptions about the neutrality of algorithmic systems. With search algorithms as a prime example, it has been shown that choices made by algorithm designers translate their values into the system, resulting in a variety of intended and unintended consequences with ethical implications (Roscoe 2015; Latzer *et al.* 2016; Mittelstadt *et al.* 2016). These dynamics not only change the relationship between state and society mediated by “big data,” software tools and algorithms, but may effectively undermine democratic governance through increased reliance on technocratic “solutionism” (Williamson 2016).

The literature also emphasizes that algorithmic governance is not a unified phenomenon. Typologies of the role of algorithms in governance distinguish between reactive and pre-emptive algorithmic systems (Yeung 2018), as well as governance through, by, and with algorithms (Campbell-Verduyn *et al.* 2017). The variety within algorithmic forms of governance has been associated with algorithmic selection, both top-down and bottom-up (searching, sorting, matching, filtering, and so on, see Just & Latzer 2017); the degree of human participation in the governance process (in the loop, on the loop, or out of the loop, see Citron & Pasquale 2014); and with the characteristics of Big Data systems that enable governance processes (Danaher *et al.* 2017, Kitchin and McArdle, 2016). These approaches allow scholars to debate how algorithms are used by private corporations and governments, and the normative consequences of algorithms in governance (see Kosta; Kharlamov and Pogrebna, this issue). Strategies have been also suggested for how to leverage the use of algorithms in governance (Just & Latzer 2017; Kitchin 2017).

Some existing studies do point to the empirical distinctiveness of “algorithmic governance.” Aneesh (2006) was first to argue that computer algorithms enable a qualitatively new form of governance, significantly different from familiar forms of hierarchy and self-governance, naming it “algocracy.” His in-depth ethnographic study of computer programmers in India working for the American software industry showed that the key organizing structure connecting globally dispersed workers was the programming code rather than a bureaucratic managerial structure. He emphasized that thinking of algocracy as “the software version of bureaucracy” is misleading since “there is no common meta-language shared by bureaucratic (legal code) and algocracy (binary code) systems of governance” (Aneesh 2009, p. 355). Following Aneesh, Yeung (2018) recognized algorithms as distinct forms of social ordering, alongside markets, and bureaucracies, showing how algorithmic forms of social ordering might be associated with certain ideological commitments. Among recent contributions, Just and Latzer (2017) conceptualized AG as governance by “software as an institution,” which, at least theoretically, permits an idea that algorithms could be placed alongside market, hierarchical, and network modes of governance. Latzer and Festic (2019) then defined AG as “the automated assignment of relevance to certain selected pieces of information and a focus on internet-based applications”.

To date, political science has not contributed significantly to this research agenda. It tends to see a set of modes of governance – hierarchy, co-governance, and self-governance – without reference to algorithms, Big Data, and the possibility of computational “design-based” governance as a distinct mode of governance, which has been discussed under the term “design-based regulation” by Law scholars (Yeung 2008, 2011; Yeung and Bygrave, this issue). Hence, it remains unclear how to situate algorithmic systems vis-à-vis dominant modes of governance, which constitutes a gap in relation to governance theory developed in political science.

2.3. Algorithmic systems and the modes of governance approach

Since collective action can be coordinated in a myriad of ways, actors who are engaged in solving collective action problems develop “go-to” ways of doing so (developing rules and norms, communicating ideas, solving conflicts, and so forth). This insight has been captured early on through the typology of hierarchy/bureaucracy, markets, and networks developed as a way of conceiving different “settings” of relationships between social actors aimed at coordinating solutions to collective action problems (Powell 1990). The scholarship applying this typology

focused on exploring how different institutional frameworks play a regulative function in governance interactions – especially in policy sciences (Rhodes 2007; Bevir 2013). The ideal types of “hierarchies,” “markets,” and “networks” played an important role in theorizing authority, effectiveness, accountability, and power relations in state–society relations within policy processes (Pahl-Wostl 2019).

The modes of governance approach has been revised and further developed, especially by Kooiman (2003) and Kooiman & Jentoft (2009). Following the sociological idea of structuration (Giddens 1984), the tradition of studying governance as coordination process sought to integrate both functional (rule/structure-based) and agential dimensions. This led to the idea of a “governance mix” – governance arrangements, such as bureaucracies, markets, and networks, are always a result of combining different logics and operating principles. Kooiman (2003) proposed to distinguish between governance modes on the basis of their “locus.” He summarized three modes of governance as follows:

- *Hierarchical governance* is where “those governing are, or see themselves as, superimposed upon those governed” (Kooiman 2003, p. 115). The key to hierarchical modes of governance is not that they are bureaucratic (also self-governance and co-governance can be bureaucratic), but that those who govern see themselves and behave like they are “governing” in a top-down manner, and privilege compliance with rules or laws they author.
- *Self-governance* is “wherever (corporate) governing actors, representing different societal domains, are able to organize networks in which they combine resources from those domains for common purposes” (Kooiman 2003, p. 83). Self-governance occurs within organizations, or among actors, to essentially “take care of themselves, outside the purview of government” (Kooiman 2003, p. 83). It often emerges in response to regulatory vacuum when, for instance, major resource users or producers in a market, make, and adopt rules over time.
- *Co-governance* is where “societal parties join hands with a common purpose in mind, and stake their identity and autonomy in the process” (Kooiman & Jentoft 2009, p. 821). There are multiple examples of co-governance but these include, in particular, collaboration, co-operation, co-management, public–private partnerships, and other governing arrangements that involve a common purpose. Co-governance is associated with sharing power, responsibility, and information. Co-governance is similar to self-governance in that networks of actors do not rely on hierarchical intervention to impose single rules on the system, but different in that co-governance can involve collaboration with government and non-government actors. The rules and norms within co-governance are also commonly designed and managed through governments “meta-governing” or coordinating from a distance, with acceptance by those involved (Sørensen & Torfing 2009).

Each mode of governance is distinct because it has a specific way of exercising control, setting incentives, and solving conflicts. In short, the rules and norms, defined as formal and informal institutions that limit and enable interactions, have a *different logic*. For hierarchy, this logic revolves around control and steering mechanisms imposed by the “governor” on the “governed,” for co-governance the logic is around cooperation and coordination based on reciprocity and social capital, and in self-governance the logic is focused on responsibility and self-respect within an organization or for the individual. Each domain within a typology remains an ideal type – not a sphere of activity or specific governance structure – and in reality, governance cases are most likely to cut across and draw upon these modes in complex ways.

The typology of “modes of governance” has been criticized for its simplistic and rigid character (e.g. Treib *et al.* 2007) and re-worked to include new forms of relationships between economic agents (Demil & Lecocq 2006), citizens, and governments (Klein 2007), emerging in the “digital age.” At heart of this new organizing logic is design, or attempts to direct individual interactions in an indirect way, by influencing their choice architecture, inducing rather than prescribing behavior (Murray & Scott 2002; Yeung 2008, 2011). Thaler and Sunstein’s (2008) now classic Nudge is an exemplar of the organizing logic of this mode of governance. Nudging techniques in the form of “informal” or “experimental” tools have long been used by state governments, private corporations, and even voluntary associations to achieve desired ends, leveraging insights from social psychology to affect the behavior of the “governed” without their knowledge. More recently, “boost” policy interventions, aiming at equipping the individual with the competence to overcome their cognitive biases, have been added to

the tool kit of design-based governance (Hertwig & Ryall 2020), as this vibrant field is being further extended to the realities of the digital age (Esmark 2019).

Crucially for our purposes, design-based governance has grown in salience as digitalization, datafication, and algorithmic systems are creating the conditions under which vast amounts of information can be mobilized to design choice architecture (Yeung 2017). Lessig (2009) persuasively argued that the code of cyberspace creates powerful architectures that control what is possible, affecting human behavior on the Web. Anesh (2006) showed that any software has a structuring effect on human behavior, importantly, on “offline” behavior. Inputting large amounts of data about human interactions into algorithmic systems, governing actors can provide governed subjects with a set of options for action that are tailored to predictions about their desires and interests in real time. This has a structuring effect similar to a hierarchy in the sense that “rules” are passed down by governing actors (although this does not have to be exclusively “the state”), but governed actors are not punished (legally) for violating the recommended courses of action. Yeung (2017, p. 119) argued that while the concept of “nudge” is very simple, algorithmic systems allow for it to rely on complex computational models for predicting and modeling interactions and to be applied in a *dynamic* manner, or what she calls a “hypernudge.” Summing up, design-based governance can be added to the hierarchical, self-governance, and co-governance devised by Kooiman as a distinct mode of governance: not simply a tool, but an organizing logic that implies both functional and agential dimensions.

Table 1 presents a two-dimensional typology of the four modes of governance interactions identified in the mainstream political science literature. This conceptual visualization of two salient dimensions of governance emerges from the literature, but is also grounded in our case analysis, elaborated in the following section. The first dimension describes how rules and norms come along – whether they are prior to the decisions taken or leave a leeway for the parties to come to terms *ex post*. The key question relates to decisions on the rules that prescribe and sanction behavior to ensure conflict is minimized. Each mode of governance offers distinct ways of addressing collective action problems via strategies for ensuring rules are in place and followed. This can be either via *ex ante* legal or administrative rules or a “choice architecture” that imposes solutions derived by a few actors (as in hierarchical and design-based governance), or *ex post* norms and practices that emerge interactively to direct actors towards common goals (in co-governance and self-governance).

The second dimension of our table captures a varying degree of *commitment of parties* to continue problem-solving in the future, a key requirement for reconciliation of individual and public/collective interests. By “commitment” we mean how much time and resources governing actors use in engaging in transactions to coordinate solutions to the collective action problem. This can be by monitoring and supervising other actors, or mediating and facilitating interactions to address a governance problem. Productive interaction between diverse actors requires not only clear rules for resolving conflicts but also credible future commitments (Scharpf 1997). Although there may be a rule for solving a disagreement, the level of commitment to solving the problem in the future has a significant impact upon the outcomes. In using “commitment,” we follow institutional economists like Ronald Coase and Douglass North who have deployed this concept to understand what leads actors to address collective problems despite having individual interests.

Table 1 allows us to illustrate the distinct logic of design-based governance. It is different from hierarchical governance because it does not involve direct orders or laws telling governed subjects what to do, choices are designed-in (Thaler & Sunstein 2008), and in the case of algorithms, we can say “programmed-in” (Brownsword,

Table 1 Revised modes of governance

		Commitment by governing parties	
		Medium/high	Low
Rules and norms	<i>...imposed ex ante</i>	Hierarchical governance <i>Monitoring by legal authority</i>	Design-based governance <i>Indirect influence on choice</i>
	<i>...emerging ex post</i>	Co-governance <i>Trust and reciprocity</i>	Self-governance <i>Self-control and exchange</i>

Source: Adapted from Powell (1990) with permission.

2005, p. 3). It is not the same as co-governance because it does not involve co-operation or collaboration in a horizontal manner. Indeed, the governed subjects in this case do not know they are being governed, or as described by Lessig, “Regulation is possible, but through different means” (Lessig 1996, p. 1406). Finally, it is not the same as self-governance because the “governor” is consciously attempting to influence the behaviors and interactions of governed subjects, albeit tacitly through organizing action space, framing information, and structuring decision processes in certain ways, exploiting common “cognitive biases” and making non-government actors “surrogate regulators” (Brownsword, 2005, p. 5). The logic here is that governance occurs through *(re-) designing actors’ choice architecture* rather than through adherence to laws, common interests and identities, or individual interest and self-control. Such manipulation may be benign or malign, but it is analytically distinct at a structural level to the top-down control or steering of society, collaboration and cooperation, and self-control.

We argue that algorithmic systems are creating the conditions under which design-based governance, seen as a separate *mode* from hierarchy, co-governance, or self-governance, brings significant changes into existing processes for solving collective action problems. Unlike speed humps, a classic example of design-based governance, the growing sophistication of algorithmic systems and scale of datafication hand substantial power to those who design and deploy algorithmic systems (Yeung 2017). The capacity of a single algorithmic “nudge” to “directly affect millions of users simultaneously” (Yeung 2017, p. 123), indicates that, down the line, algorithmically enhanced design-based governance has the potential to alter our understanding of governance modes altogether.

3. Algorithmic governance in practice

This section shows how integrating algorithmic systems in addressing particular collective action problems transforms key aspects of governance in three different cases. Defining the object of governance as a problem that requires “governing” (Mayntz 1993), we examine three examples – the governance of vehicle speeding, “disinformation” on social media, and social sharing. These are “paradigmatic” (Flyvbjerg 2006) examples of Kooiman’s three “modes” of governance (hierarchical, self-governance, and co-governance). Paradigmatic cases are selected intuitively with the aim to “develop a metaphor or establishing a school for the domain that the case concerns” – in this case, a metaphor for how including algorithmic systems transforms traditional modes of governance (Flyvbjerg 2006, p. 230).

In this instance, we deliberated over case selection with three criteria in mind: (i) each case must focus on a salient *governance problem* and the *process* through which it is addressed; (ii) the cases must have a clear locus, or set of governing actors who see the effects of using algorithmic systems, which maps onto Kooiman’s governance modes (the state, firms, and networks of individuals), and (iii) the cases must have each seen the *increased* use of algorithmic systems over the past 10–20 years, including both “simple” systems producing a series of options from which humans select actions, and machine learning (ML) systems that adapt and respond automatically to inputs. The case studies match all three criteria. In what follows we use these cases principally as examples of our theoretical argument – as paradigmatic cases are often used. We draw on the existing research base and data from existing studies to illustrate our theoretical points in each mode of governance about the shifting dynamics of power within these modes of governance. Each mode of governance is a *hybrid* mixing different combinations of hierarchical rules, market and structural incentives, and inter-organizational networks. We analyze how this “hybridity” shifts as algorithmic systems are introduced into the governance process and relate this back to our theoretical argument about design as an organizing logic resulting from introducing algorithms in each case.

3.1. Algorithms and hierarchical governance: The governance of speeding

The first case captures implications of algorithms for hierarchical governance. One of the most common examples of hierarchy – with “top-down,” *ex ante* legal regulation – is speeding on motorways. The rules are designed by the government and are known in advance by all car drivers, which is ensured through obligatory schooling. A single authority – the traffic police force – is legally empowered to monitor and sanction those who infringe the speed limit shown on motorway signs. Drivers are disciplined through fines and sanctions (like suspension of their license) enforceable through court proceedings. In this case, the logic is to set rules within which drivers are

monitored for compliance and implement a legal bureaucracy to sanction infringement. So long as drivers operate within hierarchically defined rules, they may drive where, when, and how they like.

Historically, speeding has been governed through a hybrid range of tools that are design-based and hierarchical. In terms of hierarchical governance, speeding has been an offense going back to the design of the first modern roads in the 1800s. In the United Kingdom, the 1865 *Locomotive Act* placed restrictions on traffic speed at four miles per hour, revised upwards in 1896 before the *Motor Car Act 1903* set the limit at 20 miles per hour. As it became clear that police could not continually monitor speeding offenses on the ground, experts created innovative ways of designing technological solutions. Nobel Prize-winning Physicist Arthur Holly Compton designed speed bumps in 1953 as a way of slowing down traffic close to Washington University, St. Louis, Missouri, where he worked as chancellor. Speed bumps are an archetypal example of design-based governance, as they alter the decision architecture available to the driver – either slowing down or risking damage to their car. As suggested by Yeung (2017), historical attempts at “nudging” went hand in hand with top-down bureaucratic steering.

In the past couple of decades advances in technology have allowed speeding to be monitored and even sanctioned through automated systems using algorithm-driven databases to identify speeding cars and issue fines. A 2008 study of speeding governance across Europe for the European Commission showed that in Lithuania, Czech Republic, Poland, and Greece there was a singularly manual process for registering speeding vehicles, while in Austria, Switzerland, and Sweden there was a mixed system with some automated data collection through number plate recognition systems (Larsen *et al.* 2008). In Cyprus, there was a shift from a manual to an automated system with number plate recognition. The United Kingdom had the most systematic automated system of number plate recognition, with algorithms used for integrity checks and only a minimal manual element where necessary.

Larsen *et al.* (2008) showed how under an automated system, there are significant differences to non-automated ones. In the United Kingdom and Cyprus, speeding violations are dealt with differently to those recorded by the police, compared to countries where algorithms are secondary to manual forms, or not used at all (Larsen *et al.* 2008, p. 56). In the United Kingdom, a “fixed penalty notice” – an administrative fine – is sent to the household recorded as being the registered keeper of the vehicle. This can then be contested through court if the driver has sufficient grounds for objection. By contrast, violations brought by the traffic police are tried through the court system. In Cyprus, there is a similar system whereby violations are registered by police and an automatic “fine” sent out to owners, who can “declare” the fine attributed to a different driver. Only if the driver chooses to contest the fine in court do authorities present evidence for their case.

This subtle shift is important because it demonstrates a difference between manual case-by-case and algorithmic law enforcement systems. In both United Kingdom and Cyprus, fines are issued by police based on recorded violations of rules communicated through traffic signs, but the “fixed penalty notices” (FPNs) are closer to an administrative/financial tool than legal punishment. FPNs are requests for payment that mirror an invoice from a company requesting recompense for a service. They establish a system that motorists *themselves* may contest in court, much like a market situation where exchanges can be contested if one or other actor thinks they do not conform to the rules of the market. They are not immediate “sanctions,” although such sanctions may be enforced later in a process if not recompensed immediately. In other words, supervision and enforcement is a *secondary rather than primary* goal. The first resort – determined by the algorithm – is an administrative fine that is sent out to vehicle owners, who can pass this “payment” to others in their household.

As a result, efficiency of payment is privileged over disciplinary procedure. This logic is shown, for example, in the justification of the benefits of automated systems for processing FPNs. They are justified as administrative tools for raising funds and reducing costs, as well as being aimed at creating order in the highly complex road networks common to advanced industrial democracies (Larsen *et al.* 2008, p. 15). The role of governance actors in hierarchy – establishing a structure within which the moral and legal disciplining of those who transgress rules is exercised – is not given primacy in the case of FPNs, save for in more serious cases where speeding causes accidents and even deaths.

Since 2008, there has been a further shift toward automated methods for tracking and even controlling the speed of cars (Schulz & Dankert 2016). Driverless vehicles, for example, automatically limit car speed according to not only the legal limit, but also relative distance from other vehicles, as well as environmental factors like

weather conditions and visibility (Stilgoe 2018). Were driverless cars more widespread, we would see an even more fundamental shift of this sort. The scenario would be close to an “ideal type” of top-down governance, whereby car speed is controlled almost universally by automated systems with fixed *ex ante* rules, and human involvement in addressing the problem of speeding would only be needed at the stage of algorithm design (and then if the algorithm malfunctions). Instances of speeding would therefore be attributable to the algorithm embedded within the driverless cars, either due to it malfunctioning or being overridden by human drivers.

3.2. Self-governance with algorithms: The governance of disinformation in online media

The second case focuses on the governance of disinformation in online media to analyze the effects of algorithms on self-governance. Disinformation, recently referred to as “fake news,” is increasingly seen as an important problem with the proliferation of social media (Dommett 2020). Aside from being a form of manipulation, disinformation is argued to lead to “group-think,” anxiety, and exclusion of balanced views on economic, social, and cultural questions (Allcott & Gentzkow 2017; Mihailidis & Viotty 2017). This is claimed to create a ripe environment for extremism and perils for democratic politics (McNair 2017).

A reflection on the relationship between self-governance and algorithmic systems is useful here to clarify the emergent relationship theoretically. In *Machine Dreams*, Mirowski (2002, p. 539) suggested that “markets... resemble computers, in that they take various quantitative and symbolic information as inputs, and produce prices and other symbolic information as outputs.” Mirowski stated that, in contrast to traditional “neoclassical” conceptions of market actors, who are “cognitive agents” making conscious utility maximizing choices, markets can also be seen as “automata,” preprogrammed operating systems without human agency and choice. Under conditions where computational programming is introduced, we can see a shift from the former to the latter. Once actors in a market have preselected goals (for example, profit maximization), they can move from having to cognitively assert their need for utility maximization, to allowing it to dictate all interactions particularly in complex globally interconnected market scenarios. Theoretically speaking, algorithms enable a shift in the way actors in a market scenario behave when self-governing their behavior – from the conscious, neoclassic utility maximizer to the logically automated “reproduction” of utility maximizing systems from one space to another. In so doing, algorithms may alter the way in which self-governing actors in a market scenario *control themselves* – turning previously cognitive agents into “automata”, or self-programming entities adjusting (in a predesigned fashion) seamlessly to external factors (Mirowski 2002). Algorithms have created this dynamic since the 1980s in the international stock markets. In 2006, “algotrading” – trading of stocks, shares and derivatives done purely by preprogrammed algorithms – made up as much as 80% of American stock exchange transactions, and arguably contributing significantly to the 2008 global financial crash (Dodson 2008). Our analysis focuses on a case where algorithms have risen to prominence more recently: the governance of disinformation.

The governance of disinformation in the United States is a pertinent example of self-governance complemented recently with algorithms as tools. Traditionally, the media in the United States – defined broadly to include broadcast media and newspapers – has been largely self-regulating. Journalistic standards are set by professional bodies including the US Society of Professional Journalists (SPJ), which has existed since 1909. As a voluntary standard-setting organization, the SPJ is a good example of self-governance, which Quinn (2016, p. 116) explicitly referred to as “self-governance and professional accountability”. The SPJ produces a code of ethics, including commitments to “seek truth and report it”, “minimize harm,” and “accountability and transparency.” In doing so, it aims to ensure ethical reporting and prevent the spread of willful disinformation campaigns. The application of these rules is down to individual news outlets (<https://www.spj.org/ethicscode.asp>). Crucially, these rules are not enforceable under the First Amendment to the US Constitution. Disinformation by journalists is governed by the industry itself – professional colleagues will, in the SPJ’s words, “encourage the exposure of unethical journalism as a means for rooting it out” of the market. One of the strengths of such self-governance, exemplified in transnational governing capacity (Kirton & Trebilcock 2017), is that it is rather flexible and can adapt to changing circumstances quickly because it does not require a complex legislative procedure, but is enacted through mutual adjustment. In other words, rules are embedded *ex post* rather than *ex ante*.

The entrenchment and growing sophistication of “Web 2.0” has created new dilemmas of how to ensure information consumed by the public is accurate, and willful disinformation campaigns are “rooted out” and “exposed”.

Social media networks like Facebook or Twitter and search engines like Google provide information their users have been shown to take as credible facts (Westerwick 2013), but because of their algorithm-driven “feeds” determined by user search histories or paid-for advertising, this can lead to a miasma of “news” either poorly sourced or actively aimed at misleading the “consumer.” In 2016, online disinformation campaigns during the US election campaign created completely false stories marketed as “news” and directly consumed by millions of social media users. The most infamous was “pizza-gate,” a conspiracy theory linking the Democratic party to an alleged pedophile ring run out of a New York pizza shop (Kline 2017). This story was completely fabricated, but seen by millions of users and was just one of a variety of “viral” fake stories that became known as “fake news.”

The fake news scandal ramped up pressure on social media corporations to create better ways of sifting disinformation. In their infancy, Twitter and Facebook had community “terms of use” setting broad rules about what users on the platforms were allowed to do. These rules were enforced by the companies themselves on a case-by-case basis. If a user was reported by other users as violating the “terms of use,” they would be investigated by a dedicated in-house team *post hoc*. If shown to have violated these terms, the user would be banned (their account blocked and email unable to re-register again with the site), or in more serious cases their details would be passed to relevant state authorities. Since 2015, however, and particularly in light of this scandal, more stringent *ex ante* approaches have been applied, using algorithms. Facebook, for example, has developed new “Community Standards” integrated into its “Facebook feed” algorithm to automatically remove content deemed offensive (Jørgensen 2017). In 2017, Twitter integrated details in its algorithm about the source of paid-for tweets and those relaying explicitly political messages. It has also introduced algorithm-generated warnings about potentially offensive content in tweets and accounts. Previously the companies had allowed any form of information to be disseminated on their platforms, and regulated *ex post*, whereas the use of algorithms to filter or warn users about disinformation marks a significant shift in the governance of disinformation toward *ex ante* rules embedded through algorithms.

The case of disinformation is useful for showing the effects of algorithms on self-governance, specifically because it shows how *ex post* discretion is substituted by *ex ante* design through automation. Although algorithms presuppose a great degree of formalization of rules via a single, universally applied code, self-governance involving algorithms is still about managing and ordering a complex range of individual social interactions through “soft” tools. Social networks like Facebook “attempt to steer users’ activities in a certain direction” (Van Dijck 2013, p. 144), using algorithmically determined rules that, while legally nonbinding, are “inbuilt” and very difficult to circumvent. With the shift to unified “Community Standards” integrated *within* algorithms, the companies do not have everyday discretion over what information they decide to admit on their sites – because it would require them using that discretion manually rather than through the algorithm. The algorithm, once it has been designed, filters *all* of the information seen by users on the basis of the *ex ante* rules created to do so. Rule violations are no longer addressed predominantly after the fact of disinformation emerging. Rather, outcomes are directed by the algorithms, sorting and serving news in accordance with preprogrammed company preferences.

Increased use of algorithms leads to a shift towards design-based solutions – algorithms are designed *ex ante* and include all possible scenarios while, at the same time, their rule cannot be transgressed. This has led to some incongruities with user experiences. For example, users complained Facebook’s algorithm has inappropriately blocked campaigners seeking to highlight racist language in publications on its platform (Guardian 2016a). A “chat bot” created by Microsoft to manage community standards on Twitter disintegrated into hate speech on the basis of it learning from other users (Guardian 2016b). Similar to the “high-handed” approach of a legalistic, hierarchical mode of governance, the “spooky” (Jørgensen 2017) singular authority of algorithms has evaded even the control of tech companies themselves, who have found it conflicting with their model of profit-maximizing competition. Facebook, for example, responded to criticisms of its effects on the mental well-being of users by introducing new algorithmic rules to privilege the posts of users’ friends and family, demoting paid-for content by corporate advertisers. In this sense, we can see some similarities between the conflicts experienced within self-governance and hierarchical governance when algorithms gain prominence as ways of setting rules.

3.3. Co-governance with algorithms: Sharing economy platforms

Co-governance is a type of interaction characterized by power sharing and interdependence. Structurally, it is manifested through a horizontal network in which, unlike in a vertically structured hierarchical network, multiple

nodes are connected with each other. From the agential side, this governance mode includes multiple actors sharing information and resources and ensuring collaborative goals are met. The size of the network is important: trust and reciprocity require negotiation, mutual adjustment, and knowledge about others' credibility, placing a limit on the number of actors involved. Castells (2004) showed that while there is nothing "new" about networked forms of organization per se, the scale and intensity of networks created by modern Information and Communications Technologies (ICTs) led to the emergence of a network society, a new social structure enabled by the digital technology. While it may be tempting to conceive of the "network society" as a co-governing structure, cultural theorist and philosopher Galloway (2004) has made a strong case for how control can function "after decentralization" using the case of Internet protocols. Modern ICTs, he argued, allow the existence of distributed networks that are "highly controlled despite having few if any central points of control" (p. 25). Protocols described by Galloway do not materialize their power through orders and sanctions, as laws would, but rather through the uniformity in design, that permits only a certain type of organizing. He compares protocols to speed bumps on the street – material design elements that have strong effect, and the specificity of the digital world allows speed bumps "on steroids."

Although Internet protocols affect social organization by designing communication between distributed computers, a similar logic can be revealed through the analysis of "sharing economy" platforms that enact algorithmic systems to allow communication between "distributed individuals" to solve everyday household problems, such as commuting, staying overnight, or assembling a piece of furniture. Before the advent of smartphones and other powerful hand-held devices, everyday sharing problems were typically resolved through personal networks – family and friends, colleagues, or friends of friends (an "acquaintance network") – usually on a non-monetary basis. These social favors provided on consensual basis were based on co-production and the norms of reciprocity, sharing of cost, and responsibility – all key elements of co-governance. Yet, the rise of digital platforms offering access to goods and services "on the go" of the "sharing economy," introduced a new dimension into social sharing: a transactional one.

Well-known "sharing economy" platforms include providers of transportation services (Uber, Lyft), accommodation and vacation rentals (Airbnb), mediating work (TaskRabbit, UpWork, Amazon Mechanical Turk), niche retail (Etsy, various forms of crowdsourcing), meal sharing (MamaBake), goods re-use (Freecycle, earlier Craigslist), and many more. In order to match consumers ("problem-owners") with workers for short-term tasks ("problem-solvers"), platform operators use sophisticated algorithms. Importantly, digital technology creates a possibility to extend "sharing" practices previously confined within a person's immediate social network to a large community, leading to so-called "stranger sharing" (Rowe 2017, p. 4). Instead of hosting a friend based on social norms of hospitality, you may now get money for hosting someone via Airbnb and instead of giving a lift to a colleague earn with Uber during your commute. "Sharing, bartering, lending, trading, renting, gifting, swapping" (Rowe 2017, p. 3) are conventional forms of social interaction, but algorithmic systems make "sharing economy" distinct from traditional forms of problem-solving in social networks by designing the architecture in which semi-anonymous interactions take place. While "stranger sharing" dramatically increases efficiency (one can get an Uber ride while you neighbor may be busy), the problem "object" is now a matter of exchange performed at a distance from face-to-face negotiation. Algorithms play a central role in the sharing economy aiming at making sharing economy platforms more efficient. For example, Uber combines data on customer location, driver availability, weather and traffic conditions, and more, in order to algorithmically connect a user and a driver and set the ride price. Airbnb gives more control to hosts who can decide whether to accept or decline a guest, but its algorithms still structure what they see based on previous searches (e.g. on Google). Amazon Mechanical Turk algorithmically matches humans and jobs, as well as establishing pricing.

This may create unintended consequences. Apart from allegations surrounding Uber and Airbnb (sexual assault, privacy breaching, and other serious criminal offenses), there are also cases of phishing, free-riding, and other types of social dilemmas (Martin *et al.* 2015; Hartl *et al.* 2016; Rowe 2017). Rule-based mechanisms have been created to prevent conflict emerging in these instances. Most successful cases of the "platform economy" involve a tech company acting as an intermediary, whose primary function is to manage distrust (Hartl *et al.* 2016). In addition to maintaining the platform, the platform operator may set prices (Uber), enforce rating and recommender systems (Uber, Airbnb, UpWork) and engage in other rule-making activities embedded within their algorithms. These include introducing compulsory profile pictures, authentication, and document

verification presented as trust-building. Even Wikipedia – a free online encyclopedia created by volunteers and regarded as a prime example of network at scale – has developed these arrangements. By mediating exchange, platforms not only set the “do’s and don’t’s” prior to an interaction but also reduce the need for sustained commitment to monitor and enforce these norms.

This is very different from community and user-based platforms that do not employ algorithms to enhance their performance. Platforms like MamaBake (Rowe 2017) or Freegle (Martin *et al.* 2015) preserve the primary features of co-governance. They provide a convenient structure for interested parties to meet and engage with each other directly (via messaging, posting boards), without algorithmic intermediation. While as with any software, these platforms will be to an extent guiding their users’ interactions “by design,” they do not affect commitment and do not try to algorithmically substitute trust.

4. Understanding the impact of algorithms in governance

The abovementioned examples show how algorithms install design features into the existing governance modes. When algorithmic systems are implicated in social coordination for collective problem-solving, they modify existing governance by altering the way in which rule-making and rule-taking work in practice. Algorithmic systems combine *ex ante* specified principles, the key characteristic of hierarchies, with a low level of commitment from human actors in this process. As a result, algorithms can have both complementary and adverse effects on existing governance arrangements.

4.1. Protocol, conflicts, and sanctioning

The key to solving collective action problems is reconciling individual and collective interests. Making everyone work toward the same end is no easy task and conflicts often arise, requiring rules of some form. One of the key characteristics of algorithmic tools is their capacity to design-in solutions via *protocols* (Galloway 2004). Their design incorporates various solutions mediating potential conflicts (e.g. rule-breaking) even before those conflicts arise. In this regard, algorithms promise to strengthen hierarchical governance, while an evolution toward *ex ante* rules in self-governance and co-governance affects their core mechanism of conflict mediation.

In the speeding case, we see a hierarchically defined set of rules being replaced with predefined “protocols.” With the introduction of automated speeding cameras in the United Kingdom and Cyprus, decisions over when and where to issue speeding fines became the responsibility of software programs. These essentially took the rules-based process previously undertaken by police tracking vehicles and constabulary issuing fines and court orders, and replaced it with a similar process, albeit one designing these rules more closely and into a pre-programmed computational procedure (so long as the cameras did not break) with greater efficiency.

In the case of self-governance of disinformation, potential conflicts have traditionally been mediated by private news companies self-regulating, with conflict effectively avoided via pricing strategies and competing for market share. In short, disinformation would be avoided because companies that provided inaccurate or misleading stories would lose their ethical status as “proper journalism” and go bust. With the introduction of new algorithmic tools for news feeds on social media sites flagging suspicious content, these norms of self-governance have developed toward designed-in protocols about the quality of information on news and social media sites. Large conglomerates like Facebook and Twitter now have similar algorithmic tools that regularize newsfeed content based on predetermined rules about what content might be a form of misinformation/disinformation and requires regulating.

In the social sharing case, we see a similar change from informal norms and practices to explicitly designed protocol. In co-governance, conflict is essentially mediated via interpersonal interactions, where reputation and personal connections effectively regulate who is more likely to share with whom. The reliance on informal norms typical for collaboration is substituted through code developed by platform operators. In some cases, prices are algorithmically defined rather than negotiated. In other cases, monetary sanctions become protocol for those who change their intention to interact, and cancel their Uber trip or AirBnB stay. There is thus a contrast between “real” co-governance where building and maintaining relationships is costly, and actors tend to choose voice over

exit in solving problems (Powell 1990), while in sharing economy platforms, there are few incentives to engage in informal conflict resolution; switching to another platform is easy.

In all three cases, therefore, we can see a convergence toward rules designed and automated as *protocol*, rather than informally developed through norms. In the hierarchical mode, the algorithm sits on top of a set of rules to automate existing procedures and implement the law. The challenge in translating hierarchical rule-making into algorithmic systems is that technical protocol design is never purely technical, but implies values that are at times hidden within the system (Yeung 2018). However, the most significant changes are brought to the self- and co-governance modes where information and negotiated norms and practices are being complemented – or even replaced – by formalized algorithmic solutions. The locus of rule-making changes here, and there is an attempt to avoid rule-breaking altogether by preventing transgression of rules “by design” (Treleaven *et al.* 2019). As a result, self-governance and co-governance modes, reliant on flexibility through ongoing mutual adjustment, may compromise this advantage with the introduction of algorithms. Rules are increasingly centralized in one place – the algorithmic system – and the actors involved in collective problem-solving rely on them, in extreme cases devoid of personal judgment. Unlike human actors, who “reflexively refine their decision criteria as they reason through a novel situation,” even the most sophisticated machine learning algorithms can only refine the contours of their decision boundary *ex post*, in other words, “reflexivity can only occur after the system receives feedback or additional training data” (Alkhatib & Bernstein 2019, pp. 1–2). Hence, the adaptive capacities of ML algorithms are limited by decision boundaries effectively established *ex ante*, since all currently existing algorithms by virtue of design “require feedback or additional training to update their learned models” (Alkhatib & Bernstein 2019, p. 10).

4.2. Communication, relationships, and commitment

Productive interaction between diverse actors requires not only clear rules for resolving conflicts but also credible future commitments (Scharpf 1997). While there may be a rule for solving a disagreement, the level of commitment to solving the problem in the future has a significant impact upon the outcomes. Both in hierarchical and co-governance, “relationships matter and previous interactions shape current ones” (Powell 1990, p. 302), while in self-governance interactions are free of future commitments (Powell 1990, p. 302). As a result, self-governance is best suited to address straightforward problems that require low levels of investment where speed and ease of information processing is key.

Algorithmic systems further increase the speed and ease of information processing, as they aggregate discrete transactions into Big Data that can, in its turn, be further processed to infer new system-level knowledge and even predict future interactions. The system of personal relations does not play the central role in assuring credibility, since actors interact with the platform that presents them with a collection of individual experiences as a proxy for trustworthiness. As a result, governance is “automated” and requires *less active commitment* from governing actors than would be the case without algorithmic systems. It is important not to overstate the autonomy of algorithms – they are created, deployed, monitored, and revised by human agents. Nevertheless, algorithms have the effect, and usually the explicit intent, of removing the need for direct commitment from governing actors to coordinating responses to governance problems. No one needs to rely on *someone* else for direction, but everyone needs to rely on *something*, the algorithm, to be able to coordinate their actions.

This intent of automating human discretion is particularly the case in our hierarchical and co-governance examples. In the speeding example, without algorithms to assist in catching and disciplining speeding motorists, the police have high levels of autonomy in how they act. At the same time, however, this requires intensive top-down monitoring of motorists, for example, police physically stationing themselves on roads, stopping and questioning motorists, and manually recording and passing on cases. With automated speeding cameras, this intensive commitment – time and resources – is reduced significantly.

In the case of social sharing, introducing algorithms into everyday networking reduces the extent to which individuals have to commit to building personal and professional networks. In the example of Uber, whereas previously individuals collaborated with others to, for example, find a trustworthy taxi company, by using the app they bypass the need to intensively build and maintain collaborative arrangements, and rely on the function of the algorithm-driven platform to match them with an available driver at the right price. Here, the core function of co-governance – the intensive commitment of actors networking with each other to “match” providers and

consumers – is displaced as the algorithm essentially automates this function. Future credibility does not rely on actors' commitments, but rather on algorithmic performance (and thus to passengers and drivers, e.g. having confidence in Uber or Lyft's algorithms). With the growth of algorithm-driven apps for social sharing, reputation becomes a quantitative marker – often a star “rating” – that facilitates the sharing process. Individual networked assessments of reputation do not disappear, rather they are amalgamated into a unified marker of quality. Members of the network form their preferences according to the “rating” they observe, rather than through recommendations from family or friends.

Unlike hierarchical and co-governance, in modes of self-governance relationships matter less. In the disinformation case, companies do not have a sustained, high level of commitment to stopping disinformation and each of them acts independently from the others. It is assumed that standard editorial practices within the market will prevent disinformation because success in the market depends on ethical best practices. With shifts in the market environment, and the growing issue of “fake news,” news outlets, including social media companies, have adopted efficient algorithmic tools for checking and flagging factual accuracy. In this case, companies like Facebook have sought to integrate computational solutions for checking news sources and accuracy via algorithms, rather than doing this manually. The level of commitment in this case does not change significantly: the introduction of this algorithm aims to improve the social media firm's competitiveness as a content provider. Social media companies explicitly hope the state *will not* step in to impose new rules requiring intensive manual oversight and reporting (BBC News 2019). In sum, algorithmic tools have an impact on the ways actors communicate, build, and maintain relationships, which results in declining interdependence of their choices, but an increasing dependence on technical systems for the capacity to coordinate actions. In particular, with the advent of machine learning, past interactions have an impact on the future, yet, in an opaque, invisible manner.

5. Conclusion: Convergence toward algorithmic governance?

This article has argued that introducing algorithms into traditional modes of governance creates significant, and potentially transformative, alterations in core features of those modes of governance. From comparing three cases – the governance of speeding, disinformation and social sharing – we observed that algorithms introduced design elements that created similarities between hierarchical, self-governance, and co-governance modes addressing these problems. Specifically, we have shown how deployment of algorithmic systems influences the development and application of rules that prescribe and sanction behavior to address the given problem, and changes the way actors communicate, build, and maintain relationships, thereby affecting their commitment to coordinating responses to the problem. As a result, we argue that algorithmic tools of governance provoke (i) convergence toward conflict being mediated through *ex ante* designed rules and procedures, and (ii) convergence toward lower levels of commitment by governing actors, with the designed-in algorithms taking precedence. In short – algorithms create the conditions for “design-based” governance modes in combination with hierarchical, self-governance, and co-governance modes. This core finding suggests at least two points that requiring further research, and which form the focus of this conclusion.

First, the “convergence” we identify across modes of governance does not imply equal levels of change across each mode of governance. As the discussion section suggests, introducing algorithms imply a more significant shift for some modes of governance than others. In terms of rules for conflict resolution, the changes are more significant for self-governance and co-governance because they involve informal mediation as part of their central logic. However, in a hierarchical mode of governance, where rules are implemented in a top-down manner to mediate conflict, the transformation is not as substantial. In the case of speeding motorists, state authorities already operate under a logic where rules need to be applied directly and equally across the actors being governed, and the algorithms simply make this process more efficient. In terms of levels of commitment, however, the change is more substantial for hierarchical and co-governance modes than self-governance. This is because self-governance operates under a similar logic to algorithms, in the sense that they attempt to make governance an “automatic” function of individual self-restraint under conditions of autonomy from the state – the “logic of the market” (see Hayek 1944). Hierarchical and co-governance modes attempt to govern their object problems more directly via human intervention, involving either intensive law monitoring and enforcement, or informal coordination.

Interestingly, our case relying on co-governance appears to undergo the most radical transformation as a result of algorithms being deployed. As shown earlier, algorithms not only provide market-like efficiency over less efficient collaborative arrangements, they consolidate control within platform operators who *ex ante* design how interactions will function. Hence, algorithms transform important collaborative arrangements: they substitute interpersonal trust with automated ratings and simulate reputation through data-driven analytics. A hybrid co-governance mode with algorithmic tools involves the introduction of *ex ante* rule enforcement mechanisms via automated programming, and a shift from an environment where social norms, such as hospitality, solidarity, or familiarity, all requiring significant level of *commitment*, are replaced with commodification of shared “services.” The extent of similarities and differences requires further empirical investigation, but our analysis suggests initial evidence that algorithms will have a differential impact on the central logic across different modes of governance.

Second, there are important implications for how further research into “design-based” governance is framed. Specifically, we would benefit from altering the language of design-based governance to *algorithmic governance*. Design-based governance has a long history (Lessig 1996) and is not merely concerned with technological tools, but any technique of government used to manipulate information provided to those who are governed. Nevertheless, as we have shown, algorithms create a “designed-in” element that codify rules and operate extremely efficiently to manipulate options and information available for millions of people instantaneously – what Yeung (2017) called “hypernudge.” The question, however, is how far these more radical design-based modes of governance could go, with the “design” element intricately detailed through software and distributed to millions of “users.” We argue here that a shift in language to investigating the potential for a fully-fledged mode of *algorithmic governance* may be useful, rather than subsuming an analysis of algorithms within design-based governance. This is because the ubiquity and power of algorithms can enable cognitive manipulation on an unprecedented scale. As the speeding case shows, while the idea of governing by design is old, this mode of governance is increasingly being advanced through technology – as seen in the shift from speed humps to automated speed cameras. Other scholars may suggest a unique algorithmic mode of governance is unlikely, and algorithms are always used *by somebody to achieve some outcome*.¹ We agree but also suggest that the “design” element may be increasingly of secondary importance to the *algorithmic* element – the software in and through which governance is “designed.” So, whereas the previous generation of scholars studying “nudge,” “boost,” and other design-based solutions may have focused on analogue tools – like speed humps – the centrality of algorithms to our understanding and explanation of contemporary design-based governance may require a shift in language to fully appreciate the seismic shifts they may be inducing in the nature of “design-based” modes in solving governance problems.

Acknowledgments

The authors would like to thank the participants of the Panel “Algorithmic Regulation” at the Biennial Conference of the ECPR Standing Group on Regulatory Governance 2018 for their constructive feedback, as well as the four anonymous reviewers and the Special Issue editors for their criticism and helpful suggestions on how to improve this manuscript. This work has been in part supported by the NOS-HS Exploratory Workshops Grant “Algorithms in Context” (Academy of Finland Dec. No. 327085). Wood acknowledges funding for his time from the Economic and Social Research council (ES/S00730X/1).

Endnote

1 We would like to thank the reviewers of this article for pushing us on this point.

References

- Alkhatib A, Bernstein M. (2019) Street-level Algorithms: A Theory at the Gaps Between Policy and Decisions. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–13. Glasgow, CHI: Conference on Human Factors in Computing Systems. <https://doi.org/10.1145/3290605.3300760>.
- Allcott H, Gentzkow M (2017) Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives* 31(2), 211–236.

- Anderson C, Bailey C, Heumann A, Davis D (2018) Augmented Space Planning: Using Procedural Generation to Automate Desk Layouts. *International Journal of Architectural Computing* 16(2), 164–177.
- Anthony Giddens. (1984). *The constitution of society: Outline of the theory of structuration*. Univ of California Press.
- Aneesh A (2006) *Virtual Migration: The Programming of Globalization*. Duke University Press, Durham.
- Aneesh A (2009) Global Labor: Algoratic Modes of Organization. *Sociological Theory* 27(4), 347–370.
- Athey S (2017) Beyond Prediction: Using Big Data for Policy Problems. *Science* 355(6324), 483–485.
- BBC News (2019) Facebook needs regulation as Zuckerberg 'fails' – UK MPs [Last accessed 18 June 2019.] Available from URL: <https://www.bbc.co.uk/news/technology-47255380>.
- Beer D (2017) The Social Power of Algorithms. *Information, Communication & Society* 20(1), 1–13.
- Benkler Y, Faris R, Roberts H (2018) *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. Oxford University Press, Oxford.
- Bevir M (2013) *A Theory of Governance*. University of California Press, Berkeley.
- Botton N (2018) Blockchain and trade: not a fix for Brexit, but could revolutionise global value chains (if governments let it), No. 1/2018, *ECIPE Policy Brief, Brussels: ECIPE*.
- Brownsword R (2005). Code, control, and choice: *why East is East and West is West*. *Legal Studies*, 25(1), 1–21.
- Bucher T (2012) Want to Be on the Top? Algorithmic Power and the Threat of Invisibility on Facebook. *New Media & Society* 14(7), 1164–1180.
- Burrell J (2016) How the Machine “Thinks”: Understanding Opacity in Machine Learning Algorithms. *Big Data & Society* 3(1), 2053951715622512.
- Campbell-Verduyn M, Goguen M, Porter T (2017) Big Data and Algorithmic Governance: The Case of Financial Practices. *New Political Economy* 22(2), 219–236.
- Castells M (ed) (2004) *The Network Society: A Cross-Cultural Perspective*. Edward Elgar, Cheltenham and Northampton, MA.
- Cheney-Lippold J (2017) *We Are Data: Algorithms and the Making of our Digital Selves*. New York University Press, New York.
- Citron DK, Pasquale F (2014) The Scored Society: Due Process for Automated Predictions. *Washington Law Review* 89, 1.
- Danaher J, Hogan MJ, Noone C *et al.* (2017) Algorithmic Governance: Developing a Research Agenda Through the Power of Collective Intelligence. *Big Data & Society* 4(2), 2053951717726554.
- Demil B, Lecocq X (2006) Neither Market nor Hierarchy nor Network: The Emergence of Bazaar Governance. *Organization Studies* 27(10), 1447–1466.
- Dodson S (2008) Was Software Responsible for the Financial Crisis? *The Guardian*, [Last accessed 6 November 2019.] Available from URL: <https://www.theguardian.com/technology/2008/oct/16/computing-software-financial-crisis>.
- Dommett K (2020) Introduction: Regulation and Oversight of Digital Campaigning-Problems and Solutions. *The Political Quarterly* 91(4), 705–712.
- DuPont Q (2017) Experiments in Algorithmic Governance: A History and Ethnography of “the DAO,” a Failed Decentralized Autonomous Organization. *Bitcoin and Beyond: Cryptocurrencies, Blockchains and Global Governance*. England: Routledge.
- Esmark A (2019) Communicative Governance at Work: How Choice Architects Nudge Citizens Towards Health, Wealth and Happiness in the Information Age. *Public Management Review* 21(1), 138–158.
- Eubanks V (2018) *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York, NY: St. Martin's Press.
- Flyvbjerg B (2006) Five Misunderstandings about Case-Study Research. *Qualitative Inquiry* 12(2), 219–245.
- Galloway AR (2004) *Protocol: How Control Exists after Decentralization*. Cambridge, MA: MIT press.
- Gillespie T (2014) The Relevance of Algorithms. In: Gillespie T, Boczkowski PJ, Foot KA (eds) *Media Technologies: Essays on Communication, Materiality, and Society*, pp. 167–193. The MIT Press, Cambridge, MA.
- Guardian (2016a). Facebook Blocks Shaun King, [Last accessed 15 October 2020.] Available from URL: <https://www.theguardian.com/technology/2016/sep/12/facebook-blocks-shaun-king-black-lives-matter>.
- Guardian (2016b). Microsoft Deeply Sorry for Offensive Tweets by AI Chatbot, [Last accessed 15 October 2020.] Available from Accessed 15 October 2020 URL: <https://www.theguardian.com/technology/2016/mar/26/microsoft-deeply-sorry-for-offensive-tweets-by-ai-chatbot>.
- Hartl B, Hofmann E, Kirchler E (2016) Do we Need Rules for “what’s Mine Is Yours”? Governance in Collaborative Consumption Communities. *Journal of Business Research* 69(8), 2756–2763.
- Hayek FA (1944) *The Road to Serfdom*. University of Chicago Press, Chicago.
- Hertwig R, Ryall MD (2020) Nudge Versus Boost: Agency Dynamics under Libertarian Paternalism. *The Economic Journal*.130(629), 1384–1415.
- Howlett M (2018) Matching Policy Tools and Their Targets: Beyond Nudges and Utility Maximisation in Policy Design. *Policy & Politics* 46(1), 101–124.
- Jørgensen RF (2017) What Platforms Mean When They Talk About Human Rights. *Policy & Internet* 9(3), 280–296.
- Just N, Latzer M (2017) Governance by Algorithms. *Media, Culture & Society* 39(2), 238–258.
- Kay M, Matuszek C, Munson SA (2015) Unequal Representation and Gender Stereotypes in Image Search Results for Occupations. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 3819–3828. Seoul: ACM.
- Kirton JJ, Trebilcock MJ (2017) *Hard Choices, Soft Law*. Routledge, London.
- Kitchin R, & McArdle, G. (2016). What makes Big Data, Big Data? *Exploring the ontological characteristics of 26 datasets*. *Big Data & Society*, 3(1), 2053951716631130.
- Kitchin R (2017) Thinking Critically about and Researching Algorithms. *Information, Communication and Society* 20(1), 14–29.

- Klein A (2007) Judging as Nudging: New Governance Approaches for the Enforcement of Constitutional Social and Economic Rights. *Columbia Human Rights Law Review* 39, 351.
- Kline J (2017) C.G. Jung and Norman Cohn Explain Pizzagate. *Psychological Perspectives* 60(2), 186–195.
- Kooiman J (2003) *Governing as Governance*. Sage, London.
- Kooiman J, Jentoft S (2009) Meta-Governance: Values, Norms and Principles, and the Making of Hard Choices. *Public Administration* 87(4), 818–836.
- Larsen, L., Vavakos, V. & Zaidel, D. (2008) *Comparison and analysis of traffic enforcement chains across EU Member States and in relation to EU policies*, Police Enforcement Policy and Programmes on European Roads Project Deliverable 6.
- Latour B (2005) *Reassembling the Social*. Oxford University Press, Oxford.
- Latzer M, Festic N (2019) A Guideline for Understanding and Measuring Algorithmic Governance in Everyday Life. *Internet Policy Review*, 8(2), 1–19.
- Latzer M, Hollnbuchner K, Just N, Saurwein F (2016) 19. The Economics of Algorithmic Selection on the Internet. *Handbook on the Economics of the Internet*, p. 395. University of Zurich, Zurich.
- Lee F, Björklund Larsen L (2019) How should we theorize algorithms? Five ideal types in analyzing algorithmic normativities. *Big Data and Society*, 6, 1–6. [Last accessed 15 October 2020.] <https://doi.org/10.1177/2053951719867349>.
- Lessig L (1996) The Zones of Cyberspace. *Stanford Law Review* 48, 1403–1411.
- Lessig L (2009). Code: And other laws of cyberspace. ReadHowYouWant.com.
- Madrian B (2014) Applying Insights from Behavioral Economics to Policy Design. *Annual Review of Economics* 6, 663–688.
- Margetts H, Dunleavy P (2013) The Second Wave of Digital-era Governance: A Quasi-paradigm for Government on the Web. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 371(1987), 20120382.
- Martin CJ, Upham P, Budd L (2015) Commercial Orientation in Grassroots Social Innovation: Insights from the Sharing Economy. *Ecological Economics* 118, 240–251.
- Mayntz R (1993) Governing Failures and the Problem of Governability: Some Comments on a Theoretical Paradigm. *Modern Governance: New Government-Society Interactions*, pp. 9–20. Sage, London.
- McNair B (2017) *Fake News*. Routledge, London.
- Mihailidis P, Viotty S (2017) Spreadable spectacle in digital culture. *American Behavioral Scientist* 61(4), 441–454.
- Mirowski P (2002) *Machine Dreams: Economics Becomes a Cyborg Science*. Cambridge: Cambridge University Press.
- Mittelstadt BD, Allo P, Taddeo M, Wachter S, Floridi L (2016) The Ethics of Algorithms: Mapping the Debate. *Big Data & Society* 3(2), 2053951716679679.
- Murray A, Scott C (2002) Controlling the New Media: Hybrid Responses to New Forms of Power. *The Modern Law Review* 65(4), 491–516.
- Neyland D, Möllers N (2017) Algorithmic IF...THEN Rules and the Conditions and Consequences of Power. *Information, Communication & Society* 20(1), 45–62.
- O'Reilly T (2011) Government as a Platform. *Innovations: Technology, Governance, Globalization* 6(1), 13–40.
- Pahl-Wostl C (2019) The Role of Governance Modes and Meta-governance in the Transformation towards Sustainable Water Governance. *Environmental Science & Policy* 91, 6–16.
- Pasquale F (2015) *The Black Box Society*. Harvard University Press, Cambridge, MA.
- Powell WW (1990) Neither Market nor Hierarchy: Network Forms of Organization. *Research in Organizational Behavior* 12, 295–336.
- Quinn A (2016) Co-regulation and Anti-corruption in US Journalism. *Journal of Media Ethics* 31(2), 116–129.
- Rhodes RAW (2007) Understanding Governance: Ten Years on. *Organization Studies* 28, 1243–1264.
- Roscoe PJ (2015) A Moral Economy of Transplantation: Competing Regimes of Value in the Allocation of Transplant Organs. In: Dussauge I, Helgesson C-F, Lee F (eds) *Value Practices in the Life Sciences and Medicine*. Oxford University Press, Oxford.
- Rowe PC (2017) Beyond Uber and Airbnb. *Social Media + Society* 3(2), 2056305117706784.
- Scharpf F (1997) *Games Real Actors Play. Actor-Centered Institutionalism in Policy Research*. Westwood press, Boulder, CO.
- Schulz W, Dankert K (2016) Governance by Things' as a Challenge to Regulation by Law. *Internet Policy Review* 5(2), 1–20. <https://doi.org/10.14763/2016.2.409>.
- Seaver N (2017) Algorithms as Culture: Some Tactics for the Ethnography of Algorithmic Systems. *Big Data & Society* 4(2), 2053951717738104.
- Skitka LJ, Mosier K, Burdick MD (2000) Accountability and Automation Bias. *International Journal of Human-Computer Studies* 52(4), 701–717.
- Sørensen E, Torfing J (2009) Making Governance Networks Effective and Democratic through Metagovernance. *Public Administration* 87(2), 234–258.
- Stilgoe J (2018) Machine Learning, Social Learning and the Governance of Self-driving Cars. *Social Studies of Science* 48(1), 25–56.
- Thaler RH (2015) *Misbehaving: The Making of Behavioral Economics*. W.W.Norton, New York, NY.
- Thaler RH, Sunstein CR (2008) *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven, CN: Yale University Press.
- Treib O, Bähr H, Falkner G (2007) Modes of Governance: Towards a Conceptual Clarification. *Journal of European Public Policy* 14(1), 1–20.
- Treleaven P, Barnett J, Koshiyama A (2019) Algorithms: Law and Regulation. *Computer* 52(2), 32–40.
- Van Dijck J (2013) Facebook and the Engineering of Connectivity. *Convergence* 19(2), 141–155.
- Westerwick A (2013) Effects of Sponsorship, Web Site Design, and Google Ranking on the Credibility of Online Information. *Journal of Computer-Mediated Communication* 18(2), 80–97.

- Williamson B (2014) Knowing Public Services: Cross-sector Intermediaries and Algorithmic Governance in Public Sector Reform. *Public Policy and Administration* 29(4), 292–312.
- Williamson B (2016) Political Computational Thinking: Policy Networks, Digital Governance and “Learning to Code”. *Critical Policy Studies* 10(1), 39–58.
- Wilsdon J (ed) (2001) *Digital Futures: Living in a Networked World*. Routledge, London.
- Winner L (1980) Do Artefacts Have Politics? *Dedalus* 109(1), 121–136.
- Yeung K (2008) Towards an Understanding of Regulation by Design. In: Brownsword R, Yeung K (eds) *Regulating Technologies: Legal Futures, Regulatory Frames and Technological Fixes*, pp. 79–107. Hart Publishing, Oxford.
- Yeung K (2011) Can we Employ Design-based Regulation while Avoiding Brave New World? *Law, Innovation and Technology* 3(1), 1–29.
- Yeung K (2017) “Hypernudge”: Big Data as a Mode of Regulation by Design. *Information, Communication & Society* 20(1), 118–136.
- Yeung K (2018) Algorithmic Regulation: A Critical Interrogation. *Regulation & Governance* 12(4), 505–523.
- Yeung K, Lodge M (eds) (2019) *Algorithmic Regulation*. Oxford University Press, Oxford.
- Zarsky T (2016) The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making. *Science, Technology, & Human Values* 41(1), 118–132.
- Zhou K, Yang S (2016) Understanding Household Energy Consumption Behavior. *Renewable and Sustainable Energy Reviews* 56, 810–819.