Cambridge Law Journal, Page 1 of 28 © The Authors, 2022. Published by Cambridge University Press.
doi:10.1017/S0008197322000484

PRIVATE CENSORSHIP AND STRUCTURAL DOMINANCE: WHY SOCIAL MEDIA PLATFORMS SHOULD HAVE OBLIGATIONS TO THEIR USERS UNDER FREEDOM OF EXPRESSION

STEFAN THEIL*

ABSTRACT. *Contemporary liberal accounts of free expression are almost exclusively preoccupied with the permissible exercises of state power. Influenced by this framing, free expression guarantees under the ECHR, as well as the US and German Constitutions, focus on protecting a private sphere from state interference: what happens within that sphere is only of peripheral concern. This approach is deeply unsatisfactory, especially given the significant threats emanating from private social media platforms that shape the conditions under which individuals may express themselves online. The article argues that we should take private platforms seriously as a source of significant threats, without abandoning the distinction between private actors and the state. Private platforms that are generally open to the public should have obligations to uphold free speech in their contractual relationship to users under certain conditions: if they are structurally dominant, make arbitrary decisions or significantly impact a user's societal participation.*

KEYWORDS: *freedom of expression, liberalism, constitutional law, social media, online regulation.*

I. INTRODUCTION

Alex Jones is a far-right conspiracy theorist who earns his living through an online shop attached to and promoted on his video streaming website.[1] There, one can purchase anything from apocalypse survival gear to essential oil dispensers, dietary supplements and books exposing government

* John Thornley Fellow, Sidney Sussex College, University of Cambridge. Address for Correspondence: Sidney Sussex College, Cambridge, CB2 3HU, UK. Email: st608@cam.ac.uk. I am grateful to John Adentire, Shreya Atrey, Nick Barber, Will Bateman, Jeff King, Liora Lazarus, Ewan Smith and two anonymous reviewers for helpful feedback on earlier drafts. Drafts of this paper were presented at the Bonavero Perspectives and the Oxford Working Papers in Constitutional Theory seminars – special thanks are due to all the participants for enlightening discussions.

[1] "Conspiracy Theories Made Alex Jones Very Rich. They May Bring Him Down", *New York Times*, available at https://www.nytimes.com/2018/09/07/us/politics/alex-jones-business-infowars-conspiracy.html (last accessed 15 August 2022).

conspiracies. A crucial pillar of his enterprise was the significant traffic to his shop generated on social media platforms. Through various accounts, Jones broadcast his live stream and disseminated clips and additional commentary that engaged users.

However, in late 2018 Jones was banned from most mainstream social media platforms over persistent violations of their terms and conditions of use.[2] His content included, for instance, claims that the survivors of the Sandy Hook shootings and parents of victims were crisis actors attempting to discredit the gun lobby.[3] Since his ban, attempts to create alternative accounts and upload deleted content have been steadfastly resisted by virtually all mainstream platforms. Jones and his supporters perceive his exclusion as evidence of censorship, indicative of a decaying societal appreciation of open discourse governed by freedom of expression, particularly its more robust US First Amendment incarnation.[4]

As a moral principle of liberalism, freedom of expression is highly valued. It is crucial to individual autonomy, protecting the expression of controversial ideas and information, as well as structurally important to a well-functioning democracy because it permits individuals to criticise and hold those in power to account.[5] John Stuart Mill even argued that unfettered speech ultimately leads to the discovery of truth.[6] Nonetheless, free expression is generally not thought of as a boundless guarantee, nor to have automatic priority over conflicting values. Mill famously argued that imposing restraints through law is indispensable to achieving individual liberty and therefore endorsed restricting expression to prevent harm to others.[7]

This harm principle has formed the core of the standard liberal account of free expression that has endured in the work of contemporary scholars. In adopting this framing, the standard liberal account of constitutional rights to free expression has also inherited a crucial blind spot. As Eric Barendt remarks: "[c]onstitutional rights have generally been guaranteed only against state action, because constitutionalism in its intellectual and

[2] M. Isaac and K. Roose, "Facebook Bars Alex Jones, Louis Farrakhan and Others From Its Services", *New York Times*, available at https://www.nytimes.com/2019/05/02/technology/facebook-alex-jones-louis-farrakhan-ban.html (last accessed 15 August 2022).

[3] These specific allegations are currently the subject of a defamation lawsuit; see J. Fortin, "Infowars Must Turn Over Internal Documents to Sandy Hook Families, Judge Rules", *New York Times*, available at https://www.nytimes.com/2019/01/12/us/alex-jones-infowars-lawsuit.html (last accessed 15 August 2022).

[4] The paper uses "speech" and "expression" interchangeably, because nothing of consequence in the legal and philosophical discourse turns on the employed terms, see F. Schauer, "Free Speech on Tuesdays" (2015) 34 L. & Phil. 119, 123, Footnote 8.

[5] D. Mill, "Freedom of Speech" in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Stanford 2018).

[6] J.S. Mill, *On Liberty*, 2nd ed. (London 1859), 33; whether that is empirically correct is debatable. See F. Schauer, "Social Epistemology, Holocaust Denial and the Post-Millian Calculus" in M. Herz and P. Molnar (eds.), *The Content and Context of Hate Speech: Rethinking Regulation and Responses* (Cambridge 2012), 129.

[7] Mill, *On Liberty*, ch. 2.

political origins has been concerned to limit government, not private authority".[8] The focus on state power has made many accounts of constitutional rights to free expression indifferent to the pernicious impact of private power: the power of landlords over tenants, employers over workers and large corporations over individuals and groups who depend on their products and services. The challenge to the standard account is that respect for free speech sits comfortably with a wide range of invasive private sanctions that strike at the core of the interests and liberties that freedom of expression safeguards. The economically, politically and socially marginalised are routinely subjected to private sanctions that censor, stifle and chill their ability to express themselves freely.

While Mill and the liberal thinkers who have followed him recognise private threats to liberty to varying degrees, they often ignore, or are unwilling meaningfully to restrict, the exercise of private power.

A. Private Power and Sanctions

At this point, it is helpful to clarify what exactly is meant by private power and private sanctions in this article. The term "private power" is intended to distinguish it from power exercised by the state. Private power is invariably exercised by private actors against other private actors, which can be natural persons, legal persons and other non-state entities. The exercise of private power can occur (1) through existing legal relationships (e.g. a contract for access to a social media platform), (2) at the pre-contractual stage (e.g. a job interview) or (3) in the context of mere social relationships (e.g. writing the guest list for a dinner party). As such, not all the exercises of private power are subject to formal legal regulation: being excluded from a popular social event may be hurtful, perhaps even based on morally suspect grounds, but that does not mean the hosts have committed any legal wrongs. The article is only interested in the exercise of private power in legal relationships.

The term "private sanctions" denotes the exercise of private power in a way that imposes some cost on an individual in response to their free-speech act: for instance, excluding an individual from a social media platform because they habitually share conspiracy theories about covid-19.[9] The motivation might specifically be to punish individuals for their behaviour, arise from a wish to avoid reputational damage to the platform or be to avert wider societal harm from low vaccine uptake. Regardless of the motivation, private sanctions are in principle lawful exercises of power.

[8] E.M. Barendt, *Freedom of Speech*, 2nd ed. (Oxford 2016), 22.

[9] "Facebook Deletes Accounts of German Anti-lockdown Group", *Deutsche Welle*, available at https://www.dw.com/en/facebook-deletes-accounts-of-german-anti-lockdown-group/a-59206831 (last accessed 15 August 2022); "YouTube to Remove All Anti-vaccine Misinformation", *BBC News*, available at https://www.bbc.co.uk/news/technology-58743252 (last accessed 15 August 2022).

I take it as uncontroversial that a private social media company may terminate a contract for services with users that breach its terms and conditions of use. There is evidently no employment contract in such a scenario that might lead to an enhanced level of protection (however limited),[10] and social media companies will generally avoid obligations under anti-discrimination legislation, such as the UK Equality Act 2010. The Act prohibits discrimination on the ground of protected characteristics but does not generally prevent a private platform from terminating contracts as a response to expressed beliefs and opinions. Both of these specialised regimes present exceptions to the general indifference of private law to fundamental rights and freedom of speech in particular.[11] This separates private sanctions from generally unlawful responses to free speech, such as criminal acts (e.g. assault, theft and harassment) and conduct that is actionable under private law (e.g. breach of contract, unfair dismissal, tort, discrimination on the basis of protected characteristics and breach of data protection obligations).

Private sanctions only legitimately attract state intervention under the Millian principle to the extent that they constitute harm, which, depending on the author, can encompass a wider or narrower set of sanctions (for ease of reference, this article refers to this as *relevant* harm).[12] As we explore in the article, for our purposes nothing of consequence turns on the scope of relevant harm. No matter how wide or narrow the definition, the common problem of most accounts lies in virtually ignoring private power: it leaves significant space for pernicious sanctions that raise concerns under most elaborated justifications for and commitments to freedom of expression.

This becomes clear when considering the real-world implications of relevant harm in the three jurisdictions examined for this article. Naturally, the practical protection afforded to free expression by courts does not fully map onto liberal accounts of free expression. Freedom of expression is inevitably shaped by history, societal and cultural norms, as well as other principles that limit permissible expressions beyond a strict reading of the harm principle. The point for our purposes is that the adjudication of cases, the balancing of competing interests and the development of doctrine yield important insights into the moral demands of freedom of expression and vice versa.[13] Indeed, the eccentricities of some jurisdictions often inform scholarly works and philosophical debates around appropriate limitations of freedom of expression.[14] With that in mind, the focus under the

[10] V. Mantouvalou, "I Lost my Job Over a Facebook Post – Was that Fair? Discipline and Dismissal for Social Media Activity" (2019) 35 I.J.C.C.L.I.R. 101, 5.
[11] N.J. McBride, *The Humanity of Private Law: Part II: Evaluation* (Oxford 2019), ch. 11, 120.
[12] A critique of liberal approaches along those lines is well established, see for instance the contributions of Anne M. Franks (137) and S. Chemaly (150) in S.J. Brison and K. Gelber (eds.), *Free Speech in the Digital Age* (Oxford 2019).
[13] Barendt, *Freedom of Speech*, 2.
[14] Brison and Gelber, *Free Speech in the Digital Age*; J. Rawls, *Justice as Fairness: A Restatement* (Cambridge, MA 2001), 111–14; J. Waldron, *The Harm in Hate Speech* (Cambridge, MA 2012);

ECHR, in the US and Germany is squarely on protecting a private sphere of social and legal autonomy from state interference: what happens within that sphere is only of peripheral concern.

The article argues that this approach is deeply unsatisfactory, and that human rights law should take private platforms seriously as a source of significant threats, without abandoning the generally valuable differentiation of obligations between private actors and the state. Treating state regulation of expression with suspicion is prudent, but it would be foolish to ignore private platforms that have amassed comparable levels of power. Drawing on the German constitutional law doctrine of indirect horizontal effect, the article argues that at least some private platforms that open themselves to the public ought to have obligations under freedom of expression towards their users: namely, wherever they are structurally dominant, where they arbitrarily exclude users and where an exclusion significantly impacts on an individual's societal participation.

B. Horizontal Effect

The proposed horizontal effect is indirect because fundamental rights impact the interpretation and application of private law through courts, but crucially do not become a source of direct and separately enforceable legal obligations for private actors.[15] For instance, in a dispute between a social media platform and a user, courts must have regard to free expression rights, which may preclude a platform from relying on otherwise enforceable contractual provisions to remove content or terminate an account. Fundamental rights thus limit and influence private actors in their autonomous legal conduct with others through private law but do not generally obligate them to act in accordance with fundamental rights. Indirect horizontal effect thus "enforces a limited redistribution of autonomy protection, … from agents to victims of private actions that would be unconstitutional for the government to perform".[16] What specific indirect horizontal implications a right has in a given dispute depends on the balance of competing interests and fundamental rights of the parties. It is therefore not to be confused with positive obligations, which are directed exclusively at the state and oblige it to promote the meaningful enjoyment of rights: for instance, by regulating the conduct of private actors and often under a considerable margin of appreciation.[17]

Schauer, "Social Epistemology"; R. Post, "Reconciling Theory and Doctrine in First Amendment Jurisprudence" in L.C. Bollinger and G.R. Stone (eds.), *Eternally Vigilant: Free Speech in the Modern Era* (Chicago 2018); E. Heinze, *Hate Speech and Democratic Citizenship* (Oxford 2016).

[15] S. Gardbaum, "The 'Horizontal Effect' of Constitutional Rights" (2003) 102 Mich. L. Rev. 387, 404; M. Tushnet, "The Issue of State Action/Horizontal Effect in Comparative Constitutional Law" (2003) 1 I.C.O.N. 79.

[16] Gardbaum, "Horizontal Effect", 433.

[17] Ibid., at 436.

C. Social and Other Media

Exclusion from online social media platforms is taken as the paradigmatic example that frames the elaboration of the article, but many of its conclusions could conceivably be applied beyond this context. The argument potentially has implications for traditional media outlets, such as newspapers, radio and television, but there are important differences to bear in mind. Traditional media is not open to the public at large, contributions must be submitted and accepted, and are subject to advance editorial control: ultimately, broadcast and publication are thus treated as an act of taking ownership, with all the implications media regulation provides. Social media platforms, by contrast, are open to virtually anyone and expressly disavow any form of editorial control before publication: at most, platforms remove content after the fact where it violates terms and conditions of use. I do not take a strong view on traditional mass media other than to note these differences. The communication medium is not crucial to the argument in favour of an indirect horizontal effect.

A related question in this context is whether social media platforms exercise their own freedom of expression rights in the course of content moderation decisions. EU regulation generally states that user content is not an expression attributable to the platforms themselves. They are seen as passive intermediaries which benefit from liability protection, provided they action illegal content after they become aware of its existence.[18] In a similar vein, an email service provider is not expressing itself in the course of sending and receiving messages on behalf of users. There are important nuances and complexities here that merit debate, for instance the requirements for qualifying as an intermediary,[19] as well as the adequacy of social media regulation more generally. Some elements of content moderation policies may suggest that social media platforms are entitled and indeed rely on freedom of expression to resist certain obligations towards their users. However, these discussions go beyond the interest and scope of the argument presented here. Whether or not social media platforms can rely on free expression rights, the article advocates that freedom of expression of users should be taken into account by courts. That does not preclude balancing the free expression rights of users with any relevant free-speech rights of platforms.

Section II begins by offering a standard liberal account of free expression and tracing its manifestation in the three examined jurisdictions. The

[18] See Articles 14 and 15 of Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on Certain Legal Aspects of Information Society Services, in particular Electronic Commerce, in the Internal Market (E-commerce Directive) (OJ 2000 L 178 p.1).

[19] See e.g. Judgment of 23 March 2010, *Google France v Louis Vuitton*, Joined Cases C-236/08 to C-238/08, EU:C:2010:159; Judgment of 12 July 2011, *L'Oreal and others v eBay*, C-324/09, EU:C:2011:474; Judgment of 15 September 2016, *Mc Fadden*, C-484/14, EU:C:2016:689.

section concludes that scholarship and human rights law pays insufficient attention to threats emanating from private platforms.

Section III then explores some of the implications of this focus on the state by reference to online social media platforms. The section argues that some indirect free expression obligations ought to be imposed on private platforms that have opened themselves to the public. The case law in this area is admittedly still in a nascent stage, albeit with a rich developmental history that I argue comports well with established justifications for free-speech safeguards advanced in scholarship.

II. Free Expression

Standard accounts of liberalism draw a distinction between sanctions imposed by the state and those emanating from private actors: the state, through its law-making and enforcement powers is seen as the primary threat to liberty, particularly by nineteenth-century utilitarians like John Stuart Mill. Interventions by the state are only justified on a Millian account if they meet the requirements of the harm principle. Although Mill remains rather ambivalent about what precisely constitutes harm in the context of free speech, we can infer that it is a narrow conception: in his mind, expression is generally incapable of rising to the level of relevant harm, unless it incites imminent violence.

While Mill acknowledges the threat that private sanctions may pose to individual liberty, he does not hold private actors to the demands of the harm principle. Instead, he attempts to justify the pernicious effects of private sanctions, as mere natural consequences, even where individuals harm only themselves. Likewise, contemporary liberal accounts of freedom of expression tend to sit comfortably with a wide range of private sanctions that operate below the threshold of relevant harm.

A. Harm Principle

According to the harm principle developed by Mill and now widely accepted in contemporary liberalism and human rights law, "the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others".[20] A great deal thus turns on the definition of harm. Mill argues that "self-regarding" acts should not be regulated, that is to say acts that do not harm others because they do not "violate a distinct and assignable obligation to any other person or persons".[21] Such acts only hurt the individual alone, they are an inconvenience "which society can afford to bear, for the sake of the greater good of human freedom".[22] This is contrasted

[20] Mill, *On Liberty*, 22.
[21] Ibid., at 145.
[22] Ibid., at 147.

with harm (or the risk of harm) to other individuals or the public, which according to Mill is legitimately regulated through law and the state. Mill illustrates that distinction with the example of an individual who gets drunk, for which they should not be punished, and a police officer who gets drunk *while on duty*, for which they are properly reprimanded.

Most scholars depart from the Millian account on this point by accepting, in varying degrees, that the state can (and in some cases must) regulate even self-regarding harms.[23] Rawls believed that people would choose to "protect themselves against the weakness and infirmities of their reason and will in society" in the original position.[24] Raz allows for a degree of paternalism, provided that self-harm is not as such criminalised.[25] Christopher McCrudden and Jeff King surmise in their review of the scholarship that mainstream liberal thinkers have rejected the notion that "a person's immediate desires [are] sovereign, come what may for their own wellbeing".[26] This rings especially true to those who view autonomy as a relational concept, one without clear lines between harm to self and harm to others.[27]

In the specific context of free speech, Mill implies that harm cannot arise from speech acts alone, but that something more is required to engage the harm principle. The example Mill gives is speech that incites immediate acts of violence: the difference between claiming that corn dealers starve the poor in a newspaper or doing so while addressing an angry mob outside the home of a corn dealer.[28] The fact that Mill only attributes harm to the latter, and not the former scenario, points to a narrow conception of harm: especially as he is critical of those who invoke "outrage to their feelings" as grounds for legal regulation.[29] Other harms, for instance speech advocating a loss of livelihood, the denial of equal esteem and dignity in the community, as well as societal discrimination, therefore presumably do not count.

Most accounts of free expression that broadly accept the Millian harm principle have found his conception unsatisfactory and argued for expansions that allow the state to regulate more societal ills through law.[30] Restricting speech beyond the narrow Millian notion is the accepted position in international human rights law, various regional protection regimes

[23] G. Dworkin, *The Theory and Practice of Autonomy* (Cambridge 1988), 76, 127; J. Feinberg, *Harm to Others*, vol. 1 (Oxford 1984); R. Dworkin, "Sovereign Virtue Revisited" (2002) 113 Ethics 106, 113–15.

[24] J. Rawls, *A Theory of Justice* (Oxford 1971), 249.

[25] J. Raz, *The Morality of Freedom* (Oxford 1986), ch. 15.

[26] C. McCrudden and J. King, "The Dark Side of Nudging: The Ethics, Political Economy, and Law of Libertarian Paternalism" [2015] Queen's University Belfast Law Research Paper No. 16. This is especially so given cognitive deficits and the power of inertia, which may require government assistance and intervention. See S. Conly, *Against Autonomy: Justifying Coercive Paternalism* (Cambridge 2012).

[27] C. Mackenzie, "Relational Autonomy, Normative Authority and Perfectionism" 39 Journal of Social Philosophy 512.

[28] Mill, *On Liberty*, 104.

[29] Ibid., at 150–51.

[30] S.J. Brison, "Speech, Harm, and the Mind-Body Problem in First Amendment Jurisprudence" (1998) 4 Legal Theory 39; J. Gray, *Mill on Liberty: A Defence*, 2nd ed. (London 1996) 106; G. Scarre, *Mill's* On Liberty*: A Reader's Guide* (London 2007), 39.

and virtually all domestic constitutions: free expression is treated as a right to be balanced with other considerations.[31] The debate instead focuses on the proper scope of relevant harm and its relationship to other rights and fundamental values like autonomy and dignity. The scholarship on this question is vast, with various accounts offering a wider or narrower reading of relevant harm.

Jeremy Waldron argues for an expansion of relevant harm to allow the state better to regulate and punish hate speech, which he suggests undermines assurances of equal protection and human dignity of those targeted.[32] Alexander Tsesis develops an argument along similar lines in the specific context of the US Constitution.[33] Conversely, Ronald Dworkin contends that it is restrictions on hate speech that compromise human dignity and equality of the speaker, thereby undermining anti-discrimination laws.[34] The enforcement of hate speech laws is only democratically legitimate, on Dworkin's account, if the state permits unfettered debate over the ideals being enforced. In this tradition of linking free speech to democracy, Robert Post and James Weinstein argue that relevant harm should be construed narrowly because a functional democracy requires robust free-speech protections.[35] Eric Heinze similarly views free speech as "materially constitutive" of democracy, and as such it cannot be restricted without jeopardising this overriding commitment.[36] While he only applies this argument to stable and established democracies,[37] ultimately his argument implies that the state ought to ignore at least some harm to marginalised groups in the name of democracy.[38]

The details of these and various further accounts of free speech are not immediately relevant to this article: nothing of consequence turns on the precise definition of relevant harm and its ultimate justification. The view defended in the next section is not that the free expression accounts are mistaken in whichever conception of relevant harm they adopt, but rather that they habitually ignore private sanctions as a source of credible threats.

[31] See e.g. Articles 19 and 20 of the International Covenant on Civil and Political Rights (adopted 16 December 1966, in force 23 March 1976) 999 UNTS 171 (ICCPR), and Article 10 of the Convention for the Protection of Human Rights and Fundamental Freedoms (European Convention on Human Rights, as amended) (ECHR).

[32] Waldron, Harm in Hate Speech, ch. 1.

[33] A. Tsesis, "Dignity and Speech: The Regulation of Hate Speech in a Democracy" (2009) 44 Wake Forest L. Rev. 497.

[34] R. Dworkin, *Taking Rights Seriously* (London 1977), 266–78, 364–68; R. Dworkin, "Foreword" in I. Hare and J. Weinstein (eds.), *Extreme Speech and Democracy* (Oxford 2009).

[35] R.C. Post, "Racist Speech, Democracy, and the First Amendment" (1990) 32 Wm. & Mary L. Rev. 267; J. Weinstein, "Participatory Democracy as the Central Value of American Free Speech Doctrine" (2011) 97 Va. L. Rev. 491.

[36] Heinze, *Hate Speech*, 5.

[37] Ibid., at 13.

[38] A.R. Greene and R.M. Simpson, "Tolerating Hate in the Name of Democracy" [2017] M.L.R. 746, 763.

B. Mill on Private Sanctions

It is less commonly acknowledged that Mill recognised that societies sanction disfavoured individuals through means other than those available to the state. Indeed, while he suggests that liberty requires protection primarily from the majority view in society as manifested and enforced through the state, he was also concerned with sanctions imposed by its constituent members in a private capacity: "Protection, therefore, against the tyranny of the magistrate is not enough: there needs protection also against the tyranny of the prevailing opinion and feeling; against the tendency of society to impose, by other means than civil penalties, its own ideas and practices as rules of conduct on those who dissent from them".[39]

Nonetheless, Mill, as well as a plethora of secondary literature his work spawned, is often disinterested in discussing private actors in much detail and reluctant to suggest specific restrictions on private sanctions.[40] Instead, Mill embraces a "right" to "in various ways to act upon our unfavourable opinion of any one, not to the oppression of his individuality, but the exercise of ours".[41] This "right" includes avoiding someone's company, discriminating against someone in the distribution of "optional good offices" and cautioning others against a person "if we think his example or conversation likely to have a pernicious effect on those with whom he associates".[42] Mill accepts these are significant sanctions on an individual for the sake of "faults which directly concern only himself" but brushes aside the apparent tension with his notion of liberty because "he suffers these penalties only in so far as they are the natural, and, as it were, the spontaneous consequences of the faults themselves, not because they are purposely inflicted on him for the sake of punishment".[43]

Setting aside whether Mill can convincingly sustain this distinction between spontaneous, natural consequence and purposeful infliction, it is nonetheless a surprising position for him to take. Mill argues strongly that "there ought to exist the fullest liberty of professing and discussing, … any doctrine, however immoral it may be considered".[44] In the specific context of free speech, Mill staunchly maintains that speech rarely causes relevant harm, the only legitimate ground for restriction he recognises.[45] In fairness, Mill is preoccupied with the Victorian state and hence

[39] Mill, *On Liberty*, 13.
[40] From the more recent publications, see e.g. J. Skorupski, *The Cambridge Companion to Mill* (Cambridge 1998); D.O. Brink, "Mill's Liberal Principles and Freedom of Expression" in C.L. Ten (ed.), *Mill's* On Liberty*: A Critical Guide* (Cambridge 2008), 42; J. Riley, "Racism, Blasphemy, and Free Speech" in Ten (ed.), *Mill's* On Liberty, 62; J. Skorupski, *Why Read Mill Today?* (Abingdon 2006), 58; R. Cohen-Almagor, "J.S. Mill's Boundaries of Freedom of Expression: A Critique" (2017) 92 Royal Institute of Philosophy 565.
[41] Mill, *On Liberty*, 139.
[42] Ibid.
[43] Ibid.
[44] Ibid., at 32.
[45] Ibid., at ch. 2.

understandably developed a political theory focused on limiting state power to preserve individual liberty.[46]

However, in so doing, he offers cold comfort to individuals who have offended prevailing societal norms and drawn the ire of the majority even for perfectly legal and self-regarding actions. Structural discrimination, exclusion and the various shapes that assaults on dignity and denial of equal respect take in society can dissuade individuals from expressing themselves openly no less effectively than authoritarian laws and an oppressive state. Would it really matter to an individual whether they suffer a sanction for advocacy of Covid-19 conspiracy theories as the result of state-sponsored punishment or as a spontaneous, "natural consequence" imposed by their peers? To Mill, this distinction makes all the difference.

C. Contemporary Accounts

In one sense, disparate obligations of the state and private actors are in keeping with contemporary accounts and notions of autonomy. Raz holds that individuals are entitled to a sphere of social and legal autonomy to make meaningful decisions and interact with others, to control "their own destiny, fashioning it through successive decisions throughout their lives".[47] The state is not privileged in this way; indeed it is the entity primarily responsible for facilitating individual autonomy. The blind spot of many contemporary accounts is the pernicious effect that private sanctions may have on free speech, evidenced by the almost exclusive preoccupation with the state. Waldron acknowledges the threat of private sanctions for disfavoured individuals and groups, and thus the gaps left by the Millian account, but ultimately limits his argument for legal regulation to the specific context of hate speech.[48] A similar argument is developed by Frederick Schauer when he challenges Millian truth claims and the marketplace of ideas justification in the context of holocaust denial.[49]

Edwin Baker has perhaps come closest to defending an absolute protection of free speech based on commitments to autonomy. His account is instructive because his concern with protecting free expression goes considerably further than most scholars and precludes restrictions even of hate speech on the grounds of protecting autonomy.

Respect for formal autonomy on Baker's account is crucial: it requires guaranteeing in law to an individual the ability to express the individual's views. He distinguishes it from a less important substantive autonomy which "involves a person's actual capacity and opportunities to lead the

[46] Ibid., at 7.
[47] Raz, *Morality of Freedom*, 369.
[48] Waldron, *Harm in Hate Speech*, 225–26, 232.
[49] Schauer, "Social Epistemology".

best, most meaningful, self-directed life possible".[50] Unlike formal autonomy, substantive autonomy often requires the allocation of resources and information, and thereby reduces the substantive autonomy of other individuals.

It is the formal notion of autonomy that ultimately justifies rejecting legal restrictions for hate speech on Baker's account. He readily acknowledges that hate speech does not respect others' equality or dignity but nonetheless maintains that legal restrictions would violate the formal autonomy of the speaker: because their hate speech does not interfere with the formal autonomy of others.[51] There are hence no grounds to restrict hate speech, even if it undermines substantive autonomy and regardless of "how this expressive content harms other people".[52]

However, it is less clear why Baker believes that the state can breach formal autonomy through laws that punish hate speech, but not by creating a legal environment where private actors can sanction hate speech. It would appear consistent, perhaps even necessary on Baker's conception, to argue that the state must also prevent private actors from sanctioning others for hate speech, for instance in the context of employment relationships. Instead, Baker implausibly suggests that individuals have no claim to autonomy in their legal relationships to other private actors: "prohibitions on racist or hate speech should generally be impermissible—even if arguably permissible in special, usually institutionally bound, limited contexts where the speaker has no claimed right to act autonomously—such as when, as an employee, she has given up her autonomy in order to meet role demands that are inconsistent with expressions of racism".[53]

That is far from a satisfactory or convincing justification. To start with, it appears implausible that an employee gives up any claim to autonomy, implicitly or expressly, simply by virtue of their employment contract. Certainly, there are some contractual restrictions, for instance agreeing to abide by working hours and company policies, but that is a far cry from renouncing a right to act autonomously or indeed expressing views in the workplace. Indeed, a specialised body of employment law draws considerable limitations on what employers may contractually require of their workers and what they can do to monitor and enforce compliance. Nor does relinquishing some autonomy in employment contexts provide a suitable analogy for the plethora of private law relationships individuals enter into on a daily basis. It is not at all apparent why a contract over lodging in a hotel, or for goods and services, should entail any relinquishment of the robust free expression rights Baker envisions. Carving out such a

[50] C. Edwin Baker, "Autonomy and Hate Speech" in Hare and Weinstein (eds.), *Extreme Speech and Democracy*, 143.
[51] Ibid., at 143.
[52] Ibid., at 142.
[53] Ibid., at 143.

general exception for private legal relationships exempts a significant part of everyday exercises of speech from meaningful protection.

Eric Heinze similarly allows for a curious exception to his otherwise robust protection of speech. He argues that banning hate speech damages democratic legitimacy and should generally be avoided by any well established and stable democracy. Nonetheless, Heinze is content with banning hate speech in the workplace as part of the "active democracy's prerogative to promote pluralist values".[54] He justifies this exception through educational considerations, because it teaches vulnerable groups to "answer back in public discourse", but thereby essentially concedes that curtailing hate speech in some areas of life is necessary to preserve his vision of democratic citizenship.[55] Both Heinze and Baker appear to believe that more effective responses to hate speech lie in counter-speech and private sanctions imposed by individual members of society, similar to the societal discrimination and shunning endorsed by Mill.[56]

Regardless of whether this is a sufficient and satisfactory response to hate speech, the problem is that such a view treats any state intervention as a presumptive *threat* to liberty, while viewing private power as a presumptively permissible *exercise* of liberty. The point here is not that the scope of relevant harm is too narrow. One can consistently, but in my view unconvincingly, argue that state regulation beyond Baker's account of relevant harm is an impermissible encroachment on free expression. The key problem is applying the notion of relevant harm, however broad or narrow, exclusively to threats emanating from the state.

The highly unequal distribution of wealth in society (and its associated opportunities) make startling and from the individual's perspective no less arbitrary differences to their practical enjoyment of rights.[57] To put it mildly, there are powerful inequalities in modern societies that ultimately determine whose harmful speech is the subject of private sanctions and whose is tacitly tolerated, perhaps even celebrated as a "natural" reaction to speech acts that the majority dislikes. The liberal preference for leaving private actors to their own devices operates poorly and materially reduces individual autonomy whenever grave imbalances in societal bargaining power persist: regardless of whether these are ultimately rooted in political, economic or social marginalisation.

A precariously employed worker cannot afford to offend the political sensibilities of the worker's employer and a low-income tenant will think twice before speaking out against racist remarks made by the tenant's landlord. In

[54] Heinze, *Hate Speech*, 113.
[55] Greene and Simpson, "Tolerating Hate", 765.
[56] Baker, "Autonomy and Hate Speech", 151.
[57] This is a point conceded even by committed libertarian thinkers like Robert Nozick, who accepts that inequities in the acquisition and transfer of property require wealth redistribution. See R. Nozick, *Anarchy, State and Utopia* (Oxford 1974), 152–53, 230–31.

some cases, the threat of private sanctions yields a much more restrictive environment for speech than the state could ever permissibly create through law. Meaningful autonomy thus plainly requires more than freedom from state interference; it is not a "one-way street in favour of the private instigator of action, automatically permitting her to do what it would be unconstitutional for the government to do".[58] The law need not and is not limited to restraining the exercise of state power over its citizens: it also protects weaker parties from harm inflicted by stronger private entities.[59]

To a limited extent, liberal accounts have recognised the autonomy promoting role of the state and the threats that private actors can present to individual autonomy. Joseph Raz has argued that limiting choices and frustrating autonomous decisions constitute harms.[60] This applies also in his view to instances of serious and persistent offence.[61] He views the state as best suited to promote autonomy through law and the coercion of individuals, but does not pay much attention to private actors beyond that insight.[62] Raz does not deny that private actors and organisations have "certain moral requirements" concerning the exercise of their power and that this may legitimately attract regulation, but he does not elaborate in any detail when that might be the case.[63] He notes only in passing that private coercion, say the denial of one desired option to an individual, in most cases does not interfere with autonomy because it is localised, leaving sufficient alternatives.[64] One obvious difficulty arises if there are no sufficient alternatives available.

Like Raz, Dworkin accepts the notion that private coercion is harmful and concedes that free expression can only be meaningful if it includes "some right to the opportunity to speak" and that "a society in which only the rich enjoy access to newspapers, television, or other public media" would not guarantee that right.[65] But his account, as with Raz's, is incomplete and offers us precious little in identifying when intervention against private actors might be warranted. In *Justice as Fairness*, John Rawls likewise criticises welfare capitalism because it allows wealthy private actors to dominate the economy and leverage it into disproportionate political power, thus challenging his brand of liberal egalitarianism.[66]

[58] Gardbaum, "Horizontal Effect", 459.
[59] R. West, "The Limits of Process" in J.E. Fleming (ed.), *Getting to the Rule of Law* (New York 2011), 32, 35; M. Krygier, "Four Puzzles about the Rule of Law: Why? What? Where? And Who Cares?" in Fleming (ed.), *Getting to the Rule of Law*, 64, 89; J. King, "The Social Dimension of the Rule of Law" (Jurisprudence Discussion Group, Oxford, 24 May 2018), 26.
[60] Raz, *Morality of Freedom*, ch. 15.
[61] Ibid., at 413.
[62] Ibid., at 417.
[63] Ibid., at 4–5; J. Raz, *Ethics in the Public Domain: Essays in the Morality of Law and Politics* (Oxford 1994), 89, 141.
[64] Raz, *Morality of Freedom*, 421, Footnote 1.
[65] R. Dworkin, "Women and Pornography: Book Review of *Only Words* by Catherine A MacKinnon" (1993) New York Review of Books.
[66] Rawls, *Justice as Fairness*, 136–40; for a deeper exploration of Rawls views on socialism, see W.A. Edmundson, *John Rawls: Reticent Socialist* (Cambridge 2017).

However, as Rawls presents an account of the ideal structure for a democratic society, specifically its public norms and institutions, he is less interested in possible threats arising from private power. More recent scholarship engages with the profound effects that private platforms have on free expression but have so far not gone beyond sketching possible liberal responses.[67]

D. Private Sanctions and the Courts

The limited recognition of private threats is also clearly present in the free-speech protection available under the ECHR and the US and German Constitutions. All three jurisdictions endorse a broader understanding of harm in their doctrinal approaches. This renders the corresponding scope for regulatory intervention by the state broader and deeper, leaving comparatively less space for private sanctions beneath the threshold of relevant harm. Nonetheless, courts accept a general distinction between the obligations of the state and those of private actors, being significantly more permissive on the latter.

The ECHR takes an essentially Millian view of free expression in *Palomo Sanchez*. The case concerned two workers who published a derogatory cartoon of their employer in a trade union newsletter in the context of an employment dispute. Both were subsequently dismissed for gross misconduct, and their dismissals were upheld by domestic courts.[68] A majority of the ECtHR found no violation of the Convention, emphasising the need for mutual trust in labour relations and good faith requirements. This meant that, even though the cartoon is otherwise protected under Article 10 ECHR (Freedom of expression) from state interference, in the context of private labour disputes it constituted legitimate grounds for a fair dismissal.[69]

Article 10 ECHR further generally does not require states to enable individuals to access private forums of expression, and the existence of alternative forums typically ensures that free expression rights are not violated.[70] In the case of *Appleby*, the applicants sought to set up a stand and distribute leaflets in a privately owned shopping mall.[71] The ECtHR found that freedom of expression does not "bestow any freedom of forum for the exercise of that right".[72] Rather, the denial of access must have the "effect of preventing any effective exercise of freedom of expression or it can be said that the essence of the right has been destroyed".[73]

[67] A. Bhagwat, "Free Speech Categories in the Digital Age" in Brison and Gelber (eds.), *Free Speech in the Digital Age*, 88, 89–93.
[68] *Palomo Sanchez and others v Spain* (2012) 54 E.H.R.R. 24.
[69] Ibid., at [76].
[70] *Appleby and others v United Kingdom* (2003) 37 E.H.R.R. 38, at [50].
[71] Ibid., at [40]–[48].
[72] Ibid., at [47].
[73] Ibid.

Likewise, the First Amendment to the US Constitution has little to say about threats from private actors. In the early case of *Marsh v Alabama*, the Supreme Court struck down a trespassing statute invoked to prevent the distribution of religious material on a privately owned company town sidewalk.[74] It held that the sidewalk had been dedicated to public use, reasoning that "the more an owner, for his advantage, opens up his property for use by the public in general, the more do his rights become circumscribed by the statutory and constitutional rights of those who use it".[75] This principle was expanded in *Food Employees*, where the Supreme Court found that a privately owned mall constituted "the functional equivalent of the business district" in *Marsh* and hence rejected the application of trespassing laws against peaceful picketing.[76] However, these decisions were subsequently narrowed, based on a line of cases concerned with compelled speech,[77] culminating in the finding that free-speech guarantees do not apply on private property at all.

In the case of *Lloyd Corporation v Tanner*, the US Supreme Court held that where an expression is unrelated to the operations of the private forum, the protections of the First Amendment do not apply, especially when alternative fora are available.[78] Tanner was distributing anti-war leaflets and had the opportunity to make use of the public sidewalk outside the mall for his activities. This was a key factor distinguishing this case from the *Food Employees* decision, where the expression was directly related to the shopping centre and no other venues were available to protesters. This reasoning was confirmed in *Hudgens*, where the Supreme Court held that free-speech rights do not apply to private shopping centres.[79] A unanimous court subsequently found in *Pruneyard Shopping* that First Amendment rights could not compel a private mall to allow solicitation for signatures in a political campaign.[80]

The only manner in which an ostensibly private actor may be bound by the First Amendment is where their actions are captured by the state action doctrine. According to a majority of the Supreme Court in *Manhattan Community Access Corp. v Halleck*, a private actor may be treated as a state actor where: (1) the private entity performs a traditional, exclusive public function, (2) the government compels the private entity to take a particular action or (3) the government acts jointly with the private entity.[81] Even where a private entity was specifically setup to fulfil a function assigned to the state through legislation, private actors are not viewed as

[74] *Marsh v Alabama* 326 U.S. 501 (1946).
[75] Ibid., at 506.
[76] *Food Employees v Logan Valley Plaza, Inc.* 391 U.S. 308 (1968).
[77] *West Virginia State Board of Education v Barnette* 319 U.S. 624 (1943), 633.
[78] *Lloyd Corp. v Tanner* 407 U.S. 551 (1972).
[79] *Hudgens v National Labor Relations Board (No. 74-773)* 424 U.S. 507 (1976).
[80] *Pruneyard Shopping Center v Robins* 447 U.S. 74 (1980).
[81] *Manhattan Community Access Corp. v Halleck* 139 S. Ct. 1921 (2019), 1928.

agents of the state and hence not obligated under free expression guarantees.[82] When presented with an argument that the First Amendment required broadcasters to accept paid editorial advertisements on the Vietnam war, the Supreme Court expressly rejected a First Amendment claim due to an absence of relevant state action.[83] As Stephen Gardbaum explains in his illuminating article, this limits the impact of the First Amendment: it becomes a yardstick for the constitutionality of private law invoked in disputes between private actors.[84] In practice, this protection might still go quite far, provided the Supreme Court overcomes its "complex and perplexing" doctrine of state action.[85]

What makes German constitutional law different is its embracing of a general doctrine of indirect horizontal effect. This doctrine goes back to the Constitutional Court decision in *Lüth*, which saw the Court overturn a private law injunction: it held that a lower court had failed to take proper account of freedom of expression guarantees in interpreting § 826 of the German Civil Code. This rendered unlawful the injunction granted against Erick Lüth who had called for a boycott of the work of a Nazi filmmaker.[86] The crucial difference to direct horizontal effect is the absence of a directly enforceable legal obligation of private actors to uphold fundamental rights – the German doctrine is indirect precisely because there is no such obligation, nor any procedure by which fundamental rights complaints can be brought directly by one private actor against another.[87] Fundamental rights are instead considered within the context of a court's interpretation and application of private law.

The article now introduces and defends three scenarios where a modest horizontal effect of free expression appears warranted to temper private power.

III. Tempering Private Power

The section explores three scenarios where private sanctions ought to attract intervention on free expression grounds. Generally, liberal accounts of free expression caution against state regulation and restrictions of speech for a wide variety of reasons that are too numerous to present comprehensively in this article. What is common to most of them is the Millian idea that liberty requires taking individuals seriously as autonomous moral actors. What follows is, by necessity, a broad overview and condensation of the views held in the scholarship. It is designed to illustrate a more general point:

[82] Ibid., Dissenting Opinion of Justice Sotomayor, 1936.
[83] *CBS v Democratic National Committee* 412 U.S. 94 (1973), 121.
[84] Gardbaum, "Horizontal Effect ", 391.
[85] Ibid., at 458.
[86] Federal Constitutional Court, 1 BvR 400/51, 15 January 1958, BVerfGE 7, 198 (Lüth).
[87] Gardbaum, "Horizontal Effect", 396, 404.

no account offers us compelling reasons for restricting the application and safeguards of freedom of expression to exercises of state power alone.

Based on liberal scepticism towards speech restrictions by the state, we can identify circumstances where free expression offers good normative reasons to intervene against private power by recognising a modest horizontal effect. It is modest, because it does not apply to the vast majority of exercises of private power, even where they are subject to some significant legal regulation. For instance, free expression on this account has nothing to say about an individual who is expelled from a restaurant following a rant about covid-19 conspiracy theories. Such a person experiences private sanctions and hence some level of harm, but none that should trouble us from the perspective of free expression. A horizontal effect applies only to private platforms that are open to the public at large and only under particular circumstances of (1) structural dominance, (2) arbitrary exclusion or (3) a severe impact on the societal participation of individuals, specifically their ability to express themselves publicly.

In developing this normative argument, the section draws on the rationales offered in scholarship for the protection of free expression from state interference and illustrates the argument with cases drawn from German constitutional law.

A. Structural Dominance

Structural dominance of a private actor is the scenario that is most directly related to the concerns of liberal accounts. One common recurring theme is the extent of state power and its enforcement capabilities: restrictions and proscription of free expression through law (especially criminal law) cannot be reasonably avoided. Such arguments tend to focus on the evils and (historic and current) abuses of state power in the regulation of free speech. For Waldron, the best case for protecting free expression arises from countless examples where state power was used to "suppress dissent, deflect criticism, and resist exposure of [state] malfeasances".[88] For Frederick Schauer, it is likewise this historic behaviour that warrants caution against state regulation of speech and justifies robust free expression guarantees.[89] As we have seen, Eric Heinze takes this argument further than most by suggesting that free-speech commitments are constituent elements of a democracy, such that they cannot be limited without damaging its legitimacy.[90]

The historic track record of restrictions that structurally dominant private actors impose is similarly concerning. Private power may ultimately not be as sweeping in virtually all areas of life, but it can impose heavy costs on

[88] Waldron, *Harm in Hate Speech*, 26.
[89] F.F. Schauer, *Free Speech: A Philosophical Enquiry* (Cambridge 1982), 86.
[90] Heinze, *Hate Speech*.

individuals in a manner not dissimilar to authoritarian states. On social media platforms in particular, private actors can hold a significant level of power: they set the terms and conditions of use unilaterally, alter them at will and enforce them however they see fit without any meaningful independent oversight. A structurally dominant private actor is by definition "the only game in town", to which there is no viable alternative: individuals are effectively required to contract with the private actor on its terms and accept its arbitrary enforcement power or lose access to one of the most important contemporary forums for public expression.

To address similar concerns, the German Constitutional Court developed a doctrine of indirect horizontal effect. As we saw, this indirect horizontal effect was crucial to the *Lüth* decision and is distinct from positive obligations because the focus is not on state regulatory conduct, but on the implications of fundamental rights for the interpretation and application of private law through courts.

The German Constitutional Court decision in *Fraport* specifically mused on whether an indirect horizontal effect doctrine would apply in cases of structural dominance, against the backdrop of protests in the publicly accessible areas of Frankfurt Airport, run by a corporation under a controlling majority of local, state and federal governments.[91] The Court, perhaps unsurprisingly, found that the corporation, notwithstanding its private law appearance, was directly bound by constitutional rights just as any other state entity. However, in dicta the Court suggested that, where private actors provide the infrastructure and forums previously furnished by the state (e.g. state monopolies on postal services and telecommunications or, as in this case, public forums for protest), the indirect obligations imposed through fundamental rights might well be equally exacting and indistinguishable from the direct obligations of the state.[92]

The question is whether there are currently such structurally dominant private platforms and how we would go about identifying them. One approach to determining structural dominance is to analyse market share and audience reach.

Reliable data is difficult to obtain, but estimates from July 2020 suggest there are as many as 3.96 billion active social media users, amounting to approximately 51 per cent of the global population. Platforms owned by Meta (WhatsApp, FB Messenger, Instagram), along with YouTube and WeChat make up the top five most popular global social media.[93] In specific regions, such as the US, market penetration of social media is esti-

[91] Federal Constitutional Court, 1 BvR 699/06, 22 February 2011, BVerfGE 128, 226, [46]–[48], [53].
[92] Ibid., at [59], subsequently confirmed in Federal Constitutional Court, 1 BvQ 25/15, 18 July 2020, NJW 2015, 2485, at [6]–[7].
[93] "Global Social Media Overview July 2020", available at https://datareportal.com/reports/digital-2020-july-global-statshot (last accessed 15 August 2022).

mated to be significantly higher, at 70 per cent of the population in 2019.[94] Following that approach to dominance, the demands of free expression would likely vary considerably according to the jurisdiction and the user being excluded: a platform may well be structurally dominant in one country, but not another, or may be highly relevant to only a certain type of user.

Political figures and parties are of particular interest because they occupy a role of considerable public and constitutional significance: they rely on effective communication to secure support at the ballot box and a significant part of (traditional) media regulation is focused on providing equitable access to all candidates and thus ultimately ensure free and fair elections. If private actors are, at least conceivably, entitled to access structurally dominant platforms, then political figures and parties would be the most likely candidates.

A preliminary ruling gave the German Constitutional Court the opportunity to apply its reasoning in *Fraport* specifically to social media platforms and political parties. The Court required Facebook to reinstate the page of a far-right party in the lead up to the European Parliament election.[95] The party had shared content to its page which Facebook classified as violating provisions of its community standards prohibiting hate speech. Facebook first curtailed the visibility of the posts and prevented the sharing of further content, before ultimately deleting the page entirely. As the party had sought a preliminary injunction, the Court did not interrogate the merits of the case in detail. Pursuant to settled case law, the requirements for an injunction are that a violation of constitutional rights is at least arguable and that granting the injunction (and later finding that the complaint was unfounded) would produce less harm than refusing the injunction (and later finding that the complaint was well founded). On this basis, the Court determined a violation of fundamental rights was arguable, especially as the content shared by the party was not unambiguously illegal: this was a crucial finding as otherwise Facebook's responsibility to delete under the German Network Enforcement Law (NetzDG) would have been triggered.[96] Given the close proximity to the European elections, the Court required Facebook to reinstate the page.

Admittedly, the case had several distinct features that are unlikely to apply to the vast majority of cases: there was an impending election, hence deleting the page at that particular time put the political party at a unique disadvantage. Crucially, it would have also weighed heavily in the favour of an injunction that it is virtually impossible to repair the

[94] "Social Media Fact Sheet", available at https://www.pewresearch.org/internet/fact-sheet/social-media/ (last accessed 15 August 2022); "Digital 2019: United States of America", available at https://datareportal.com/reports/digital-2019-united-states-of-america (last accessed 15 August 2022).

[95] German Federal Constitutional Court, 1 BvQ 42/19, 22 May 2019, unpublished.

[96] Ibid., at [21]. For a more detailed account, see S. Theil, "The Online Harms White Paper: Comparing the UK and German Approaches to Regulation" [2019] J.M.L. 1.

electoral disadvantage of losing a primary campaigning platform. This is in no small part due to the significance that social media platforms have attained for political campaigning. Data from the Electoral Commission in the UK analysed by Katharine Dommett and Sam Power suggest that, in the 2017 General Election, parties spent 42.8% of their overall advertising budget on online advertising, the vast majority on Facebook.[97] It may therefore be difficult in practice for a dominant social media platform to exclude political figures and parties in the run-up to an election.

One could argue that Facebook, as an individual platform at least, is not structurally dominant in Germany. However, the various companies under direct control of the Facebook corporation, now rebranded as Meta, account for three out of the top five most used social media platforms.[98] This presents a more nuanced picture: a corporation may well be structurally dominant if and to the extent that its decisions to exclude a political party are coordinated and jointly enforced by the platforms it controls.

Typically, a ban from one social media platform does not entail removal from another, even if they are formally within the same corporate structure. Individuals also often maintain multiple accounts, with estimates suggesting that the average user had up to seven accounts on different platforms in 2016.[99] Therefore, at least for the average user, exclusion from one social media platform still leaves open the possibility of expression on other platforms with comparable audiences and societal participation or indeed the creation of several accounts on one and the same platform.[100]

This may be less viable for political parties who rely on uniformly communicating their message to bolster support. That does not mean that political figures and parties cannot be punished for violations of terms and conditions or for breaches of relevant criminal laws. It merely means that a temporary removal at a critical time before an election or a permanent exclusion is likely to be impermissible. This is so because the typical sanction for breaches of criminal laws is punishment and for breaches of private law is compensation (for instance in cases of economic damage caused by publicly questioning the creditworthiness of a company)[101] or specific

[97] K. Dommett and S. Power, "The Political Economy of Facebook Advertising: Election Spending, Regulation and Targeting Online" (2019) 90 The Political Quarterly 257.

[98] Facebook, WhatsApp and FB Messenger; Instagram is sixth; see "Digital 2020: Germany", available at https://datareportal.com/reports/digital-2020-germany (last accessed 15 August 2022).

[99] J. Mander, "Internet Users Have Average of 7 Social Accounts", available at https://blog.globalwebindex.com/chart-of-the-day/internet-users-have-average-of-7-social-accounts/ (last accessed 15 August 2022).

[100] Although the latter practice is not permitted on some platforms, it is tolerated and encouraged by others; see "How to Manage Multiple Accounts", available at https://help.twitter.com/en/managing-your-account/managing-multiple-twitter-accounts (last accessed 15 August 2022), and "Can I Create Multiple Facebook Accounts?", available at https://www.facebook.com/help/975828035803295 (last accessed 15 August 2022).

[101] See e.g. Higher Regional Court Munich, 5 U 2472/09, 14 December 2012, ZIP 2013, 558, a civil case revolving around Kirch Media Group, which sued Deutsche Bank after an executive questioned their creditworthiness in an interview; the case was eventually settled. See "Deutsche Bank Settles

performance (refraining from repeating certain utterances deemed defamatory). It is not losing the ability to express oneself indefinitely.[102] But should we brush aside concerns over free expression simply because an individual was not excluded by a structurally dominant platform (or corporation) and, following the reasoning of Raz, alternatives remain open to that individual in principle?

B. Arbitrary Exclusion

Even if a private actor is not structurally dominant, something is nonetheless lost when speech expressing dissenting views is removed from a platform. The enactment of private sanctions can be tainted by arbitrariness and have significant detrimental impact on individuals.

Arbitrariness can be generally defined as decision-making based on "random choice or personal whim, rather than any reason or system".[103] Key features of an arbitrary exercise of power are that it is (1) unilateral, the affected individual has no say in the decision,[104] (2) lacking intelligible supporting reasons (for instance because it is inconsistent with previous decisions or unreasonable) and thus based primarily on the whim of the decision maker and (3) not constrained or guided by a system of rules (for instance, conditions that must be met for its exercise or limitations placed on subject matter and jurisdiction).[105] This makes such decisions generally difficult to predict in advance and liable to abuse, especially by powerful actors.

Arbitrary private power is suspect under a liberal commitment to free expression chiefly due to commitments to free and open debate within a pluralistic society. These tend to be instrumentalist liberal arguments, which suggest the value of free expression lies in some benefit to other values, like truth, democracy or pluralism.[106] The core justification for free expression on Mill's account invokes epistemic humility. Given that Mill believes we cannot ever be entirely certain about what is true and what is false, we should allow for the broadest possible scope for free expression to arrive ultimately at the discovery of truth through unimpeded discourse.[107] Even if we are virtually certain that a view is false, society nonetheless benefits on balance from permitting the expression.

Dispute over Kirch Media Group Bankruptcy", *Deutsche Welle*, available at https://p.dw.com/p/1BCNB (last accessed 15 August 2022).

[102] Notable exceptions include the loss of certain civil rights, for instance the right to stand for public office, for offences against the state.

[103] *Oxford English Dictionary*, "Arbitrary".

[104] G. Postema, "Fidelity in Law's Commonwealth" in L.M. Austin and D. Klimchuk (eds.), *Private Law and the Rule of Law* (Oxford 2014), 18.

[105] King, "Social Dimension", 19–20.

[106] W.A. Galston, *Liberal Pluralism: The Implications of Value Pluralism for Political Theory and Practice* (Cambridge 2002).

[107] Mill, *On Liberty*, ch. 2; Cohen-Almagor, "J.S. Mill's Boundaries", 576; for a contemporary defence, see Gray, *Mill on Liberty*, ch. 5.

Raz justifies free expression as serving the public good because it allows citizens to be better informed and freely articulate their preferences without governmental manipulation.[108] He emphasises that the articulation, presentation and toleration of often incompatible conceptions of the good life, and thus free expression, forms "a foundational part of the political and civic culture of pluralistic democracies".[109] Raz appears to accept that the law can and should protect not just from state interference but also from private actors when he acknowledges that "freedom of expression can be supported as part of a pluralist argument for using the law to promote pluralism in the society".[110] Underlining many of these arguments is a general concern that expression should not be stifled without good reason (or on some accounts should never be) and that, in any case, mere disagreement with an expressed view cannot justify its suppression through the state. There are good reasons to extend this protection to private power as well. In this respect, a German case concerning a ban from football stadiums is illuminating.

In the decision, the German Constitutional Court addressed access disputes between private actors following so-called "stadium bans": the practice of German football clubs to exclude collectively certain individuals from attending any Bundesliga and lower tier games.[111] The excluded individual was part of a group of "ultra-fans" that frequently engaged in acts of vandalism and brawled with rival supporters. The individual was investigated by police for breach of the peace at a match venue, but charges were ultimately dropped. In the interim, the concerned club nonetheless banned the individual on behalf of the German Football Association from attending any game for the duration of two years. The measure was based on reciprocal agreements between the private football clubs collectively to recognise and enforce such bans and ultimately relies on the property rights of each individual venue to deny an individual entry.

The applicant challenged his stadium ban before the Constitutional Court under equality rights but was denied relief. However, the Court elaborated that private venues that open themselves to a mass public audience cannot arbitrarily exclude individuals. They must offer a good reason and ensure basic procedural safeguards, basing their decision on specific and provable facts of sufficient weight.[112] The court found such a good reason in the past behaviour and association of the individual with violent fans, and the inference reasonable that his presence constituted a risk for the security of players and spectators.[113] The judgment expressed some doubts whether

[108] J. Raz, "Free Expression and Personal Identification" (1991) 11 O.J.L.S. 303, 306.
[109] Ibid., at 321, 324.
[110] Ibid., at 323.
[111] Federal Constitutional Court, 1 BvR 3080/09, 11 April 2018, BVerfGE 148, 267.
[112] Ibid., at [45], [46].
[113] Ibid., at [55].

a ban would be upheld without giving an individual reasons for the decision and an opportunity to be heard.[114] However, an intervening change in Football Association rules on stadium bans specifically to require hearings rendered the legal point moot.[115]

Following the *Stadium Ban* reasoning, excluding an individual user would therefore require at least some good reason and a right to be heard. As we discussed in the introduction, Alex Jones is effectively barred for the purpose of expressing his views on social media platforms. However, he was ultimately removed based on what would likely be considered good reasons. Jones is known for Sandy Hook Shooting conspiracy theories and other content that is potentially punishable under criminal law and that platforms may have a separate legal duty to remove under the German NetzDG. Moreover, the internal appeal process employed on most platforms ensures that individuals have an opportunity to be heard on content moderation decisions. There are some questions as to how robust the internal processes are in practice, but the German Constitutional Court did not formulate exacting requirements.[116] It would therefore seem that the decision to exclude Jones was not arbitrary in the sense considered here. He was heard before the ban was imposed (and indeed had received several warnings), there were intelligible reasons supporting the decision, and it was ultimately guided by a system of rules that was at least generally predictable, namely the terms and conditions outlining prohibited behaviour and sanctions.

However, an exclusion from social media might nonetheless fall afoul of freedom of expression if it significantly impacts an individual's societal participation, specifically the ability to express oneself. This might be the case here because Jones was banned from virtually all major social media platforms indefinitely.

C. Societal Participation

Many scholars view freedom of expression not simply as a good in service of other values, but as a good in and of itself, a crucial element of individual self-fulfilment. On this account it is not just some other value, like the search for truth, democracy or pluralism that is impacted when speech is silenced, but individuals themselves are wronged: for instance, in the exercise of their autonomy and respect for their dignity. Such arguments are therefore not necessarily instrumentalist,[117] and hold in principle even if an individual uses free expression to undermine truth, democracy or pluralism.[118]

[114] Ibid., at [46].
[115] Ibid., at [58].
[116] Ibid., at [58].
[117] Barendt, *Freedom of Speech*, 13.
[118] J. Waldron, "A Right to do Wrong" (1981) 92 Ethics 21; O.J. Herstein, "Defending the Right to do Wrong" (2012) 31 Law and Philosophy 343.

As we saw, Dworkin saw the enforcement of laws as legitimate only if the state permits unfettered debate over the ideals being enforced.[119] This point is perhaps taken furthest by Edwin Baker, who argues that any restriction on speech is a violation of formal autonomy: a violation so grave that no amount of harm caused by speech could on its own justify intervention.[120]

Some authors reach similar conclusions by focusing on the intended audience of the speaker. For instance, Senea Shiffrin ultimately justifies an individual right to free expression through the human interest in authentically disclosing our thoughts and feelings to others.[121] In a similar fashion, Thomas Scanlon holds that "the powers of a state are limited to those that citizens could recognize while still regarding themselves as equal, autonomous, rational agents".[122] As a person is only autonomous in Scanlon's estimation if they are free to weigh the arguments for various courses of action independently, governments are not entitled to suppress speech, even on the grounds that it may instil harmful beliefs or provoke harmful actions.[123] The suppression of speech therefore harms individuals because it denies them access to ideas and information that they need to develop their own views.

Imposing some free-speech obligations on private actors may therefore be justified by the intrinsic value of free speech and the importance of protecting social relationships between speakers and the audience.

The German Constitutional Court has commented on the impact on societal participation in the context of free expression and political beliefs. The Court rejected the complaint of a far-right politician who was denied access to a privately owned wellness hotel.[124] The hotel had initially confirmed the politician's four-day reservation, but then cancelled and directed the applicant to alternative accommodations in the area. Upon receiving a request for clarification, hotel management banned him from entering the premises indefinitely, explaining that his vocal and prominent political beliefs would inconvenience other patrons and damage the reputation of the establishment. The applicant sued the hotel in an attempt to overturn the ban and reinstate his reservation. The Federal Court of Justice, the highest civil court in Germany, accepted this claim insofar as it pertained to reinstating the reservation, but confirmed the hotel could legally exclude him from future bookings. The applicant challenged this decision before the Constitutional Court, claiming that the exclusion from future bookings amounted to a violation of

[119] Dworkin, *Taking Rights Seriously*, 266–78, 364–68.
[120] Baker, "Autonomy and Hate Speech", 146.
[121] S.V. Shiffrin, *Speech Matters* (Princeton 2014).
[122] T. M. Scanlon, "A Theory of Freedom of Expression" in T. M. Scanlon (ed.), *The Difficulty of Tolerance: Essays in Political Philosophy* (Cambridge 2003), 15.
[123] Barendt, *Freedom of Speech*, 16.
[124] German Federal Constitutional Court, 1 BvR 879/12, 27 August 2019, unpublished.

his rights. He argued that the exclusion from one hotel would effectively lead to exclusion from others in the region on essentially the same discriminatory basis, namely his political beliefs.

The general constitutional equality provision (*Grundgesetz für die Bundesrepublik Deutschland* (GG), art. 3(1)) and the specific prohibition on discrimination (GG, art. 3(3)) both apply in principle to private contractual relationships under the indirect horizontal effect doctrine. However, the Court held that the general and specific enumerated prohibitions on discrimination, notably on the basis of political beliefs, could not be interpreted as requiring comprehensive political neutrality of private actors.[125] The hotel owners could avail themselves of the right to property (GG, art. 14) and occupational freedom (GG, art. 12) to justify the exclusion, which on balance prevailed over the competing interests of the applicant.[126] The Court therefore found it generally permissible for the hotel to exclude individuals not simply on the basis of criminally punishable expression (which is fairly uncontroversial on the standard liberal account provided one accepts a sufficiently broad notion of relevant harm) but even based on merely disfavoured opinions and unwelcome, but constitutionally protected, political expression. In an important caveat, the Constitutional Court accepted that a ban from *all* hotels would lead to a different outcome in the balancing of competing rights, because this would exclude the individual from societal participation. However, it did not find this established on the facts: the applicant had been expressly offered accommodation in alternative hotels.[127] The question is whether Alex Jones was subjected to a functionally equivalent violation when his accounts were removed and whether his audience was therefore in some meaningful sense deprived of hearing his views.

As we have seen, Jones experienced terminations of his accounts across multiple platforms, and subsequent attempts to create fresh accounts have been thwarted: this has clear implications for his participation in online discourse and deprives his audience on these platforms from hearing his views directly. Does the cumulative effect then give Jones a claim for continued access to mainstream social media platforms? Possibly. It is important to note here that, in contrast to the ban from football stadiums and exclusion from the wellness hotel we discussed, the removal of Jones' accounts is indefinite and comprehensive across social media platforms. This may ultimately prove unreasonable: Jones has no viable avenue to regain access.

At least for now, Jones has not, however, been deprived of his ability to express himself online to his viewers. His audience referrals from social media have deteriorated significantly. Mainstream social media websites

[125] Ibid., at [11].
[126] Ibid., at [13].
[127] Ibid., at [12].

now account for just 1.2 per cent of traffic to the website.[128] Likewise, the daily average of 1.4 million visits to his website in September 2018 has dropped to approximately 283,000 in April 2022.[129] This means the website retains its position within the top 11,000 websites globally and even ranks in the top 3,000 in the United States.

It is important to stress that this assessment is contextual. The position would change if Jones found himself destitute and no longer able to communicate his views through his own website. Under such circumstances, maintaining an indefinite exclusion from virtually all mainstream social media platforms would likely no longer be justified.[130] That does not rule out sanctioning a reinstated user account for violations of terms and conditions of use, including removing content and temporary bans, but it does bar platforms from excluding an individual indefinitely.

IV. Conclusion

Ultimately, we are presented with a nuanced picture of freedom of expression. Most private platforms are legally entitled to sanction the expression of users as they see fit, unless a specialised legal regime provides regulation (for instance employment, anti-discrimination legislation or possibly even parts of some countries' data protection legislation): this includes excluding them from social media platforms. Individuals are not generally entitled to address any particular audience on any particular private platform.

While standard liberal accounts recognise in principle the dangers of private sanctions, they appear unwilling to hold private actors to anything approaching the harm principle. This leaves a wide and troubling gap, especially in an age where the exercise of free expression has never been more reliant on private platforms. In arguing that obligations should be imposed on some private platforms, the article promised to address private threats to free expression, while not abandoning the liberal commitment to a sphere of individual social and legal autonomy. Notably, a private platform must be open to the public and either (1) be structurally dominant, (2) arbitrarily exclude individuals or (3) significantly impact their societal participation. Under those circumstances, it appears justified to curtail contractual autonomy through an indirect horizontal effect.

Determining structural dominance of a platform requires a context-sensitive evaluation in any given jurisdiction, including of the broader corporate structure. Currently, there may well be dominant corporations, but it

[128] "April 2022 Overview: Infowars.com", available at https://www.similarweb.com/website/infowars.com/ (last accessed 15 August 2022).

[129] J. Nicas, "Alex Jones Said Bans Would Strengthen Him: He Was Wrong", *New York Times*, available at https://www.nytimes.com/2018/09/04/technology/alex-jones-infowars-bans-traffic.html (last accessed 15 August 2022).

[130] This is a salient point particularly in the ongoing debate about the exclusion of former US President Donald Trump from mainstream social media platforms.

is worth noting that most sanctions for most individuals have been account and platform specific. Political parties may well present the likeliest candidate for claims to access under these circumstances.

Arbitrary exclusion turns our attention to the procedures of an individual platform. It requires platforms to offer good reasons for an exclusion and give the user an opportunity to be heard. Nonetheless, it leaves open the possibility for longer term, albeit not indefinite, exclusion. This is ultimately justifiable because the platform is by definition not in a structurally dominant position and alternatives are available.

The matter may be different where an individual is barred from virtually all mainstream platforms. For the wealthy, even an indefinite ban may not impact significantly on their ability to express themselves online. Former US President Donald Trump comes to mind, as he maintains a high public profile that attracts much media attention, as does Alex Jones, who runs an overall exceptionally well-frequented website hosting his content and online shop.

However, all of these assessments are contextual. For most individuals, imposing an indefinite exclusion from mainstream social media platforms would likely not be justified, even for flagrant violations of contractual terms or breaches of criminal laws. This does not mean that such individuals cannot be sanctioned, for instance by removing their posts, instituting temporary bans or bringing relevant criminal charges. It merely means that indefinite exclusion is impermissible because the typical punishment for breaches of the law is a civil remedy or criminal sanction, not the loss of the ability to express oneself altogether.