EDITOR'S NOTE

In the waning summer months of 2022, likely due to the boredom induced by preparing for 3L, the Journal of Law, Economics & Policy Editorial Board formed the idea to bring JLEP into the 21st century. Our plan was to update the format and style of our peer-reviewed, student-run Journal.

First on our agenda was to remove the "double space after the terminal punctuation" from our formatting requirements. Thus, our rule is now **one space after the sentence punctuation**. The rule applies to the body and the footnotes. Our new rule is consistent with the Texas Manual of Style (see rule 1.39), the default settings on Word, and what our authors and the public expect. Second on our agenda was to update the font on our masthead. Thus, our rule is now **Times New Roman 11-point font for the Journal of Law, Economics & Policy Masthead**. This rule applies to published mastheads and those issued by the Journal using other mediums. Our new rule standardizes the appearance of the Journal masthead and makes the masthead consistent with the font requirements of published articles.

Sara Cutter
Editor-in-Chief
Arlington, VA
November 9, 2022

MEMBERS

CHRISTOPHER BARNEWOLT
KEVIN P. FENCHAK
CODA J. KEHL
TIMOTHY KILCULLEN
JULIET LOMEO
LUKE MARSTON
KONNOR TERNUS
TEJO S. TUNUGUNTLA

44682-gme_17-4 Sheet No. 2 Side A 11/29/2022 09:39:42

CONTENTS

ARTICLES

44682-gme_17-4 Sheet No. 2 Side B 11/29/2022 09:39:42

THE ESSENTIAL MEANING OF THE RULE OF LAW

Thomas W. Merrill[*]

INTRODUCTION

We have heard much in recent times about the rule of law. Everyone seems to be in favor of it. Everyone seems to think that those with whom they strongly disagree are violating it. Let me remind you of a few examples.

President Obama, frustrated by Congress's failure to adopt immigration reform, stated at a cabinet meeting that he still had a "pen and a phone."[1] He proceeded to announce a policy called DACA, short for Deferred Action for Childhood Arrivals, which effectively adopted a type of amnesty for some 700,000 persons who had arrived in the country as children without legal authority.[2] This was denounced by political opponents as "executive legislation" and a violation of the rule of law.[3]

His successor, President Trump, accumulated a record of sorts for being charged with flaunting the rule of law.[4] To cite just one episode, he demanded that Congress appropriate funds for construction of a wall along the border between Mexico and the U.S. When Congress enacted an appropriations bill that specifically excluded any such appropriation, Trump refused to sign it,

---

[*] Charles Evans Hughes Professor, Columbia Law School. Earlier versions of this essay were presented at conferences of academics and judges in the fall of 2021 sponsored by the Law and Economics Center of the Antonin Scalia Law School at George Mason University. I am grateful to Madhav Khosla and Ben Liebman for drawing my attention to literature on legal philosophy and the dual state and China, respectively. Avi Weis and Kyle Oefelein provided invaluable research assistance.

[1] Jennifer Epstein, *Obama's Pen-and-phone Strategy*, POLITICO (Jan. 14, 2014, 11:49 am), https://Politico.com/story/2014/01/Obama-state-of-the-union-2014-strategy-102151.

[2] Press Release, White House, Obama Administration, Remarks by the President on Immigration (June 15, 2012), https://obamawhitehouse.archives.gov/the-press-office/2012/06/15/remarks-president-immigration. *See also* Memorandum from Janet Napolitano, Sec'y of Homeland Sec., to David V. Aguilar, Acting Comm'r, U.S. Customs and Border Prot., Alejandro Majorkas, Dir., U.S. Citizenship and Immigr. Servs. & John Morton, Dir., U.S. Immigr. and Customs Enf't (June 15, 2012), https://www.dhs.gov/xlibrary/assets/s1-exercising-prosecutorial-discretion-individuals-who-came-to-us-as-children.pdf.

[3] *See, e.g.*, Michael W. McConnell, *Michael McConnell on Executive Orders, DACA, and the Constitution*, SLS BLOGS (Sept. 6, 2017), https://law.stanford.edu/2017/09/06/michael-mcconnell-on-executive-orders-daca-and-the-constitution/. When the Trump Administration rescinded the DACA program on the ground that it was not based on lawful authority, the Supreme Court, in a divided decision, remanded the rescission order for a further explanation of how this would affect the reliance interests of DACA recipients. Dep't of Homeland Sec. v. Regents of California, 140 S. Ct. 1891 (2020). Justice Thomas complained in dissent that the executive needed no further explanation when seeking to restore the "rule of law." *Id.* at 1931 (Thomas, J. dissenting in part).

[4] *See, e.g.*, Peter L. Strauss, *The Trump Administration and the Rule of Law*, 170 REVUE FRANCAISE D'ADMINISTRATION PUBLIQUE 433 (2019).

triggering a 35-day shutdown of the government.[5] Trump eventually relented, but instructed subordinates to scrounge for other pots of money to build the wall. One identified source was Section 8005 of the Defense Appropriations Act of 2019, which authorized the re-transfer of up to $4 billion in Defense Department "working capital funds" based on "unforeseen military requirements," as long as transfer of the funds had not been "denied by the Congress."[6] The Sierra Club obtained an injunction against this use of these funds, which was hailed by leading Democrats as a vindication of the rule of law.[7] A divided Supreme Court stayed the injunction, noting that the government had made a "sufficient showing" that the Sierra Club had "no cause of action to challenge compliance with Section 8005."[8] So construction of the wall was allowed to proceed, until the Biden Administration brought it to an end.[9]

Not to be left behind, President Biden has also been condemned for lack of fidelity to the rule of law. As part of the federal response to the COVID pandemic, the Centers for Disease Control ordered a nationwide moratorium on evictions of tenants for nonpayment of rent. This was later ratified and extended for a short period of time by Congress. As the expiration of the moratorium approached, the CDC decided it could extend the moratorium without additional congressional authority, based on a 1944 statute that authorized it to issue orders for fumigation, pest extermination, and "other measures, as in [its] judgment may be necessary" to prevent "sources of dangerous infection to human beings."[10] When the Supreme Court, acting on a stay application, expressed skepticism about whether the statute authorized an eviction moratorium, the Biden Administration asked Congress to enact emergency legislation extending the moratorium. But when Congress failed to act, the CDC extended the moratorium anyway. Speaking to reporters, Biden admitted that the extension order was on shaky legal ground, but he said it was worth doing because "by the time it gets litigated, it will probably give some additional time."[11] This too was decried as a violation of the rule

---

5 Mihir Zaveri et al., *The Government Shutdown Was the Longest Ever. Here's the History*, N.Y. TIMES (Jan. 25, 2019), https://www.nytimes.com/interactive/2019/01/09/us/politics/longest-government-shutdown.html.

6 Department of Defense Appropriations Act of 2019, Pub. L. No. 115-245, § 8005, 132 Stat. 2981, 2999 (2018).

7 Press Release, Nancy Pelosi, Speaker, House of Representatives, Pelosi Statement on Court Ruling Against President Trump's Border Wall (Dec. 10, 2019), https://Pelosi.house.gov/news/press-releases/Pelosi-statement-on-court-ruling-against-president-trumps-s-border-wall/.

8 Trump v. Sierra Club, 140 S. Ct. 1, 1 (2019) (per curium).

9 Press Release, White House, Biden Administration, Fact Sheet: Department of Defense and Department of Homeland Security Plans for Border Wall Funds (June 11, 2021), https://Whitehouse.gov/omb/briefing-room/2021/06/11/fact-sheet-department-of-defense-and-department-of-homeland-security-plans-for-border-wall-funds/.

10 Public Health Service Act of 1944 § 361(a), codified at 42 U.S.C. §264(a).

11 Press Release, White House, Biden Administration, Remarks by President Biden on Fighting the COVID-19 Pandemic (Aug. 3, 2021), https://www.whitehouse.gov/briefing-room/speeches-remarks/2021/08/03/remarks-by-president-biden-on-fighting-the-covid-19-pandemic.

of law, as the Supreme Court effectively held when the issue returned on another stay application.[12]

As these examples suggest, recent Presidents have become increasingly bold in taking action that has not been authorized by law or is at best only dubiously authorized by law. The examples also reveal that political opponents of the President have not been shy about condemning these actions as violating the rule of law. These and other episodes suggest that the idea of the rule of law is centrally concerned with the understanding that the executive must confine its actions to what is authorized by law and must abide by the limits prescribed by law.

Charges of rule of law violations also increasingly appear in controversial judicial decisions. Consider, as one example of several, the Supreme Court's recent *Dobbs* decision overruling *Roe v. Wade*.[13] The majority, speaking through Justice Alito, decried *Roe* as having no basis in the Constitution or in practice deeply rooted in the Nation's history and traditions. He declared it was "time to heed the Constitution and return the issue of abortion to the people's elected representatives" and "[t]hat is what the Constitution and the rule of law demand."[14] The dissenters saw the very act of overruling *Roe* an affront to the rule of law. They claimed the decision was "based on nothing more than the new views of new judges" and "[t]he majority thereby substitutes a rule by judges for the rule of law."[15] The Justices, one can safely conclude, regard the rule of law as vitally important to the legitimacy of the judicial enterprise. Some tend to equate the rule of law with fidelity to the original understanding of the Constitution and other

---

[12] Ala. Assn. of Realtors v. Dept. of Health and Hum. Services, 141 S. Ct. 2485, 2490 (2021) (lifting stay of injunction against the order). A similar fate awaited an emergency regulation of the Occupational Safety and Health Administration requiring all firms with 100 or more employees to require that they be either vaccinated or regularly tested. *See* Nat'l Fed'n of Ind. Bus. v. Dep't. of Labor, 142 S. Ct. 661, 664–65 (2020).

[13] Dobbs v. Jackson Women's Health Org., 142 S. Ct. 2228, 2242 (2022), overruling Roe v. Wade, 410 U.S. 113 (1973).

[14] *Id.* at 2243.

[15] *Id.* at 2335 (dissenting opinion of Justices Breyer, Kagan, and Sotomayor). Justice Kagan expanded on the importance of stare decisis to preserving the rule of law in another dissenting opinion:

> Adherence to precedent is "a foundation stone of the rule of law." "[I]t promotes the evenhanded, predictable, and consistent development of legal principles, fosters reliance on judicial decisions, and contributes to the actual and perceived integrity of the judicial process." …. [T]he entire idea of *stare decisis* is that judges do not get to reverse a decision just because they never liked it in the first instance. [T]hey need a reason *other than* the idea "that the precedent was wrongly decided." For it is hard to overstate the value, in a country like ours, of stability in the law.

Knick v. Twp. Of Scott, Pennsylvania, 139 S. Ct. 2162, 2189–90 (2019) (Kagan, J. dissenting) (citations omitted).

enacted laws; others tend to equate the rule of law with adherence to principles established by prior judicial decisions.[16]

Academics I am obliged to report—some of them at any rate—seem to think the idea of rule of law is either very slippery or meaningless. Jeremy Waldron, after reviewing the charges and counter-charges in the Florida electoral recount in 2000, pronounced the rule of law an "essentially contested concept."[17] Judith Shklar would not even dignify the rule of law by calling it a concept, labeling it "just another one of those self-congratulatory rhetorical devices that grace the public utterances of Anglo-American politicians. No intellectual effort need therefore be wasted on this bit of ruling class chatter."[18]

Against the admonition of Professor Skhlar, I propose to expend some intellectual effort in trying to make sense of the idea. Does the rule of law have an essential meaning, such that it is possible to say that a judge or other government official or even an entire polity is or is not faithful to the rule of law? Is it possible to disentangle the rule of law from other aspects of the modern liberal state, such as respect for individual rights? If the rule of law has a core meaning, are there certain institutional features, like an independent judiciary, that are critical to establishing a legal system governed by the rule of law? Is it possible to have a society with a large administrative apparatus exercising great discretionary power and still speak of it as one that it characterized by the rule of law?

My claim is that the rule of law does have an essential meaning: it describes a political ideal in which the executive arm of the state exercises coercive power against individuals only when this is authorized by settled principles of law. As such, the rule of law describes a state of affairs—predictability about the prospect of government coercion—which is widely regarded as desirable. But it does not describe everything that we might want to include in the description of a good or just society. My further claim is that it is important to maintain a distinct understanding of the meaning of the rule of law, in order to know when it is appropriate to criticize an official or condemn a government regime for violating the rule of law. If we begin to use the phrase to refer to an enlarged set of attributes that we regard as good or desirable, then the phrase is apt to degenerate into a slogan, as Professor Shklar suggests, and its value in pinpointing a certain type of legal behavior that may warrant correction is compromised.

---

16 *See generally* Richard H. Fallon, Jr., *"The Rule of Law" as a Concept in Constitutional Discourse*, 97 COLUM. L. REV. 1, 24–36 (1997) (discussing different "ideal types" of the rule of law in Supreme Court opinions).

17 Jeremy Waldron, *Is the Rule of Law an Essentially Contested Concept (in Florida)?*, 21 L. & Phil. 137 (2002).

18 Judith Shklar, *Political Theory and the Rule of Law* at 1, *in* THE RULE OF LAW: IDEAL OR IDEOLOGY (A. Hutchison and P. Monahan eds. 1987).

I. TWO TRADITIONS ABOUT THE MEANING OF THE RULE OF LAW

Commentators have discerned that there are two distinct traditions in terms of explicating what we mean by the rule of law.[19] These traditions go by different names. Sometimes the distinction is framed in terms of a "formal" versus a "substantive" view of the rule of law. Others have contrasted a "thin" as opposed to a "thick" version. These labels seem designed to subtly disparage the first version at the expense of the second.[20] To avoid any such implication, I will speak of the "predictable law" conception of the rule of law as opposed to the "rights conception" of what it means to refer to the rule of law.

A. *The Predictable Law Model*

Certain anticipations of the predictable law conception of the rule of law can be found in Aristotle and in coronation oaths that prevailed in the middle ages.[21] But the modern form of this understanding originates with the Magna Carta, a list of concessions that a group of nobles extracted from King John in 1215. Clause 39 famously stated:

> No free man shall be taken or imprisoned or disseised or outlawed or exiled or in any way ruined, nor will we go or send against him, except by the lawful judgement of his peers or by the law of the land.[22]

If we take this as the original expression of the predictability understanding, it is worth unpacking some of its features. Note that the threat against which the promises are extracted is a certain kind of tyranny emanating from the executive, in this case the King, who speaks with the royal "we." There is no mention of general lawlessness or criminality, of which we can surmise there was plenty in 1215. Note too that there is no mention of courts other than the allusion to the jury, if that is what is meant by "the lawful judgment of his peers." The promise by the King is to avoid taking a variety of coercive actions against "free men" except when this accords with "the lawful judgment of his peers" or is consistent with "the law of the land." The latter phrase is a bit mysterious, because in 1215 the King himself was generally thought to be the principal expounder of the law. But perhaps the "law of the land" was understood to mean something like the

---

19 *E.g.*, BRIAN Z. TAMANAHA, ON THE RULE OF LAW: HISTORY, POLITICS, THEORY 91-113 (2004).

20 For a critique of the distinction between "formal" and "substantive" conceptions of the rule of law, see John Gardner, *The Supposed Formality of the Rule of Law*, in LAW AS A LEAP OF FAITH 195-220 (2012).

21 TAMANAHA, *supra* note 19, at 7–25.

22 J.C. HOLT, MAGNA CARTA 461 (2d ed. 1992).

settled or customary law of the realm.[23] If so, we can interpret clause 39 to mean that the executive will not take adverse action against citizens of the realm unless those actions are consistent with settled legal understandings about when it is permissible to do so. So understood, Magna Carta captured the predictable law meaning of the rule of law.

Later English commentators reinforced the predictable law understanding of the rule of law.[24] We can perhaps jump forward to the British constitutional theorist A.V. Dicey, who authored an influential treatise in 1885 generally credited with coining the phrase, "the rule of law."[25] Dicey offered this as his primary definition of the rule of law: "No man is punishable or can lawfully be made to suffer in body or goods except for a distinct breach of law established in the ordinary legal manner before the ordinary Courts of the land."[26] This is close to a restatement of clause 39 of Magna Carta, although the reference to the lawful judgment of one's peers has been expanded to law "established in the ordinary legal manner before the ordinary Courts of the land." Dicey went beyond predictability, however to posit additional elements of the rule of law, including that no person is above the law. Under the English constitution, he wrote, "every man, whatever be his rank or condition, is subject to the ordinary law of the realm and amenable to the jurisdiction of the ordinary tribunals."[27] Thus, for Dicey, the rule of law also means that administrative officials, if they act against persons without legal authority, can be brought into court and charged with a dereliction of duty.

A third commentator commonly credited with developing the predictability understanding is Friedrich Hayek. "Stripped of all technicalities," Hayek wrote in his most popular work, the rule of law "means that government in all its actions is bound by rules fixed and announced beforehand—rules which make it possible to foresee with fair certainty how the authority will use its coercive powers in given circumstances and to plan one's individual affairs on the basis of this knowledge."[28] Hayek's principal contribution to understanding the rule of law was in precisely identifying as its central purpose predictability about how the state will use its "coercive powers in given circumstances." If it is possible to "foresee with fair certainty" what the state will permit and what it will prohibit, individuals will adjust their plans and actions so as to advance their aspirations within the

---

23 In 1215 and for many years afterwards the law was conceived to be something that was discovered not made. *See* F.A. HAYEK, THE CONSTITUTION OF LIBERTY 234–36 (1960) [hereinafter HAYEK, CONSTITUTION]; TAMANAHA, *supra* note 19, at 56.

24 Hayek's intellectual history of the rule of law includes statements from John Locke, David Hume, and Edmund Burke, among others. *See* HAYEK, CONSTITUTION, *supra* note 23, at 251–60.

25 A.V. DICEY, INTRODUCTION TO THE STUDY OF THE LAW OF THE CONSTITUTION (9th ed. 1945) (orig. 1885).

26 *Id.* at 110.

27 *Id.* at 193.

28 F.A. HAYEK, THE ROAD TO SERFDOM 112 (2007) (orig. 1944) [hereinafter HAYEK, SERFDOM].

space left open for individual initiative. The rule of law, so understood is an essential condition of liberty.[29]

Like Dicey, Hayek offered additional attributes of the idea of the rule of law. It was imperative "that the discretion left to the executive organs wielding coercive power should be reduced as much as possible."[30] He readily conceded that there must be rules of conduct, such as laws of contract, property, tort, and crime. But he insisted that the rules should be as formal as possible. They should be expressed in general or abstract terms not aimed at any particular group or individual, apply equally to all, and be highly certain.

The predictability version of the rule of law did not stop with Hayek. The contemporary legal philosopher Joseph Raz should be counted as an adherent of this school. Following Hayek, Raz identified the "basic intuition" behind the rule of law as the idea that "the law must be capable of guiding the behavior of its subjects."[31] In order to perform this guidance function, Raz wrote, the law must be prospective, general, clear, and relatively stable. Raz identified several institutional features that are likely to be important in generating predictable legal rules, including an independent judiciary, fair and impartial hearings when individuals are accused of violating the law, judicial review of administrative action to assure that it conforms with delegated power, and limits on the discretion of the police to assure compliance with the law.[32]

Raz provocatively argued that any legal system that shares these features can claim to be one based on the rule of law. It need not be a democracy but could also be a "non-democratic legal system, based on the denial of human rights, on extensive poverty, on racial segregation, sexual inequalities, and religious persecution."[33] Indeed, he went so far as to argue that the law may "institute slavery without violating the rule of law."[34]

Another relatively contemporary proponent who shares the predictability perspective was Harvard Law Professor Lon Fuller. In a book originally published in 1964 entitled *The Morality of Law*,[35] Fuller sought to explicate the general features of a legal system that give it moral authority. He did not claim the title "rule of law" for his conception. But the features of a legal system that command this moral respect, as he identified them, bear a close resemblance to what Dicey, Hayek and Raz called "the rule of law." Specifically, Fuller listed the following features of a legal system that give rise to a claim to what he called an inner morality: generality, clarity, public promulgation, stability over time, consistency between official rules and the

---

29 Hayek reaffirmed this basic position in a more scholarly volume published in 1960. *See* HAYEK, CONSTITUTION, *supra* note 23, at 215–31.

30 HAYEK, SERFDOM, *supra* note 28, at 112.

31 JOSEPH RAZ, *The Rule of Law and Its Virtue*, *in* THE AUTHORITY OF LAW 210, 214 (1979).

32 *Id.* at 214–18.

33 *Id.* at 211.

34 *Id.* at 221.

35 *See generally* LON L. FULLER, THE MORALITY OF LAW (rev. ed. 1969).

actual conduct of legal actors, and prohibitions against retroactive law, contradictions in the law, and against requiring the impossible of those subject to the law.[36] Each of these features operates to make the law more predictable to those to whom it applies. Fuller's exegesis of the characteristics that entitle a legal system to claim moral authority can therefore be said to encapsulate what would be required by the rule of law, if that is understood in the sense of predictable law.

What is one to make of the fact that the various thinkers I have grouped together as endorsing a predictable law conception have all discussed factors in addition to predictability as being constitutive of the rule of law? Am I being excessively reductive in characterizing their views as variations on this one theme? I think not, for three reasons.

First, many of the additional features cited by these thinkers reduce to more specific features of a legal system that make the law predictable. Lon Fuller is perhaps the clearest example of this. His eight desiderata of a legal system that aspires to inner morality are all characteristics that allow persons to predict when the law will be applied in a coercive or disabling fashion. This is quite clear with respect to clarity, public promulgation, consistency between the law on the books and the law in action, prospectivity, and noncontradiction. These are all features that allow those subject to the law to predict how it will be applied. The requirement of generality is less obviously related to predictability, but by this Fuller meant that the law is sufficiently rule-like or contains principles that are sufficiently broad that they can be perceived as applying to a variety of circumstances as opposed to being purely case-specific or ad hoc.[37] The requirement of stability performs, in Fuller's explication, an epistemic function: if the law changes too frequently it is less likely to be known and if not known will not be predictable.[38] And the requirement that the law not require the impossible can be seen as related to predictability, since an impossible legal duty is likely to be interpreted and enforced in an unpredictable manner.[39] The list of ancillary features offered by Raz – prospectivity, generality, clarity and stability – can be explicated in a similar way.

---

[36] *Id.* at 46–91. Fuller's work has recently featured in a debate over whether the modern administrative state is or is not consistent with Fuller's conception of the morality of law. Cass Sunstein and Adrian Vermeule have argued in a long article in the Harvard Law Review and a follow-on book that the modern administrative state, with its commitment to judicial review and its general requirement of reason-giving, largely conforms to Fuller's vision of the morality of law. Cass R. Sunstein & Adrian Vermeule, *The Morality of Administrative Law*, 131 HARV. L. REV. 1924 (2018); CASS R. SUNSTEIN & ADRIAN VERMEULE, LAW AND LEVIATHAN (2019). Richard Epstein, in a response, disagrees, citing the vagueness of statutory constraints under which agencies operate and doctrines of deference to agencies even on questions of law. *See generally* RICHARD A. EPSTEIN, THE DUBIOUS MORALITY OF ADMINISTRATIVE LAW (2020).

[37] FULLER, *supra* note 35, at 46-49 (explaining that to avoid "directing . . . every action" a ruler "may find it essential to articulate and convey . . . certain general principles of conduct.").

[38] *Id.* at 79-80.

[39] *Id.* at 70-79.

Second, many of the additional features stressed by thinkers I have grouped in the predictable law school are in fact institutional conditions that are necessary (or at least desirable) in producing a system of predictable law. A central weakness of the rule of law literature is the failure to distinguish between what we mean by the rule of law (the concept) and what sort of institutional arrangements are likely to generate a system characterized by the rule of law (a question of causation). When Raz identifies an independent judiciary, fair and impartial hearings, judicial review of administrative action, and limits on police discretion as important to the rule of law, he has slipped from explicating the meaning of the rule of law to offering some views about the types of institutional arrangements important in generating it. Similar points can be made about Dicey and Hayek insofar as they make an independent judiciary or a common law baseline features of a regime characterized by the rule of law. I do not mean to belittle the significance of the question of what conditions are likely to produce or sustain the rule of law–it is of paramount importance. But it may be easier to make some headway on the question of causation if the meaning of the rule of law is kept distinct from ruminating about the conditions that make it possible.

Third, some of the thinkers I have associated with the predictable law meaning of the rule of law have also been concerned more broadly with what it means to live in a free or just society. Hayek is perhaps the most prominent example here. His later work, such as *The Constitution of Liberty*, includes extensive material about the rule of law, but was concerned more broadly with determining the conditions of a society characterized by individual freedom, as he understood it. So for example, he endorsed a written constitution and a bill of rights as part of legal system that secures liberty.[40] But I think it would be a mistake to interpret these broader reflections as qualifying his understanding of the rule of law. He continued to believe that the rule of law was a critical feature of any system of government that promotes individual freedom, and he continued to define the rule of law to mean that "government must never coerce an individual except in the enforcement of a known rule."[41]

What can be said by way of generalization about the predictability version of the rule of law? First, the target of concern is the exercise of government power over individuals. Little attention has been given by adherents of this view to the possibility of private lawlessness, such as rampant crime or intimidation by gangs. The rule of law is not synonymous with "law and order."[42] Neither is the focus on the possibility of legally

40 HAYEK, CONSTITUTION, *supra* note 23, at 275-86, 324-28.

41 HAYEK CONSTITUTION, *supra* note 23, at 309-10.

42 One could argue that there is a connection, in that predictable enforcement by the government of laws against private lawlessness will tend to reduce the incidence of private lawlessness. But additional conditions must be met for the equation to hold, including the enactment of sufficiently clear and comprehensive legal prohibitions on private lawlessness, sufficient deterrents in these laws to discourage

authorized subordination of groups, such as the disenfranchisement of women or even apartheid, if such classifications are sanctioned by general laws and are enforced in a consistent and even-handed manner.

Second, the dominant virtue that animates this view is predictability. In particular, predictability about what the government can do to individuals in the way of taking their liberty and property. "Law," a word of many meanings, refers in this conception to regularity, as in the laws of physics.

Third, the proponents of the predictability view have advanced only sketchy ideas about the institutional features of a legal system that generate predictable law, even as an incompletely realized ideal. The importance of an independent judiciary is a recurring theme. But little attention has been given to the conditions that would lead an independent judiciary to develop predictable rules or would incline executive actors to enforce the orders of an independent judiciary. In this sense, the literature associated with the predictability understanding is incomplete.

B. *The Rights Conception*

What is here called the "rights conception" of the rule of law has been given a variety of names.[43] Some commentators speak of a "substantive" (as opposed to the formal) version of the rule of law; others of a "thick" (as opposed to the thin) version. What distinguishes the rights conception from the predictability view can be easily identified: the rights version does not reject the idea that law should be predictable, but adds that it must also predictably protect. This includes a list of rights deemed to be fundamental, such as the freedom of speech, the equal treatment of persons without respect to race or gender, the right to a fair trial, the prohibition of slavery, freedom from torture, and so forth. In effect, the understanding—that law must be knowable and predictable—is augmented in the rights conception by a list of rights, such as is found in the U.S. Bill of Rights, the Universal Declaration of Human Rights of 1948, or the European Convention on Human Rights.

A lucid expression of the rights conception has been provided by the esteemed British judge, Lord Bingham, in a book aptly titled *The Rule of Law*.[44] Bingham begins his book with an account of the predictability conventional view, which includes Magna Carta and A.V. Dicey (but oddly does not mention Hayek). Then, in a few short pages, he announces that this

---

private lawlessness, adequate funding of the police and courts to enforce these laws, and accurate identification by the police and courts of those who have violated the laws. One could argue that private lawlessness, in the form of assaults, robberies, and frauds, is as much or more of a threat to individual liberty, as conceived by Hayek and other proponents of the predictability conception of the rule of law. But, for whatever reasons, this has not been the focus of this tradition in the literature on the rule of law.

43 The "rights" conception is the name given to the substantive conception of the rule of law by Ronald Dworkin. *See* RONALD DWORKIN, A MATTER OF PRINCIPLE 11 (1985).

44 *See* TOM BINGHAM, THE RULE OF LAW (2010).

view must be supplemented with a conception that embraces the protection of fundamental human rights. Lord Bingham explains:

> A state which savagely represses or persecutes sections of its people cannot in my view be regarded as observing the rule of law, even if the transport of the persecuted minority to the concentration camp or the compulsory exposure of female children on the mountainside is the subject of detailed laws duly enacted and scrupulously observed. So to hold would, I think, be to strip 'the existing constitutional principle of the rule of law' . . . of much of its virtue.[45]

For Bingham, in other words, the "rule of law" includes not just predictable law but all the protections of individuals associated with the modern liberal constitutional order, especially those adopted by various international conventions on human rights.

No doubt there are a variety of explanations for this evolution in the understanding of the rule of law. A pivotal event was likely the Nuremberg Trials conducted by the victorious allies in 1945-46 to try leading figures of the defeated Nazi regime. The Nazis of course were guilty of enormous atrocities. Churchill at Yalta argued that the Nazi leaders should just be rounded up and shot.[46] He was outvoted by Roosevelt and Stalin, the latter on the ground that a criminal prosecution would make a valuable show trial.[47] The ex-Nazis were accordingly charged with various offenses, including the crimes of engaging in offensive war and genocide, which at that time had a weak provenance in terms of customary international law.[48]

The Nuremberg defendants naturally took the position that their actions were legal under German law as it existed when they took them. They were just following orders, and the orders came from higher authorities empowered to give them. Thus, they argued, the charges of the Nuremberg prosecutors, including Robert Jackson, on leave from the U.S. Supreme Court, were a form of ex post facto criminal liability inconsistent with the rule of law.[49] A careful dissection of the facts would probably have revealed that many, perhaps most, of the atrocities were not expressly authorized by duly enacted German law—certainly not by any law that was publicly known and applied in a consistent fashion.[50] But the tribunal did not have the time or the patience to undertake such a careful refutation of the defense.

---

45 *Id.* at 67.

46 Ian Cobain, *Britain Favoured Executive Over Nuremberg Trials for Nazi Leaders*, THE GUARDIAN (Oct. 25, 2012, 7:05 pm), https://www.theguardian.com/world/2012/oct/26/britain-execution-nuremberg-nazi-leaders.

47 *Id.*

48 As the chief prosecutor, effectively conceded. *See* Second Day, Wednesday, 21 November 1945, *in* 2 TRIAL OF THE MAJOR WAR CRIMINALS BEFORE THE INTERNATIONAL MILITARY TRIBUNAL, NUREMBERG 14 NOVEMBER 1945–1 OCTOBER 1946 98–102 (1947).

49 *See* Charles E. Wyzanski, *Nuremberg: A Fair Trial? A Dangerous Precedent*, THE ATLANTIC, Apr. 1946, at 60–65 (voicing similar concerns).

50 For an account of the deterioration of the Nazi legal system as the regime went on, see INGO MULLER, HITLER'S JUSTICE: THE COURTS OF THE THIRD REICH (Deborah Lucas Schneider, Trans., 1991).

The primary response, both at the trials and afterwards, was that the conduct of the defendants was contrary to customary international law, which in turn included inchoate ideals about the importance of fundamental human rights. Which is to say, if we assume that fundamental human rights are incorporated into the concept of the rule of law, the prosecutions were consistent with the rule of law.

The rights conception of the rule of law, as articulated by Lord Bingham, is largely a continuation and expansion of the version that grew out of the Nuremberg prosecutions. It is not surprising that he would come to see universal human rights as a critical aspect of the rule of law. His career as a senior British judge coincided with the enactment of the Human Rights Act of 1998, which was intended to give at least some direct effect in the United Kingdom of the European Convention on Human Rights and its interpretation by the European Court of Justice.[51] Later in his career, he also witnessed the adoption by Parliament of the Constitutional Reform Act of 2005, which transformed a committee of the House of Lords into a Supreme Court, and laid the groundwork, in the eyes of some, for more rigorous enforcement of individual rights notwithstanding traditional British norms of legislative supremacy.[52] As a senior judge charged with the faithful enforcement of these provisions, it is not surprising that he came to regard them as a critical component of the rule of law.

More generally, the rights version of the rule of law is clearly the dominant conception today among the spokespersons of the United Nations and leaders of the European Union. Particularly notable in this regard are the efforts of the Venice Commission, established by the European Council. In its *Report on the Rule of Law*, issued in 2011, the Commission concluded that the rule of law was "indefinable," but could be captured by a "checklist" of attributes, against which individual legal regimes or actors can be judged.[53] The checklist includes features commonly associated with the predictability conception, such as the principles that executive actors are bound by existing law, the law should be certain, and individuals should be able to challenge

---

51 Erin F. Delaney, *Judiciary Rising: Constitutional Change in the United Kingdom*, 108 NW. U. L. REV. 543, 556–57 (2014).

52 For these developments, see *id.* at 555–73.

53 EUROPEAN COMMISSION FOR DEMOCRACY THROUGH LAW (VENICE COMMISSION), REPORT ON THE RULE OF LAW (2011). The statement that the rule of law is undefinable appears in a summary document released on the internet. Venice Comm'n, *Rule of Law*, COUNCIL OF EUROPE, https://www.venice.coe.int/WebForms/pages/?p=02_Rule_of_law&lang=EN (last visited Sept. 19, 2022) ("While drafting the report, the Venice Commission reflected of [sic] the definition of the Rule of Law and reached the conclusion that the Rule of Law was undefinable… Rather than searching for a theoretical definition, it therefore took an operational approach and concentrated on identifying the core elements of the Rule of Law."). The report itself, with the checklist, can be found at VENICE COMM'N, REPORT ON THE RULE OF LAW, COUNCIL OF EUROPE (Mar. 25--26, 2011), https://www.venice.coe.int/webforms/documents/?pdf=CDL-AD(2011)003rev-e [hereinafter Venice Comm'n, Report on the Rule of Law].

government actions before independent and impartial courts.[54] But it also includes "respect for human rights" and the principle of non-discrimination on grounds such as "race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status."[55] Applying this checklist, the Venice Commission later released reports condemning the member states of Hungary and Poland for violations of the rule of law, particularly for measures interfering with the independence of the judiciary.[56] These reports ultimately led to efforts by the EU Commission to impose monetary sanctions on Hungary and Poland in an effort to deter these transgressions.[57]

What can be said by way of generalization about the rights conception of the rule of law? First, the rights conception does not reject the value of predictability in the law. Rights proponents always endorse features that predictability theorists applaud as designed to "guide the behavior" of those subject to the law: clarity, generality, publicity, prospectivity, congruence between the law on the books and the law as enforced, the right to judicial review by independent courts, etc. In this respect, there is an important overlap between predictability theorists and rights theorists like Bingham and the Venice Commission.

Second, rights theorists reject the idea that the rule of law promotes liberty only in the Hayekian sense that what is not clearly prohibited is permitted. They fear, perhaps rightly, that a conception of the rule of law limited to predictability can be enlisted on behalf of a legal regime that uses the law to exploit minorities, suppress dissent, destroy free elections, or worse. Accordingly, rights theorists would build into the concept of the rule of law a set of stops, in the form of a list of individual rights, designed to eliminate the risk of the rule of law being invoked as an instrument to achieve ends that violate the norms of the modern liberal constitutional order.

Third, rights theorists generate a conception of the rule of law that is far more complex than that associated with predictability theorists. Once one

---

54 Venice Comm'n, Report on the Rule of Law, *supra* note 53, at 10–12.

55 *Id.* at 12–13.

56 DUNJA MIJATOVIC, COUNCIL OF EUROPE, REPORT FOLLOWING HER VISIT TO HUNGARY FROM 4 TO 8 FEBRUARY 2019 (May 21, 2019), https://rm.coe.int/report-on-the-visit-to-hungary-from-4-to-8-february-2019-by-dunja-mija/1680942f0d; DUNJA MIJATOVIC, COUNCIL OF EUROPE, REPORT FOLLOWING HER VISIT TO POLAND FROM 11 TO 15 MARCH 2019 (June 28, 2019), https://rm.coe.int/report-on-the-visit-to-poland-from-11-to-15-march-2019-by-dunja-mijato/168094d848.

57 Case C-157/21, Republic of Poland v. European Parliament, ECLI:EU:C:2022:98, ¶ 17 (Feb. 16, 2022); Case C-156/21, Hungary v. European Parliament, ECLI:EU:C:2022:97, ¶ 111--112 (Feb. 16, 2022). See also Gabriella Baczynska & Gergely Szakacs, *In a First, European Union Moves To Cut Hungary Funding Over Damaging Democracy*, REUTERS (Sept. 18, 2022), https://www.reuters.com/world/first-eu-seen-moving-cut-money-hungary-over-damaging-democracy-2022-09-18/; Daniel Tilles, *After Moving To Cut Hungary's Funds, EU Is "Analysing Poland", Which Has "Many Problems"*, NOTES FROM POLAND (Sept. 19, 2022), https://notesfrompoland.com/2022/09/19/after-moving-to-cut-hungarys-funds-eu-is-analysing-poland-which-has-many-problems/.

works through the list of procedural and substantive rights endorsed by rights theorists, it is difficult to distinguish the rule of law from a catalogue of features that characterize liberal constitutionalism more generally. This complexity also makes it difficult if not impossible to define the "rule of law" in any easily comprehensible manner, which in turn invites skepticism about whether the rule of law is more than a slogan or a charge to hurl against decisions or regimes that one dislikes.

II. WHICH IS THE BETTER VERSION?

The rule of law is a normative vision. It describes an ideal state of affairs—not any existing legal system. Of course, some legal systems come closer to realizing this ideal than others. But everyone seems to agree that no system, either historically or presently existing, has completely incorporated the ideal of the rule of law.[58]

If the rule of law simply describes a kind of ideal state, it might seem that the rights conception is superior to the predictability version. After all, the rights version contains more attractive normative features than does the predictability one. The rights version incorporates a panoply of individual rights, such as freedom of expression and freedom from race and gender discrimination. These rights are regarded today as essential aspects of a liberal constitutional order. And because it incorporates these restrictions on state action, the modern version does not suffer the embarrassment that something like the Nazi regime, or the apartheid regime of South Africa, or a system that permits slavery, might be deemed to be consistent with the rule of law. If one conception incorporates more valuable things than the other, and by incorporating more valuable things avoids the embarrassment of being potentially compatible with evil, why adhere to the narrow predictability ideal?

It must be stressed here that the issue is one of definition. Should the concept of the "rule of law" be defined relatively narrowly, in terms of the capacity of a system to allow individuals to predict when state officials will act coercively against them, or should it be defined relatively expansively, as entailing not just predictability but also a cluster of individual rights deemed to be fundamental? As a matter of policy, this is not an either/or choice. All or nearly all proponents of the predictable law conception also favor individual rights and all or nearly all proponents of the rights conception

58 *E.g.*, HAYEK, CONSTITUTION, *supra* note 23, at 311 (describing the rule of law as "a meta-legal doctrine or a political ideal"). *See* Jeremy Waldron, *The Concept and the Rule of Law*, 43 GA. L. REV. 1, 44–45 (2008) ("[T]he requirements associated with the Rule of Law are all matters of degree. . . . Moreover, this feature of the Rule of Law seems essential to the work that it does as a political ideal. We use it to make nuanced and qualified assessments as well as all-or-nothing condemnations or commendations of systems of governance.").

favor predictability. [59] The question is whether predictability and rights should be considered as separately as ideals, or should be merged into a single concept.[60]

A. *The Case for the Predictability Definition*

There are multiple overlapping reasons to prefer the narrow, predictability understanding of the rule of law. One is analytical clarity. The narrow definition identifies a state of affairs that is relatively easy to comprehend: is the polity so structured so that individuals know what to expect from government officials? The rights conception, while not repudiating the desirability of predictability, creates a much more complex menu of desirable governmental attributes. The government must not only be predictable, it must also adhere to a list of individual rights–a list which may be sharply contested and continually evolving. The rights conception thus "ranges over highly diverse subject matter…arguments and criticism purportedly in the name of the 'rule of law' tend to be arguments and criticisms in the name of too many things at once."[61]

Confining the rule of law to the narrow definition also helps highlight whether a particular regime is adhering to widely-shared conceptions of individual rights. Under the predictability conception, it is possible to say that a particular regime adheres to the rule of law but does not respect one or more individual rights. The criticism can be targeted to the failing. Under the rights conception, the judgment that a particular regime does or does not adhere to the rule of law is a composite of its record in terms of predictability and its respect for a list of individual rights. Thus, in denouncing a regime for its failure to adhere to the rule of law, it becomes unclear which failing is being singled out for condemnation. Is the regime coming up short because it acts capriciously, or because it fails to respect one or more rights?

It is also plausible that it is easier to reach general agreement about the desirability of legal predictability than it is about the relevant set of rights. As any lawyer will tell you, the law is never entirely predictable. New issues continually arise, which often generate conflicting answers when attempts are made to extrapolate from authoritative texts or what has been decided in the past. But there is nearly always a significant core of propositions that are settled.[62] Courts generally have little trouble identifying and adhering to such

---

[59] Hayek, for example, was a strong proponent of negative individual rights like freedom of speech even though he condemned socialism and egalitarian redistribution. HAYEK, SERFDOM, *supra* note 28, at 112–23.

[60] HAYEK, CONSTITUTION, *supra* note 23, at 112–23.

[61] Robert S. Summers, *A Formal Theory of the Rule of Law*, 6 RATIO JURIS 127, 137 (1993). *See also* TAMANAHA, *supra* note 19, at 113 ("The rule of law cannot be about everything good that people desire from government. The persistent temptation to read it this way is a testament to the symbolic power of the rule of law, but it should not be indulged.").

[62] *E.g.*, FREDERICK SCHAUER, THINKING LIKE A LAWYER 18-23 (2009).

settled propositions. The same holds for executive actors charged with bringing enforcement actions. Also, legal predictability can be married to a wide variety of regimes, ranging from laissez faire to socialism, from dictatorship to democracy. Although it would be misleading to describe the predictability norm as "politically neutral,"[63] it is likely to receive widespread assent on the ground that predictable law makes a legal regime—whatever its other value commitments—more effective.

The relevant set of rights that should be incorporated into the rights conception of the rule of law is more likely to be a matter of disagreement, as illustrated by unending controversy over questions like abortion and the death penalty. And the set of recognized rights is always changing, as witnessed by the emergence within a matter of two decades of strong support for LGBT rights. Many regimes (not as much the U.S.) have also come to embrace not only negative rights—rights to be free of certain kinds of government intrusion—but also positive rights, such as rights to basic subsistence, shelter, medical care, and an education.[64] To the extent that individual rights come to be understood to include such positive rights, this inevitably generates complicated questions about the definition of such rights, how their provision will be funded, and how they will be enforced, all of which makes the "rights" perspective more controversial.

Finally, it is significant that proponents of the rights conception all concur (at least implicitly) that it is important that the set of fundamental rights they endorse should be predictability enforced. In effect, predictable enforcement is a necessary condition of securing a regime that respects individual rights. Proponents of the rights perspective often skip quickly over the point, or assume that once the relevant rights are recognized, they will be enforced. But in an era in which authoritarian government is on the rise, enforcement of individual rights should not be taken for granted. Which suggests that predictable enforcement is worthy of attention separate and distinct from discourse about what rights should or should not be included in the system of law that governs the polity.

B. *Rights as a Shield from "The Prerogative State"*

An important counter-argument is that the rule of law, in the narrow sense of predictability, may be impossible to sustain unless the governing regime also respects at least a core of basic rights. The counter-argument presents an urgent question, as it has become clear that autocratic regimes like Putin's Russia and Communist China are increasingly prone to violate the rights of dissidents and other "enemies" of the regime. One-party rule, it seems, creates a kind of downward spiral in terms of respect for rights (at

63 Summers, *supra* note 61, at 136 (describing the formal conception of the rule of law as "more or less politically neutral").

64 *See, e.g.*, Jamal Greene, 12 Law & Ethics Hum. Rts. 37 (2018).

least the rights of anyone that the party in power deems to be a threat to its continued dominance). The question becomes: Does the progressive erosion of rights in one-party states eventually infect the entire legal system, such that the predictability of law also inevitably collapses? If so, then it would seem that the concept of the rule of law must include the protection of at least some basic rights, if it is to have any enduring (that is, predictable) aspect.

Discussions of this important question are often framed in terms of Ernest Fraenkel's analysis of the "dual state."[65] Fraenkel took as his object lesson Germany after the Nazi Party seized control pursuant to a decree of emergency powers in 1933.[66] Fraenkel described the resulting regime (up to 1941, when he fled to England and then the U.S.) as consisting of a dual state: the "prerogative state" consisting of any issues that the Nazis unilaterally decided were important to their continued absolute rule, and the "normative state," consisting of issues that the Nazis decided could be left to be governed by the legal norms that prevailed before they seized power.[67] The prerogative state was truly terrifying for those regarded as enemies according to Nazi ideology: communists, socialists, Jehovah's Witnesses, mentally impaired persons, and of course Jews. These persons were stripped of all rights, and it would seem they were denied any legal predictability, other than the prediction that they had no rights and their property, jobs, and very lives existed at the sufferance of Nazis.

Fraenkel also concluded, however, that the Nazis allowed the normative state, characterized by a significant degree of legal predictability, to continue to operate in critical realms like heavy industry and agricultural production. This was to facilitate the rearmament of Germany in preparation for war.[68] The Nazis wanted to create a powerful military machine and to do so before their potential adversaries caught on and joined in an arms race. This required leaving property rights, corporate governance principles, and contract law as it applied to heavy industry and agriculture largely undisturbed. Workers were stripped of their rights of collective bargaining. But the full employment produced by the rearmament campaign evidently quieted any mass movement of workers in opposition to the regime.[69]

Fraenkel's characterization of the Nazi regime as a dual state has played an important role in assessing the evolution of Communist China. Eva Pils has compared the regime of Chairman Xi Jinping to Fraenkel's prerogative state, with the gloomy assessment that the rights of dissenters have been almost completely crushed.[70] Taisu Zhang and Tom Ginsburg counter that even if China is becoming more of a dictatorship, it is doing so while also

---

65 ERNST FRAENKEL, THE DUAL STATE: A CONTRIBUTION TO THE THEORY OF DICTATORSHIP (1941).

66 *Id.* at 3-56.

67 *Id.* at 5, 65.

68 *Id.* at 175–87.

69 *Id.* at 179-85.

70 *See generally* Eva Pils, *China's Dual State Revival* (unpublished draft) (on file with the author).

being increasingly committed to "bureaucratic legalization," which is regarded by the regime as an important tool for maintaining centralized party control over the country.[71] In other words, China, for its own reasons, has fostered a kind of normative state in the face of a prerogative state that eliminates nearly all vestiges of liberal rights.

These ruminations on China seem to confirm Fraenkel's assessment of Nazi Germany (at least up to 1941): The dual state can effectively obliterate any semblance of human rights (and with it the rule of law) for any group deemed to be an "enemy" of the prerogative state. Yet at the same time, the prerogative state may continue to honor something like the rule of law in sectors of society deemed to be critical to the goals of the prerogative state. This conclusion would seem to support Raz's thesis that the rule of law in the sense of legal predictability can at least partially co-exist with regimes that otherwise engage in extreme violations of basic rights. It also suggests that protection of rights is not a necessary condition of preserving the rule of law, at least when it is in the interests of a prerogative state to do so. This last conclusion is obviously highly tentative, and it remains possible that a regime that holds individual rights of "enemies" in utter contempt will eventually degenerate into a pure terror state. But Fraenkel's account and assessments of contemporary China suggest that it is at least theoretically possible to distinguish the rule of law in the sense of predictability from the protection of basic individual rights.

C. *Summing Up*

None of the foregoing should be taken as a criticism of the importance of individual rights, whether embodied in a bill of rights or in various documents generated by bodies of international lawyers. Individual rights are unquestionably of great importance, and every decent government should strive to respect them. The point is simply that having a government that acts in predictable ways in the exercise of coercive power is also a good thing. And there is reason to keep the good that derives from government predictability distinct from the good that comes from government respecting individual rights.

The principal takeaway from this discussion is that the predictability understanding of the rule of law is preferable for several related reasons. The narrow understanding provides a more precise basis for criticizing (or praising) different government regimes. By adhering to the predictability version, we assure that the good that comes from government predictability is not submerged—and thus in danger of being forgotten—in a general approbation of the ideals associated with individual rights. Maintaining the predictability understanding also highlights that the good that comes from government predictability may be in tension with some of the more expansive

[71] Taisu Zhang & Tom Ginsburg, *China's Turn Toward Law*, 59 VA. J. INT. L. 306, 312–13 (2019).

conceptions of individual rights. It may not be possible to pursue both goals simultaneously, making it necessary to consider trade-offs. Finally, affirming the predictability understanding does not disparage the ideals associated with individual rights. It simply requires that they be justified on their own terms and not by generalized assertions that they are required by "the rule of law."

III. THE INSTITUTIONAL FOUNDATIONS OF THE RULE OF LAW

Our task is incomplete, because there are several versions of the predictability conception of the rule of law. Although the different conceptions of the rule of law are united in their understanding that a key purpose of the rule of law is to make the use of government coercion predictable, they offer divergent ideas about the institutional arrangements that provide this kind of certainty about the legal system. Some stress the importance particular substantive rules like the non-retroactivity of criminal law (*nulla poena sine lege*) and the principle of res judicata that bars re-opening final judicial judgments. Others stress the importance of judicial review of executive or administrative action. Still others stress the importance of access to open hearings and the requirement of reason-giving before coercive action is taken.[72]

Perhaps the most common theme in the literature on the rule of law, in the sense of predictability about government coercion, is the importance of an independent judiciary. So I propose to start there. Beginning with Dicey, theorists of the rule of law have emphasized the importance of an independent judiciary. Indeed, Dicey went so far as to write that a legal regime is characterized by the rule of law when its principles are "the result of judicial decisions determining the rights of private persons in particular cases brought before the courts."[73] But others, including Hayek, Raz, Jeremy Waldron, Bingham, and the Venice Commission have also put great stress on the importance of an independent judiciary. Unfortunately, there is less agreement on the institutional conditions that give rise to an independent judiciary.

American authors are likely to cite separation of powers, and the related idea of checks and balances, as necessary conditions supporting an independent judiciary. Several of the historic expositors of separation of powers, including Montesquieu and the authors of the Federalist Papers, associated the principle of separation of powers with the preservation of "liberty," which they understood as freedom to act within the constraints of law—an idea not dissimilar from Hayak's ideas about the relationship

---

72 *See* Waldron, *supra* note 58, at 60 (arguing that the rule of law must include "the opportunities for argumentation that a free and self-possessed individual is likely to demand").

73 DICEY, *supra* note 25, at 115.

between predictability of the law and individual liberty.[74] Montesquieu wrote that the legislature and executive must be kept in separate hands, lest the same body of persons "should enact tyrannical laws, to execute them in a tyrannical manner."[75] He made a similar point about the judicial power:

> Again, there is no liberty, if the judiciary power be not separated from the legislative and the executive. Were it joined to the legislative, the life and liberty of the subject would be subject to arbitrary control; for the judge would then be legislator. Were it joined to the executive power, the judge might behave with violence and oppression.[76]

The U.S. Constitution, reflecting in significant part the ideas of Montesquieu, contains a more fully worked out system of separation of powers. The document contemplates that the Congress will be the source of legal change. The executive branch was apparently conceived to be a rather skeletal operation, headed by the President, who was subject to an express duty to see that the laws are "faithfully executed."[77] And the federal judiciary was to have life tenure and secure compensation, assuring a high degree of independence from both the Congress and the President.[78]

The theory that would explain why such a separation of powers would lead to a preservation of liberty was not clearly spelled out, either by Montesquieu or the Framers of the Constitution. The best explanation is based on checks and balances. Coercion of citizens by the government requires the concurrence of all three institutional actors: the legislature must pass a law, the executive must enforce it, and the judiciary must interpret it. If kept separate, each branch of the government will be checked by the knowledge that the other two must go along in order to achieve the desired coercion. If the legislature knows that the laws it enacts will be enforced and interpreted by other actors, it will be deterred from enacting oppressive laws. If the executive knows that its enforcement policies will be subject to interpretation by the judiciary, and to override by the legislature, it will be deterred from enforcing the law in a partial or uneven fashion. And if the judiciary knows that its interpretations will be subject to enforcement by the executive, and override by the legislature, it will be deterred from adopting interpretations strongly opposed by the other divisions of government.

---

74 BARON DE MONTESQUIEU, 1 THE SPIRIT OF THE LAWS [8] (Timothy Dwight et al. eds., Thomas Nugent trans., The Colonial Press, rev. ed., 1899) (1748) ("Liberty is a right of doing whatever the laws permit[.]").

75 *Id.* at 151–52.

76 *Id.* at 152.

77 U.S. CONST. art. II, § 3; THE FEDERALIST NO. 48, at 256 (James Madison) (George W. Carey & James McClellan eds., 2001).

78 U.S. CONST. art. III, § 1; THE FEDERALIST NO. 78, at 401 (Alexander Hamilton) (George W. Carey & James McClellan eds., 2001).

In the abstract, the idea that an institutional separation of powers will sustain an independent judiciary is not without merit.[79] Yet whether it consistently works this way in practice is open to doubt. With the extension of popular democracy to the method of selection of the President and the Senate, the prospect of swing elections producing unified control of the legislative and executive branches increases. And if unified control persists for a significant time, as it did during the post-Civil War period and during the New Deal, unified control can extend to the composition of the judiciary. If the same party controls all three branches, there is much less potential for checks and balances working.

We also have the example of parliamentary systems, where the legislative and executive branches are effectively controlled by the parliamentary majority. Some of these systems, most prominently those of the United Kingdom and former colonies modelled on the government of the UK, rate high on the scale of respect for the rule of law in the sense of predictability, as well as individual liberty more generally. On the other side of the coin, history is littered with examples in which countries have adopted constitutions with highly articulated separation of powers provisions, only to collapse into one form or another of authoritarian dictatorship.[80]

The inconsistent correlation between constitutions based on separation of powers principles and respect for the rule of law has led some commentators to conclude that commitment to the rule of law is a function of cultural factors. The rule of law exists in countries, like England, "owing to a widespread and unquestioned belief in the rule of law, in the inviolability of certain fundamental legal restraints on government, not to any specific legal mechanism."[81] Dicey was apparently of this view, characterizing the respect for the rule of law in his country as ultimately grounded in "political" or "moral" factors, not legal institutions.[82] Hayek too ultimately settled for this explanation. "In a democracy," he wrote, the rule of law will not prevail "unless it forms part of the moral tradition of the community, a common ideal shared and unquestionably accepted by the majority."[83] As for the United States, Tocqueville suggested that a relatively strong commitment to the rule of law could be explained in part by the high percentage of lawyers in

---

79 Kim Scheppele, in an interesting essay, points out that the recent rise in autocratic rule in Hungary, Russia, Turkey, and Venezuela has occurred in countries in which the constitution imposes fewer institutional checks on a popularly-elected executive than the United States. Kim Lane Scheppele, *Autocratic Legalism*, 85 U. CHI. L. REV. 545 (2018). The first step taken by most of these autocrats, once they have secured enough power to direct constitutional changes, is to eliminate the independence of the judiciary. *Id.* at 551–53.

80 *See* ZACHARY ELKINS, TOM GINSBURG & JAMES MELTON, THE ENDURANCE OF NATIONAL CONSTITUTIONS 25 (2009) (noting that the U.S. Constitution with its separation of powers was widely copied in Latin American in the nineteen century, and nearly all these constitutions have disappeared).

81 TAMANAHA, *supra* note 19, at 58.

82 DICEY, *supra* note 25, at 26–35.

83 HAYEK, CONSTITUTION, *supra* note 23, at 311.

government.[84] Explication of existing principles of law is the stock in trade of lawyers, so a government of lawyers can be expected to be sensitive to preserving the rule of law.

Again, there is clearly something to the case for a cultural explanation.[85] But these explanations ultimately beg the question. How does a cultural commitment to the rule of law get started, unless the institutions of society make settled principles of law a significant factor in determining the distribution of rights and obligations in the society? A taste for legal predictability presumably gets started because the law is sufficiently predictable to generate significant payoffs to understanding the law. And lawyers are presumably attracted to government service in the U.S. because a capacity for explicating the law gives lawyers a comparative advantage relative to those who do not share such training. So we are back to searching for institutional variables that generate and sustain the rule of law, understood to mean a system characterized by a high degree of predictability about the requirements of the law.

There is also the problem of explaining the other half of the rule-of-law equation: the practice of the executive following the judgments of the courts before coercing members of the public. An independent judiciary may be wholly committed to the principled explication of the law. But if executive authorities regard themselves at liberty to disregard judicial judgments they consider inconvenient, individuals will have no confidence about predicting what they can and cannot do under the law.

In order to explain how an independent judiciary can give rise to legal predictability, we need a more complete theory of how an independent judiciary is likely to function given other limitations on its power. The salient point here, which has been largely ignored by commentators on the rule of law, is that the judiciary, even if assured of complete independence in rendering judgments in the cases that come before it, is in other respects completely *dependent* on the other branches of government.[86] These other respects include the scope of the courts' jurisdiction over particular categories of cases, appropriations for courthouses and ancillary personnel like clerks and bailiffs, and the power of the legislature to override judicial decisions, at least those not based on the Constitution. Perhaps most importantly, these other respects include enforcement of the judgments the courts have rendered. Few if any constitutions (including the U.S.

---

84 ALEXIS DE TOCQUEVILLE, DEMOCRACY IN AMERICA 272–80 (Phillips Bradley ed., Henry Reeve & Francis Bowen trans., Alfred A. Knopf 1990) (1835).

85 *See* TAMANAHA, *supra* note 19, at 4 ("If it is not already firmly in place, the rule of law appears mysteriously difficult to establish.").

86 For a model explaining in an analogous fashion how judicial restraint emerges from a balance between independence and dependence, see John A. Ferejohn & Larry D. Kramer, *Independent Judges, Dependent Judiciary: Institutionalizing Judicial Restraint*, 77 N.Y.U. L. REV. 962 (2002).

Constitution) have conferred independent authority on courts to enforce their judgments; instead, they charge the executive branch with doing so.[87]

The role of judicial independence is the more familiar half of the picture. As Alexander Hamilton wrote in *The Federalist* No. 78, judicial independence is necessary in order "to guard the Constitution and the rights of individuals from . . . dangerous innovations in the government, and serious oppressions of the minor party in the community."[88] Yet if a legal system did no more than create a body of judges completely independent of the political process, this would raise something of the opposite concern—the prospect of appointing "a bevy of Platonic Guardians" that would rule society in the name of its own vision of the good.[89] Hamilton in writing *The Federalist* No. 78 was anxious to assure his audience that this would not happen. He emphasized that the federal judiciary would be "beyond comparison the weakest of the three departments of power."[90] It would have "no influence over either the sword or the purse" and thus would exercise "neither FORCE nor WILL, but merely judgment."[91] Indeed, he specifically noted that the judicial branch "must ultimately depend upon the aid of the executive arm even for the efficacy of its judgments."[92]

Extrapolating a bit, Hamilton was arguing that a body of judges which enjoys complete *independence* from the political branches in rendering judgments, but complete *dependence* on the political branches in other respects, will necessarily exercise "judgment" as opposed to "will" in deciding disputes that come before it. We can say that such judges would thus naturally gravitate to a strategy of resolving disputes in a manner consistent with settled law.

Why is that? Hamilton did not spell out why a combination of independence-in-judging with dependence-in-enforcing would produce judges inclined to enforce settled understandings of the law. Perhaps the explanation would go something like the following.[93] If judges are insulated from direct political pressure in how they decide individual cases, they will eschew favoritism toward politically favored parties or hostility toward

---

87 Enforcement of judgments by the U.S. Marshall's Service, a division of the Justice Department, is required by statute. See 28 U.S.C. § 566(a). (A predecessor of the statute dates from the Judiciary Act of 1789, 1 Stat. 73, 87.) Interestingly, the statute requires the Service to enforce "all orders of the United States District Courts, the United States Courts of Appeals, the Court of International Trade, and the United States Tax Court," but does not require that it enforce orders of the Supreme Court. *Id.*

88 THE FEDERALIST NO. 78, supra note 78, at 405.

89 *See* LEARNED HAND, THE BILL OF RIGHTS: THE OLIVER WENDELL HOLMES LECTURES 73 (1958) ("For myself it would be most irksome to be ruled by a bevy of Platonic Guardians, even if I knew how to choose them, which I assuredly do not.").

90 THE FEDERALIST NO. 78, *supra* note 78, at 402.

91 *Id. (emphasis in original).*

92 *Id.*

93 *See* Thomas W. Merrill, *Legitimate Interpretation – Or Legitimate Adjudication*? 105 CORNELL L. REV. 1395, 1412–17 (2020) [hereinafter Merrill, *Legitimate Interpretation*].

politically disfavored ones.[94] One can say they will resolve questions of fact and law-application in the disputes that come before them in an impartial manner. At the same time, if judges know they are ultimately dependent on the political branches in other respects, they will not stray very far, certainly not on a consistent basis, from settled expectations about the nature of the decisional norms they use in assessing the conduct of the parties that come before them. They will adopt legal principles that are generally understood to conform to the rule of law in the sense of predictability. If they engage in too much legal innovation, they will find that their interpretations of the Constitution have been overridden by amendment, or their interpretations of federal statutes have been overridden by new legislation, or perhaps that Congress has curtailed their jurisdiction, limited the funds appropriated for courthouse renovations, or decided to pack the courts with more compliant judges.[95] In the extreme case, they will find that the executive has declined to enforce the judgments they have reached.[96]

The important point is that the independent judiciary must be aware of the potential for *pushback* by the political branches if the judges stray too far or too long from what the other political branches regard as acceptable legal norms. Probably the best form of pushback takes the form of discrete overriding of judicial decisions, as when the Sixteenth Amendment was adopted overriding a Supreme Court decision suggesting that an income tax was impermissible,[97] or when Congress adopted legislation overturning a decision holding that discrimination based on pregnancy was not sex discrimination.[98] These sorts of actions remind the courts that the legal system presupposes a "continuing dialogue among the three branches of government" about acceptable legal norms.[99] Jurisdiction stripping, funding cutoffs, and court packing are blunter instruments, and carry with them the risk of impairing judicial independence as well as enforcing appropriate judicial modesty. Nevertheless, if independent judges become convinced that there is little or no possibility of pushback, an important constraint that inclines them toward adhering to settled law will disappear.

To these general constraints under which courts operate, we can add the importance of stare decisis to judicial legitimacy. Historically, arguments

---

94 *Id.*

95 For a more complete account of the many ways in which the courts are subject to "a wide array of controls by the political branches" see Ferejohn and Kramer, *supra* note 86, at 977 and more generally at 976-994.

96 For an instructive study of episodes in which federal officials have declined to comply with district court judgments, see Nicholas R. Parrillo, *The Endgame of Administrative Law: Governmental Disobedience and the Judicial Contempt Power*, 131 HARV. L. REV. 685, 712–35 (2018) (the courts generally backed down).

97 See U.S. Const., amend XVI, overriding Pollock v. Farmers' Loan & Trust Co., 158 U.S. 601 (1895).

98 Pregnancy Discrimination Act of 1978, 92 Stat. 2076 (codified at 42 U.S.C. § 2000e(k) (2018)), overriding General Electric Co. v. Gilbert, 429 U.S. 125 (1976).

99 United States v. Home Concrete & Supply, LLC, 566 U.S. 478, 503 (2012).

from precedent have been closely associated with private law, where common law has played a prominent role. In recent decades, however, private law has become increasingly dominated by statutes, including uniform laws, model state laws, and federal and state regulatory enactments. Meanwhile, amendments to the Constitution and to many framework statutes have become increasingly difficult to obtain. With the "statutorification" of private law[100] and gridlock afflicting legislated changes in public law, arguments from precedent have receded in private law and surged to the fore in public law.[101] This is especially pronounced in federal constitutional law, where nearly every contested case is resolved by following, distinguishing, or qualifying existing precedent.[102] Arguments from precedent are also increasingly prevalent in cases governed by statute, especially where the statute has been around for a long time and frequently litigated. In general, the older the text, and the more frequently it has been interpreted in the past, the greater the likelihood that we will find legal argumentation based on precedent, rather than interpretation of the instructions of the enacting body. Public law, and especially constitutional law, has become the new common law.

The logic that impels courts to follow precedent when it exists, especially when the precedent is considered settled, is straightforward: if courts do not adhere to their own precedents, no one else will either. Having "no influence over either the sword or the purse," as Hamilton pithily observed, and no independent authority to enforce their judgments, the power of courts is a function of the predictability of their judgments.[103] And the predictability of their judgments is increasingly a function of their fidelity to their own precedents.

Generalizing, one can say that courts are driven to the predictability conception of the rule of law because they have so little power. Such power as they possess is a function of their fidelity to settled law. Sometimes this requires that they enforce the plain meaning of a recently enacted statute or a rarely interpreted constitutional provision. More often, it means that they enforce directly applicable judicial precedent.

What then about the other half of the picture: the willingness of the executive to comply with judicial judgments about the law? The rule of law, under the predictability conception, requires that the *executive* conform its actions to settled understandings of the law. Even if the judiciary, for its own reasons just considered, decides to adopt a strategy of adhering to settled

---

100 *See* GUIDO CALABRESI, A COMMON LAW FOR THE AGE OF STATUTES 1–10 (1982).

101 *See* Thomas W. Merrill, *Interpreting an Unamendable Text*, 71 VAND. L. REV. 547, 578–83 (2018).

102 *See, e.g.*, Ethan Buenode Mesquita & Matthew Stephenson, *Informative Precedent and Intrajudicial Communication*, 96 AM. POL. SCI. REV. 755, 764 (2002) (noting that arguments from precedent "vastly outnumber all other kinds of arguments in attorney's written briefs, the Court's written opinions, and the justices' arguments in conference discussions."); *see generally* Jack Knight & Lee Epstein, *The Norm of Stare Decisis*, 40 AM. J. POL. SCI. 1018 (1996) (providing empirical data).

103 THE FEDERALIST NO. 78, *supra* note 78, at 402.

principles of law, why does the executive follow suit? Why doesn't the executive branch set up an office, called the Enforcement Bureau, which considers whether to enforce particular judicial judgments based on their compatibility with current executive policy preferences?[104]

To some extent, of course, executive discretion to decline to enforce settled law already exists, in the form of prosecutorial discretion. If Congress enacts a statute making the sale of marijuana a crime, and the courts uphold convictions based on the statute, a subsequent administration can nevertheless decide to decline to prosecute such offenses based on changing social attitudes about recreational pot.[105] Note, however, that this type of executive discretion involves a decision to decline to exercise coercive authority against particular individuals. Although it may be inconsistent with settled law, it does not involve the use of executive power to coerce individuals in a manner inconsistent with settled law.

The more interesting question is why the government refrains from attempting to coerce individuals when settled law indicates that it is impermissible to do so. Let us take an imaginary hypothetical to see how this would play out. Suppose the Supreme Court holds that a particular drug, call it Euphoria, is not a controlled substance subject to criminal prosecution under the federal drug laws. The incumbent administration disagrees, believing for good reasons that Euphoria is highly addictive and dangerous. It directs the Justice Department to prosecute Elmer for selling Euphoria. What happens next? The public defender appointed to represent Elmer files a motion to dismiss, citing the relevant Supreme Court precedent. The district judge promptly agrees and dismisses the indictment. If the government appeals, the court of appeals affirms. If the Justice Department takes the matter to the Supreme Court, asking the Court to overrule the controlling precedent, the Court rejects the request, pointing out the importance of adhering to settled law and urging the Department to seek an amendment to the drug laws from Congress.

Suppose the Department then decides to seize Elmer and throw him in a federal detention center without a trial. What is going to be the likely response? The organized bar, the mainstream media, and even social media will erupt in protest. Former Justice Department lawyers will file admonitory letters. Congress will hold hearings. Editorials will be written suggesting that this is an impeachable offense. If the administration has any sense, it will

104 *Cf.* Kate Andrias, *The President's Enforcement Power*, 88 N.Y.U. L. REV. 1031, 1037–38 (2013) (proposing the creation of an office within the White House to set guidelines for discretionary enforcement authority across the administrative state).

105 Memorandum from James M. Cole, Deputy Att'y Gen., Dep't of Just. to All U.S. Att'ys (Aug. 29, 2013), https://www.justice.gov/iso/opa/resources/3052013829132756857467.pdf (advising the exercise of prosecutorial discretion in enforcing the Controlled Substances Act when state laws legalize "small amounts of marijuana"). *See generally* MARIJUANA FEDERALISM: UNCLE SAM AND MARY JANE (Jonathan Adler, ed. 2020).

promptly release Elmer and issue a statement explaining that the action was based on a miscommunication.

What explains this reaction? The proximate cause is the behavior of the courts. They have acted with unanimity and without hesitation, grounding their action in prior controlling authority. They have said unequivocally that the behavior of the government is contrary to the rule of law. The prestige of the courts, built up over many years of more-or-less faithfully adhering to settled law, is such that the judgment of the courts is fatally damning of the behavior of the executive. This is enough for almost everyone to reach the judgment that the executive action must be condemned.

The ultimate cause runs much deeper. There is a deep desire, on the part of almost everyone, to live in a society that adheres to settled principles of law. Almost no one wants to live in a society where the police engage in random acts of brutality, tax collectors demand bribes, and the courts base their rulings on what the party in power tells them to do. When ordinary citizens learn that the government has defied the courts and acted contrary to what the courts say the law requires, they will be alarmed. Opinion polling will run strongly against the government. It has little choice but to back down.

The hypothetical is admittedly stylized and extreme, but it explains why the executive nearly always defers to the courts on the requirements of the law, at least when the courts have spoken clearly and consistently. People do not want to live under a rapacious state. Courts, as an institution that spans multiple administrations, will have built up a reputation for being more faithful to settled law than the executive. Thus, when the executive appears to be acting coercively in a manner than has been condemned by the courts as contrary to law, the people take this as a cue that the rule of law is jeopardy. This explains why the courts—the least powerful branch in terms of conventional sources of power—have been able to force the executive to adhere to the rule of law. They have done so only because of their own adherence to predictable law, at least to a greater extent than that of the executive.

IV. SOME IMPLICATIONS AND CONCERNS

The foregoing account, if correct, has several important implications for understanding the rule of law, as conventionally understood to mean predictability about government coercion. I will start with the implications, and then mention some concerns.

A. *Implications of the Predictability Understanding of the Rule of Law*

1. Is the Rule of Law a Virtue?

Perhaps the most far-reaching question raised by the predictable law conception is whether we can continue to describe the rule of law as a virtue. This is not a problem for the rights conception, since it posits that the individual rights associated with the liberal constitutional order are good things, and therefore the rule of law is a good thing. Nor is it a problem for libertarians like Hayek, who understand the rule of law as requiring a small set of general or abstract rules that leave large swathes of human activity unregulated. If one implicitly assumes that the rule of law, in the sense of predictability, will emerge in a society governed by a minimal state, then predictable law leaves large space for individual initiative, and hence liberty as thinkers like Hayek understand it. But if we define the rule of law as the predictable use of coercion by the executive, and exclude from the definition any conception of the permissible scope or ends of the state, then the rule of law can be married with all sorts of political regimes, including authoritarian or totalitarian states, and even states that practice apartheid or genocide. How can the rule of law be described as a virtue, if its only consequence – in some applications – would be to make the exercise of evil more predictable?

I think the argument can be made that the rule of law is a virtue even in these dire circumstances. The basic point is that if a regime is dedicated to perpetrating evil, those who may be victims are better off if this is predictable, than they will be if the designs of the state are concealed or obscure or are implemented in a random fashion by gangs of thugs allowed to commit atrocities without legal recourse. If the evil intentions of the state are publicly promulgated, clearly stated, consistently upheld that the courts, etc., then those who may fall victim to those designs will know that they must somehow try to avoid such an outcome, whether it be by escaping, resisting, or going into hiding. Forewarned is better than being fooled, or succumbing to wishful thinking.[106] This is hardly the kind of ringing endorsement of the rule of law familiar from various Law Day speeches. But it is enough to suggest that predictability about the possibility of coercion is always a virtue, even if the other conditions of the polity utterly fail to meet basic standards of human dignity.

---

[106] As John Gardner wrote:

> If the relevant populations are lucky enough to live under the rule of law, [they] should be able to find out, before and afterwards, what the law has to say about their actions, and the law should be such that, once they know what it says, they can judge when they are violating it and find a way to avoid doing so. They should be able to rely on what the law has to say to predict and plan for the official response.

Gardner, *supra* note 20, at 213.

2. Judicial Dependence

The previous sketch of the institutional features that give rise to the rule of law, in the sense of predictability, validates the common intuition that an independent judiciary is an important element in creating and sustaining the rule of law.[107] Courts must be sufficiently independent from the political branches that they can fearlessly enforce settled principles of law in resolving the cases that come before them. What most commentators on the rule of law have missed is the second half of Hamilton's essay in *The Federalist* No. 78: that courts must also be *dependent* on the political branches for their general efficacy, including enforcement of the judgments they reach.[108] If the political branches perceive that the courts apply decisional norms based on a good faith extrapolation from settled law, they will continue to abide by judicial understandings of the law. If they come to believe that courts are developing decisional norms based on their own policy preferences, support for the decisions of the independent judiciary will evaporate. If this happens, the rule of law may evaporate too.

3. Judicial Review

The account developed here also suggests that a broad right of judicial review of executive action is critical in creating and sustaining the rule of law. If the ultimate purpose of the rule of law is to make the use of coercive force by the executive predictable, and if the judiciary's penchant for enforcing settled law is the lynchpin in creating such a condition, then the executive must be answerable to the courts. The examples of recent presidential behavior set forth at the beginning of this paper suggest that Presidents are most tempted to disregard the settled understanding of the law when they think their action will not be judicially reviewable—either because it involves an unreviewable enforcement policy,[109] or it involves compliance with appropriations statutes thought to be enforceable only by Congress,[110] or because they imagine judicial proceedings will drag out long enough to make the matter moot.[111] That Presidents are increasingly tempted to skirt the

107 *E.g.*, DICEY, *supra* note 25, at 3; HAYEK, CONSTITUTION, *supra* note 23, at 319-20; Raz, *supra* note 31, at 214; Bingham, *supra* note 43, at 91-96; Venice Comm'n Report on the Rule of Law, *supra* note 53, at 10-12; Waldron, *supra* note 58.

108 *See* THE FEDERALIST NO. 78, *supra* note 78, at 402, 404.

109 Executive decisions not to enforce the law are generally regarded as unreviewable. *See, e.g.*, Heckler v. Chaney, 470 U.S. 821 (1985).

110 *See* Gillian Metzger, *Taking Appropriations Seriously*, 121 COLUM. L. REV. 1075. 1109–11 (2021) (describing doctrines limiting litigation over appropriations).

111 *See supra* note 11.

law when they think they can get away with it is deeply troubling.[112] Which suggests that there should be a strong presumption in favor of judicial review, as the courts have generally held.[113]

4. Settled Law

Hayek's insistence in his early writing that the rule of law requires bright-lines rules that admit of no discretion was overstated.[114] What is critical is that the executive be constrained by settled law. Settled law sometimes means relatively formal rules. But settled law also includes general standards that must be applied in a case-by-case fashion. Consider, for example, liability for negligence as it applies in the law of torts. Negligence liability is not open-ended: one must prove duty, causation and actual injury in addition to unreasonable behavior on the part of the defendant. And identifying what amounts to unreasonable behavior is also somewhat predictable: trial lawyers refer to published reports of jury verdicts in negotiating settlements, and insurance companies use rules of thumb in making payments.[115] In the end, however, liability turns on a standard, which necessarily applies ex post in a case-by-case fashion. This is settled law, and is not inconsistent with a regime characterized as one governed by the rule of law.

5. Common Law or Codified Law

There is no necessary connection between the common law and the rule of law, other than the historically contingent fact that the judicial habit of following precedent originated in the common law. One can easily imagine a code system, such as exists in civil law countries, which conforms to the ideal of the rule of law. In implementing a code, courts start with close interpretation of the text. Over time, however, these interpretations become

---

112 *See* Thomas W. Merrill, *Presidential Administration and the Traditions of Administrative Law*, 115 COLUM. L. REV. 1953, 1968–77 (2015) (noting the rise of "Presidential Administration" in areas of public law largely immune from judicial review).

113 *See* Guerro-Lasprilla v. Barr, 140 S.Ct. 1062, 1069 (2020); Kucana v. Holder, 558 U.S. 233, 251-53 (2010); Abbott Laboratories v. Gardner, 387 U.S. 136, 141 (1967).

114 Hayek's thought evolved significantly in this regard. In *The Constitution of Liberty*, published in 1960, he associated predictable law with the law that emerges a common law system. HAYEK, CONSTITUTION, *supra* note 23, at 215-31. By the time he wrote his final major work he conceived of law as a set of abstract principles that reflect a spontaneous order which is continuously modified by judges (and legislatures) in an evolutionary fashion. *See* 1 F.A. HAYEK, LAW, LEGISLATION, AND LIBERTY 94–123 (1973). The touchstone, however, always remained vindicating the expectations about the law held by individuals.

115 *See* H. Laurence Ross, *Settled Out of Court*, in THE SOCIAL PROCESS OF INSURANCE CLAIMS ADJUSTMENT 98-99 (2d ed. 1980).

settled, and adherence to the settled meaning of the text is a form of the rule of law. In the U.S. and other so-called common law jurisdictions, we see a strong movement toward codes over time, either restating or substituting for the common law. Given their common law background, courts in these countries readily adapt to treating precedents interpreting codes in a manner similar to the way they treat common law precedents. If anything, the conventions of stare decisis are stronger in matters of statutory interpretation than they are in common law interpretation.[116] Thus, although the common law tradition was arguably causally responsible in significant part for the development of judicial respect for the rule of law, the existence or persistence of such a tradition is not a necessary condition of achieving a system based on the rule of law.

6. No One is Above the Law

The idea that the rule of law requires that all persons be treated equally, found in Hayek's work and repeated in other commentary, is confusing. The more accurate proposition is that the rule of law ideal requires that no one is above the law. Thus, for example, if the regime is a monarchy, the king is likely to enjoy certain prerogatives that no other citizen of the realm can claim. But as long as the royal person enjoys only those prerogatives that are authorized by settled law, there is no violation of the rule of law. Thus, strict equality of treatment of persons under the law is not a necessary element of rule of law. If the law says that left-handed persons are entitled to a special government bonus, it would violate the rule of law to deny the bonus to a person who is provably left-handed. But it would not violate the rule of law to deny the bonus to a right-handed person. Obviously, one can argue that the left-handed bonus program is stupid, irrational, and perhaps that it even violates the Equal Protection Clause or one or more provisions of fundamental human rights. But it does not violate the rule of law, as long as it is enforced in accordance with settled expectations about the law.

116 The Supreme Court has repeatedly affirmed "that stare decisis has 'great weight…in the area of statutory construction' but 'is at its weakest' in constitutional cases." Thomas R. Lee, *Stare Decisis in Historical Perspective: From the Founding Era to the Rehnquist Court*, 52 VAND. L. REV. 647, 703-04 (1999). The relative strength of stare decisis in common law cases is hard to specify, since the common law is the province of 50 different state courts. My general impression is that fidelity to precedent is somewhat weaker in common-law cases than in statutory cases, perhaps because courts are both the author and the expounder of the relevant sources of authority, and therefore regard themselves as having greater liberty to revise legal rules. *See generally* KARL N. LLEWELLYN, THE COMMON LAW TRADITION: DECIDING APPEALS 62-120 (1960) (discussing the "leeways of precedent").

B. *Some Concerns*

Grumpy judges and commentators have been decrying a decline in the rule of law for as long as the concept has been around.[117] The account developed here suggests several grounds for concern going forward, which will be briefly noted.

1. The Administrative State

A particularly difficult problem is posed by the continued growth of the administrative state.[118] The administrative state was originally conceived as a way to consolidate all governmental functions in a single body, thereby eliminating the roadblocks to an expansion of government power associated with the separation of powers.[119] Specifically, it was designed to reduce the influence of the judiciary, which was seen as a reactionary force impeding progressive reforms. These objectives obviously posed a distinctive threat to the rule of law, insofar as judicial review for compliance with settled law is the keystone of the rule of law. A corrective was adopted in the form of the Administrative Procedure Act (APA) in 1946, which was interpreted as creating a presumption in favor of judicial review and instructed reviewing courts to "decide all relevant questions of law."[120] But the APA acquiesced in allowing agencies to resolve fact disputes, subject to a deferential standard of review.[121] This compromised the impartiality of adjudication associated with an independent judiciary.[122]

More recently, courts adopted the *Chevron* doctrine for addressing agency interpretations of law.[123] The *Chevron* decision itself might have been developed in a manner consistent with the rule of law.[124] But the typical formulation of the standard—requiring courts to accept reasonable agency interpretations of ambiguities in the statutes they administer—provides no obvious place for courts to consider whether an agency interpretation is consistent or inconsistent with settled expectations about the law. This has led to episodes in which the law oscillates from one administration to another,

117 TAMANAHA, *supra* note 19, at 60 ("It is an odd paradox that the unparalleled current popularity of the rule of law coincides with widespread agreement among theorists that it has degenerated in the West.").

118 *Cf.* Gary Lawson, *The Rise and Rise of the Administrative State*, 107 HARV. L. REV. 1231 (1994).

119 *See, e.g.*, JAMES M. LANDIS, THE ADMINISTRATIVE PROCESS (1938).

120 5 U.S.C. § 706.

121 *Id.* § 706(2)(E) (allowing courts to set aside the factual basis of agency decisions only if "unsupported by substantial evidence" considering the record as a whole).

122 Thomas W. Merrill, *Fair and Impartial Adjudication*, 26 GEO. MASON L. REV. 897 (2019).

123 Chevron U.S.A. Inc. v. Nat. Res. Def. Council, Inc., 467 U.S. 837 (1984).

124 *See* Thomas W. Merrill, *Re-Reading* Chevron, 70 DUKE L. J. 1153, 1178 (2021).

injecting significant instability into the law.[125] There are signs that the Court is alert to the problem, and may make appropriate adjustments in the *Chevron* doctrine at some point in the future.[126] But in its most aggressive applications, *Chevron*-style review must be regarded as a threat to rule of law values.

2. Congressional Weakness

For a variety of reasons, Congress has become a relatively weak institution. The causes range from extreme partisan division in the country, to the rise of one-party districts in which potential primary challenges from extremists discourage cross-party cooperation, to the demands of raising large campaign contributions from interest groups. The relative weakness of Congress means that power has flowed to the executive and the judiciary, in both cases with adverse implications for the rule of law.

On the executive side, recent presidents have turned to issuing Executive Orders to effect significant goals of the administration, directing subordinate officials to pursue one or another preferred policies.[127] These orders are inherently impermanent. They do not bind anyone outside the administration, and many are designed simply to reverse the executive orders of the previous administration.[128] The result is a lot of soft law that fluctuates from one administration to the next, leaving little that can form the basis of reliance on the future conduct of the government.

On the judicial side, the perception that Congress is largely incapable of overriding judicial decisions, the lapse of any serious concern with curtailing

---

125 Some examples: (1) the regular flip-flopping between Republican and Democratic Administrations as to whether family planning clinics can provide the names of abortion providers to pregnant women; (2) the expansion and contraction in the scope of wetlands subject to federal permitting requirements as part of the "waters of the United States"; (3) rejection, adoption, rejection, and adoption of the so-called "net neutrality" requirement for internet service providers, depending on the party affiliation of the Chair of the Federal Communications Commission; and (4) the oscillation between skepticism and conviction about the need for urgent action to reduce to the risk of climate change associated with the accumulation in the atmosphere of greenhouse gases. The examples are discussed in THOMAS W. MERRILL, THE CHEVRON DOCTRINE: ITS RISE AND FALL, AND THE FUTURE OF THE ADMINISTRATIVE STATE 171-72, 173-75, 209-14, 317 n.28 (2022).

126 In *West Virginia v. EPA*, 142 S.Ct. 2587, 2608 (2022), the Court carved out an apparent exception from *Chevron* for agency decisions that expand agency authority in an "unprecedented" fashion based on "cryptic" statutory authority. *Cf.* Kisor v. Wilkie, 139 S. Ct. 2400, 2400 (2019) (recognizing an exception to the principle of deference to agency interpretations of their own regulations when the interpretation upsets reliance interests created by previous agency action).

127 For an influential defense of Presidential direction of policy, see Elena Kagan, *Presidential Administration*, 114 HARV. L. REV. 2245 (2001). For the turn to executive orders by the Obama Administration, see Binyamin Appelbaum & Michael D. Shear, *Once Skeptical of Executive Power, Obama Has Come to Embrace It*, N.Y. TIMES, (Aug. 13, 2016), https://www.nytimes.com/2016/08/14/us/politics/obama-era-legacy-regulation.html.

128 *See* KENNETH R. MAYER, WITH THE STROKE OF A PEN: EXECUTIVE ORDERS AND PRESIDENTIAL POWER (2001).

judicial jurisdiction, and the remote prospect of other forms of retaliation like Court-packing, encourages courts to become more aggressive about asserting their own policy preferences. At the Supreme Court level, this occasionally takes the form of a dramatic overruling like the *Dobbs* decision.[129] More commonly, it takes the form of what I have called "Scrabble Board precedentialism," in which the Justices aggressively manipulate quotations from prior decisions to justify changes in the law.[130] The result is opinions that, in form, appear to following existing precedent, but more accurately considered represent a change in the law. At the district court level, it takes the form of "nationwide" or (more accurately, universal) injunctions against executive policy initiatives.[131] Partisan litigators have little trouble identifying conservative judges who will enjoin Democratic initiatives, or liberal judges who will enjoin Republican ones. This embroils the judiciary in partisan disputes that must be resolved on stay applications, first to the circuit courts of appeals and then to the Supreme Court. These developments reflect a disturbing erosion of judicial commitment to enforcing settled law.

Ironically, the weakness of Congress—the branch of government designed to introduce changes in the law—has in fact worked to undermine the rule of law. This is because it has encouraged the executive to rule by Executive Order and other relatively impermanent types of authority that are inherently unstable. And it has tempted courts to enter the partisan fray, either through manipulation of precedent or by issuing injunctions that determine national policy, at least until they can be overturned.

3. Too Much Law

A different sort of concern is that too much law has accumulated over time—both through the periodic enactment of blockbuster statutes by Congress (when it can find the will to legislate) and of course through the process by which the administrative agencies continue to pump out regulations at an extraordinary rate. The result, as Judge Stephen Williams argued, is that predictability about the law becomes almost impossible for all but the largest corporations that can afford the services of big law firms.[132] For ordinary individuals and small businesses, the requirements of the law that affects them becomes increasingly incomprehensible. As a result, the space in which they have the freedom to exercise initiative becomes uncertain at best.

---

129 Dobbs v. Jackson Women's Health Org., 142 S. Ct. 2228 (2022), overruling Roe v. Wade, 410 U.S. 113 (1973).

130 *See* Merrill, *Legitimate Interpretation*, *supra* note 93, at 1450–56.

131 *Id.* at 1456–61.

132 *See generally* Stephen Williams, *The More Law, The Less Rule of Law*, 2 GREEN BAG 2D 403 (1999).

This may be the most serious threat to the rule of law, because there is no obvious solution. In a sense, it represents a kind of prisoners' dilemma, in which every Congress and every agency has good reasons to enact more laws, but no one has an incentive to consider the cumulative effect on society of incrementally adding to the great weight of law. The implications are dire insofar as the mounting pile of law points toward an increasingly oligopolistic structure of society—compliance with law creating a de facto barrier to entry. One wonders if the response of many individuals and small businesses may be to give up trying to comprehend the requirements of the law altogether—pushing us unto a society where law exists for oligopolistic firms, and everyone else follows the norms that prevail in their peer group, without regard to whether they conform to the law.

CONCLUSION

The essential meaning of the rule of law is that the government forbears from coercing individuals except when it is predictable that they will do so. Predictability about coercion opens up space for individuals to pursue their own aspirations—that is to say, it promotes freedom. This kind of predictability requires that courts find it in their interest to adhere to settled principles of law. If courts render predicable decisions, the executive will generally comply with judicial judgments, which means that the system will largely conform to the ideal of the rule of law. The rule of law is not the only desirable value that a system of government should seek to realize. But it is important that legal predictability be recognized as a good in and of itself, distinct from other important values.

Whether the rule of law can be preserved in the face of a burgeoning administrative state, a weakened legislature, and a system that is suffused with so much law that predictability becomes increasingly difficult, are extremely serious concerns. But a starting point in addressing these concerns is to recognize the essential meaning of rule of law.

STAKEHOLDERS, SHAREHOLDERS, AND PURPOSE OF THE CORPORATION

Sanjai Bhagat
University of Colorado

Glenn Hubbard
Columbia University and the National Bureau of Economic Research

INTRODUCTION

The modern corporation has been an enormously productive societal and organizational invention. Large corporate organizations have figured prominently in the economic growth in the industrial world of the past century. Core ideas of the corporate construct, including respect for private property and pursuit of profits within the framework of the rule of law, have lifted billions of individuals from poverty toward middle-income levels in recent decades in developing countries.

In this paper, we analyze the current debate among policymakers, corporate leaders, institutional investors, and social activists on the purpose of the modern corporation. We conclude that the modern corporation should maximize shareholder value, while conforming to the law of the land. To strengthen the prospects for success of long-term shareholder value maximization, we suggest steps to align shareholder wealth maximization with stakeholder interests: First, antitrust public policies should be vigorously enforced to maintain and enhance competition in product markets and labor markets. Second, management and board compensation should focus on creating and sustaining long-term shareholder value. Third, the Business Roundtable and other organizations should reconsider their efforts of applying direct and indirect pressure on corporations to focus on non-shareholder priorities. Because public corporations are more susceptible to such pressure, it could be an impetus to companies to go private or not go public. Recent evidence suggests fewer public companies leads to more concentrated product markets, with the increased likelihood of diminished competition in these product markets. Finally, for many of society's more serious problems, corporations do not represent the appropriate level of action. Climate change, for example, poses significant challenges for societies and businesses. But significant changes to combat climate change require public policy changes in the United States *and* abroad. Turning more to corporations because the political process seems broken won't do.

The remainder of the paper is organized as follows. The next section discusses the historical and contemporary debate on the corporate purpose. Also, it discusses conditions under which shareholder wealth maximization

and stakeholder interests lead to identical and different corporate policies. Next, we examine arguments related to managerial myopia and the need for stakeholder primacy. In section four, we discuss the relation between ESG and corporate performance, and between ESG and mutual fund performance.[1] We find no consistent evidence that ESG ratings are positively related to long-term shareholder value. Furthermore, funds investing in companies that publicly embrace ESG sacrifice financial returns without gaining much, if anything, in terms of actually furthering ESG interests. We highlight the role of the corporate board in section six, and some limits to value-maximizing corporate actions in section seven. The final section includes our conclusions.

1. PURPOSE OF THE CORPORATION

1.1 Debate on the corporate purpose: The past half-century

Just over 50 years ago, Milton Friedman, later to be a Nobel laureate in economics, famously argued that corporations should focus solely on shareholder value maximization, while following the law of the land.[2] That view was controversial then. After years of dominance as a governance idea, shareholder primacy is once again controversial. We argue that it is more or less still correct.[3]

Of course, fifty years is a long time in a changing economy. Calls for a reexamination of the purpose and goals of the modern business corporation reflect many cross-currents in economic and political thinking. These cross-currents include underlying questions about the value of managerial capitalism itself, the role of business in an era in which government action is difficult, calls for corporate 'purpose' by leading institutional investors, and concerns that many gains to shareholders come at the expense of other 'stakeholders' of the corporation. These debates are not simply part of

---

[1] ESG is the acronym for environment/social/governance policies. The literature uses CSR and ESG interchangeably. Ionnis Ioannou & George Serafeim, *The Consequences of Mandatory Corporate Sustainability Reporting*, 1, 2 n.1 (2017), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2657712 ("The terms 'sustainability', 'environmental, social and governance' (ESG), 'non-financial' or 'corporate social responsibility' (CSR) reporting are all been used interchangeably, to describe reports with different degrees of focus on environmental, social or corporate governance issues."; Florian Berg et al., *Aggregate Confusion: The Divergence of ESG Ratings*, REV. FIN. 8 (forthcoming 2022), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3438533 ("ESG ratings are also referred to as sustainability ratings or corporate social responsibility ratings. We use the terms ESG ratings and sustainability ratings interchangeably.").

[2] Milton Freedman, A Friedman doctrine-*The Social Responsibility of Business Is to Increase Its Profits*, N.Y. TIMES (Sept. 13, 1970), https://www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html.

[3] Sanjai Bhagat & Glenn Hubbard, *Rule of Law and Purpose of the Corporation*, 30 CORP. GOVERNANCE: INT'L REV. 10, 22 (2021).

popular discourse, but reflect arguments from the core of managerial capitalism, including the signing on by 181 chief executive officers of major U.S. corporations to a *Statement of the Purpose of the Corporation* from the Business Roundtable and letters to the U.S. corporations' managements and boards from Laurence Fink, the chief executive officer of BlackRock, a major institutional investor. Analysis from legal and economic sources[4] adds fuel to the stakeholder fire.[5]

The Business Roundtable's list of tasks for corporate managers and boards of directors is an expansive one. After rightly emphasizing the central role of business corporations in producing goods and services and generating wealth, the Roundtable's statement commits the board to customers, employees, suppliers, communities, and finally, shareholders.[6] The Business Roundtable statement of 2019 garnered attention in large part because of its departure from its statement in 1997 that: "the principal objective of a business enterprise is to generate economic returns to its owners."[7] Hence, many commentators have argued that the Roundtable executives are walking back from an earlier principle of shareholder primacy.[8]

On the investor side, the letters from BlackRock's Mr. Fink in 2018 observed:

> Without a sense of purpose, no company, either public or private, can achieve its full potential. It will ultimately lose the license to operate from key stakeholders… Companies must ask themselves: What role do we play in the community? How are we

---

4 Symposium, Corporate Governance "Counter-Narratives": On Corporate Purpose and Shareholder Value 31 J. APP. CORP. FIN. 9 (2019); ALEX EDMANS, GROW THE PIE: HOW GREAT COMPANIES DELIVER BOTH PURPOSE AND PROFIT (2020); COLIN MAYER, PROSPERITY: BETTER BUSINESS MAKES THE GREATER GOOD (2018).

5 The popular media is increasingly focused on how corporations and investors should integrate ESG in their decision-making. *See* The Wall Street Journal Editorial Board, *The ESG Investing Backlash Arrives*, WALL ST. J. (August 15, 2022, 6:47 PM E.T.), https://www.wsj.com/articles/the-esg-backlash-arrives-blackrock-mark-brnovich-strive-asset-management-attorneys-general-11660600459?mod=Searchresults_pos1&page=1; Andy Kessler, *The Many Reasons ESG Is a Loser*, WALL ST. J. (July 10, 2022, 11:57 AM E.T.), https://www.wsj.com/articles/esg-loser-funds-costs-basis-points-blackrock-500-environment-green-sec-11657461127?mod=Searchresults_pos1&page=2; Matt Wirz, *Russia War in Ukraine Exposes Weakness in ESG Fund*, WALL ST. J. (April 1, 2022, 7:00 AM. E.T.), https://www.wsj.com/articles/russia-war-exposes-weakness-in-esg-fund-11648765019?mod=Searchresults_pos16&page=2.

6 *See Business Roundtable Redefines the Purpose of a Corporation to Promote 'An Economy That Serves All Americans,'* BUSINESS ROUNDTABLE (Aug. 2019), https://www.businessroundtable.org/business-roundtable-redefines-the-purpose-of-a-corporation-to-promote-an-economy-that-serves-all-americans.

7 *See Statement on Corporate Governance*, Business Roundtable (Sept. 1997), http://www.ralphgomory.com/wp-content/uploads/2018/05/Business-Roundtable-1997.pdf.

8 However, in 2022, Mr. Fink observes, "[Stakeholder capitalism] is not 'woke'… Make no mistake, the fair pursuit of profit is still what animates markets; and long-term profitability is the measure by which markets will ultimately determine your company's success." *Larry Fink's 2022 Letter to CEOs: The Power of Capitalism*, BLACKROCK (2022), https://www.blackrock.com/corporate/investor-relations/larry-fink-ceo-letter.

> managing our impact on the environment? Are we working to create a diverse workforce? Are we adapting to technological change? Are we providing the retraining and opportunities that our employees and our business will need to adjust to an increasingly automated world?[9]

Politicians have also weighed in. U.S. Senator and 2020 Democratic presidential candidate Elizabeth Warren has proposed an Accountable Capitalism Act, which would enshrine a role for workers on U.S. public corporate boards of directors and a legal requirement for directors to consider the interests of all relevant stakeholders when making decisions.[10] Given the interest in the topic we noted above, Senator Warren's proposal is not likely to be the last salvo from elected officials or political candidates.

Indeed, this political concern is not one-sided. Republican Senator Marco Rubio, himself a candidate for the presidency in 2016, explicitly moved away from "shareholder primacy theory" in his study *American Investment in the 21st Century: Project for Strong Labor Markets and National Development*.[11] Even still, Senator Rubio argued that American capitalism has produced more prosperity for more people than any economic system in the history of the world.[12]

What is the 'purpose' of the corporation? Whose interests are corporate leaders — managements and boards — supposed to advance and maximize? How do corporate leaders trade-off interests of stakeholders — shareholders, employees, customers, suppliers, and broader — business and society — in making business decisions? Is the corporation the right level of collective action for all of the concerns outlined in some groups' statements of corporate 'purpose'?

Thinking about these questions has evolved over time with the ownership and governance of business corporations. Earlier in time, large corporations that grew from entrepreneurial or owner-managed enterprises generally maximized the value of the owner's equity. Given the long-term interest of such owners, that maximization incorporated expanding the firm's going-forward productivity, profitability, and equity value to the greatest extent possible. That focus still required attention to constraints — willing participation by other stakeholders (employees, customers, and suppliers) in the firm's activities and limits imposed by government through taxation, regulation, and rules governing business practices. The corporate owner could use his or her funds to advance private or social purposes as he or she saw fit.

The subsequent rise of managerial capitalism in the United States and other industrial economies featured manager leaders of corporations with

---

9 *See id.* and accompanying text.

10 Accountable Capitalism Act, S. 3348, 115th Cong. (2018).

11 Marco Rubio, *AMERICAN INVESTMENT IN THE 21ST CENTURY: PROJECT FOR STRONG LABOR MARKETS AND NATIONAL DEVELOPMENT* 22 (May 15, 2019), https://www.rubio.senate.gov/public/_cache/files/9f25139a-6039-465a-9cf1-feb5567aebb7/4526E9620A9A7DB74267ABEA5881022F.5.15.2019.-final-project-report-american-investment.pdf.

12 *Id*. at 7.

many diverse shareholders. Such a situation presents in economic terms a 'principal-agent problem,' in which managers as *agents* run the corporation on behalf of shareholder *principals* represented by a board of directors. Without proper monitoring and rules of the road, agents may pursue other objectives than the long-term productivity, profitability, and equity value of the enterprise. Such concerns rose to prominence among both economists and business leaders in the 1960s and 1970s.[13]

It was against this backdrop that, seeking to re-center the debate over the corporation's objectives given the structure of corporate governance, Milton Friedman, fired a broadside not from an abstruse academic publication but from the *Sunday Magazine* of the *New York Times* in 1970.[14] Evaluating the objective of the corporation is straightforward: Essentially, the business of business is business.[15] Managers and boards owe a duty to *shareholders*, full stop, to maximize the equity value of the firm.[16] Broader corporate responsibility to achieve social objectives was not an issue in Friedman's view, shareholders could use the profits of the corporation for social purposes if they wished, less wastefully than if management pursued such activities with, perhaps, more self-interest on its part.[17] For example, Bill Gates, having founded a very successful company, Microsoft, now uses his share of the profits from this company to engage in significant philanthropic activities across the globe.[18]

Of course, the same external constraints described earlier still apply. Other stakeholders — workers, customers, suppliers, and the broader business world and society — must be treated equitably to ensure their participation; this is illustrated in Figure 1.[19] And government policy still constrains the corporation's use of assets and business practices.[20]

---

13 *E.g.*, Michael Jenson & William Meckling, *Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure*, 3 J. FIN. ECON. 306 (1976).

14 *See* Friedman, *supra* note 2, at 17.

15 *See id.*

16 *See id.*

17 *See id.*

18 The Bill and Melinda Gates Foundation, https://www.gatesfoundation.org/.

19 *See supra* note 8 and accompanying text.

20 *See id.*

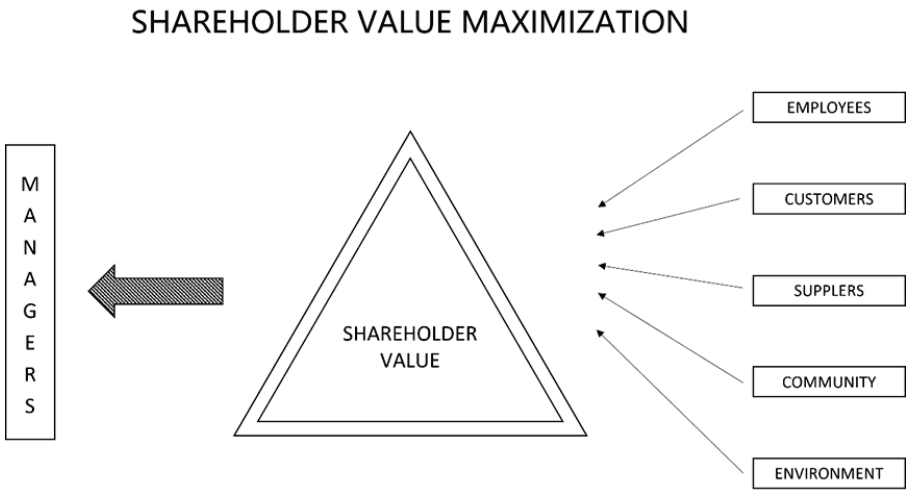


Figure 1: Shareholder Primacy: Managers consider the priorities of employees, customers, suppliers, community, and environment via the prism of long-term shareholder value.

Friedman's succinct analysis was reflected in economists' work on agency costs in the business corporation and in many business and legal practitioner views that the duties of loyalty and care of a board of directors are to shareholders.[21] Despite this beachhead, Friedman's delineation of corporate objectives has come under criticism, both fair and unfair, from an economic perspective.[22]

In one sense, Friedman insisted that the corporation should focus on doing just one thing.[23] But the economist in us see a more nuanced picture. There are complex trade-offs that must be considered. Sometimes that one thing is composed of several parts — a bundle of offerings, if you will — if an organization can accomplish them better together. For example, what if a corporation has one goal, but that goal is multi-faceted? What if shareholders want the corporation to maximize profit, but they also care about corporate social responsibility? For example, corporate sponsorship of non-profit and charitable activities in neighborhoods populated by the company's employees. Arguably, the corporation should embrace that corporate social responsibility (CSR) if by doing so it can obtain a more favorable trade-off between profits and social good than shareholders can achieve on their own shareholders given their preferences. But that is a minor change to Friedman's basic point.

21 *Id.*

22 Margaret M. Blair & Lynn A. Stout, A Team Production Theory of Corporate Law, 85 VA. L. Rev. 247, 321 (1999).

23 *See* Friedman, *supra* note 2.

Less fairly, Friedman's shareholder focus has been taken to mean a focus on 'short-term' value alone.[24] One can imagine actions a firm may take to bolster its short-term value at the expense of the long-term value or viability of the enterprise. We do not believe this reading of Friedman's intent is fair. If one interprets Friedman as emphasizing the maximization of the long-term value of the corporation for its shareholders, that emphasis grants boards of directors the leeway to evaluate management on its stewardship and growth of the firm's value over time, with attendant trade-offs. While excessive managerial 'short-termism' in corporations is debatable, nothing in Friedman's dictum precludes a focus on the long term for shareholder value maximization.

Also unfair is a critique that a focus on long-term shareholder value maximization implies a tolerance for poorly treating other stakeholders or, worse, expropriating value from them.[25] Even putting aside any legal or regulatory constraints on actions by corporations' boards of directors, managers and directors acting in shareholders' interest will want to preserve valuable relationships with stakeholders to the extent that the corporation's value can internalize the value of those relationships. This observation can help square the circle of Friedman and the Business Roundtable's parries from an economic perspective. Notwithstanding the arguments against managerial myopia, public policy-makers and large institutional investors are giving serious consideration to this concern which we discuss below.[26]

But therein also can lie the rub: What if the value of these relationships cannot be fully internalized by the firm? For example, what if an employer makes major investments in training its employees only to see them leave to accept a job at a competing firm? What if a corporation's investments in its community are met with free riding by other corporations? These are not unsolvable problems, as we discuss in more detail later, with remedies

---

24 *See, e.g.*, Steve Denning, *The Dumbest Idea In The World: Maximizing Shareholder Value*, FORBES (Nov. 28, 2011), https://www.forbes.com/sites/stevedenning/2011/11/28/maximizing-shareholder-value-the-dumbest-idea-in-the-world/?sh=78c2f6012287; *A Free Market Manifesto That Changed the World, Reconsidered*, N.Y. TIMES (Sept. 11, 2020), https://www.nytimes.com/2020/09/11/business/dealbook/milton-friedman-doctrine-social-responsibility-of-business.html.

25 *See* Lucian A. Bebchuk & Roberto Tallarita, *The Illusory Promise of Stakeholder Governance*, 106 CORNELL L. REV. 91, 176–77 (asserting that "stakeholderism" could lead to substantial societal costs by reducing accountability of corporate leadership, increasing insulation from shareholder oversight, and impeding or delaying other mechanisms that could deliver meaningful reform to stakeholders).

26 In March 2022, the United States Securities and Exchange Commission requested public comments on its proposal for major regulations regarding climate related disclosures by publicly-listed firms. The Enhancement and Standardization of Climate-Related Disclosures for Investors, 87 Fed. Reg. 21334 (proposed Mar. 21, 2022) (to be codified at 17 C.F.R. pts. 210, 229, 232, 239, and 249), https://www.sec.gov/rules/proposed/2022/33-11042.pdf. For different perspectives on the proposal, *compare* Jill E. Fisch et al., Comment Letter of Securities Law Scholars on the SEC's Authority to Pursue Climate-related Disclosure (June 6, 2022), https://papers.ssrn.com/sol3/papers.cfm?abstract_id =4129614 *with* Lawrence A. Cunningham et al., Comment Letter of Law and Finance Professors on the Proposal on Climate-Related Disclosures for Investors (Apr. 25, 2022), https://www.sec.gov/comments/ s7-10-22/s71022-20126528-287180.pdf.

ranging from public-private partnerships to tax policy changes, but they are not problems completely solvable by an individual shareholder-value-maximizing corporation on its own.

Calls for firms to address social concerns were at the heart of Friedman's 1970 statement of shareholder primacy.[27] Less talked about fifty years hence is the original target of Friedman's intellectual attack — calls for corporations to act in a socially responsible manner toward issues of the day.[28] At that time, the call was for business to participate in wage and price guidelines or controls. Friedman's argument was that such a shift could undermine the market system when the problem lay elsewhere (in this case, inflationary policies of the Federal Reserve).[29]

That not all trade-offs lead to internalized value does not imply that corporate boards and managers should weaken a commitment to stakeholders, communities, and broader social concerns. The market underpinnings of the modern corporation and its freedom to engage in commerce and finance are in large part social constructs. Caring about broader social concerns is not just an exercise in 'corporate social responsibility,' as meritorious as that activity may be, but a realization that a lack of such concerns can weaken social support for the economic foundation of the corporation. It is in this role that organizations like the Business Roundtable (or the Chamber of Commerce or Committee for Economic Development) can lead in coordination among business corporations.

1.2 Shareholder primacy and the stakeholder paradigm

Let's return to the basic question: What should be the objective of a corporation? The shareholder primacy viewpoint argues for maximizing the company's long-term value, while conforming with the laws of the land.[30] The corporate objective of long-term shareholder value maximization has a rich intellectual tradition in corporate finance and corporate law.

As we noted above, the Business Roundtable has embraced a new paradigm for corporate objectives – Managers consider the priorities of employees, customers, suppliers, community, environment, and shareholder value independent of each other, and each important in its own way.[31] Bebchuk and Tallarta refer to this viewpoint as *stakeholderism*.[32] We illustrate this concept in Figure 2.

---

27 Friedman, *supra* note 2.

28 *Id.*

29 *Id.*

30 *Id.*

31 *See* Lucian A. Bebchuk & Roberto Tallarita, *The Illusory Promise of Stakeholder Governance*, 106 CORNELL L. REV. 91 for an insightful and informative analysis of the debate surrounding shareholder value maximization and stakeholder capitalism. Their conclusions, and some of their analysis, are very similar in spirit and substance to ours.

32 *Id.* at 94.

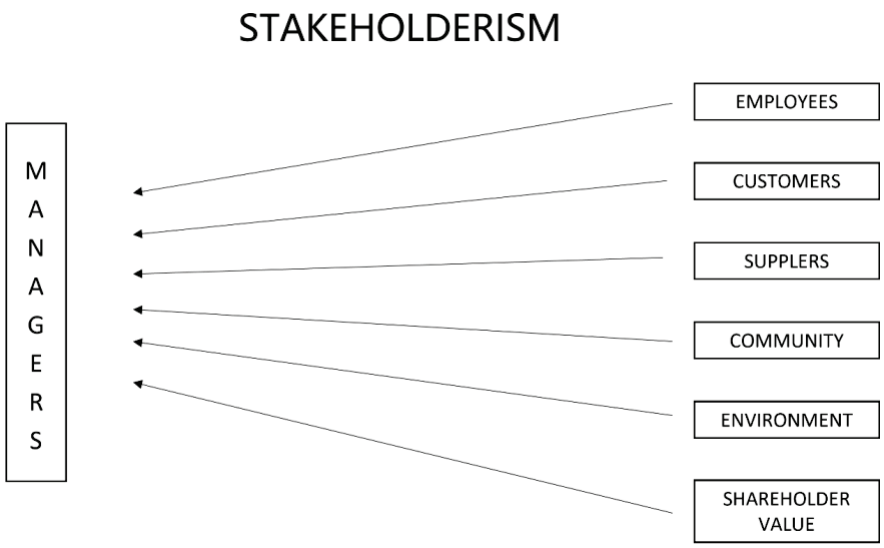


Figure 2: Stakeholderism: Managers consider the priorities of employees, customers, suppliers, community, environment, and shareholder value independent of each other, and each important in its own way.

Stakeholderism has to address both theoretical and practical challenges. Theoretically, it is quite difficult to maximize (optimize) more than a single objective. From a practical viewpoint, there are two challenges. First, stakeholderism suffers from internal inconsistencies. Recently, GM closed several auto plants manufacturing gasoline/diesel vehicles in Michigan and Ohio laying off thousands of workers.[33] These plants were later retrofitted to manufacture electric vehicles, but employed many fewer workers.[34] How should GM managers consider the trade-off between jobs and environment-friendly vehicles, if not for the underlying discipline of long-term shareholder value?[35]

The second practical challenge of stakeholderism is the operationalization and measurement of commitment to customers, employees, suppliers, communities, and to long-term shareholder value.[36]

---

33 Chris Isidore & Vanessa Yurkevich, GM's Mexican factories – Why ending the UAW strike will be so hard, CNN BUSINESS, Oct. 9, 2019, https://www.cnn.com/2019/10/09/business/uaw-gm-plant/index.html.

34 Paul A. Eisenstein, GM to invest $2.2B in first all-electric vehicle plant, create 2,200 jobs, NBC NEWS (Jan. 27, 2020), https://www.nbcnews.com/business/autos/gm-invest-2-2b-first-all-electric-vehicle-plant-create-n1124086.

35 See, e.g., Keith Naughton & Edward Ludlow, Ford Plans Up to 8,000 Job Cuts to Help Fund EV Investment, Bloomberg (July 20, 2022), https://www.bloomberg.com/news/articles/2022-07-20/ford-plans-up-to-8-000-job-cuts-to-help-fund-investments-in-evs (explaining how Ford is confronting a similar problem).

36 Bebchuk, s*upra* note 31, at 127.

This measurement challenge raises two separate, though related, problems. The first is the problem for shareholders through the board of directors of a company in providing a clear directive to corporate managers, and then holding them accountable to it. Under stakeholderism, almost any management expenditure of corporate resources, short of outright fraud, can be justified as consistent with addressing the priorities of *some* stakeholder group. Hence, the lack of managerial accountability becomes a problem. Second, this measurement challenge causes confusion in the public policy debate among investors, policymakers, and scholars regarding the performance of a corporation in the stakeholder governance paradigm. In the next section, we detail the theoretical and empirical problems in defining and measuring governance indices attempting to measure stakeholder value.

Stakeholderism assumes that shareholder value maximization will have an adverse impact on the other stakeholders. For example, managers will not compensate their employees fairly, provide them substandard set of benefits, and/or encourage or tolerate a hostile workplace for employees with certain demographic characteristics. However, this version of stakeholderism does not incorporate the argument that employees who are unfairly treated have avenues of redress in markets or through non-market means. In a competitive labor market, the employee can resign and seek alternative employment. If the labor market for the particular employee is not very competitive, or if the search and moving costs are very high, the employee can seek redress via litigation or regulatory relief. Managers, aware of the employees' redress options, are less likely or unlikely to treat the employees unfairly. Hence, managers with a focus on shareholder wealth maximization have a strong incentive to treat their employees fairly. Stakeholderism assumes that shareholder value maximization will have an adverse impact on the other stakeholders. For example, managers will not compensate their employees fairly, provide them substandard set of benefits, and/or encourage or tolerate a hostile workplace for employees with certain demographic characteristics. However, this version of stakeholderism does not incorporate the argument that employees who are unfairly treated have avenues of redress in markets or through non-market means. In a competitive labor market, the employee can resign and seek alternative employment. If the labor market for the particular employee is not very competitive, or if the search and moving costs are very high, the employee can seek redress via litigation or regulatory relief. Managers, aware of the employees' redress options, are less likely or unlikely to treat the employees unfairly. Hence, managers with a focus on shareholder wealth maximization have a strong incentive to treat their employees fairly.

Figure 3 illustrates the shareholder primacy versus stakeholder paradigm. Everything above the horizontal (x-axis) line has positive shareholder value. Everything to the right of the vertical (y-axis) line has positive stakeholder value. In the upper-right quadrant and the lower-left quadrant (highlighted in green), there is no disagreement between the

shareholder primacy and stakeholderism viewpoints. In the upper-left quadrant and the lower-right quadrant (highlighted in orange), disagreement emerges between the shareholder primacy and stakeholderism viewpoints.

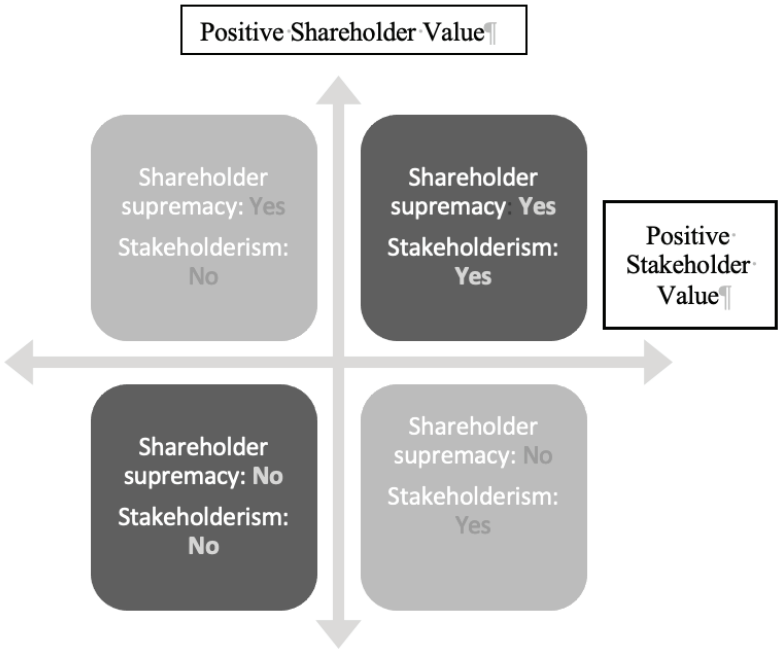


Figure 3: Corporate Actions: Shareholder primacy versus stakeholderism.

If most corporate projects fall in the two green quadrants, the traditional net present value rule for evaluating projects would be fine. Consider the upper-right quadrant Figure (highlighted in green). This quadrant represents operational and human resource investments that many U.S. corporations made during spring 2020 in response to the Covid-19 pandemic. For example, sixty-nine percent of the one-hundred largest U.S. corporations made drastic changes to their employee work schedules to safeguard employee health.[37] Sixty-four percent of the one-hundred largest companies made significant changes to accommodate customer concerns about health safety and logistical convenience.[38] Sixty-two percent increased the size and scope of their contributions to their communities.[39] These activities are clearly consistent with the stakeholder paradigm. Given the voluntary nature

[37] *See Surveying the American Public on Corporate America's Response to Covid-19*, JUST CAPITAL, https://justcapital.com/reports/surveying-the-american-public-on-corporate-americas-response-to-covid-19/ (last visited Oct. 23, 2022).

[38] *Id.*

[39] *Id.*

of these activities, these managers must consider them consistent with the interest of their long-term shareholders. Relatedly, Gartenberg, Prat, and Serafeim find that firms with better-motivated employees, who maintain strong beliefs in the meaning of their work, experience better performance.[40]

The shareholder primacy versus stakeholderism debate is about the projects that fall in the upper-left or lower-right (orange) quadrants. There are three systematic reasons for projects to fall in the upper-left or lower-right (orange) quadrants: lack of competitive product or labor markets, differential time horizons, and differential risk exposures. Competitive labor and product markets would locate most corporate projects on the upper-right or lower-left (green) quadrants. Differential time horizons and risk exposures could locate projects in the upper-left or lower-right (orange) quadrants. We explain these scenarios below.

Lack of competitive product or labor markets. If labor markets are competitive, companies have to compensate their employees fairly at market levels and treat them well, else the employee will seek alternative employment. Companies with a reputation for treating their employees well (poorly) will find it easier (more difficult) to attract and retain higher caliber employees. Similarly, in competitive product markets, companies will lose customers to competition (and, ultimately face bankruptcy), if they are not able to provide those customers with attractive or at least, satisfactory, products or services in the price-quality space. It is important to note that Friedman had in mind competitive labor and product markets when recommending shareholder value maximization as the corporate objective.

Differential time horizons. In principle, shareholders, or managers acting on behalf of shareholders, may place greater emphasis on short-term corporate performance. Stakeholders, especially those concerned about the environment, may be more focused on the long-term impact of a company's actions on the environment. The long-term negative impact on the environment could lead to negative impact on the company's long-term share price via costs of litigation and adverse reputation effects. But, if management compensation is appropriately focused on the long-term share price, then managers will be discouraged from engaging in actions that have a negative impact on the environment to the extent those actions also impair long-term shareholder value. Hence, stakeholder concerns about the long-term negative impact of a company's actions on the environment can, at least, partially, be addressed by management incentive compensation that is focused on the long term.[41]

---

40 *See* Gartenberg, C., Prat, A., & Serafeim, G., *Corporate Purpose and Financial Performance*, 30 ORG. SCI. 1, 1–18 (2019), https://doi.org/10.1287/orsc.2018.1230.

41 Long-term value maximization could be enhanced through longer shareholder investment horizons. For example, long-term institutional investors could contribute to align shareholder wealth maximization with stakeholder interests. For more information and evidence on this topic, see Ian R. Appel et al., *Passive Investors, Not Passive Owners*, 121 J. FIN. ECON. 111 (2016).

It is possible to design a managerial incentive compensation plan that will align manager incentives with long-run firm value.[42] For example, the incentive compensation of senior corporate executives could consist only of restricted equity (that is, restricted stock and restricted stock options). The compensation is restricted in the sense that the individual cannot sell the shares or exercise the options for one to two years after their last day in office.

The incentives generated by this restricted equity compensation plan structure would be relevant for maximizing long-term shareholder value. Consider the case of certain corporate misbehavior somewhat different in form but similar in the negative long-term share-price impact from the environmental actions noted above. Specifically, consider the cases of Enron, WorldCom, Qwest, many of the big-banks circa 2007-2008, and Wells Fargo more recently. Senior executives in these companies made misleading public statements regarding the operations or earnings of their respective companies. In legal proceedings, some have alleged that these misleading statements led to a temporary rise in the share prices of these companies.[43] These executives liquidated significant amounts of their equity positions during the period while their companies' share price was temporarily inflated. If these executives' incentive compensation had consisted of only restricted stock and restricted stock options that they could not liquidate for one to two years after their last day in office, they would not have had the financial incentive to make misleading statements.

During the past decade, many U.S. companies, especially in the financial sector, have increased the vesting period for at least some components of compensation to three to five years after they are granted.[44] We applaud this move on the part of board compensation committees; however, we would suggest the vesting period extend for a period after the executive's last day in his or her leadership role. This type of managerial incentive compensation plan would be a significant improvement in focusing managerial attention on the long-term. However, even these plans may not be fully effective in discouraging corporate actions for which negative impact becomes evident after a much *longer* period, say, one or two decades – irresponsible environmental actions would fall in this category. Hence, as we argue below, there remains a possible role for public policy and regulation (though, as we also argue, such policy interventions can create their own problems).

---

42 *See* SANJAI BHAGAT, FINANCIAL CRISIS, CORPORATE GOVERNANCE, AND BANK CAPITAL, 85–90 (Cambridge University Press, 1st ed., 2017).

43 For a particularly illustrative example, review Jeffrey Skilling's indictment in the Second District of Texas. United States of America v. Jeffrey K. Skilling and Richard A. Causey, Cr. No. H-04-25 (Feb. 18, 2004). *See generally Appeals Court Restores Qwest Insider Trading Conviction* (Feb. 26, 2009), http://www.nytimes.com/2009/02/26/business/26qwest.html.

44 Steve Lovett et al., *Stock Options, Restricted Stock, Salary, or Bonus? Managing CEO Compensation to Maximize Organizational Performance*, 65 BUS. HORIZONS 115 (2022).

There are additional market mechanisms to align stakeholder interests with shareholder wealth-maximization — that is, reasons why corporate policies will be in the upper-right or lower-left quadrants in the above figure. An extensive body of studies by scholars in finance, accounting, and economics documents that financial markets do impose a significant penalty on companies that violate environmental regulations, engage in employee discrimination, or financial misconduct.[45] The penalty imposed by the financial markets tends to be several times larger than administrative or legal fines the companies pay for above misconduct. The types of corporate misconduct that have been studied include litigation involving corporations as defendants in environmental regulations, product liability, antitrust, financial disclosure violation, financial fraud, employment, and age discrimination.[46] This evidence suggests that, in some important cases, *market forces* align stakeholder interests with shareholder wealth-maximization.[47]

Differential risk exposures. To the extent shareholders hold a diversified portfolio, they do not bear the idiosyncratic risk of the company, bearing only market or systematic risk.[48] Also, the feature of limited-liability protects shareholders from severe adverse impacts on the company's value beyond their investment.[49] However, there are stakeholders more likely than others to be adversely impacted by such a destruction of the company's value — employees through decrease in compensation or job loss, customers through loss of implicit and explicit quality assurances, and society through an inability to rely on the company's assurances of acting in a socially-responsible or environmentally-friendly manner.[50] While most of these shareholders are protected from this adverse event by the limited-liability feature of common stock, managers are not.[51] Managers would be negatively impacted through job loss and/or negative impact on their career via reputation effects.[52]

---

45 *See* Bhagat Sanjai & Romano Roberta, EMPIRICAL STUDIES OF CORPORATE LAW, for HANDBOOK OF LAW AND ECONOMICS 30 (A. Mitchell Polinsky & Steven Shavell eds., 2004).

46 Hans B. Christensen et al., *Economic Analysis of Widespread Adoption of CSR and Sustainability Reporting Standards* 35-36, 1-166 (Sustainability Acct. Standards Bd. Working Paper, 2018).

47 Christian Friedrich, *Corporate Misconduct and the Impact of Market Forces, Regulatory Change, and Auditor-Provided Services* (Oct. 16, 1991) (Ph.D. dissertation, Technischen Universität Darmstadt) (on file with university).

48 John Amour & Jeffrey N. Gordon, *Systemic Harms and Shareholder Value* 36, 36-85 (Colum. L. & Econ., Working Paper No. 452, 2014).

49 Jonathan R. Macey, *The Limited Liability Company: Lessons for Corporate Law*, 73 WASH. U. L. REV. 451, 454 (1995).

50 Bonnie Glantz Fatell & Michael B Schaedle, *Management compensation and incentive programmes after bankruptcy reform*, *in* AMS. RESTRUCTURING & INSOLVENCY GUIDE 2008/2009 1,1 (Blanke Rome LLP, 2009).

51 *Id.*

52 Pat Auger et al., *How Much Does a Company's Reputation Matter in Recruiting?*, MIT SLOAN (March 19, 2013), https://sloanreview.mit.edu/article/how-much-does-a-companys-reputation-matter-in-recruiting/.

The manager and director compensation proposals, we have noted, which will result in managers and directors holding significantly more stock than they have held in the past, helps to partially address this problem by providing these managers and directors powerful incentives to not invest in projects that could lead to bankruptcy or a significant reduction in long-term shareholder value.[53] However, these compensation plans will not fully address this problem where the negative impact of corporate actions (e.g.,severe environmental damage) are large.[54] Hence, as we argue below, there remains a possible role for public policy and regulation; though, as we also argue, such policy interventions can create their own problems.

2. MANAGERIAL MYOPIA AND THE NEED FOR STAKEHOLDER PRIMACY

If managers have a short-term value orientation in their leadership of the firm, they are likely to underinvest in research and development (payoffs from which are uncertain and in the distant future), and not be concerned about the reputational costs of treating their employees unfairly or producing poor quality products (the quality of which is only known over time).[55] Given the negative impact of myopic managerial actions on customers, employees, and even shareholders, calls for stakeholder primacy as the paradigm for evaluating business decisions has become louder.[56] The Business Roundtable proposal, which we highlighted earlier, offers one such example.[57]

There is substantial empirical evidence in the finance research consistent with managers and investors focusing on long-term shareholder value; Kaplan finds that the evidence on venture capital investments, private equity investments, and corporate valuations are inconsistent with the predictions of the short-term critics.[58] Relatedly, more than eighty percent of the Initial Public Offerings ("IPOs") in the United States have negative earnings during the twelve-month period prior to the IPO.[59] Yet, at the time of the IPO, the market value of these companies is often in the hundreds of millions of dollars — indeed, for some, in the billions. Investors in these IPOs and managers of these companies have to be thinking of something other than losses in valuing these companies.[60] Investors and managers in these

---

53 Amour, *supra* note 48, at 51, 55, 58-9.

54 *See* M. Jin Lee et al., *Public-private differences in incentive structures: a laboratory experiment on work motivation and performance*, INT'L PUB. MGMT. J. 183, 197-98 (2020).

55 Steven N. Kaplan, *Are US Companies Too Short-Term Oriented? Some Thoughts*, 18 INNOVATION POL'Y & ECON. 107, 109 (2018).

56 Business Roundtable Letter on Reforming Shareholder Proposal Process from Business Roundtable President Joshua Bolten to Comptroller Thomas P. DiNapoli, (on file with author).

57 *Id.*

58 Kaplan, *supra* note 55, at 107.

59 Sanjai Bhagat et al., *IPO Valuation: The International Evidence* 18-21, U 1-51 (Dec. 2018) (unpublished manuscript) (on file with the author).

60 Jay R. Ritter, *Initial Public Offerings: Updated Statistics* 3-4 (U. FLA. ED., 2022).

companies have an expectation that the company's earnings (i.e., cash flows) will grow over time, turn positive, and keep growing. The present value of these future earnings extending a decade and longer justifies the IPO valuation of hundreds of millions of dollars, or billions of dollars in some cases.

In addition to the Business Roundtable's proposal, there are other proposals to address managerial myopia. Some have suggested formation of a stock exchange in which investor voting rights increase with their holding periods.[61] This segregation, though, will lead to several problems from an economic perspective. First, it will create a wedge between those who receive dividends and other cash flow from the company and those who have the legal authority to direct the company's strategies and resources – similar to the problem of dual voting class shares in some of the tech unicorns. Second, the corporate myopia problem is about not sufficiently weighing *future* cash flows from promising new investments; the holding-period proposal gives more weight to the *past* holding period.

Regardless of whether managers behave myopically, public and political pressure on corporations could coerce managers to focus on ESG priorities over their relevance to long-term shareholder value.[62] As a result, managers and their bosses, the board of directors, could agree that they (managers) should be focused on ESG priorities.[63] But is there a way to hold managers accountable to this objective of maximizing ESG priorities?

3. ESG, CORPORATE PERFORMANCE, AND MUTUAL FUND PERFORMANCE

3.1 ESG and Corporate Performance

One way to evaluate the economic importance of ESG is to measure the impact of ESG activities on firm value.[64] Corporate ESG actions can be voluntary or involuntary, and this distinction is important in understanding the true impact of ESG on company value. But new government mandates to pursue ESG goals are likely to prove costly to American shareholders and workers.[65]

---

61 Roger L. Martin, *What If Investors Who Held Their Shares Longer Got More Voting Power?*, HARV. BUS. REV. (April 27, 2017), https://hbr.org/2017/04/what-if-investors-who-held-their-shares-longer-got-more-voting-power.

62 *See* George Serafeim, *Social-Impact Efforts That Create Real Value*, HARV. BUS. REV. (Sept. 2020), https://hbr.org/2020/09/social-impact-efforts-that-create-real-value.

63 *See Id.*

64 Tensie Whelan et al., *Uncovering the Relationship by Aggregating Evidence from 1,000 Plus Studies Published between 2015 – 2020* 7-9 (unpublished manuscript) (on file with author).

65 Elad L. Roisman, Former Comm'r, Speech at SEC: *Can the SEC Make ESG Rules that are Sustainable?* (June 22, 2021).

Voluntary ESG actions, such as operational and human resource investments that many U.S. corporations made during spring 2020 in response to the Covid-19 pandemic, are observationally equivalent to what we would expect to find if managers were focused on maximizing long-term shareholder value.[66] For example, as we noted earlier, sixty-nine percent of the hundred largest U.S. corporations voluntarily made drastic changes to their employee work schedules to safeguard employee health, and sixty-four percent of these companies voluntarily made significant changes to accommodate customer concerns about health safety and logistical convenience.[67] These examples highlight an important point – in a competitive labor market, managers (voluntarily) have to compensate their workers fairly and treat them well if they wish to hire and retain them. Similarly, in a competitive product market, managers (voluntarily) have to be responsive to customer concerns.

Meta-analysis of financial studies indicates that the overall relationship between adopting ESG practices and firm performance is either zero (statistically insignificant) or marginally positive. [68] Moreover, the economic magnitude of any effect is rather small (in the sense of Cohen (1992), compared to, for example, when a corporation becomes the target of an acquisition.[69] Even this small effect becomes smaller when researchers include basic controls like industry and size of the company.[70]

More interesting, these studies indicate that *prior* positive financial performance does predict the adoption of ESG policies (such as, recruiting diverse executives and board members, making large charitable donations), but a firm's adoption of ESG policies does not produce similar positive *future* financial performance.[71] This difference suggests that profitable companies can afford to engage in ESG activities that do not increase shareholder value; however, just engaging in ESG activities without changes in corporate business operations do not enhance financial performance.

When ESG activity is driven not by the company, but by government mandate, however, the impact on firm value is mostly negative. For example, Grewal, et al, find negative stock market reaction to the European Union's

---

66 Steve Denning, *Why Maximizing Shareholder Value Is Finally Dying*, FORBES (Aug. 19, 2019), https://www.forbes.com/sites/stevedenning/2019/08/19/why-maximizing-shareholder-value-is-finally-dying/?sh=72b62e1e6746.

67 *The COVID-19 Corporate Response Tracker: How America's Largest Employers Are Treating Stakeholders Amid the Coronavirus Crisis*, JUST CAPITAL, https://justcapital.com/reports/the-covid-19-corporate-response-tracker-how-americas-largest-employers-are-treating-stakeholders-amid-the-coronavirus-crisis/.

68 Christensen, *supra* note 46, at 57; Luca Berchicci & Andrew A. King, *Corporate Sustainability: A Model Uncertainty Analysis of Materiality* 18 (Jan. 10, 2022) (unpublished manuscript) (on file with author).

69 Jacob Cohen, *A Power Primer*, 112 PSYCH. BULL. 155, 156-59 (1992).

70 *Id. at* 157-58.

71 *See* Luca Berchicci & Andrew A. King, *Corporate Sustainability: A Model Uncertainty Analysis of Materiality* 32-34 (2022); Markus Kitzmueller & Jay Shimshack, *Economic Perspectives on Corporate Social Responsibility*, 50 J. OF ECON. LITERATURE 51, 70-71 (2012) (discussing the development of Corporate Social Responsibility initiatives and their economic impacts).

mandates to increase nonfinancial disclosure related to ESG activities.[72] Another recent paper documents negative market response to the California senate bill that mandated board gender diversity quotas for companies headquartered in California.[73] This negative response to board diversity quotas has also been documented for Norwegian companies when Norway passed a law requiring that forty percent of board members be women.[74] In a different continent and on a different version of ESG requirements, Manchiraju and Rajgopal find that companies in India experienced a negative stock market reaction to their government passing a law that two percent of corporate income be spent on ESG initiatives.[75]

Why does it matter if pursuing ESG goals is voluntary or required? One explanation is that much voluntary ESG behavior is largely symbolic, such as the widespread practice of "greenwashing" - misleading corporate communications that aim to form overly positive beliefs among stakeholders about a company's environmental practices.[76] Furthermore, as we discussed below, these greenwashing practices are facilitated by the notoriously vague and inconsistent metrics used by ESG rating firms that make it possible for companies to pick the most favorable rating agency as their judge.[77] Involuntary mandates, by contrast, typically are more objective and harder to evade, with higher costs for shareholders.[78]

Going beyond shareholder value, how should one measure a company's ESG policy, practice, and commitment and compare them across time and with those of other companies? There are several commercial vendors that compile ESG -related data on thousands of U.S. and international companies. These vendors use this ESG-related data to report ESG ratings of companies and they claim higher-rated companies have adopted and are acting upon more responsible ESG policies. Concerns have been raised about the bias and lack of objectivity of these ratings, and the lack of transparency of these ratings providers. For example, Doyle reports that large-cap companies consistently get higher ratings than mid-cap or small-cap companies;

---

72 Jody Grewal, et al., *Market Reaction to Mandatory Nonfinancial Disclosure*, 65 MGMT. SCI. 3061, 3070-74 (2017).

73 Daniel Greene et al., *Do board gender quotas affect firm? Evidence from California Senate Bill No. 826*, 60 J. OF CORP. FIN. 1, 2 (2020).

74 Kenneth R. Ahern & Amy K. Dittmar, *The Changing of the Boards: The Impact on Firm Valuation of Mandated Female Board Representation*, 127 Q. J. OF ECON., 137, 157 (2012).

75 Hariom Manchiraju & Shivaram Rajgopal, *Does Corporate Social Responsibility (CSR) Create Shareholder Value? Evidence from the Indian Companies Act 2013*, 55 J. OF ACCT. RSCH. 1, 20-22 (2017).

76 Ricardo Torelli et al., *Greenwashing and Environmental Communication: Effects on Stakeholders' Perceptions*, 29 BUS. STRATEGY AND THE ENV'T 1, 2 (2019).

77 Additionally, serious questions have been raised as to what constitutes a socially responsible investment: https://www.wsj.com/articles/u-s-authorities-probing-deutsche-banks-dws-over-sustain abil ity -claims-11629923018?mod=djemalertNEWS; https://www.wsj.com/articles/sec-is-investigating-goldman-sachs-over-esg-funds-sources-say-11654895917?mod=Searchresults_pos18&page=2.

78 Grewal, *supra* note 73, at 8.

European companies tend to receive higher ratings than U.S. companies.[79] Additionally, ratings across industries are counter-intuitive; for example, one prominent vendor rates the utilities industry highest, and healthcare the lowest among industries.[80]

Three widely used ESG ratings are: (1) MSCI produced by KLD Research and Analytics focusing on seven areas (environment, community, human rights, employee relations, diversity, product, and corporate governance) and rating companies on 280 data points using a proprietary rating system, and produce 50 ESG indicators); (2) ASSET4 provided by Thomson Reuters (considering over 750 data points on a company and combine them into 250 key performance indicators); and (3) Sustainanalytics (including 60 to 100 indicators weighted by industry-specific considerations).[81] While these ratings providers consider an extensive set of variables, there is lack of consistency in some of their key overall ratings.[82] For example, Berg, Kölbel, and Rigobon note a negative correlation in some components of the ESG ratings in the cross-section of large U.S. companies.[83] This pattern suggests that using a particular vendor's rating, company A would be rated higher than company B; however, quite the opposite would be the case using a different vendor's rating.

To summarize, there is no consistent evidence that ESG ratings are positively related to long-term shareholder value. But, also important, there is no evidence that ESG ratings are consistently related to an emerging popular understanding of issues related to the environment, communities, human rights, employee relations, and diversity.

3.2 ESG and mutual fund performance

As of July 2022, assets under management at global exchange-traded "sustainable" funds that publicly set environmental, social, and governance (ESG) investment objectives amounted to more than $2.7 trillion; 81% were in European-based funds, and 13% in U.S. based funds.[84] How have investors in these sustainable funds fared? Not that well, it seems.

---

79 Timothy M. Doyle, *Ratings That Don't Rate: The Subjective World of ESG Ratings Agencies*, AMERICAN COUNCIL FOR CAPITAL FORMATION, 9-10 (2018), https://accfcorpgov.org/wp-content/uploads/2018/07/ACCF_RatingsESGReport.pdf.

80 James Mackintosh, *ESG Is All The Rage. Big Investors Can't Agree on Why*, WALL ST. J. (Mar. 4, 2021), https://www.wsj.com/articles/everyone-sees-esg-investing-differently-but-they-all-want-to-buy-11614866558.

81 Doyle, *supra* note 80 at 7–8.

82 Berg, *supra* note 1, at 8. *See* Monica Billio et al., *Inside the ESG Ratings:(Dis)Agreement and Performance*, 1 (Sustainable Architecture for Fin. in Eur., Working Paper No. 284, 2020).

83 Berg, *supra* note 80, at 11-13.

84 Sanjai Bhagat, An Inconvenient Truth About ESG Investing, HARV. BUS. REV. (Mar. 31, 2022), https://hbr.org/2022/03/an-inconvenient-truth-about-esg-investing.

To begin, ESG funds perform poorly in financial terms. Hartzmark and Sussman analyzed the Morningstar sustainability ratings of more than 20,000 mutual funds representing over $8 trillion of investor savings.[85] Although the highest-rated funds in terms of sustainability certainly attracted more capital than the lowest-rated funds, none of the high-sustainability funds outperformed any of the lowest-rated funds.[86]

Some commentators have suggested that the above result might be expected since it is possible that investors would be happy to sacrifice financial returns in exchange for better ESG performance. But researchers are skeptical that ESG funds deliver better ESG performance. Raghunandan and Rajgopal, for example, compared the ESG record of U.S. companies in 147 ESG fund portfolios and that of U.S. companies in 2,428 non-ESG portfolios.[87] They found that the companies in the ESG portfolios had worse compliance records for both labor and environmental rules.[88] They also found that companies added to ESG portfolios did not subsequently improve compliance with labor or environmental regulations.[89] That result is not isolated. Gibson-Brandon, et al, compare the ESG scores of companies invested in by 684 US institutional investors that signed the United Nation's Principles of Responsible Investment (PRI) and 6,481 institutional investors that did not sign the PRI during 2013-2017.[90] They do not detect any improvement in the ESG scores of companies held by PRI signatory funds subsequent to their signing.[91] Furthermore, the financial returns are lower and the risk higher for the PRI signatories.[92]

What factors could explain ESG funds' underperformance? Part of the explanation may simply be that an express focus on ESG is redundant: As we argued above, in competitive labor markets and product markets, corporate managers trying to maximize long-term shareholder value would *of their own accord* prioritize employee, customer, community, and environmental interests. On this basis, setting ESG targets may actually distort corporate decision making.

There is also evidence that companies publicly embrace ESG as a cover for poor business performance. Flugum and Souther report that when managers underperform the earnings expectations (set by analysts following

---

85 Samuel M. Hartzmark & Abigail B. Sussman, *Do Investors Value Sustainability? A Natural Experiment Examining Ranking and Fund Flows*, 1-2 (EUR. CORP. GOVERNANCE INST., Working Paper No. 565, 2018.

86 *Id.*

87 Aneesh Raghunandan & Shivaram Rajgopal, *Do ESG Funs make stakeholder-friendly investments?*, 27 REV. OF ACCT.STUD. 1, 3 (2022).

88 *Id.*

89 *Id.*

90 Rajna Gibson et al., *Do Responsible Investors Invest Responsibly?* 1,41 (Eur. Corp. Governance Inst., Finance Working Paper No. 712, 2020).

91 *Id*. at 46.

92 *Id.* at 50.

their company), they often publicly talked about their focus on ESG.[93] But when they exceeded earnings expectations, they made few, if any, public statements related to ESG.[94] Hence, sustainable fund managers who direct their investments to companies publicly embracing ESG principles may be over-investing in financially underperforming companies.

To summarize, recent empirical research suggests that funds investing in companies that publicly embrace ESG sacrifice financial returns without gaining much, if anything, in terms of actually furthering ESG interests.

4. NON-SHAREHOLDER-VALUE MAXIMIZATION, PUBLIC COMPANIES, AND COMPETITION

A concern with moving board and management focus away from long-term shareholder value maximization is the potential adverse effects of such a move on the desirability of the public corporation as an organizational form. Such a concern is not abstract. Recent evidence in Doidge, Kahle, Karolyi, and Stulz indicates that the number of public firms in the United States has been decreasing over the past three decades.[95] Public firms are likely to feel greater direct and indirect pressure to focus on non-shareholder priorities. This pressure would come from institutional investors, and media via the required public disclosures. The Business Roundtable's stakeholder initiative would give additional impetus to firms to go private in order to maximize shareholder value, or not go public.[96]

While there are fewer public firms during the past three decades, some remaining public firms are larger, leading to increased concentration within some industry sectors. Increased concentration raises concerns about diminished competition; Bonaime and Wang document the effects of concentration in the pharmaceutical industry) and a potential reduction in

---

[93] Ryan Flugum and Matthew Southern, *Stake Holder Value: A Convenient Excuse for Underperforming Managers?* 4 (2022), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3725828#:~:text=Stakeholder%20language%20appears%20to%20influence%20the%20evaluation%20of,excuse%20that%20reduces%20accountability%20for%20poor%20firm%20performance.

[94] *Id.* at 17.

[95] Doidge et al, *Eclipse of the Public Corporation or Eclipse of the Public Market?,* 30 J. APPLIED CORP. FIN. 8 (2018)

[96] Bae et al., *Does CSR Matter in a Time of Crisis? Evidence from the Covid-19 Pand*emic, 67 J. OF CORP. FIN. 1, 1 (2021). Public and political pressure on corporations could force managers to focus on ESG priorities over and beyond their relevance to long-term shareholder value. Bae, et al, (2021) study the impact of ESG policies on corporate performance during the Covid-19 pandemic-induced market crash and the post-crash recovery. Interestingly, they find that Business Roundtable companies, which committed to protecting stakeholder interests prior to the pandemic, do not outperform other companies during the pandemic crisis.

consumer welfare, and diminished competition in the labor markets.[97] This consideration is important because, as we discussed above, a greater level of competition in the product and labor markets will align the interests of shareholders with that of other stakeholders (notably, customers and employees). Conversely, a diminished level of competition in the product and labor markets will likely draw a wedge between shareholder and stakeholder interests. Hence, proponents of stakeholderism, in their good faith attempts to pressure public companies to focus on non-shareholder priorities, may undermine consumer and employee choices and welfare.

5. ROLE OF THE BOARD IN BALANCING CORPORATE OBJECTIVES

It is important to complement a discussion of the objectives of the corporation and corporate governance with an analysis of at what level of action is needed to address a particular objective. To begin, the management of many stakeholder interests, including shareholder interests, can be accomplished by a board of directors engaged in maximizing *long-term shareholder value*.

To make this point clear, consider the responsibilities of boards of directors proposed by prominent corporate lawyer Martin Lipton in his work "The New Paradigm."[98] The proposal begins with the foundation we have suggested — that the purpose of the corporation is long-term business success and long-term increase in the corporation's value. Achieving that purpose requires a meaningful business and relationships among the corporation and its stakeholders. At the next level, the firm's governing board of directors are elected by shareholders. Those directors' fiduciary duty to shareholders (or, for that matter, to other specific stakeholders) only requires that those directors act to advance the corporation's long-term business success and increase in value. In "The New Paradigm," directors may use their business judgment to allocate value to non-shareholder stakeholders if directors believe that so doing will enhance long-term corporate business success and value.[99] Also in so doing, directors should consider reputational, legal, and regulation risks that could affect the business success or value of the corporation.[100] Notably, Lipton urges directors to focus on the long term and not the short term for shareholder value maximization. We agree, as Friedman would have agreed. In particular, directors could use their business judgment to reject an offer to acquire the corporation at a premium or a

---

97 Alice Bonaime and Ye Wang, *Mergers, Product Prices, and Innovation: Evidence from the Pharmaceutical Industry*, SSRN (Aug. 12, 2022), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3445753.

98 Martin Lipton, *Corporate Governance: The New Paradigm*, HARV. L. SCH. F. ON CORP. GOVERNANCE (Oct. 3, 2022 9:51 PM).

99 *Id.*

100 *Id.*

shareholder demand that could generate a short-term gain in equity value if they believe the long-term value would be higher under an alternative strategy.[101]

Lipton's view, as with ours, does not imply a lack of challenges shaping the contours of shareholder value maximization by corporate boards. The management of trade-offs among stakeholders and a determination of long-term value maximization imply a time-consuming role by directors, with a focus on both internal and external information. Large investors, and occasionally activist investors, will play the challenger's role in answering the familiar '*Quis custodiet ipsos custodes?'* question. Both boards and managements will need to pay more attention to political and social currents challenging business even in the context of long-term shareholder value maximizations. One way to get corporate boards to focus more on the long-term is to align their compensation toward creating and sustaining long-term shareholder value.

In *The Modern Corporation and Private Property*, Adolf Berle and Gardiner Means described an evolving phenomenon, as the twentieth century progressed, of the domination of the large public corporation by professional management as the separation of ownership and control.[102] The firm's nominal owners, the shareholders, in such companies exercised virtually no control over either day to day operations or long-term policy.[103] Instead, control was vested in the professional managers who typically owned only a very small portion of the firm's shares.[104]

One consequence of this phenomenon identified by Berle and Means was the filling of board seats with individuals selected not from the shareholding ranks, but chosen instead because of some prior relationship with management.[105] Boards could be composed either of the managers themselves (the inside directors) or associates of the managers, not otherwise employed by or affiliated with the enterprise (the outside or non-management directors). This new type of outside director often had little or no shareholding interest in the enterprise. However, as the shareholders' legal fiduciaries, the outside directors were still expected to expend independent time and effort in their roles, and, consequently, it began to be recognized that they must now be compensated directly for those activities.

The consequences of this shift in the composition of the board was to exacerbate the manager-shareholder conflict of interest — the principal agent problem in economic terms — inherent in the large-scale corporate form. Without the direct economic incentive of substantial stock ownership, directors, given the substantial reputation enhancement and monetary

101 *Id.*

102 ADOLF BEARLE AND GARDINER MEANS, THE MODERN CORPORATION AND PRIVATE PROPERTY 4-8 (1933).

103 *Id.*

104 *Id.*

105 *Id.*

compensation board service came to entail, had little incentive other than their legal fiduciary duties to engage in active managerial oversight.[106] Some analysts have also observed that the compensation received for board service may have actually acted as a disincentive for active management monitoring in this period, given management control over the director appointment and retention process.[107]

But, this problem was also not unsolvable, and resolving it still gives power to the argument for long-term shareholder value maximization.[108] Since the identification of this phenomenon, both legal and finance scholars have struggled to formulate effective solutions.[109] Numerous legal reforms have been proposed, often involving such steps as the creation of the professional independent director, strengthened board fiduciary duties, or stimulation of effective institutional shareholder activism.[110] None is a bulletproof fix[111]: Shareholders, mindful of disasters at Enron, WorldCom, Qwest, the big banks circa 2008, and more recently, Wells Fargo, Equifax, General Electric, Johnson & Johnson, and Boeing are keenly aware of this problem.[112]

But further aligning incentives of the board with long-term shareholder value maximization offers a stronger route.[113] Director compensation could be structured along the lines of the executive compensation proposal we noted earlier.[114] Specifically, compensation of corporate directors could be concentrated more in restricted equity (restricted stock and restricted stock options) – restricted in the sense that the director cannot sell the shares or exercise the options for a period following their last board meeting.[115] Such a director compensation structure is further motivated by the following empirical findings of links between director compensation structure and corporate performance[116]: companies in which directors own more stock

---

106 Bhagat & Hubbard, *supra* note 3, at 22. Sanjai Bhagat, Dennis C. Carey & Charles M. Elson, *Direct Ownership, Corporate Performance, Corporate Performance, and Management Turnover*, 54 BUS. LAW. 885, 887 (1999).

107 Bhagat & Hubbard, *supra* note 3, at 22; Bhagat, Carey & Elson, *supra* note 107, at 887-88; *See* Charles M. Elson, *Direct Compensation and the Management-captured Board – The History of a Symptom and a Cure*, 50 SMU L. REV. 127, 132–133 (1997).

108 Bhagat & Hubbard, *supra* note 3, at 22.

109 *Id*.

110 Bhagat & Hubbard, *supra* note 3, at 22; Elson, *supra* note 108, at 129-30.

111 Bhagat & Hubbard, *supra* note 3, at 22.

112 *Id*.; *See* Elson, *supra* note 108, at 128.

113 Bhagat & Hubbard, *supra* note 3, at 22.

114 *Id*.

115 Bhagat & Hubbard, *supra* note 3, at 22; GLENN HUBBARD, THE WALL AND THE BRIDGE: FEAR AND OPPORTUNITY IN DISRUPTION'S WAKE 140 (2022) [hereinafter HUBBARD, THE WALL]. *See* Sanjai Bhagat & Glenn Hubbard, *Should the Modern Corporation Maximize Shareholder Value?* 10 AEI ECON. PERSPS. (Sept. 10, 2020), https://www.aei.org/research-products/report/should-the-modern-corporation-maximize-shareholder-value/ [hereinafter Bhagat, *Modern Corporation*]; Elson, *supra* note 108, at 133-34, 165.

116 Bhagat & Hubbard, *supra* note 3, at 22.

perform better in the future years;[117] directors who own more stock are more likely to discipline or fire the CEO when the stock price performance of their company has been sub-par in the previous two years;[118] and directors who own more stock are less likely to let their company engage in value-destroying acquisitions.[119]

6. LIMITS TO SHAREHOLDER-VALUE-MAXIMIZING CORPORATE ACTIONS

That long-term shareholder value maximization can balance trade-offs among other corporate stakeholders — assuming a value-maximizing focus by corporate managers and boards — does *not* imply that long-term shareholder value maximization will address *all* problems faced by the firm.[120] First, for many problems, corporations do not represent the appropriate level of action to solve the problem.[121] We mentioned earlier the free-rider problem for certain corporate investments in employee training or communities.[122] Training support could be enhanced by tax credits to offset the externality and by partnerships among companies with community colleges or universities in a geographic area.[123] Support for communities could be enhanced by corporate partnerships coordinated by local business organizations or economic development agencies.[124] Some aid for communities and particular workers in those communities is better accomplished by government though economic development funds or augmenting support for earnings from work in areas with high levels of structural unemployment.[125]

Other social problems are even more complex.[126] Climate change, for example, poses significant challenges for societies and businesses to reduce carbon in the atmosphere and greenhouse gas emissions and adapt to

---

117 *Id. See* Sanjai Bhagat & Brian Bolton, *Corporate Governance and Firm Performance: The Sequel*, 58 J. CORP. FINANCE, 142, 161 (2019).

118 Bhagat & Hubbard, *supra* note 3, at 22. *See* Sanjai Bhagat & Brian Bolton, *Corporate Governance and Firm Performance*, 14 J. CORP. FIN. 257, 270 (2008).

119 Bhagat & Hubbard, *supra* note 3, at 22; Sanjai Bhagat & Brian Bolton, *Director Ownership, Governance, and Performance*, 48 J. OF FIN. & QUANTITATIVE ANALYSIS, 105, 132-33 (2013).

120 Bhagat & Hubbard, *supra* note 3, at 22; Bhagat, *Modern Corporation*, *supra* note 116, at 10.

121 *Id.*

122 Bhagat & Hubbard, *supra* note 3, at 22; Bhagat, *Modern Corporation*, *supra* note 116, at 10; *See* HUBBARD, THE WALL *supra* note 116, at 145.

123 Bhagat & Hubbard, *supra* note 3, at 22; Bhagat, *Modern Corporation*, *supra* note 116, at 10; *See* HUBBARD, THE WALL, *supra* note 116, at 145-46 (Citing Managing the Future Work, *Unpacking Amazon's development strategy*, HARVARD BUS. SCH. (Jan. 08, 2020), https://www.hbs.edu/managing-the-future-of-work/podcast/Pages/podcast-details.aspx?episode=12651971.)

124 Bhagat & Hubbard, *supra* note 3, at 22; Bhagat, *Modern Corporation*, *supra* note 116, at 10; *See* HUBBARD, THE WALL, *supra* note 116, at 146.

125 Bhagat & Hubbard, *supra* note 3, at 22; Bhagat, *Modern Corporation*, *supra* note 116, at 10.

126 *Id.*

evolving changes in surface temperatures.[127] Investors could and should press corporations to disclose more information about the exposure of their long-term value to climate change and corporations may act to reduce emissions and increase their adaptability in service of a focus on long-term value maximization.[128] That step is an extension of a market process and response.[129] But that step alone will not resolve negative externalities from effects of business activities on climate change.[130] But significant changes to combat climate change require *public policy* changes in the United States *and* abroad — a carbon tax or alternative-energy technology subsidies, for example.[131] Turning more to corporations because the political process seems broken and makes little progress won't do.[132]

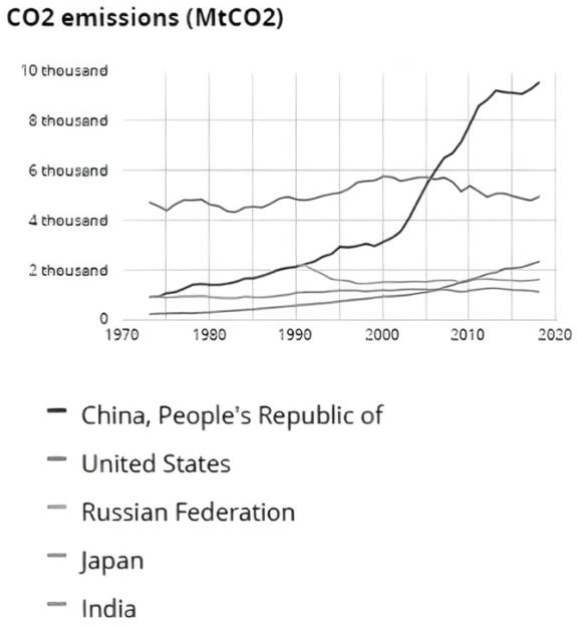


Figure 4: CO2 Emissions of Major Countries[133]

127 Bhagat & Hubbard, *supra* note 3, at 22; Bhagat, *Modern Corporation*, *supra* note 116, at 10; HUBBARD, THE WALL, *supra* note 116, at 147.

128 *Id.*

129 *Id.*

130 *Id.*

131 As Figure 4 highlights, during the past two decades – incremental carbon emissions in the United States have been *decreasing*, whereas carbon emissions in China has tripled. Bhagat, *Modern Corporation*, *supra* note 116, at 10; HUBBARD, THE WALL, *supra* note 116, at 147.

132 Bhagat & Hubbard, *supra* note 3, at 23; Bhagat, *Modern Corporation*, *supra* note 116, at 10; HUBBARD, THE WALL, *supra* note 116, at 148.

133 Data from the international agency energy atlas http://energyatlas.iea.org/.

While corporations control substantial financial resources and can significantly impact the welfare of their employees, customers, and communities they operate in, even these large and well-resourced corporations are limited in two important ways[134] First, a behavior changing/modifying tool they lack, which all state and county governments have, is the power of criminal prosecution and legal coercion.[135] Second, corporations can impact the behavior of their employees/customers/suppliers, but not of citizens who choose not to be their employees/customers/suppliers.[136] The larger social issues would need compliance from most/all citizens.[137] Only the state's criminal prosecution and coercive powers can lead to compliance by the broader citizenry.[138]

Who should decide on a corporation's stand on a particular social issue?[139] Some have suggested that shareholders could vote on the particular issue.[140] Two concerns regarding shareholder voting on public policy issues[141]: First, this assumes shareholders know what is in their best interest and would vote accordingly.[142] A long tradition of papers in corporate finance suggest neither of the above assumptions might be valid.[143] For example, acquisitions of public U.S. targets by public U.S. acquirers is a well-recognized case of acquiring shareholders voting for a transaction that, on average, has a small negative or zero value impact on themselves (acquiring shareholders).[144] Given costly information, such behavior might not be irrational.[145] The point is many shareholders do not put the effort to understand the implications of their vote, or even care to vote.[146] Second, perhaps more important, there could be a mismatch between the geographical location of the voting shareholders and impacted citizens.[147] Some of these

134 Bhagat & Hubbard, *supra* note 3, at 23; Bhagat, *Modern Corporation*, *supra* note 116, at 10.

135 *Id.*

136 *Id.*

137 *Id.*

138 *Id.*

139 Bhagat & Hubbard, *supra* note 3, at 23; Bhagat, *Modern Corporation*, *supra* note 116, at 10; HUBBARD, THE WALL, *supra* note 116, at 141.

140 *Id.*

141 Bhagat & Hubbard, *supra* note 3, at 23; Bhagat, *Modern Corporation*, *supra* note 116, at 10.

142 *Id.*

143 Bhagat & Hubbard, *supra* note 3, at 23; HUBBARD, THE WALL, *supra* note 116, at 141. *See* Sanjai Bhagat & Richard H. Jefferis, *Voting Power in the Proxy Process: The Case of Antitakeover Charter Amendments*, 30 J. Fin. Econ. 193, 222 (1991).

144 Bhagat & Hubbard, *supra* note 3, at 23; *See generally* Sara B. Moeller, Frederik P. Schlingemann & Rene M. Stulz, *Wealth Destruction on a Massive Scale? A study of Acquiring-Firm Returns in the Recent Merger Wave*, 60 J. FINANCE 757 (2005).

145 Bhagat & Hubbard, *supra* note 3, at 23.

146 *Id*. *See* Frank H. Easterbrook & Daniel R. Fischel, *Voting in Corporate Law*, 26 J. L. & ECON. 395, 396-97 (1983).

147 Bhagat & Hubbard, *supra* note 3, at 23.

shareholders could be citizens of other countries residing abroad.[148] Would we want these overseas shareholders dictating the behavior of citizens here?[149]

There are instances in which direct policy action is required to alter shareholder value maximization.[150] Examples include antitrust laws to limit exploitation of product market power, anti-monopsony rules to enhance competition for employees, and corporate tax policy to affect levels of corporate profitability, location decisions, wages paid to workers or incentives to invest. These explicit policies address social objectives that would not, in some cases, be achievable by individual firms and would not otherwise receive the same level of attention in an unconstrained long-term shareholder value maximization by the board of directors.[151] Even in such cases, it is important to note that the above public policy interventions are a complement, not a substitute for, long-term shareholder value maximization.[152]Will public policy play this needed role? While some of this concern reflects populist opposition to big business in the aftermath of policy responses to the 2007-2009 financial crisis, the concern's roots may deepen given the unprecedented and devastating economic impact of the ongoing Covid-19 pandemic. With political dysfunction, social challenges are going unmet, so why not turn to corporations and their leaders for progressive action? Simply put, broad social challenges involving externalities and spillovers require government intervention. While business leaders can individually or collectively champion such interventions, corporate actions alone will not meet the challenges.

CONCLUSION

The modern corporation has been an enormously productive societal and organizational invention. Large corporate organizations have figured prominently in the economic growth in the industrial world of the past century. Core ideas of the corporate construct, including respect for private property rights within the framework of the rule of law, has lifted billions of individuals around the globe in recent decades from poverty toward middle-income, *and* decreased income inequality.

Why? What has guided corporations to enhance growth and well-being? In recent decades, as Milton Friedman argued, the contemporary corporate objective was encapsulated in the shareholder primacy viewpoint which

148 *Id.*

149 *Id.*

150 Bhagat & Hubbard, *supra* note 3, at 23; Bhagat, *Modern Corporation*, *supra* note 116, at 10; *See* HUBBARD, THE WALL, *supra* note 116, at 146.

151 Bhagat & Hubbard *supra* note 3, at 23; Bhagat, *Modern Corporation*, *supra* note 116, at 11.

152 Bhagat & Hubbard, *supra* note 3, at 23; Bhagat, *Modern Corporation*, *supra* note 116, at 11; HUBBARD, The Wall, *supra* note 116, at 147.

argued for maximizing the company's long-term shareholder value, while conforming to applicable laws and regulations.[153] Fifty years after he articulated it, his statement remains a good guide under some broad assumptions. Recently, the Business Roundtable's *Statement of the Purpose of the Corporation* commits the board to other stakeholders as well — including customers, employees, suppliers, and communities the company operates in.[154] But importantly, if labor markets and product markets are competitive, the shareholder primacy and stakeholder paradigms would lead to identical corporate policies. When these markets are not competitive, differential time horizons and differential risk exposures of various stakeholders could lead to differing corporate policies; these are bright lines justifying of public policy interventions. But such public policy interventions are a complement to, not a substitute for, long-term shareholder value maximization.

To strengthen the prospects for success of long-term shareholder value maximization, we suggest steps to align shareholder wealth maximization with stakeholder interests: First, antitrust public policies should be vigorously enforced to maintain and enhance competition in product markets and labor markets. Second, management and board compensation should be reformed to focus on creating and sustaining long-term shareholder value. Third, the Business Roundtable and other organizations should reconsider their efforts of applying direct and indirect pressure on corporations to focus on non-shareholder priorities. Because public corporations are more susceptible to such pressure, it could be an impetus to companies to go private or not go public. Recent evidence suggests fewer public companies leads to more concentrated product markets, with the increased likelihood of diminished competition in these product markets.

Also, we discuss the relationship between ESG and corporate performance, and between ESG and mutual fund performance. We find no consistent evidence that ESG ratings are positively related to long-term shareholder value. Furthermore, funds investing in companies that publicly embrace ESG appear to sacrifice financial returns without gaining much, if anything, in terms of actually furthering ESG interests.

Clarity in objectives and governance facilitates accountability of corporate leaders and boards of directors and shapes the contours of corporate activities. Pursuing objectives too broad and diffuse complicates accountability, limits the value creation by corporate businesses, and may lead to a decline in the use of the corporate organizational form for business enterprises. Pursuing objectives too narrow can lead to conflicts among shareholders and between shareholders and other non-shareholder stakeholders. Corporate focus on long-term shareholder value maximization,

153 Friedman, *supra* note 2, at 32.

154 *Statement on the Purpose of a Corporation*, BUS. ROUNDTABLE (Aug. 2019), https://www.businessroundtable.org/business-roundtable-redefines-the-purpose-of-a-corporation-to-promote-an-economy-that-serves-all-americans.

remains the best way to enhance value and the broader corporate contribution to society.

Where externalities or non-market objectives are important, other coordinating mechanisms or public policy interventions can expand or limit the choices of actions by a value-maximizing corporation. And most externalities in current debates require external policy action. Government intervention can also be constructive in preparing workers and places for economic shifts from globalization and technological change. But these are government interventions. Political polarization does not make the corporation the logical alternative for action instead, a point Friedman actually addressed in the context of corporate responsibility in his 1970 *New York Times* article.[155] Frustration with the state of political discourse around social problems does not change this point. And altering the purpose of the corporation away from long-term shareholder value maximization risks vagueness that can disrupt the wealth-producing and job-creating power we take for granted from the modern corporate enterprise.

More than a half-century in, Friedman's shareholder primacy idea, while not the entire story, remains the right place to start.

155 Friedman, *supra* note 2.

SHOULD TRADE'S 'LOSERS' BE COMPENSATED?: AN EXPLORATION OF THE WELFARE ECONOMICS OF THE LOSSES AND COSTS OF ECONOMIC CHANGE

Donald J. Boudreaux[*]

The law evidently cannot prohibit all actions which may harm others, not only because no one can foresee all the effects of any action, but also because most changes of plans which new circumstances suggest to some are likely to be to the disadvantage of some others. The protection against disappointment of expectation which the law can give in an ever-changing society will always be only the protection of some expectations but not of all.[1]

- F.A. Hayek

INTRODUCTION

Economists as a group famously support free trade and do so overwhelmingly.[2] This support is not determined by party lines.[3] Surveys of economists' opinions on trade policy typically find that around four in five economists endorse a policy of free trade.[4] Such support is perhaps at least partly explained by the fact that modern economics was launched with the publication of a book that featured at its core a veritable disembowelment of mercantilism and all of its protectionist fallacies. That book, of course, is the

[*] Professor of Economics, George Mason University, Getchell Chair at George Mason University's Mercatus Center, and Research Fellow at the Law & Economics Center at George Mason University Antonin Scalia Law School. E-mail: dboudrea@gmu.edu.

Research for this writing was funded by a grant from the Law & Economics Center. The opinions expressed here are mine alone and do not represent the official position of the Law & Economics Center, of George Mason University, or of any other organization. I thank for their helpful comments Veronique de Rugy, Richard Epstein, Elisa Philip Gentry, Andrew Morriss, Paulo Saguato, John Yun, and participants in the Law & Economics Center's September 2021 Research Roundtable on Capitalism and the Rule of Law. Despite this immensely productive feedback, errors and weaknesses remain, which are all my fault.

[1] 1 F. A. HAYEK, *Rules and Order*, *in* LAW, LEGISLATION AND LIBERTY 101–02 (1973).

[2] *See* Gregory N. Mankiw, *Economists Actually Agree on This: The Wisdom of Free Trade*, N.Y. TIMES (Apr. 24, 2015), https://www.nytimes.com/2015/04/26/upshot/economists-actually-agree-on-this-point-the-wisdom-of-free-trade.html.

[3] *See id.*

[4] *See* Robert Whaples, *The Policy Views of American Economic Association Members: The Results of a New Survey*, 6 ECON J. WATCH 337, 340, 343 (2009).

second of only two written by Adam Smith, his 1776 *An Inquiry Into the Nature and Causes of the Wealth of Nations*.

Although nearly all of the arguments that Smith marshaled in support of free trade had been offered earlier by a variety of writers, never before were the arguments against protectionism assembled in one place, presented so systematically, and delivered with such an arresting combination of power and eloquence. Support for a policy of unilateral free trade, largely without conditions, has ever since been hard wired into the DNA of economic science.[5]

Some economists, alas, have mutant genes. Among the most scholarly of these today is Harvard's Dani Rodrik, a prominent trade skeptic. In his 2018 book, *Straight Talk On Trade*, Rodrik explains that the typical academic economist, when asked by a reporter about trade policy, "will be enthusiastic in his support of free trade."[6] But, continues Rodrik, if the reporter were to sneak into this economist's classroom, that person would encounter a quite different message:

> The professor would then launch into a long and tortured exegesis that will ultimately culminate in a heavily hedged statement: "So if the long list of conditions I have just described are satisfied, and assuming we can tax the beneficiaries to compensate the losers, freer trade has the potential to increase everyone's well-being."[7]

Rodrik is correct that when we academic economists do trade theory for fellow economists, and when we teach it to our students, we consider a broad range of theoretical possibilities that we seldom mention in public-policy discussions with non-economists. And it is true that any one of these exceptions, were it to occur regularly in the real world, would indeed render reckless all professions of unqualified support for a policy of unilateral free trade. But Rodrik is mistaken to conclude that economists who advocate enthusiastically for a policy of unilateral free trade are being untrue to their professional learning or reckless in conveying its implications for public policy. The theoretical exceptions to the finding that a policy of unilateral free trade will maximize material living standards in the home country are just that: theoretical exceptions.

The standard response of pro-free-traders to Rodrik and others who argue against a policy of unilateral free trade is to point out that, in the real world, policy is designed and applied by fallible human beings. These policymakers and enforcers typically have no reliable way of knowing when

---

[5] Adam Smith was no free-trade absolutist. In *The Wealth of Nations* he explicitly identified four exceptions to the case for a policy of unilateral free trade. Yet both the substance and tenor of his work make clear that, for Smith, the burden of persuasion is squarely on those who would impose protectionist policies. On the nature and reach of Smith's exceptions to the case for free trade see Donald J. Boudreaux, *Today's Relevance of Adam Smith's Wealth of Nations*, 24 INDEP. REV. 487, 491, 493–95 (2020).

[6] DANI RODRIK, STRAIGHT TALK ON TRADE: IDEAS FOR A SANE WORLD ECONOMY 119 (2018).

[7] *Id.*

theoretical exceptions arise in reality. Nor do these government officials have access to the detailed knowledge that is necessary to ensure that these exceptions are addressed in ways that do not make matters worse. In addition, policymakers and enforcers are subject to often-perverse political incentives.[8] And such perverse incentives, as amply documented by Douglas Irwin, have for all of American history played a prominent role in the setting of U.S. trade policy.[9]

All of the theoretical exceptions to a policy of free trade are built, if usually only implicitly, on the presumption that policymakers and enforcers possess both the knowledge and the motivation of gods. Strip away one or the other—or, especially, both—of these presumptions and the wisdom of the *rule* of a policy of unilateral free trade becomes evident. Ordinary human beings simply have neither the capacity nor the political courage to be trusted with the discretion to intervene in trade in ways that make the theoretical exceptions relevant in reality.

Although this response to free-trade skeptics is commonplace[10], repeating it is nevertheless important. The reason is that even more commonplace is the habit of ignoring these features of real-world policymaking.

There is, however, at least one particular prong of the trade-skeptics' case that warrants a response that is more foundational. This prong is the one that, as Rodrik says in the above quotation, insists that the theoretically sound case for free trade rests on the assumption that "we can tax the beneficiaries to compensate the losers."[11] Because in the real world the "losers" from free trade are not compensated with tax revenues paid by the "winners," free trade

---

[8] These often-perverse political incentives, and their consequences, are most consistently revealed and studied by scholars working in the public-choice tradition. Classic works in public choice include ANTHONY DOWNS, AN ECONOMIC THEORY OF DEMOCRACY (1st ed. 1957); JAMES M. BUCHANAN & GORDON TULLOCK, THE CALCULUS OF CONSENT (1st ed. 1962); WILLIAM H. RIKER, THE THEORY OF POLITICAL COALITIONS (1st ed. 1962); MANCUR OLSON, THE LOGIC OF COLLECTIVE ACTION: PUBLIC GOODS AND THE THEORY OF GROUPS (1st ed. 1965); WILLIAM A. NISKANEN, BUREAUCRACY & REPRESENTATIVE GOVERNMENT (1st ed. 1971); GEORGE J. STIGLER, THE CITIZEN AND THE STATE: ESSAYS ON REGULATION (1st ed. 1975); JAMES M. BUCHANAN & RICHARD E. WAGNER, DEMOCRACY IN DEFICIT: THE POLITICAL LEGACY OF LORD KEYNES (1st ed. 1977); BRYAN CAPLAN, THE MYTH OF THE RATIONAL VOTER: WHY DEMOCRACIES CHOOSE BAD POLICIES (1st ed. 2007). An accessible introduction is GORDON TULLOCK & GORDON L. BRADY, GOVERNMENT: WHOSE OBEDIENT SERVANT? A PRIMER IN PUBLIC CHOICE (2000). On the relevance of public choice to the study of law, see MAXWELL L. STEARNS, TODD J. ZYWICKI & THOMAS J. MICELI, LAW AND ECONOMICS: PRIVATE AND PUBLIC (1st ed. 2018).

[9] *See generally* DOUGLAS A. IRWIN, CLASHING OVER COMMERCE: A HISTORY OF US TRADE POLICY (1st ed. 2017).

[10] *See, e.g.*, Boudreaux, *supra* note 5, at 496–97 (quoting ADAM SMITH, AN INQUIRY INTO THE NATURE AND CAUSES OF THE WEALTH OF NATIONS 456 (Liberty Fund 1981) (1776).

[11] RODRIK, *supra* note 6.

in reality is missing a key pillar on which rests the theoretical case for free trade. Or such is the argument of Rodrik and many other trade skeptics.[12]

My purpose in this paper is to explain that even the theoretical case for free trade does not, in practice or even in principle, require that trade's "winners" compensate trade's "losers." Nor does it require even that trade's "winners" be *capable* of compensating "losers" such that at least one person is left better off with no one made worse off.

On this particular point, economists themselves – including many unapologetic pro-free-trade economists – have unwittingly contributed to the confusion. By misapplying to trade the Pareto and Kaldor-Hicks welfare criteria, economists have made the case for a policy of free trade appear more tentative than it really is.

I. THE WELFARE CRITERIA

Economists' core criterion for assessing the welfare effects of policies and economic activities is the Pareto criterion, named for the Italian economist Vilfredo Pareto. Adoption of a policy or the undertaking of some activity satisfies the Pareto criterion if that policy or activity makes at least one person better off without making anyone worse off. This Pareto criterion is simultaneously quite weak and impossibly strong.

The Pareto criterion is weak in the sense that few people with modern, liberal sensibilities would find any policy or activity that satisfies it to be objectionable. Accepting the Pareto criterion obviously involves a value judgment, but one that is almost universally shared in modern society.[13]

But the Pareto criterion is (nearly) impossibly strong in the sense that it is extremely difficult in reality to identify a policy or action that does not cause at least one or a handful of people to be made worse off. If satisfaction of the Pareto criterion were required for any real-world action to be taken, very few actions would be taken in the real world. Ironically, *everyone* over time would be made immeasurably worse off as the resulting inactivity would

---

[12] Rodrik is not alone. Another famous trade economist who insists that the case for free trade is valid only if trade's 'winners' compensate trade's 'losers' is Nobel laureate Paul Krugman:

> the conventional case for trade liberalization relies on the assertion that the government could redistribute income to ensure that everyone wins - but we now have an ideology utterly opposed to such redistribution in full control of one party, and with blocking power against anything but a minor move in that direction by the other.
>
> So the elite case for ever-freer trade is largely a scam….

Paul Krugman, *A Protectionist Moment?*, N.Y. TIMES: THE CONSCIENCE OF A LIBERAL (Mar. 9, 2016, 4:32 PM), https://archive.nytimes.com/krugman.blogs.nytimes.com/2016/03/09/a-protectionist-moment/.

[13] *See* David Gordon, *The Pareto Criterion and Ethics*, ORG. & MKTS. (Sep. 22, 2006, 7:31 PM), https://organizationsandmarkets.com/2006/09/22/the-pareto-criterion-and-ethics/*; but see* Pareto Concepts Handout from Ethan Bueno de Mesquita, Sydney Stein Professor & Deputy Dean, Harris Sch. of Pub. Pol'y, Univ. of Chi. 43 (2019) (https://home.uchicago.edu/bdm/pepp/pareto_handout.pdf).

bring on humanity's extinction. Obviously, a welfare criterion of such stringency is not merely impractical; it is apparently severely flawed.

Actually, though, what is flawed is not the criterion *per se* but the domain in which it is often carelessly applied. As James Buchanan argued in 1962, the relevance of the Pareto criterion applies not at the level of actions but, instead, at the level of rules.[14] Not only to thrive, but even to survive, the rules prevailing in society must leave each to individual a great deal of freedom to act in ways that take account of local knowledge, which frequently changes. Such freedom will very often result in the taking of actions that, examined in isolation, make some persons better off while making other persons worse off. The relevance of the Pareto criterion, therefore, is found not in the assessment of individual actions but, instead, in the assessment of the *rules* under which actions are taken.[15] A change in rules that makes at least one person better off without harming anyone else not only satisfies that criterion but is far more likely to exist in reality than is any individual *action* that satisfies the criterion.

Fifty thousand years ago a rabbit unexpectedly appears in the brush. If the person who spots it must, before taking action to capture the creature, confirm that his capturing it will not make anyone else worse off – say, a hungry neighbor who would have spotted the uncaptured rabbit a few moments later – the rabbit will hop merrily away, never to be captured by a human being. We can be sure that any human ancestors who insisted on governing their band or tribe by applying the Pareto criterion to each and every action were long ago out-competed by humans who followed more realistic strategies.

But we can be equally sure that our long-ago ancestors used *some* welfare criterion for choosing (or settling upon) the *rules* by which they lived. The human bands that chose (or stumbled upon) better rules – rules that better promoted the survival of members of the bands – outcompeted bands that chose (or stumbled upon) worse rules. 'Survival' here implies rules that at least discourage actions that diminish the bands' access to resources. The rule, for example, that prohibits the spotter of the rabbit from capturing the animal until and unless the spotter first confirms that his capturing the rabbit will harm no one else in the band is a rule that would reduce the band's access to resources. So, for the band to survive the spotter must be allowed to capture the rabbit without first having to consult anyone else.

II. THE RANGE OF PROPERTY RIGHTS

Central among the rules used by all successful societies are those that can be classified under the label "private property rights" (or "several

14 *See* James M. Buchanan, *The Relevance of Pareto Optimality*, 6 J. CONFLICT RESOL. 341, 341 (1962).

15 *Id.*

property rights").[16] While the concrete details of the law of property vary across time and countries, a feature of all such law is the distinction between actions that are permissible and those that are impermissible.[17] In Anglo-American common law, property owners are generally permitted to use their properties in ways that do not interfere with other property-owners' abilities to *physically* use their properties.[18] But the law does not protect property-owners' market values in their properties.[19]

This feature of the law does not reflect some arbitrary choice. Even less does it reflect any 'pro-capitalist' bias. Instead, this feature of the law reflects nothing more – or less – than the law's cognizance of reality. Because material facts of the world change in unexpected ways, as do individuals' subjective tastes and preferences, protecting property owners against diminutions in the values of their properties is impossible.

Begin by noting that the law cannot simultaneously protect the physical integrity of individuals' properties and the market values of those properties. To allow individuals to physically use their properties as they wish—on condition that these uses do not interfere with other property-owners' physical uses—necessarily is to allow property owners to act in ways that will change the market values both of their own properties and that of other persons' properties.

To allow Smith to erect a new residential-apartment building on his land is to allow Smith to change his land's market value. If his new building is a success – as determined by the eagerness of tenants to rent space in it – the market value of his property rises. If the building is a failure, the market value of his property falls. More importantly for our discussion, to allow Smith to erect a new apartment building on his land is to allow him to change the market values also of *other people's* properties. If Smith's building is a success, the market values of nearby apartment buildings might fall as owners of these other buildings must lower their rents or increase their maintenance expenses in order to compete for tenants.[20] Implicit in – indeed, inseparable from – the law's 'decision' to protect people's physical uses of property is the law's 'decision' *not* to protect properties' market values.[21]

---

16 *See, e.g.*, Armen A. Alchian, *Property Rights*, ECONLIB, https://www.econlib.org/library/Enc/PropertyRights.html (last visited on Oct. 3, 2022); TOM BETHELL, THE NOBLEST TRIUMPH: PROPERTY AND PROSPERITY THROUGH THE AGES 10-11 (1998); RICHARD PIPES, PROPERTY AND FREEDOM, 3–63 (1999).

17 *See* STEARNS ET AL., *supra* note 8, at 222.

18 *Id.*

19 *See* Alchian, *supra* note 16.

20 Smith's new building might also cause the market values of nearby apartment buildings to *rise*. In this paper, however, I ignore this possibility in order to focus on the negative impacts that economic change nearly always has on some particular market values. People complain chiefly about negative impacts, seldom about positive ones.

21 *See generally* U.S. CONST. amend. V (protecting private property from being taken and only compensating for taking such property rights). *See also* Alchian, *supra* note 16.

The law's 'choice' between protecting physical integrity and uses or protecting market values is not a toss-up. It is not the case that the law could as easily, or almost as easily, have 'chosen' instead to protect market values while not protecting owner's physical uses of property. This truth is difficult to see if focus is limited to any one instance considered in isolation, such as Smith and the apartment building. It appears obvious that the existing value of Smith's own land, as well as the values of nearby apartment buildings, can be protected by prohibiting Smith from physically changing the uses to which he puts his land.

But appearances here are deceiving. To protect the market values of all of these pieces of land requires far more than restrictions on what Smith and other existing landowners may do with their properties. Such protection of market values would require, in addition, an unfathomably large number of other physical restrictions. Tenants must be prohibited from moving out of their spaces (or at least from failing to continue to pay their monthly rents). Yet because tenants cannot be prohibited from dying, government must have the power to somehow ensure that all units once occupied by now-deceased tenants are immediately rented by others.

How would government achieve this outcome? Several different schemes are imaginable, but none that does not involve a high probability of bringing about precisely what the law aims to avoid, namely, reductions in the market values of some people's properties. The law could, for example, compel residents of other neighborhoods to move into apartments that are vacated due to death. But this move would reduce the market value of apartment buildings in these other neighborhoods. Alternatively, government could simply take over responsibility for paying the rents previously paid by tenants now dead. This scheme would reduce the real incomes of those who are taxed to supply the necessary funds, and this reduction in real incomes reduces the market values of those assets, including human capital, owned by the affected taxpayers.

We need not extend the example further. Sober consideration of what the law would have to do in a determined pursuit of the goal of protecting market values rather than protecting physical uses quickly reveals the goal of protecting market values to be impossible. In happy contrast, the goal of protecting physical uses is not only possible, but one that better ensures that real market values of assets *generally* rise over time. That is, the protection of physical uses rather than of market values, while resulting in many specific instances in decreases in the market values of particular assets, enables individuals to use their assets in ways that, over time, have the best prospect of ensuring the highest possible value of these assets.

This outcome is not ironic. In a world of scarce resources, changes in market values are necessary to attract resources to those particular uses in which they will be of greatest use – of greatest 'value' – to humankind and,

hence, away from those particular uses that are not as useful.[22] In a world in which property owners are protected from suffering reductions in the values of their resources and assets, owners have little incentive to release their resources and assets from relatively unproductive uses, and are severely restricted by law from seeking different, more-productive uses for their resources and assets. The result over time is economic stagnation. Per-capita incomes fall. Standards of living fall. And property values in the aggregate are lower than they would be in an alternative world in which the law does not protect market values.[23]

III. COMPENSATION?

Of course, a policy in which trade's 'winners' compensate trade's 'losers' is imaginable without the law switching over entirely to the protection of all market values instead of physical uses. And if such compensation is practically possible, perhaps it is also advisable as a means of further ensuring that the rising tide of prosperity created by free trade does indeed lift all boats. At the very least we can ask: Why *shouldn't* trade's 'winners' compensate trade's 'losers'?

The short response is to ask: Why *should* trade's 'winners' compensate trade's 'losers'? More fully: Why should property owners have the values of their properties protected from one specific source of potential decline if these values are not protected from other sources of decline? Unless there is something unique about economic change fueled by commerce that crosses political borders, the case for the 'winners' from *international* trade compensating the 'losers' from *international* trade is a loser.

And there is indeed nothing about commerce that crosses political borders that differs in any essential economic respects from commerce that occurs exclusively within the confines of a country's national borders.[24] This conclusion, it can be fairly said, is a sound summary of economists' teachings about trade and trade policy since Adam Smith first put quill to parchment.[25]

There simply is nothing unique about trade that crosses political borders at creating what are called 'winners' and 'losers.' *All* economic change does

---

22 *See* Alchian, *supra* note 16 ("No matter who the owner is, the use of the resource is influenced by what the rest of the public thinks is the most valuable use. The reason is that an owner who chooses some other use must forsake that highest-valued use—and the price others would pay him for the resource or for the use of it.").

23 *See* BETHELL, *supra* note 16, at 10 (Observing that "Property sets up fences, but it also surrounds us with mirrors, reflecting back upon us the consequences of our own behavior. Both the prudent and the profligate will tend to experience their deserts." A system of property law that aimed to protect market values would shatter these mirrors.) *Id.* at 11 elaborates on the connection between economic prosperity and private property.

24 *See* Boudreaux, *supra* note 5, at 496–97.

25 *Id.*

so, and in a modern commercial economy it does so incessantly.[26] Implementation in California of a newly discovered lower-cost process for producing aluminum will destroy some jobs in Ohio's steel factories.[27] The harms suffered by American workers who thus lose their jobs, as well as the harms suffered by American steel-company shareholders whose portfolios thus take a hit, are no less real or harsh than are the harms they would have suffered had the demand for their outputs instead been reduced by a lowering of American tariffs on steel imports from Brazil or China.[28] Therefore, to say that international trade has winners and losers – or that a lowering of trade barriers creates winners and losers – is to imply incorrectly that there is something unique about international trade at destroying jobs and businesses in the home country.

This fact alone renders unsupportable the claim, such as made above by Dani Rodrik, that the economic case for free trade rests on the assumption that trade's 'losers' be compensated by trade's 'winners.'[29] Because no one believes that every source of economic change, even those that are purely domestic, are justified only if 'winners' compensate 'losers,' to single out economic change that is manifested through international trade as alone having to satisfy this requirement is unwarranted both as a matter of economics and of ethics.[30]

Those who are reluctant to accept this conclusion might ask themselves this question: Should the lowering of legal and cultural barriers to women working in the market have been conditioned on the resulting 'winners'

---

26 *See* Donald J. Boudreaux, *Free Trade and How It Enriches Us*, INST. OF ECON. AFFS. 32 (2018); Daron Acemoglu, David Autor, David Dorn, Gordon H. Hanson & Brendan Price, *Import Competition and the Great US Employment Sag of the 2000s*, 34 J. OF LAB. ECON. 141, 144-45, 183 (2016). Boudreaux documents that the number of jobs typically destroyed each *month* in the United States today is not much smaller than the total number of jobs destroyed during the so-called "China Shock" period – that is, the 13-year stretch from 1999 through 2011 in which American trade with China is said to have inflicted unusually intense harm on American manufacturing workers.

27 *See* DOUGLAS A. IRWIN, FREE TRADE UNDER FIRE 203–04 (5th ed. 2020).

28 *See* DONALD J. BOUDREAUX, GLOBALIZATION 144–45 (1st ed. 2007).

29 *See* RODRIK, *supra* note 6, at 3.

30 This paper is not a primer in trade theory. So I here only summarize economists' dual conclusion that (1) foreigners buy from, and sell to, us for the very same reasons that we buy from, and sell to, each other; and (2) all positive economic benefits created by purely domestic trade are of a type created also by foreign trade, and all negative economic consequences created by foreign trade are of a type created also by purely domestic trade. Although there are inessential economic differences, such as the need to make currency exchanges, that typically distinguish foreign trade from domestic trade – and although foreign trade sometimes has *non*-economic consequences (such as ones affecting national security) that are much more rare with purely domestic trade – nothing in either economic theory or economic history reveals foreign trade to differ in any economically essential WAYS from purely domestic trade. *See* e.g., BOUDREAUX, *supra* note 26, at 57; Boudreaux *supra* note 5, at 33; Daniel Griswold, MAD ABOUT TRADE: WHY MAIN STREET AMERICA SHOULD EMBRACE GLOBALIZATION 31–32 (2009); PIERRE LEMIEUX, WHAT'S WRONG WITH PROTECTIONISM?: ANSWERING COMMON OBJECTIONS TO FREE TRADE 79 (2018); IRWIN, *supra* note 27, at 38, 42.

compensating the 'losers'?[31] After all, the entry of more women into the market certainly destroyed some particular jobs held by men and worsened the bottom lines of some particular businesses owned by other men. Everyone who would reject the assertion that the case for allowing women to work in the market is premised on the assumption that the resulting 'winners' will compensate the 'lowers' should also reject, and for the very same reasons, the assertion that allowing freer trade is premised on the requirement of paying such compensation.

It is no good objection to observe that the case of women entering the market, unlike the case of making trade freer, involves only fellow citizens. The 'winners' who scholars such as Rodrik believe should compensate the 'losers' are only those domestic citizens who directly gain from freer trade.[32] Although some foreigners also gain when trade at home is made freer, these foreigners are not in the pool of persons who are regarded as taxable for the purpose of compensating trade's 'losers.'[33] The pool of 'winners' whose gains should be taxed in order to get the funds necessary to compensate trade's 'losers' is believed to consist only of those domestic citizens – those domestic producers and domestic consumers – who can be identified as gaining from a freeing of trade.

But there is an even deeper reason to reject the claim that the case for free trade requires that 'winners' compensate 'losers.' It is a reason that returns us to the above discussion of the Pareto criterion for assessing the welfare consequences of actions and policies. This reason can be stated boldly: A *policy* of unilateral free trade *does* in fact satisfy the Pareto criterion. A *policy* (or a *rule*) of unilateral free trade does in fact make everyone better off without making anyone worse off.

This claim initially seems to be incurably mistaken. When a cut in tariffs prompts Mike in Michigan to buy more steel from Brazil and, hence, less steel from Ohio, Joe and Jennifer, who worked for years at an Ohio steel mill, lose their jobs. Their incomes fall. This freeing of trade clearly seems to inflict losses upon them. Joe and Jennifer appear to be among trade's losers. Workers such as these are obviously the kinds of people who Rodrik, Krugman, and others have in mind when they talk of trade's "losers."[34]

Here as earlier, appearances are deceiving. When we consider the larger, time-embedded market process in which Joe and Jennifer (and Mike) are active participants, it becomes clear that none of these Americans are losers from trade. Indeed, all are winners. Big time.

---

31 The same question might also be asked about the removal of legal and cultural barriers that resulted in more black employees working in the market.

32 *See* RODRIK, *supra* note 6, at 204.

33 Nonresident aliens can only be taxed for income earned from domestic sources, not economic benefits they accrue abroad. *See, e.g.*, *Taxation of Nonresident Aliens*, IRS, https://www.irs.gov/individuals/international-taxpayers/taxation-of-nonresident-aliens (June 2. 2022).

34 *See* RODRIK, *supra* note 6, at 230; Krugman, *supra* note 12.

IV. A (TRULY) GREAT GAME

Do the following mental experiment. Suppose you can join a special poker game in which every time a new hand is dealt the house gives each player more money with which to play. The amount of new money the house adds to your earnings is random: Sometimes you get a lot; other times you get very little. Also, sometimes the amount of new money that you get will be more than what the person sitting beside you gets, while other times it will be less. But each time you get from the house at least some additional new money.

Here, though, is what makes this poker game really special: The amount of new money that the house distributes over time to each player is so great that every player who remains in the game long enough is guaranteed to become richer over time, regardless of how well or poorly he or she plays the game. And the longer a player remains in the game, the richer that player becomes.

Keep in mind that you can also add to your earnings by playing the game well. Your skill and care at playing the game matter; they positively affect your prospects of winning at each 'play' of the game as well as your winnings over time. In each round you can bet whatever amount of your earnings you wish. If you win a particular round, you obviously grow richer – you get the table's winnings from that round *and* whatever amount of new money the house distributes to you at the conclusion of that round.

Of course, you will not win every round. Indeed, if you bet a great deal in a round that you lose, you might be poorer when that round ends than you were before that round started, despite the new house money that you received at the end of that round. But again, the house guarantees that, regardless of your fate after each round, if you play the game long enough your real wealth will grow.

In this game, therefore, there are no losers. There are, of course, individuals who suffer losses in each round. But because the house guarantees to each player that he or she will grow richer over time the longer he or she remains in the game, this game has only winners. A person who, after losing a particular round, quits this game or calls for its rules to be fundamentally altered, would be foolish.

Free trade is very much like this special poker game. The house – a policy of free trade in bourgeois society[35]—ensures that over time every participant grows richer than that person would be without free trade. But there is no guarantee that in every 'round' of economic 'play' no one suffers losses. Indeed, as explained above, because economic growth requires that people be free to use their properties in ways that cause unexpected and unwanted changes in the market values of other people's properties, in each

35 *See generally* DEIRDRE N. MCCLOSKEY AND ART CARDEN, LEAVE ME ALONE AND I'LL MAKE YOU RICH: HOW THE BOURGEOIS DEAL ENRICHED THE WORLD (2020).

and every economic 'round' some people will be made worse off than they were at the start of that 'round.' And so, while today some peoples' incomes do fall because of free trade, it makes no sense to describe these people as losing from trade itself.[36] Their very participation in the global economy makes them – makes *everyone* in modern society– richer over time.[37]

Do I exaggerate by claiming that literally *everyone* in modern society is made better off over time by trade? No. Empirical evidence supports this conclusion. In a 2017 study, Gary Clyde Hufbauer and Zhiyao (Lucy) Lu report that "the payoff to the United States from trade expansion" from 1950 to 2016 "is roughly $2.1 trillion (measured in 2016 dollars)… US GDP per capita and GDP per household accordingly increased by $7,014 and $18,131, respectively (both measured in 2016 dollars)."[38]

Is it possible that, even though this study spans nearly two-thirds of a century and finds significant per-capita income gains, some individual Americans were nevertheless made worse off over this span of time by trade? Yes, it is possible, but only barely.

Of course, some workers who lost jobs to imports during this period never found new jobs, while some others found jobs that, over the remainder of their working lives, paid them, in inflation-adjusted dollars, less in total than these workers would have been paid had they never lost their jobs to imports. But during this time span most of even these workers gained increased access, year after year, to the fruits of global trade – fruits such as novel consumer electronics and other new goods from abroad, as well as lower prices and higher quality – because of foreign competition – of many domestically produced consumer goods.[39] These benefits accumulate over time.

Consider a hypothetical American who lost a job to imports relatively early – say, 1970 – in the period studied by Hufbauer and Lu. If he or she survived for 40 more years, that person gained access to (among other benefits) automobiles made much better by foreign competition, to cell phones and other consumer electronics that in some cases were made possible, and that in nearly all cases were made more affordable, by foreign competition, and to clothing the real prices of which fell substantially,[40]

---

36 *See* Donald J. Boudreaux, *Trade and Romance*, CAFÉ HAYEK (January 7, 2007), https://cafehayek.com/2007/01/trade_and_roman.html (analogizing to the claim that a man whose wife left him is not a 'loser' at the game of love, but instead won in the past, lost now, and might win in the future).

37 *See, e.g.*, The Copenhagen Consensus Centers Post 2015 Consensus, *Free Trade Could Make Everyone in the World Richer*, Global Policy (Dec. 1, 2014), https://www.globalpolicyjournal .com/blog/01/12/2014/free-trade-could-make-everyone-world-richer.

38 Gary Clyde Hufbauer and Zhiyao (Lucy) Lu, Peterson Inst. for Int'l Econ., *Globalization: A Fresh Look with a Focus on Costs to Workers* 1, PB17-16 (May 2017).

39 *See, e.g.*, The Copenhagen Consensus, *supra* note 37.

40 E-mail from Mark Perry, Senior Fellow, American Enterprise Institute. to Donald J. Boudreaux, Professor of Economics, George Mason University (Oct. 19, 2021) (on file with author) (His calculation that between 1970 and 2010 the real price of clothing in the U.S. fell by 64%).

driven largely by the off-shoring of much clothing production. The fall in this worker's income in 1970 was real, but so too are the additions that trade contributed to this person's real income over the next 40 years. It is highly unlikely that, all things considered, trade inflicted on this person a net loss from the time he or she lost the job until his or her death in 2010. To classify this person as a "loser" from trade, therefore, would be a mistake.

The mistakenness of such a classification becomes even more evident when we look, not only forward, but also backwards. Looking backwards reveals that the job this hypothetical worker lost in 1970 was itself very likely the result of international trade. If the job this worker lost was, say, in a television-manufacturing plant, that job was likely created years earlier because of foreign demand for American-made television sets. Almost certainly, the wages paid to workers employed in that plant, including wages paid to the hypothetical worker whose job in 1970 was 'destroyed' by trade, were higher than these wages would have been had mid-20th-century American television manufacturers not had an export market for their outputs – an export market that would not have existed were Americans unable to import.[41]

The direct impact in earlier years of trade in creating jobs, and in raising the pay of workers who held these jobs, must be reckoned against the direct impact in later years of trade then destroying these same jobs. Workers who today, because of trade, lose jobs that were created yesterday *by* trade are not legitimately classified as "losers" from trade.

V. COSTS ARE NOT LOSSES

Participation in the modern economy signals consent to play by the rules of the modern economy. As stated above, among the most central of these rules is that the law protects the physical integrity of people's properties but not properties' market values.[42] And so when market forces cause a fall in the market value of some property – including that which a worker has in his or her labor – that property owner is more accurately described as paying the *cost* of participating in a market economy rather than suffering a *loss*.[43] I

41 *See generally* Boudreaux, *supra* note 5, at 490 (describing how free trade benefitted US manufacturing from the 1940s to the 1970s).

42 *See, e.g.*, Alchian, *supra* note 16.

43 Such costs are the sort that are classified by Anglo-American tort law as *damnum absque injuria*—that is, damage without injury. *See Damnum Absque Injuria*, THE WOLTERS KLUWER BOUVIER LAW DICTIONARY DESK EDITION (2012), https://plus.lexis.com/ (search "damnum absque injuria" in quotation marks under "Secondary Materials" and it is the third result) ("There are cases when the act of one man may cause a damage or loss to another, and for which the latter has no remedy; he is then said to have received damnum absque injuria; as, for example, if a man should set up a school in the neighborhood of another school, and, by that means, deprive the former of its patronage; or if a man should build a mill alongside of another, and consequently reduce his custom"). For alerting me to this fact I thank Richard Epstein.

propose that the word "loss," in economic and legal policy discussions, be reserved to describe losses of possessions or benefits to which a person is legally entitled. A thief who steals $100 from you inflicts on you a loss of $100. A trespasser inflicts on you loss of your right to control your land. Your unfortunate encounter in the intersection of Main St. and Elm Blvd. with a negligent driver inflicts on you a loss of some of your property. Contract breach brings loss of a bargained-for anticipated benefit. In each of these examples, you had a legitimate expectation of being legally entitled to some property interest that was unexpectedly taken from you. For these losses, you are rightfully entitled to be made whole by compensation from the party who wronged you.

Costs differ from losses categorically. Unlike losses, costs are what choosers voluntarily sacrifice in exchange for benefits. Both losses and costs, each standing alone, are decrements from individuals' welfare. But only losses spring from a series of human interactions that decrease that welfare on net. When someone suffers a loss, that person is made worse off. In contrast, when someone incurs a cost, that person is made better off.

This odd-sounding conclusion about costs follows from the fact that costs are the inescapable consequences of choices.[44] I choose to pay $15 for a pizza because I expect that the satisfaction that I will get from the pizza exceeds the satisfaction that I would get were I to keep that $15 or to spend or invest it in some other way. Because I can get satisfaction from the pizza only by incurring the cost of sacrificing $15 for it, incurring this cost makes me better off. This reality is not changed by the fact that the net increase in my welfare would be even greater were the restaurant owner to give me the pizza free of charge. The bottom line here is that no one would call the $15 that I spend for the pizza a "loss."

Consider another example. Suppose that, in order to buy a home, I borrow $200,000 from a mortgage lender and agree to repay the loan, in monthly installments, at a certain rate of interest over the course of fifteen years. I move into the home today and commence living in it. A man from Mars, with no previous knowledge of earthly conventions, pops down to earth two years from now and observes me sending a check each month to the mortgage holder. After a few months of observation, the Martian reports to his leaders on the red planet that each month some mysterious force inflicts on me a loss in the amount of my mortgage payment. The Martian and his leaders conclude that I would be better off were I not obliged to suffer a loss each month in the form of these monthly amounts.

But no knowledgeable earthling would describe me as each month suffering a loss. When asked to describe the meaning of my mortgage payments, the earthling would instead say that I am paying the *cost* of having

44 *See generally,* JAMES M. BUCHANAN, COST AND CHOICE: AN INQUIRY IN ECONOMIC THEORY (1969).

borrowed money to purchase a home.[45] The earthling would be correct. Of course, I would be delighted, after obtaining the borrowed the funds, for the mortgage holder to then relieve me of my obligation to repay. My welfare would be raised by such relief. But I clearly would be made worse off if, as a result of the mistaken conclusion that the institution of mortgage lending imposes *losses* on borrowers, creditors had long been prohibited from demanding repayment from debtors. While in my ideal world the mortgage lender would simply give me the $200,000 with no strings attached, I am nevertheless better off in a world in which mortgage lenders can demand repayment of loans than I would be in a world in which such demands are unlawful.

Now consider an American steelworker who loses a job because fellow Americans start buying more steel from Brazil and, hence, less steel from Ohio. This worker suffers no loss. Instead, this worker is paying a cost of participating in the modern commercial global economy. It is indisputable that this worker would prefer not to have to pay this cost, just as it is indisputable that I would prefer not to have to repay my mortgage lender. But this reality does not transform the worker's (or my) cost into a loss.

Each worker in a modern commercial economy is very much like the above-mentioned mortgage holder. Each such worker voluntarily participates in this economy because of the enormous benefits he or she reaps from doing so. But these benefits are possible only because producers must compete for consumers' dollars—only because consumers are generally free to spend their incomes as they choose and are not regarded as contractually binding themselves, with their purchases, indefinitely to each producer that he or she patronizes. These benefits are possible, in other words, because the law protects the physical uses and integrity of property and not properties' market values.[46]

VI. DO WE REALLY CHOOSE TO PARTICIPATE IN THE MODERN ECONOMY?

It is fair to ask if I am correct to claim that workers in the modern commercial economy are akin to persons who take out mortgages. To take out a mortgage, after all, is a voluntary act.

But so too is participation in the modern commercial economy a voluntary act. No one is forced to participate. Each of us has the option of

---

45 The situation would differ categorically if, owning my home outright, I am accosted by mafioso who credibly inform me that if I do not pay to them a certain sum each month they will forcibly remove me from my home. In this case my monthly payments are indeed appropriately described as losses. (Technically, I incur the full loss the moment the mafia threatens to remove me from my home if I refuse to send them monthly payments. Nothing essential is changed by the fact that these particular mafioso kindly allow me to endure my loss in the form of monthly installment payments.)

46 *See, e.g.*, Alchian, *supra* note 16.

withdrawing from commercial society as a tiny handful of people have actually done.[47] A great deal of rural land is available for purchase.[48] Each of us is free to buy this land on which we can scratch out a living, either literally alone or with whatever small band of individuals we persuade to accompany us.

Of course, the resulting material standard of living of economically isolated individuals would be desperately low compared to the standard of living available even to the poorest of those who participate in the modern commercial economy.[49] Indeed, the colossal difference between the maximum standard of living achievable by those who divorce themselves from modernity and the minimum standard of living available those who remain integrated into modernity is what makes my claim that it is possible to abandon commercial society seem so far-fetched.

What *is* far-fetched – in fact, what is impossible – is to divorce oneself from modernity while simultaneously maintaining a standard of living close to what we moderns today regard as minimally acceptable. Realization of the crushing material deprivation of a life detached from the modern economy is what gives people the impression that such a life is impossible. Yet such a life is not impossible; it is just extremely costly given the option of remaining in modernity.[50] The cost of freeing oneself from the consequences of the decisions of financiers in New York and London, of steel producers in China and India, of farmers in France and Florida, of entrepreneurs in California and Denmark, and of governments in Washington and Beijing is simply too high to be remotely appealing to all but the most anti-social recluses.

And so nearly everyone born into modernity *chooses* to remain in modernity. The fact that this choice is such an easy one does not make it less of a choice. Each of us who chooses to earn an income by serving strangers in the market, and to spend that income on goods and services produced by

---

47 *See* DAVID SCHMIDTZ, PERSON, POLIS, PLANET 199 (2011).

48 A majority of US land is privately owned and thus available for purchase. *See* Stacey Vanek Smith & Cardiff Garcia, *The U.S. Has Nearly 1.9 Billion Acres Of Land. Here's How It Is Used*, NPR (July 26, 2019, 4:28 PM), https://www.npr.org/2019/07/26/745731823/the-u-s-has-nearly-1-9-billion-acres-of-land-heres-how-it-is-used. *See also* Rural Land Prices for Texas, TEXAS REAL ESTATE RESEARCH CENTER AT TEXAS A&M UNIVERSITY, https://www.recenter.tamu.edu/data/rural-land#!/state/Texas (last visited Oct. 2, 2022) (showing land sales in several states); Undeveloped Land for Sale, LANDWATCH, https://www.landwatch.com/undeveloped-land (last visited Oct. 2, 2022) (showing undeveloped land available for purchase).

49 *See* J. Paul Newell, *Rural Healthcare: The Challenges of a Changing Environment*, 47 MERCER L. REV. 979, 982 (1996) (discussing issues with healthcare services in rural communities and inadvertently showing how well off they are compared to hunter-gatherers); Allison Lacko, Shu Wen Ng & Barry Popkin, *Urban vs. Rural Socioeconomic Differences in the Nutritional Quality of Household Packaged Food Purchases by Store Type*, INT'L J. OF ENV'T RSCH. & PUB. HEALTH 39, 50 (2020) (discussing issues with food product quality in rural communities and inadvertently showing how well off they are compared to hunter-gatherers).

50 *See* Daniel Mark Schwartz, *How Much Money It Takes To Live Off The Grid | Five Hidden Costs*, OFF GRID PERMACULTURE, https://offgridpermaculture.com/Beginners/How_Much_Money_It_Takes_to_Live_Off_the_Grid_5_Hidden_Costs.html (last visited Oct. 3, 2022).

strangers and put at our disposal by the market, thereby *chooses* to play by the rules of the market. Among the most crucial of these rules is the law's protection of property-owners' physical rights to their properties but not of the market values in these properties.[51]

As posed above, the choice is stark: remain in, or withdraw from, modernity. Fortunately, modernity itself supplies options along a spectrum, with divorcing from modernity being only an endpoint. Within the modern market are options for different degrees of protection from the risks of market adjustments, although nothing close to complete protection against such risk. Each of us can, without abandoning market connections completely, choose economic lives with more or less risk of being subject to market adjustments.

Modern workers make current choices in the labor market with eyes to the future and, hence, many workers *do* make estimates about the risks of job loss. Nearly every young person in the developed world confronts various broad career options. Attend college or not? Major in electrical engineering, in accounting, in education, or in gender studies? Pursue an MBA, an MD, or a JD? Accept this job offer or that other job offer? Although the relevant variables for making such decisions are many, job security is certainly among these variables.[52]

It is also unrealistic to imagine that individuals have no or only inaccurate information to rely upon when assessing the security of different jobs. At one extreme, young men and women who enlist in the military know the terms of enlistment and can be confident that the option to reenlist will be available in the future.[53] At the other extreme, entrepreneurs who launch new firms selling never-before-available goods or services understand that the probability of failure is high.[54] In between these two extremes stretches a wide spectrum of different jobs and career paths, nearly every one of which is accompanied by some public knowledge about its pay, amenities, prestige, difficulty, and security.

This public knowledge does not have to be perfect for workers with different preferences for job security to sort themselves into different jobs with different degrees of security.[55] For example, public-sector employers perform fewer layoffs than private sector employers during recessions, a fact

---

51 *See, e.g.*, Alchian, *supra* note 16.

52 A May 2017 poll conducted in the United States by Yahoo! Finance found that workers do indeed assess the job-security aspects of different employment options. Rick Newman, *Labor Shortages? Blame Corporate America*, YAHOO! NEWS (June 13, 2017), https://www.yahoo.com/news/labor-shortages-blame-corporate-america-173604463.html.

53 *See Joining the Military: What You Should Know Before Committing*, MILITARY.COM, https://www.military.com/join-armed-forces/making-commitment.html (Last visited Oct. 3, 2022).

54 *See* Bobby Chernev, *What Percentage of Small Businesses Fail?*, REVIEW42, https://review42.com/resources/what-percentage-of-small-businesses-fail/ (Feb. 1, 2022).

55 *See* Susan Adams, *Job Security A Top Priority For Global Workers*, FORBES (July 31, 2014, 10:19 AM), https://www.forbes.com/sites/susanadams/2014/07/31/job-security-a-top-priority-for-global-workers/?sh=5b5b8a581acc.

that might attract workers who put a high priority on job security.[56] Such workers also tend to choose jobs in large, long-established firms over jobs in small upstart firms funded with venture capital. These individuals are more likely to move, in search of work, away from geographical locations that are in economic decline to geographical locations that are growing and seem likely to continue to grow. People with different preferences for job security choose differently.

No less importantly, employers have incentives to discover just how strongly workers want job security. All other things equal, employers who offer greater job security can pay their workers lower money wages than are paid by employers who offer less job security.[57] Workers who do not value additional job security enough to accept lower wages will work in jobs that pay more but offer less security, while workers who have higher demands for job security will be attracted to jobs that offer greater security but pay lower wages.

Employers have every incentive to supply additional units of such security as long as the costs of supplying additional units of job security fall short of the benefits employers receive in return in the form of lower wage bills. Therefore, the greater are workers' demands for job security, the greater are the reductions in wages and fringe benefits that workers are willing to accept in exchange for increased job security – and the greater are employers' abilities and willingness to supply such security.

Given that labor markets are generally competitive, we can presume that the market supplies different optimal levels of job security for different workers. The amount of such security in any given job is not greater than it is now because the cuts in wages and fringe benefits that workers would have to accept in exchange for more security are judged by workers to be too great. But it cannot be denied that different jobs supply different degrees of job security. Nor can it be denied that many jobs, perhaps most, supply at least *some* degree of job security, even if such security does not appear as a formal term in employment contracts (clerks in a McDonald's restaurant are not laid off after one day of lower-than-expected sales volume.). If nothing else, workers form plausibly correct expectations about just how secure each job is, and wages adjust upwards or downwards to reflect these expectations.

---

56 *See* ALICIA H. MUNNELL & REBECCA CANNON FRAENKEL, CTR. FOR RET. RSCH., PUBLIC SECTOR WORKERS AND JOB SECURITY 2–3 (2013), https://crr.bc.edu/wp-content/uploads/2013/05/SLP31_508.pdf.

57 *See* Marios Michaelides, *A New Test of Compensating Differences: Evidence on the Importance of Unobserved Heterogeneity*, 11 J. OF SPORTS ECON. 475, 475–79 (2010); Matthew Dimick, *Compensation, Employment Security, and the Economics of Public-Sector Labor Law*, 43 U. TOL. L. REV. 533, 545–46 (2012).

CONCLUSION

One conclusion is straightforward. Anyone who wishes to own assets, including human capital, possessing market values must subject those assets to market-valuation processes – which is to say, to competitive markets. Those persons who intensely fear the fate of suffering any declines in the market valuation of their assets can escape that fate, but only by choosing to divorce themselves from the market by scratching out a subsistence existence unattached to commercial society. Those persons who instead choose to remain attached to commercial society thereby choose to play by the rules of competitive markets. They choose to incur the risk of declines in the market valuations of their assets as the cost of enjoying the prospect of increases in the market valuations of their assets. Any declines in asset values caused by market processes – including those caused by changes in the pattern of international trade – are therefore not losses that in justice ought to be compensated but, instead, are costs that in justice must be paid by each person who enjoys the benefits of participating in modern commercial markets.

A second, related conclusion is warranted: Economists' concept of "pecuniary externalities"[58] is worse than useless. It is misleading. Changes in the market values of properties occur incessantly and are indeed very often caused by the activities of fellow human beings who change their commercial behaviors, consumers who change their patterns of spending, and entrepreneurs who change the availability of goods and services. These changes typically arise out of self-interested ("utility-maximizing" and "profit-maximizing") actions without the actors taking account of the consequences of their actions on the market values of other people's properties. Yet in no sense are these changes in market valuations externalities. They are a routine, expected, *and essential* feature of the competitive market process. To label them "externalities," even when such labeling is followed by an explanation of why policymakers should ignore them, is therefore inaccurate. Such labeling creates only confusion. The term should be completely jettisoned.

[58] *See, e.g.*, *Pecuniary Externality*, LIQUISEARCH, https://www.liquisearch.com/pecuniary_externality (last visited on Oct. 17, 2022).

CREATIVE DESTRUCTION: HOW CAPITALISM UNDERMINES RULE OF LAW

Randall G. Holcombe[*]

THE EVOLUTIONARY NATURE OF CAPITALISM

Joseph Schumpeter described capitalism as a system of creative destruction. Innovative new ideas, new products, and new production methods displace the old. This system works to the advantage of entrepreneurial individuals who bring profitable innovations to market, enabling them to get ahead by producing more value for consumers. This same process that works to the advantage of entrepreneurial individuals who bring innovations to market threatens those who have succeeded by doing so in the past. Their past successes can be undermined by the same entrepreneurial forces they employed to displace those who came before them. Some individuals are in a position to benefit from the forces of creative destruction. Those who have benefited in the past from those forces eventually find themselves in a position of being threatened by the forces of creative destruction.

Rule of law benefits entrepreneurial individuals because it creates a level playing field that gives entrepreneurs the opportunity to introduce innovations that can displace the products of established firms. Once they become established, however, those same individuals have an incentive to undermine rule of law, because a level playing field enables potential rivals to challenge their established positions in the market. Once established, firms have an incentive to use the connections that come with the economic power they have accumulated to change the rules of the game to favor themselves—to create barriers to entry to potential rivals. In capitalism, where the new replaces the old, and the new benefit from a competitive environment where they can challenge established firms, while the old look for ways to change the rules to create stability rather than progress. Those who have risen to the top in the competitive environment of capitalism want a legal framework that stabilizes the status quo, rather than legal institutions that enable the creative destruction that is an integral part of capitalism.

Schumpeter says:

> The essential point to grasp is that in dealing with capitalism we are dealing with an evolutionary process . . . . Capitalism, then, is by nature a form or method of economic change and not only never is but never can be stationary . . . . The fundamental impulse that sets and keeps the capitalist engine in motion comes from the new consumers' goods,

* Florida State University. The author appreciates helpful comments from Jeremy Kidd, Alex Tabarrok, and participants at a George Mason University Law School workshop.

> the new methods of production or transportation, the new markets, the new forms of industrial organization that capitalist enterprise creates. . . . . This process of Creative Destruction is the essential fact about capitalism. It is what capitalism consists in and what every capitalist concern has got to live in.[1]

Those who have the greatest amount of economic power have the most to lose in the capitalist environment of creative destruction, so they act to change the rules to create stability rather than progress. By undermining that process of creative destruction, they are acting to undermine what Schumpeter calls an essential element of capitalism.

I. CAPITALISM AND PROSPERITY

This point is obscured in much economic analysis because of its focus on equilibrium outcomes in an economy, and because of its focus on income growth rather than economic progress. Economic models that focus on equilibrium concepts overlook the evolutionary nature of capitalism emphasized by Schumpeter.[2] Once an economy arrives at equilibrium, as these models depict it, the economy tends to stay there unless disturbed by outside forces, in which case it tends to reequilibrate. But the force that is always at work to disturb the existing state of affairs is entrepreneurship. When entrepreneurial innovations are introduced into an economy, they disturb the status quo, and the economy never returns to its former state of affairs. A capitalist economy is characterized by progress, not equilibrium.

When automobiles displaced horse-drawn transportation, stables and farriers were replaced by gas stations and auto mechanics. Similarly, personal computers replaced typewriters, and the internet made hardcover encyclopedias obsolete. The new replaced the old. That is the evolutionary process Schumpeter described. To try to shoehorn this into some vision of equilibrium is misleading. One might say the equilibrium is constantly changing, but in that case it is not really an equilibrium. It is a continually evolving system.[3]

To understand the relationship between capitalism and rule of law requires an understanding of how the capitalist economy operates, as a decentralized system of voluntary exchanges. Capitalism is a market system in which economic activity takes place through mutually advantageous

---

[1] JOSEPH A. SCHUMPETER, CAPITALISM, SOCIALISM, AND DEMOCRACY 82-83 (2d ed. 1947).

[2] For a good critique of this equilibrium approach to economic analysis, see generally Meir Kohn, *Value and Exchange*, 24 CATO J. 303 (2004).

[3] Economists often refer to equilibrium prices and quantities in a supply and demand framework, but a more appropriate way to look at it is that the model shows market-clearing prices and quantities. The model shows how prices adjust so that the quantity supplied equals the quantity demanded at any point in time, but in a continually evolving capitalist economy, to refer to an equilibrium is misleading. Markets clear, but never arrive at equilibrium, if equilibrium refers to a situation to which the system returns if disturbed. Equilibrium describes a stagnant economy, not a continually evolving one.

transactions. In a market, people transact only when they voluntarily agree to do so, which means that all parties to a transaction must believe the transaction is advantageous to them. Otherwise, they would decline to transact. If entrepreneurs want to enter new markets or increase their share of old ones, they must entice people to transact with them, and they do that by offering potential customers more value than they can get elsewhere.

Kirzner says that entrepreneurship is the discovering of previously unnoticed profit opportunities.[4] Those previously unnoticed opportunities add value to the economy, and profit is an indicator of the added value. Entrepreneurial innovation is the activity that takes place when entrepreneurs discover ways in which they can create value for others. This means changing the status quo, not clinging to an equilibrium.

When economists discuss economic growth, they typically measure it in terms of income growth. In a simple framework, output, Q, which is measured in terms of its monetary value, is a function f of capital, K, and labor, L, so Q = f(K,L). In this framework, economic growth, which is depicted as growing Q, results from increases in capital and labor. Capital growth results from investment, and labor growth can be an increase in the number of workers, but more significantly, following Lucas, an increase in human capital.[5] Input growth accounts for only a small share of output growth, so most growth is a result of increases in f, productivity increases that allow a given quantity of inputs to produce more output.

This formulation omits the creative destruction that Schumpeter says is the essence of capitalism. In the twentieth century per capita income in the United States increased by seven times. In that century when Q per capita grew by seven times, people did not own seven times as many horses or eat seven times as much food. Rather, that growth in Q was driven by new consumption opportunities that entrepreneurs continued to make available to people. That growth was driven by progress—new and improved goods and services, and income growth is only a small component of the progress that capitalism has created. Prosperity is created by this type of progress, and income growth would not be possible without the progress initiated by entrepreneurial innovation. Economic models that depict growth as an increase in Q cannot capture the innovation that is the actual source of growth.

The reason capitalism is characterized by this innovation is that in an economy where people can only prosper by enticing others into voluntarily trading with them, entrepreneurs have an incentive to look for ways to create more value for people. People get ahead in market economies by finding ways to create value. The voluntary nature of market exchange is an essential feature of capitalism, because with voluntary exchange, people can only get ahead by creating value for others. Should people be forced to engage in

---

[4] *See* ISRAEL M. KIRZENER, COMPETITION & ENTREPRENEURSHIP 30-43 (1973).

[5] *See generally* Robert E. Lucas, Jr., *On the Mechanics of Economic Development*, 22 J. MONETARY ECON. 3-42 (1988).

economic activities, the opportunity would exist for people to use force to take from others, rather than create value, and when that opportunity exists, there is every reason to think that opportunistic individuals will take advantage of it.

Capitalism rests on an institutional framework that enables voluntary transactions among individuals and prevents the forcible transfer of resources from some to others. These institutions can be summarized briefly as protection of property rights, rule of law, and limited government. If this institutional framework (discussed further below) is compromised, the capitalist engine of economic progress is compromised along with it.

II. THE NEW IS BETTER THAN THE OLD

In the process of creative destruction, what is created is more valuable than what is destroyed. This must be the case because people voluntarily choose to consume the new rather than the old, demonstrating that they value it more. In the evolution of economic systems, like biological systems, "survival of the fittest" determines how that evolution takes place, as those who are more "fit" displace those who are less. In biological systems, there is no metric to say that what has evolved is in any way better or worse than what went before. Indeed, sometimes when "invasive species" enter an ecosystem, people judge what has evolved to be inferior to what came before. Capitalism does have a metric to determine whether what has evolved is better: profitability. The fact that people choose the new over the old shows that, by their own revealed preferences, they place a higher value on it, and those who succeed in introducing innovations into the economy reap their reward as profit. Profit is an indication that the value of the output produced is greater than the value of the resources used to produce it.

The prosperity that capitalism has generated is the result of entrepreneurial innovations that have produced new and improved goods and services, and more efficient production processes. Entrepreneurial innovation is the driver of economic progress, and entrepreneurs have the incentive to create value for others in a system where they can only profit if people agree that they are better off engaging in those transactions. These fundamental ideas explain why capitalism is more productive than central economic planning. The decentralized knowledge of entrepreneurs throughout the economy leads to innovations that could not be comprehended by economic planners.[6]

---

[6] Much more could be said to explain the advantages of the decentralized economic decision-making that occurs under capitalism when compared to central planning. LUDWIG VON MISES, SOCIALISM: AN ECONOMIC AND SOCIOLOGICAL ANALYSIS (1951) explains why rational economic calculation requires markets and market prices. Friedrich A. Hayek, *The Use of Knowledge in Society*, 35 AM. ECON. REV. 519-30 (1945) makes a significant contribution, and DON LAVOIE, NATIONAL ECONOMIC PLANNING: WHAT IS LEFT? (1985) offers an excellent discussion.

All of this relies on an economic system in which individuals interact with each other through voluntary exchange, and whereas a result individuals profit from those exchanges by creating value for others. Rule of law is an essential element in that institutional foundation, in which everyone deals with each other only through voluntary action, and in which property rights are protected. Capitalism creates prosperity because it is a social system in which people improve their own welfare when they produce value for others. It requires an institutional foundation in which individuals deal with each other through mutual agreement. The voluntary nature of interpersonal interactions prevents opportunistic individuals from taking from others by force.

III. TWO VIEWS ON PROFITS

Economic analysis typically views profit as a sign of economic inefficiency. According to Marshall, profit is the result of either monopoly power or market disequilibrium, both of which are inefficient.[7] In contrast to Marshall, Schumpeter says "without development, there is no profit, without profit, no development."[8] Schumpeter's view is that profit is necessary for economic efficiency.[9] In the static neoclassical view, competitive markets allocate resources efficiently, and all profits are competed away in competitive equilibrium, so profit (beyond a "normal" profit) is a sign of inefficiency.[10] Schumpeter's evolutionary view of competition depicts profit as the lure that entices entrepreneurs to introduce innovations into the economy, resulting in economic progress and a continual increase in economic well-being.[11] Profit is an indicator of the additional value entrepreneurs create in an economy.

Schumpeter's view of profit ties directly to his description of capitalism as an evolutionary process of creative destruction.[12] In contrast to the static view of efficiency inherent in neoclassical welfare economics, where efficiency is an outcome (Pareto optimality) that maximizes welfare, an evolutionary view depicts welfare maximization as an ongoing process, continually increasing prosperity as a result of economic progress.[13] Profit is an indicator of the additional value that innovative entrepreneurs bring to the

---

7 ALFRED MARSHALL, PRINCIPLES OF ECONOMICS (8th ed. New York, MacMillan & Co. 1948) (1890).

8 JOSEPH A. SCHUMPETER, THE THEORY OF ECONOMIC DEVELOPMENT 154 (Redvers Opie trans., Cambridge, Harvard University Press 1949) (1934).

9 *Id.*

10 Leferis Tsoulfidis, Pub. No. 43999, Classical vs. Neoclassical Conceptions of Competition 7 (2013).

11 SCHUMPETER, *supra* note 8, at 45.

12 SCHUMPETER, *supra* note 1, at 81.

13 RICHARD R. NELSON & SIDNEY G. WINTER, AN EVOLUTIONARY THEORY OF ECONOMIC CHANGE 31 (1982).

economy. Profit arises because customers value what the profitable firm brings to market more than the alternatives.

Consider the antitrust case the Justice Department brought against IBM in 1969, asserting that IBM was monopolizing the computer industry.[14] From an economic standpoint, the case rested on the facts that IBM was very profitable (more than "normal" profits) and that it had an overwhelming market share, both indicators of monopoly power in the Marshallian neoclassical framework.[15] IBM gained its profitability and its market share by developing its IBM 360 computer, the first computer with a multitasking operating system.[16] Before the IBM 360, people wrote computer programs to run on the computer, and when one program finished running, another could be loaded.

One innovation in its new system was that programs were written to run on the operating system rather than directly on the computer, and multiple programs could be loaded and run at the same time. The IBM 360 was so much more advanced than competing computer systems that customers flocked to it, and the profitability and market share IBM gained were the result of introducing a better product into the market. The 360 project was not without its risks. It cost so much to develop that people inside IBM referred to the 360 project as "you bet your company," because they thought IBM could go bankrupt if the project failed.[17]

The profit IBM earned from its 360 computer was an indicator of the extra value the computer provided to customers over competing computers. Nobody was forced to use a 360 computer, so the money users paid IBM showed the value of that computer when compared to alternatives. The profit was an indicator of the extra value of the 360 computer, but as Holcombe notes, not a complete measure of the extra value, because some of that extra value went to the consumers who used the computer.[18] The extra value was shared between IBM and its customers.

The lawsuit was dropped in 1982, before it went to trial.[19] By that time, competitors had caught up with IBM's innovations and both the market share and the profitability of the company fell. This is what happens in a dynamic economy. Successful innovators get ahead of their rivals, giving them larger market shares and increased profitability. That drives economic development, as Schumpeter notes.[20] Eventually, rivals catch up, and the profits that were an indicator of the extra value entrepreneurs brought to the economy dissipate. New innovations displace the old. The value of the

14 United States v. IBM, 69 Civ. 200 (S.D.N.Y.).

15 John E. Lopatka, United States v. IBM*: A Monument to Arrogance*, 68 ANTITRUST L.J. 145, 146-48 (2000).

16 *Id.* at 147.

17 T.A. Wise, *IBM's $5,000,000,000 Gamble*, FORTUNE, Sept. 1966, at 118.

18 Randall G. Holcombe, *The Common Poll of Transitional Profits*, 27 REV. AUSTRIAN ECON. 387, 389-90 (2014).

19 Lopatka, *supra* note 15, at 145.

20 SCHUMPETER, *supra* note 8, at 45.

innovation remains, but as Holcombe explains, competition transforms that value from profits to the innovator into consumer surplus for those who use the innovations.[21]

The neoclassical model of competitive equilibrium concludes that profit is an indicator of economic inefficiency.[22] Profit is a sign of monopoly power or disequilibrium in a market, both of which are inefficient. Schumpeter's evolutionary view of capitalism concludes that profit is necessary for economic efficiency. Profit is the engine that drives economic progress, and guides entrepreneurs toward the efficient allocation of resources.

As Schumpeter notes, the potential for profit drives economic development, and economic development produces profits.[23] This requires a market economy in which economic activity takes place through voluntary transactions, so that entrepreneurial individuals have the incentive to create value for others. That is the dynamic nature of capitalism—an economic system that is always changing as it generates economic progress.

IV. RULE OF LAW AND THE INSTITUTIONAL FOUNDATION FOR CAPITALISM

Government plays an essential role in preserving an institutional foundation that enables capitalism. For a market economy to function, property rights must be protected, and people must be accountable to a generally applicable and objectively enforced set of laws. Government provides the institutions that create these conditions. In the absence of government, people profit from plunder rather than productivity. Nobody has an incentive to be productive if what they produce can just be taken away from them.[24] If government law enforcement favors individuals who have connections to those in power, undermining rule of law, entrepreneurial individuals have an incentive to seek profits by developing connections to the politically powerful rather than by producing products that add value to the economy. People's entrepreneurial impulses lead them to engage in destructive rather than productive activities, as Baumol explains.[25] Rule of law creates a level playing field that channels entrepreneurial individuals to

---

21 Holcombe, *supra* note 18, at 389-90.

22 Tsoulfidis, *supra* note 10 at 7–8.

23 SCHUMPETER, *supra* note 8, at 45.

24 Friedman and Rothbard both argue that government activity is unnecessary and even immoral, but a detour into those ideas will not be taken here because government is ubiquitous. DAVID D. FRIEDMAN, THE MACHINERY OF FREEDOM: A GUIDE TO RADICAL CAPITALISM, 1 (2d ed. Chicago: Open Court Publishing Co. 1973); MURRAY N. ROTHBARD, FOR A NEW LIBERTY: THE LIBERTARIAN MANIFESTO (2d ed. Auburn: Ludwig von Mises Institute 2006) (1973). For an alternative view, see Randall G. Holcombe, *Government: Unnecessary But Inevitable*, 35 AM. ECON. REV., 325-42 (2004).

25 William J. Baumol, *Entrepreneurship: Productive, Unproductive, and Destructive*, 98, J. POL. ECON. 893, 894 (1990).

seek profits by producing value for others rather than by using the force of government to plunder others.[26]

Until recently in human history, governments were both the main institutions that protected people from outside invaders and the main institutions that plundered their own citizens. Diamond and Pinker both attribute the orderly and relatively peaceful condition of modern societies to strong government,[27] but government must be more than strong. It must be powerful enough to protect the rights of its citizens, but its power must be constrained so that it does not violate the rights of its citizens.[28] Rule of law represents one constraint. Without rule of law, some people are treated more favorably than others, and entrepreneurial individuals have an incentive to use their entrepreneurial instincts to gain membership in the group favored under the law rather than to produce value for others. Absence of rule of law leads to cronyism and corruption, which undermines the innovative entrepreneurship that drives economic progress.

Capitalism is founded on institutions that enforce rule of law, protect property rights, and do not interfere with mutually advantageous exchanges. A good summary of those institutions is given by Gwartney, Lawson, and Hall, under the heading of economic freedom.[29] Small government, as measured by low taxes and government expenditures, low regulatory barriers to exchange, and freedom to trade internationally promote capitalism because people can make their own economic decisions rather than having government make economic choices for them. Two key elements in their analysis, which describes the institutional foundation that is necessary for capitalism, are protection of property rights and rule of law.

A government that enforces rule of law enables the creative destruction that characterizes capitalism. In the absence of rule of law, entrepreneurial individuals look for ways to profit themselves by using the force of government to gain favorable treatment under the law. Government serves a useful function in a capitalist economy when it protects property rights and enforces rule of law, but it also interferes with and undermines capitalism when it interferes with mutually agreeable exchanges.[30]

---

26 *Id.*

27 JARED. DIAMOND, THE WORLD UNTIL YESTERDAY: WHAT CAN WE LEARN FROM TRADITIONAL SOCIETIES?, 75 (2012); STEVEN PINKER, ENLIGHTENMENT NOW: THE CASE FOR REASON, SCIENCE, HUMANISM, AND PROGRESS, 212–17 (2018).

28 JAMES M. BUCHANAN, THE LIMITS OF LIBERTY: BETWEEN ANARCHY AND LEVIATHAN (1975); DARON ACEMOGLU & JAMES A. ROBINSON, THE NARROW CORRIDOR: STATE, SOCIETIES, AND THE FATE OF LIBERTY xi – xvi (2019).

29 GWARTNEY, JAMES GWARTNEY, JOSHUA HALL ROBERT LAWSON & RYAN MURPHY. ECONOMIC FREEDOM OF THE WORLD: 2019 ANNUAL REPORT (Fraser Institute 2019).

30 De Soto notes that government interference with market exchange can push economic activity to underground markets where people have limited protection from rule of law, resulting in lower productivity as entrepreneurs try to stay one step ahead of government regulations. HERNANDO DE SOTO, THE OTHER PATH: THE INVISIBLE REVOLUTION IN THE THIRD WORLD, 146-51 (June Abbott trans. 1989) (1988).

V. GOVERNMENT AND BUSINESS

There are two sets of ideas in the academic literature, rarely viewed together, that shed light on the relationship between government and business. One is the public choice literature that explains how concentrated interests are able to interact with government to benefit themselves at the expense of the general public.[31] The other originates in sociology and political science and explains why institutions are designed to benefit the elite over the masses.[32] Elite theory explains *who* benefits from the relationship between government and business, and public choice theory explains *how* they benefit.

C. Wright Mills says, "The powers of ordinary men are circumscribed by the everyday world in which they live . . .. But not all men are in this sense ordinary. As the means of information and power are centralized, some men come to occupy positions in American society from which they can look down upon, so to speak, and by their decisions mightily affect, the everyday world of ordinary men and women."[33] Mills calls these individuals the power elite.[34] They are the people who have the connections that enable them to benefit themselves through special treatment unavailable to the people Mills calls ordinary men.[35]

Joseph Stiglitz offers a similar message, noting how an elite few are able to avoid being constrained by rule of law, because they make and enforce the law. "It's one thing to win a 'fair' game. It's quite another to be able to write the rules of the game—and write them in ways that enhance one's prospects of winning. And it's even worse if you can choose your own referees."[36] David Stockman echoes the same message: "In truth, the historic boundary between the free market and the state has been eradicated, and therefore anything that can be peddled by crony capitalists … is fair game."[37] These are a few examples of a more extensive literature on elite theory, which is

---

31 *See, e.g.*, Gordon Tullock, *The Welfare Cost of Tariffs, Monopolies, and Theft*, 5 W. ECON. J. 224, 224-32 (1967); Anne O. Krueger, *The Political Economy of the Rent-Seeking Society*, 64 AM. ECON. REV. 291, 291-303 (1974); Fred S. McChesney, *Rent Extraction and Rent Creation in the Economic Theory of Regulation*, 16 J. LEGAL STUDIES, 101-18 (1987) [hereinafter *Rent Extraction*]; FRED S. MCCHESNEY, MONEY FOR NOTHING: POLITICIANS, RENT EXTRACTION, AND POLITICAL EXTORTION (1997) [hereinafter MONEY FOR NOTHING]; PETER SCHWEIZER, EXTORTION: HOW POLITICIANS EXTRACT YOUR MONEY, BUY VOTES, AND LINE THEIR OWN POCKETS 108-18 (2013); George J. Stigler, *The Theory of Economic Regulation*, 2 BELL J. ECON. AND MGMT. SCI. 3, 3-21 (1971).

32 *See* C. WRIGHT MILLS, THE POWER ELITE 3 (1956).

33 *Id.*

34 *Id.*

35 *Id.*

36 JOSEPH E. STIGLITZ, THE PRICE OF INEQUALITY: HOW TODAY'S DIVIDED SOCIETY ENDANGERS THE FUTURE 59 (2012).

37 DAVID STOCKMAN, THE GREAT DEFORMATION: THE CORRUPTION OF CAPITALISM IN AMERICA, 606 (2013).

discussed more thoroughly by Holcombe.[38] Elite theory explains why there is one group of people—the elite—who are favored by the institutional structure so that they have inherent advantages over the masses.

Elite theory explains who benefits; public choice theory explains how they benefit. Public choice describes institutions that undermine rule of law and enable some to use political institutions to benefit themselves at the expense of others. Rent-seeking,[39] rent extraction,[40] regulatory capture,[41] and interest group politics[42] are broad headings within the public choice literature that explain the mechanisms by which some people are able to negotiate favorable treatment to benefit themselves while imposing costs on the general public. Essentially, these theories describe the mechanisms through which some are able to influence the government to create public policies that provide concentrated benefits for themselves by imposing dispersed costs on the masses.[43]

The public choice literature omits an important detail that the literature on elite theory fills in. The people who are able to benefit themselves through their interactions with government are a small and well-connected group. The public choice literature on rent-seeking, for example, makes it appear that rent-seeking is a competitive contest that anyone can enter to compete for rents. Successful rent-seeking requires political connections, and people gain political connections only when they have something to offer politicians. Most people do not have these connections and cannot engage in successful rent-seeking. Successful rent-seekers are members of the economic elite, who are well-connected and can obtain political favors from the political elite.[44]

Well-established businesses have the financial resources to buy political support for regulatory barriers to entry, subsidies and tax breaks, and trade barriers that protect them from international competitors. Organizations that represent large groups of voters can offer politicians endorsements and potential voter support in addition to "contributions" to their parties and campaigns. Those people are the elite, the well-connected, the cronies of crony capitalism. Any business could benefit by receiving special treatment by the government, but only an elite few are in a position to negotiate with the government to get special treatment.

---

38 RANDALL G. HOLCOMBE, POLITICAL CAPITALISM: HOW ECONOMIC AND POLITICAL POWER IS MADE AND MAINTAINED 44-71 (2018).

39 *See* Tullock, *supra* note 31; Krueger, *supra* note 3.

40 *See* McChesney, *Rent Extraction*, *supra* note 31; MCCHESNEY, MONEY FOR NOTHING, *supra* note 31; SCHWEIZER, *supra* note 31.

41 Stigler, *supra* note 31, at 3 –21.

42 MANCUR, JR. OLSON, THE LOGIC OF COLLECTIVE ACTION: PUBLIC GOODS AND THE THEORY OF GROUPS 5–10 (1965).

43 For detailed descriptions of these theories, see generally RANDALL G. HOLCOMBE, ADVANCED INTRODUCTION TO PUBLIC CHOICE (2016).

44 Randall G. Holcombe, *Political Incentives for Rent Creation*, 28 CONST. POL. ECON. 62, 63-76 (2017).

VI. CREATIVE DESTRUCTION: WINNERS AND LOSERS

An entrepreneurial economy rewards those who introduce new and improved products and more efficient production methods. Rule of law creates a level playing field so that those with the best ideas can bring them to market and get ahead. The winners in this system are the creators, the innovators, the people who find new ways to create value for others. But the system of creative destruction that creates opportunities for some destroys the profitability of others. Those who want to get ahead in a capitalist economy benefit from the creative destruction that rule of law enables, while those who want to stay ahead are threatened by it.[45]

Many writers have credited the institutions of capitalism for the remarkable economic prosperity that has emerged since the beginning of the Industrial Revolution,[46] but fewer have noted the way that the evolutionary nature of capitalism threatens those who are well-established, as the new displaces the old. Part of this threat arises because new ideas are sometimes better than old ones, so those who have found a formula for success can be displaced by newcomers who have a better formula. Even if established firms recognize the merit in new ideas, they can be reluctant to adopt them because in the process their old sources of profitability are destroyed.

This is referred to as the innovator's dilemma.[47] Creative firms establish themselves in the market by introducing innovative new products that create value for consumers. Once they become established, their incentives change from introducing disruptive products into the market—they already did that—to preventing other firms from introducing disruptive products that compete with their previous and now well-established innovations.[48] They may be able to do that by using their political connections to create regulatory barriers to entry, as the railroads did in the late 1800s by capturing the Interstate Commerce Commission,[49] or the airlines did after World War II with the help of the Civil Aeronautics Board (established in 1938; abolished in 1985),[50] but these examples show that regulatory barriers may not be effective in the longer run at preventing firms from being displaced by innovators.

---

45 *See generally* Randall G. Holcombe, *Creative Destruction: Getting Ahead and Staying Ahead in a Capitalist Economy*, REVIEW OF AUSTRIAN ECONOMICS (forthcoming).

46 *See e.g.* JOEL MOKYR, THE LEVER OF RICHES 81-83 (1990); DAVID LANDES, THE WEALTH AND POVERTY OF NATIONS: WHY SOME ARE SO RICH AND SOME SO POOR 186-87, 194 (1998).

47 *See generally* CLAYTON M. CHRISTENSEN, THE INNOVATOR'S DILEMMA: WHEN NEW TECHNOLOGIES CAUSE GREAT FIRMS TO FAIL (1997).

48 *Id.* at xiii-xiv, 24.

49 GABRIEL KOLKO, RAILROADS AND REGULATION, 47-56 (Princeton University Press 1965); RICHARD WHITE, RAILROADED: THE TRANSCONTINENTALS AND THE MAKING OF MODERN AMERICA 355, 357, 359 (W.W. Norton 2011).

50 Records of the Civil Aeronautics Board [CAB], National Archives, (Aug. 15, 2016), https://www.archives.gov/research/guide-fed-records/groups/197.html.

The case of IBM and the computer industry, discussed earlier, provides an excellent example of the innovator's dilemma.[51] IBM, as noted above, was able to achieve huge market share and substantial profitability with the introduction of their innovative 360 computer. Moving into the 1980s, companies began introducing minicomputers that could undertake many of the business functions that customers were then doing on mainframes, but IBM was slow to get into the minicomputer business because they did not want their minicomputers to cannibalize their very profitable mainframe business. The result was eroding market share and eroding profitability as their customers turned to minicomputers.

The leading minicomputer manufacturer in the 1980s was Digital Equipment Corporation (DEC)—the innovator that was able to dominate that market. But DEC faced the same innovator's dilemma as personal computers grew in power. DEC was slow to move into the PC market because it did not want to cannibalize its minicomputers. Eventually DEC was bought by Compaq which was later merged into Hewlett-Packard.

These examples show how firms that benefit from the creative destruction of capitalism, as IBM and DEC did, can eventually become victims of that same creative destruction. DEC, as a firm, was destroyed, and IBM has survived only to drastically change its business model. When getting ahead, those firms benefited from the creative destruction of the market mechanism, supported by rule of law, that enables people to prosper by bringing innovations to market that create more value for others. Once they get ahead, innovators want to preserve the profitability of their innovations, which means undermining the process of creative destruction and the rule of law to try to protect their current market positions. Public choice theory explains how this is done: through rent-seeking, regulatory capture, and policies that benefit their special interests, often by imposing costs on the general public.

Creative individuals are the winners in this game, but once they have emerged on top their incentives change, and their interest is to stop the game—to undermine that process of creative destruction so they can remain in their dominant positions. They want to undermine rule of law and replace it with a set of institutions that preserve the status quo.

VII. POLITICAL CAPITALISM

Capitalism rests on an institutional environment which specifies that individuals interact voluntarily with each other in their economic activities. Rule of law and protection of property rights are essential elements in that institutional environment. Capitalism's institutions level the playing field so that those who are best able to create value for others profit from bringing their innovations to market and increase the general standard of living in the

51 CHRISTENSEN, *supra* note 47, at xiii-xiv, 24.

process. Those who are most successful accumulate wealth, and along with wealth comes economic power. Those who are trying to get ahead benefit from capitalist institutions which allow them to profit from their innovations. But once they achieve a certain level of success, they become the people who are threatened by a new group of innovators who are trying to displace them.

For market leaders, even innovations they might bring to market themselves present a threat to their current positions.[52] The incentives of well-established firms shift from gaining advantages through the continually evolving nature of capitalism toward preserving the status quo in which they are the economic elite. They increasingly view progress as a threat; they want stability.

Economic analysis largely focuses on the way that people make decisions subject to constraints. However, many of the constraints on people's actions are institutional constraints that are the result of human design and can be redesigned by human action.[53] Rule of law is this type of constraint. These are the rules of the game that are designed and enforced by the elite for their own benefit. Those who have achieved economic success have an incentive to preserve their positions at the top of the economic hierarchy by undermining rule of law, modifying the institutional environment to make it more difficult for entrepreneurial individuals to displace the already-successful. Their incentives shift from playing to win within the rules of the game to changing the rules of the game.[54]

The institutional mechanisms the economic elite use have been referenced already: the rent-seeking, regulatory capture, and the power of concentrated special interests. The economic elite are able to profit in these ways because they have connections to the political elite, and because they can use their economic power to buy political influence. Their cronyism and corruption produce a political capitalism in which profitability increasingly is the result of political connections rather than the production of value for consumers. The evolutionary progress that results from the creative destruction of capitalism is undermined by the erosion of rule of law. When this happens, stagnation displaces progress, and legal favoritism displaces rule of law.

The economic and political elite are able to cooperate with each other in this way because they face low transaction costs and are able to negotiate institutional changes that are mutually beneficial. Most people—the masses—face high transaction costs and are unable to participate in public policy negotiations. The Coase theorem is perhaps more relevant to political exchanges than to market exchanges.[55]

---

52 *Id.*

53 *See generally* James M. Buchanan, *The Domain of Constitutional Economics*, 1 CONST. POL. ECON. 1 (1990).

54 JOSEPH E. STIGLITZ, THE PRICE OF INEQUALITY: HOW TODAY'S DIVIDED SOCIETY ENDANGERS THE FUTURE 29-32 (2012).

55 *See generally* Ronald H. Coase*, The Problem of Social Cost*, 3 J. LAW & ECON. 1-44 (1960).

Consider the common example of a smoke-polluting industry that creates an externality for those who live nearby and breathe polluted air. Steel mills, for example, produce steel to sell to automobile manufacturers and others. The steel companies and auto companies are able to engage in exchanges that maximize the value of resources to themselves because they face low transaction costs. Those who breathe the polluted air are unable to bargain with the polluters because they face high transaction costs. The polluters are not trying to impose costs on their neighbors. Those costs are just a by-product of their actions as they engage in transactions with others to maximize the value of resources to themselves. High transaction costs prevent those who suffer external costs from being able to bargain to mitigate them.

The same thing occurs with political exchanges. The economic and political elite face low transaction costs and bargain among themselves to maximize the value of legislation to themselves, just like the steel mills and auto manufacturers. Legislators bargain among themselves, and lobbyists have access which enables them to transact with legislators to reach mutually beneficial agreements. The by-product is that the masses, who face high transaction costs, bear the costs of the bargains negotiated between the economic and political elite, just like those who breathe the polluted air. When rent-seekers are able to use the force of government to transfer concentrated benefits to themselves, the dispersed costs fall on the masses in a manner analogous to the air pollution people suffer from nearby industries. The elite negotiate with themselves to maximize the value of their transactions, and the masses face high transaction costs and are left to bear the costs of the policies created by the elite.

To preserve their positions in the economic hierarchy, the economic elite secure for themselves regulatory barriers to entry, tax preferences, subsidies, and government contracts, undermining rule of law and making their profitability increasingly dependent on their political connections rather than on producing value for consumers. Rule of law evolves into cronyism and favoritism—referred to as political capitalism.[56] Entrepreneurs who are trying to get ahead benefit from rule of law. Capitalists, once they are well-established, attempt to undermine rule of law, using their economic power to buy political favors to protect themselves against the creative destruction of capitalist institutions. Over time, special interests become more firmly entrenched in political institutions, undermining rule of law and replacing it with cronyism that benefits those who have political connections, leading to the decline of nations.[57]

This does not imply that up-and-coming entrepreneurs would not accept that same type of political favoritism that the cronies receive, but they are not in a good position to get it. They do not have the established connections,

[56] GABRIEL KOLKO, THE TRIUMPH OF CONSERVATISM: A REINTERPRETATION OF AMERICAN HISTORY, 1877-1916 (1963).

[57] *See generally* MANCUR OLSON, THE RISE AND DECLINE OF NATIONS (1982).

and have limited resources they can use to compensate the political elite. Given the constraints they face, emerging entrepreneurs have no alternative but to try to profit by creating value for consumers, while well-established cronies who are threatened by the creative destruction of capitalism work to undermine rule of law to retain their positions in the economic hierarchy.

VIII. CAPITALISTS ARE THE BIGGEST THREAT TO CAPITALISM

Schumpeter pessimistically wrote "Can capitalism survive? No, I do not think it can."[58] Those who benefit most from it, Schumpeter said, will not stand up to defend it. Rather than supporting a level economic playing field, capitalists lobby for special interest benefits that undermine free markets and rule of law. Schumpeter saw socialism as the impending threat to capitalism. Also in the 1940s, Hayek concluded that socialism is the road to serfdom.[59] A mixed economy leads to increasingly more interventionism.[60] Government interventions create negative unintended consequences, which result in demands for more government intervention to deal with those consequences, producing still more unintended negative consequences and another round of demands for government to address them. Thus, a mixed economy edges away from capitalism toward socialism as public policy responds to those ongoing unintended consequences.

While the defenders of a market order see socialism as the looming threat to capitalism, capitalism is being undermined from within, by the actions of capitalists. Capitalists are the biggest threat to capitalism. When capitalists interact with the political system, they do not ask for a level playing field or the protection of free markets, they ask for special interest benefits, for regulatory protections, for tax breaks, for subsidies, and barriers to foreign competitors. They do not say they want these policies to benefit themselves at the expense of others, of course. They argue that the institutional changes they seek are "pro-business." In general, pro-business means anti-free market.

This is true of even the most innovative firms. Walter Isaacson. in his biography of Apple Computer CEO Steve Jobs, relates,

> President Clinton's Justice Department was preparing a massive antitrust case against Microsoft. Jobs invited the lead prosecutor, Joel Klein, to Palo Alto. Don't worry about extracting a huge remedy against Microsoft, Jobs told him over coffee. Instead simply keep them tied up in litigation. That would allow Apple the opportunity, Jobs explained, to make an 'end run' against Microsoft and start offering competitive products.[61]

---

58 JOSEPH A. SCHUMPETER, CAPITALISM SOCIALISM & DEMOCRACY 61 (2003).

59 *See generally* FRIEDRICH A HAYEK, THE ROAD TO SERFDOM (1944).

60 SANFORD IKEDA, DYNAMICS OF THE MIXED ECONOMY: TOWARD A THEORY OF INTERVENTIONISM 133 (1997).

61 WALTER ISAACSON, STEVE JOBS 323 (2011).

As Lamoreaux notes, when big firms engage in the political process, it is almost always to attempt to provide special interest benefits for themselves or to impose costs on their competitors.[62]

Capitalist institutions create a competitive environment in which profitable firms can survive and thrive, but unprofitable firms ultimately must fail. The quest for profits in capitalist economies does not directly give entrepreneurs an incentive to support free markets or rule of law, but rather to engage in activities that enhance their profits. One way to enhance a firm's profitability is to seek advantages over the firm's rivals by using government connections, tilting the playing field in their direction by undermining rule of law.

Neoclassical economics assumes that in capitalist economies, firms act to maximize profits, and for good reason. Profitable firms survive and prosper while unprofitable ones die off.[63] Under capitalist institutions—protection of property rights, rule of law, and limited government—firms maximize profits by producing value for their customers. But the ultimate goal of firms is not to produce value for their customers, but to produce profits for themselves. Another strategy firms can pursue to enhance their profitability is to engage in the political process to undermine rule of law to change the rules to favor themselves and create impediments to their rivals. In a competitive economy where firms' very survival is at stake, profit-maximizing firms have every incentive to alter the rules to favor themselves whenever possible. Capitalists do not have the incentive to preserve capitalism, but rather to undermine it to protect themselves from being victims of creative destruction. If capitalism is to defy Schumpeter's expectations and survive, it will be despite the actions of capitalists, because capitalist institutions themselves give capitalists strong incentives to undermine capitalism.

62 Naomi R Lamoreaux, *The Problem of Bigness: From Standard Oil to Google,* J. OF ECON. PERSP. Summer 2019, at 16.

63 Armen A. Alchian, *Uncertainty, Evolution, and Economic Theory*, 58 J OF POL. ECON., no. 3, June 1950, at 211-21.

WHAT IS A LAW AND POLITICAL ECONOMY MOVEMENT WITHOUT LAW AND ECONOMICS OR POLITICAL ECONOMY?

Justin (Gus) Hurwitz[*]

INTRODUCTION

The Law and Political Economy (LPE) project is an initiative at Yale Law School that "brings together a network of scholars, practitioners, and students working to develop innovative intellectual, pedagogical, and political interventions to advance the study of political economy and law."[1] Since 2017, it has led to the establishment of the Journal of Law and Political Economy; the formation of student groups at schools including Harvard, Columbia, Cornell, NYU, Penn, Georgetown, and Berkeley; an ongoing series of regular conferences, workshops, symposia, and other events; and a network of over 50 academics at top law schools who speak and publish under the banner of the Law and Political Economy Project.

In recent years, academics writing under this banner have published articles with titles such as *The Law and Political Economy of Workplace Technological Change*;[2] *"There Is No Such Thing as An Illegal Strike": Reconceptualizing the Strike in Law and Political Economy*;[3] *The Law and Political Economy of a Student Debt Jubilee*;[4] *Taxation and Law and Political Economy*;[5] and *"Social Engineering": Notes on The Law and Political Economy of Integration*.[6] A dozen articles referring to Law and Political Economy as an "emerging," "growing," or even "burgeoning" movement have been published in the past two years.

---

[*] Professor of Law and Menard Director of the Governance and Technology Center, University of Nebraska. As always, this project has benefited from the input of many. Thanks in particular to my colleagues Elana Ziede and James Teirney, and to Tomas Merrill, Judge Doug Ginsburg, and Terry Anderson for their comments on earlier drafts of this paper, and to Robert Drust for outstanding research assistance. Many errors, of which I expect that there are many, are my own; other errors of which I will be accused may or may not be attributable to me or to my accusers.

[1] The LPE Project, *About*, https://lpeproject.org/about (last visited Oct. 24, 2022).

[2] Brishen Rogers, *The Law and Political Economy of Workplace Technological Change*, 55 HARV. C.R.-C.L. L. REV. 531 (2020).

[3] Diana S. Reddy, *"There is No Such Thing As An Illegal Strike": Reconceptualizing the Strike In Law and Political Economy*, 130 YALE L.J. F. 421 (2021).

[4] Luke Herrine, *The Law and Political Economy of a Student Debt Jubilee*, 68 BUFF. L. REV. 281 (2020).

[5] Bearer-Friend et al., *Taxation and Law and Political Economy*, 83 OHIO ST. L.J. 471 (2022).

[6] Olatunde C.A. Johnson, Symposium, *"Social Engineering": Notes on the Law and Political Economy of Integration*, 40 CARDOZO L. REV. 1149 (2019).

These are small numbers in absolute terms – but they represent the vanguard of a movement organized to promote its ideas and the work of scholars associated with them. The growing influence of the field is perhaps best seen in individual fields. For instance, LPE-affiliated scholars have published several antitrust-related articles in recent years, including several in the Yale Law Journal.[7] A superficial review of the papers, and topics covered in them that are of interest to the LPE project – from antitrust to labor law, tax policy, and housing regulation – make clear an interest in topics, if not methods, shared between the LPE project and law & economics (L&E) scholars.

This paper does two things. It begins with a critical discussion of the Law and Political Economy Project. This discussion casts the LPE project as reactionary to, and misrepresentative of, Law & Economics. But the purpose of this paper is not to criticize for the sake of criticism. To the contrary, the second part of the paper argues that the LPE and L&E schools are better read together than apart. They are animated by many of the same questions, and nearly every paper that self-identifies as part of the LPE initiative is on a topic of interest to L&E scholars. While they are certainly in many ways antagonistic, their perspectives are better understood in dialogue with each other than as dismissive of each other.

The LPE Project is interesting in that it has a clear ur-text: Britton-Purdy, Grewal, Kapczynski, & Rahman's *Building a Law-and-Political-Economy Framework: Beyond the Twentieth-Century Synthesis.*[8] This, in turn, builds upon an earlier "manifesto" for the LPE Project.[9] It also has a clear epexegesis in the editors' introduction to the first volume of the Journal of Law and Political Economy.[10]

In their "manifesto" and subsequent article, Britton-Purdy, Grewal, and Kapczynski – the founders of the LPE project – lay out their view that we are in "a time of crisis," characterized by rising inequality, racialized and gendered injustice, ecological disaster, and eroded democracy. They continue that "Law is central to how these crises were created, and will be central to any reckoning with them" and propose "a new orientation to legal scholarship that helps illuminate how law and legal scholarship facilitated these shifts, and formulates insights and proposals to help combat them."[11] Their central argument is that over the course of the 20th Century the law –

---

7 *See, e.g.*, Lina Khan, *Amazon's Antitrust Paradox*, 126 YALE L.J. 710 (2017); Sanjukta Paul, *Recovering the Moral Economy Foundations of the Sherman Act*, 131 YALE L.J. 175 (2021); Ramsi Woodcock, *The Obsolescence of Advertising in the Information Age*, 127 YALE L.J. 2270 (2018).

8 Jedediah Britton-Purdy et al., *Building a Law-And-Political-Economy Framework: Beyond the Twentieth-Century Synthesis*, 129 YALE L.J. 1784 (2020).

9 Jedediah Britton-Purdy et al., *Law and Political Economy: Toward a Manifesto*, LPE PROJECT BLOG (Nov. 6, 2017), https://lpeproject.org/blog/law-and-political-economy-toward-a-manifesto [hereinafter *LPE Manifesto*].

10 Angela Harris & James Varellas, *Introduction: Law and Political Economy in a Time of Accelerating Crises*, 1 J.L. & POL. ECON. 1, 1–2 (2020).

11 *LPE Manifesto, supra* note 9.

largely driven by law & economics – created an artificial divide between questions of economy and questions of politics.[12] This division correlates roughly into the delineation between private law (dealing with economic matters, to the exclusion of political ones) and public law (dealing with political matters, to the exclusion of economic considerations). Their central thesis is that addressing the various crises that they say characterize our current moment requires a reunification of the economic and political domains. As they describe it, "At the core of this reconstruction is a renewed commitment to questions of political economy," which "investigates the relation of politics to the economy, understanding that the economy is always already political in both its origins and its consequences."[13]

Both these authors and Harris & Varellas situate LPE in relation to, and as a successor to, traditions of legal realism. Britton-Purdy, Grewal, Kapczynski & Rahman frame many of the questions prompted by LPE in terms of legal realism and critical legal theories.[14] Harris & Varellas expressly call LPE a "reinvention" of Critical Legal Studies, saying that it "might be deemed critical legal studies for the age of neoliberalism."[15]

This paper proceeds in three parts. Part I describes LPE's claims and deconstructs what its founders refer to as, "the Twentieth-Century Synthesis." This "synthesis" refers to what the authors identify as, "two interrelated moves" that together came to define American legal scholarship and pedagogy in the latter 20th Century: a focus on efficiency in private law matters, and an insulation of distributional and other private economic impacts from public law matters.

Part II offers a critical assessment of that synthesis. This assessment is generally negative: the LPE characterization of both sides of its purported synthesis is founded in fundamental misunderstandings of history, argument, and theory. At the same time, Part III also acknowledges an uncomfortable truth: while LPE synthesis offers an account of law & economics that is wholly unrecognizable to most L&E scholars, it is also unquestionably an account of law & economics that is widely held by most lawyers, academics, and law students.

Part III marks a transition in the discussion, from criticism to construction. It starts by taking a moment to reflect on the relationship between competing schools of legal thought. Both L&E and LPE (and its ancestors in the Critical Legal Studies camp) have a regrettable history of contempt towards one another, which would profitably be replaced with a better understanding of one another. As things stand, many proponents of each are fairly and lamentably criticized as ideologues. Part III then turns to the LPE Project's curious invocation of "Political Economy." It accepts the

---

12 *Id.* ("Much of legal scholarship and practice in recent decades has held politics and economics apart, abstracting away from, or actively denying, their interdependence").

13 Britton-Purdy et al., *supra* note 8, at 1792.

14 *Id.* at 1793.

15 Harris & Varellas, *supra* note 10, at 8.

LPE founders' invitation to return to "the older and more foundational usage [of political economy] familiar to nineteenth-century audiences."[16] This concluding section offers a synthesis of my own, looking at the work of three central scholars in any serious study of political economy – Adam Smith, Elinor Ostrom, and Derrick Bell – to argue that the more important "synthesis" to be found in any law and political economy movement is one that embraces the questions and methods of both. This is the "political economy" of Adam Smith, and his radical appeals to liberal and enlightenment values. Taking LPE as a successor to the legal realists and critical legal theorists, I conclude by arguing that there is a valuable synthesis to be had between LPE and L&E, as seen through a lens of public choice.

I. THE LAW & POLITICAL ECONOMY PROJECT'S "SYNTHESIS"

The discussion below presents a description of the LPE Project, drawing from the founders of the LPE Project's work launching and articulating the views and goals of the project. It starts in Part I.A by describing the project broadly. Parts I.B and I.C then present the project's characterization of private and public law, respectively, which together form what the founders describe as "the twentieth-century synthesis"[17] that defines how law is taught and practiced in the United States. Part I.D then discusses the normative goals articulated for the LPE Project.

In all cases, the discussion in this part of this paper attempts to be merely descriptive, explaining the arguments made by the LPE Project founders in their own terms. I critique this these views in Part II.

A. *LPE and the Public-Private Law "Synthesis" that Excludes Consideration of "Political Economy"*

The public origins of the LPE project are documented in a series of blog posts from November 2017, on the LPE Project blog. The inaugural day of this blog includes four posts by the LPE Project's founders, with the following titles: *Law and Political Economy: Toward a Manifesto*;[18] *Why Law and Political Economy?*;[19] *Law & Neoliberalism*;[20] and *Why*

16 Britton-Purdy et al., *supra* note 8, at 1792.

17 *Id.* at 1790.

18 *LPE Manifesto*, *supra* note 9.

19 Jedediah Britton-Purdy & David Singh Grewal, *Why Law and Political Economy?*, LPE PROJECT BLOG (Nov. 6, 2017), https://lpeproject.org/blog/why-law-and-political-economy [hereinafter *Why Law and Political Economy?*].

20 Jedediah Britton-Purdy & David Singh Grewal, *Law & Neoliberalism*, LPE PROJECT BLOG (Nov. 6, 2017), https://lpeproject.org/blog/law-neoliberalism.

"Intellectual Property" Law?.[21] Collectively, they set the tone and subject matter of the project, discussing Thomas Piketty,[22] framing the need to rethink the law in the wake of the 2008 market collapse,[23] positing the role of the law in the rise of global inequality,[24] and discussing the role of "efficiency talk" in all of this.[25]

The most important of these blog posts is, unsurprisingly, the one that labels itself a step "toward a manifesto."[26] This post was the first one on the blog, is still highlighted on the LPE Project's web page,[27] and lays out the foundation for the LPE Project's founders' subsequent academic article, *Building a Law-and-Political-Economy Framework: Beyond the Twentieth-Century Synthesis*.[28]

This "manifesto" explains, for instance, that:

> The approach we call law and political economy is rooted in a commitment to a more egalitarian and democratic society. Scholars working in this vein are seeking to reconnect political conversations about the economic order with questions of dignity, belonging, or "recognition" and to challenge versions of "freedom" or "rights" that ignore or downplay social and economic power.[29]

More usefully, it explains LPE's focus on economic and political questions:

> Much of legal scholarship and practice in recent decades has held politics and economics apart, abstracting away from, or actively denying, their interdependence.
>
> Law schools and legal scholarship are divided along an implicit divide between "public" and "private" fields of law, which is constructed in significant part by the role that economics is thought to play in these respective fields. Many fields are thought of as being "about the economy" – contracts, torts, anti-trust, intellectual property, trade, consumer protection are examples. For the past several decades, scholarship in these fields has been dominated by law and economics approaches that have downplayed considerations of distribution and elevated questions of efficiency. This approach treats efficiency as a "neutral" value, yet construes the term in a manner that reproduces a constitutive priority for the privileged.
>
> Public-law scholarship, in turn, has tended to make questions of economy foreign. To learn and practice constitutional law today, for example, is often to assert that constitutional values have no purchase on questions of economy or class: these, after all,

---

21 Amy Kapczynski, *Why "Intellectual Property" Law?*, LPE PROJECT BLOG (Nov. 6, 2017), https://lpeproject.org/blog/why-intellectual-property-law [hereinafter *Why "Intellectual Property" Law?*].

22 *Why Law and Political Economy?*, *supra* note 19.

23 *Id.*

24 *Why "Intellectual Property" Law?*, *supra* note 21.

25 *Id.*

26 *LPE Manifesto*, *supra* note 9.

27 *Id.*

28 Britton-Purdy et al., *supra* note 8, at 1792.

29 *LPE Manifesto*, *supra* note 9.

> are the received lessons of Lochner and Carolene Products, of San Antonio and McRae. More generally, scholars in these public-law fields rarely devote themselves to the normative question of what kind of economic order might be necessary to make democracy real and vindicate constitutional principles such as equality.[30]

Both private and public law, in other words, are accused of having excised consideration of the sociopolitical impacts of their respective doctrines. And the effect of these fields' respective myopias taken together – what the authors refer to as "the twentieth-century synthesis"[31] – is alleged to be the law's complicity in what the authors call our "time of crisis":

> This is a time of crises. Inequality is accelerating, with gains concentrated at the top of the income and wealth distributions. This trend – interacting with deep racialized and gendered injustice – has had profound implications for our politics, and for the sense of agency, opportunity, and security of all but the narrowest sliver of the global elite. Technology has intensified the sense that we are both interconnected and divided, controlled and out of control. New ecological disasters unfold each day. The future of our planet is at stake: we are all at risk, yet unequally so. The rise of right-wing movements and autocrats around the world is threatening democratic institutions and political commitments to equality and openness. But new movements on the left are also emerging. They are challenging economic inequality, eroded democracy, the carceral state, and racism, sexism, and other forms of discrimination with a force that was unthinkable just a few years ago.[32]

The authors say that "Law is central to how these crises were created," and they draw from this their normative goal – and that of the LPE Project – that the law "be central to any reckoning with them."[33]

The authors develop these ideas further, most notably explaining their concept of "the twentieth-century synthesis" in *Building a Law-and-Political-Economy Framework: Beyond the Twentieth-Century Synthesis*.[34] The rest of this Part draws from that paper to explain the LPE Project's self-described predicates and normative goals, as presented in that paper.

B. *The Private Law Side of the "Synthesis"*

The LPE Founders articulate their view of "the twentieth-century synthesis" in *Building a Law-and-Political-Economy Framework: Beyond the Twentieth-Century Synthesis*.[35] The two parts of this synthesis are, on the one hand, the retreat of private law fields from consideration of the political effects of legal doctrine to focus instead solely on questions of efficiency,[36]

---

30 *Id.*

31 Britton-Purdy et al., *supra* note 8, at 1790.

32 *LPE Manifesto*, *supra* note 9.

33 *Id.*

34 Britton-Purdy et al., *supra* note 8, at 1790.

35 *Id.*

36 *Id.* at 1797.

and on the other, the exclusion of economic consideration from public law to focus instead on neutral principals irrespective of class, economic, or other characteristics or effects.[37] Their paper discusses these two sides of the synthesis in its parts I.A and I.C. This same delineation of focus between private and public law is followed in this and the next section of this paper.

Part I.A focuses squarely on the effects of the rise of law & economics on the private law.[38] It argues that law & economics "defined law's goals and methods in … what we might call 'market supremacy,' or the necessary subordination of the political to the economic."[39] It goes on to identify "[t]hree theories [that] are key to the market supremacy model of law and economics."[40] It first identifies the use of efficiency and wealth maximization to guide decision makers (e.g., judges and legislators).[41] It then pairs externalities and transaction costs together, as "bridg[ing] the core account of the market in neoclassical economics … with the traditional institutional focus of the law."[42]

Taking these in turn: *Beyond the Twentieth-Century Synthesis* asserts that the law & economics view holds that "the law should be oriented to ensure the greatest aggregate 'wealth,'"[43] giving as an example that a rich man should be able to take bread from a starving poor man because the rich man, having more money, has a greater "willingness to pay" for that bread.[44] It notes that few defend this as a normative theory, but that "law and economics almost invariably reverts to wealth maximization as a criterion"[45] out of a need to reduce costs, benefits, and transfers to "a common denominator like money."[46]

The discussion continues, noting that "[e]levating efficiency as a value also marginalized questions of distribution."[47] They explain that "the Coase Theorem *as commonly understood*" erases questions of distributional and endowment effects.[48] With respect to both wealth maximization and efficiency, it notes dismissively that the problems with them are "many and

37 *Id.* at 1806.

38 *Id.* at 1795 ("The first move of the Synthesis can be best understood by charting the rise in the 1970s and 1980s of modern law and economics . . .").

39 *Id.* at 1796.

40 *Id.*

41 *Id.*

42 *Id.* at 1798.

43 *Id.* at 1796.

44 *Id.* at 1797.

45 *Id.*

46 *Id.*

47 *Id.* at 1797.

48 *Id.* at 1797–98 (emphasis added).

by now quite well understood"[49] and that "sophisticated critiques … revealed the many problems"[50] of efficiency arguments.

Turning to the paper's discussion of externalities and transaction costs. After defining the concepts in single sentences, it offers the following characterization:

> Various features of institutional life could thus be homogenized into one capacious concept that also helped to naturalize market exchange: as long as transaction costs were low, Coase implied, it could be assumed that entitlements would naturally flow to their "best" or "highest value" use through voluntary exchange. Where they were high, planners could reengineer entitlements to approximate the most efficient (i.e., hypothetical market) outcome. This assumption epitomizes law and economics: it simultaneously recognizes and embraces the fact that law makes markets, while demanding that the satisfaction of markets becomes the aim of politics.[51]

Building from this characterization, *Beyond the Twentieth-Century Synthesis* offers that law and economics "was inevitably itself a form of planning,"[52] and that "[t]he role of scholarship was to identify transaction costs and externalities … and to recommend a shift in entitlements."[53] From this follows one of the more peculiar statements in the paper, which will be discussed in Part III, below, that:

> The agent of law reform imagined here [by law and economics] was not the people but the technician: the judge, economist, or bureaucrat who would calculate hypothetical consumer and producer surplus to order law and policy to serve the aims of wealth maximization.[54]

C. *The Public Law Side of the "Synthesis"*

Turning to the public law side of the LPE Project's twentieth-century synthesis, Part I.C of *Beyond the Twentieth-Century Synthesis* argues that "in areas regarded as 'essentially about' the liberty and equality of citizens, the last half-century has seen withdrawal from questions of economic distribution and structural coercion."[55] Unlike the private law-side of the synthesis, whose "genealogy is essentially one of economics-informed legal theory," "[i]n the public-law half of the Synthesis, the situation is very

---

49 *Id.* at 1796–97.

50 *Id.* at 1798. Authors cite to Richard S. Markovits, *Why Kaplow and Shavell's "Double-Distortion" Articles Are Wrong*, 13 GEO. MASON L. REV. 511, 519–523 (2004), and Zachary Liscow, Note, *Reducing Inequality on the Cheap: When Legal Rule Design Should Incorporate Equity as Well as Efficiency*, 123 Yale L.J. 2478, 2478–2483 (2014) to show other critiques.

51 *Id.* at 1799.

52 *Id.*

53 *Id.* at 1799–1800.

54 *Id.* at 1800.

55 *Id.* at 1806.

different"[56], one that took shape "around a particularly thin version of key liberal values: freedom, equality, and state neutrality" and "the decisions of increasingly conservative judges drove the change."[57]

As with the private law discussion, here, too, the LPE founders divide their discussion into three main subjects: "exil[ing] matters of class and material, structural inequality from the reach of constitutional law," expansion of First Amendment speech protections to include certain economic concepts, and an embrace of public choice.[58] The first two of these focus on the development of public law – primarily 14th and 1st Amendment jurisprudence – over the second half of the 20th century. The third is more driven by theory and principle.

The arguments rehash well-trodden territory, with which many agree and other disagree.[59] Starting with the 14th Amendment, the arguments track longstanding debates over the whether the amendment guarantees positive or negative rights. Across a range of categories during the 1960s and 1970s the Supreme Court declined to heighten legal protections, benefits, or rights to particular classes of citizens, specifically the poor, racial minorities, and women.[60] In other words, the Court embraced a negative rights view of the 14th Amendment, that it is protection from government discrimination, as opposed to a positive rights view under which the Government is to work to ensure equality of outcomes. The effect, in the LPE view, was that "[j]ust when the achievement of formal equality meant that the major threats to an egalitarian society lay in structural inequality, the Court approved policies that compounded inherited forms of inequality."[61] The authors lament the Court's embrace of "[a] conception of equality that ignored material deprivation and focused on improper intent," that "encased the most pressing sources of inequality from constitutional review, *even when they were reproduced and amplified by state action*, and went so far as to invalidate policies that sought to mitigate structural inequality by taking explicit account of characteristics such as race."[62]

Beyond the Twentieth-Century Synthesis offers a similar account of 1st Amendment jurisprudence, lamenting "the merging of First Amendment

---

56 *Id.*

57 *Id.* at 1806–07.

58 *Id.* at 1807.

59 *See* RANDY E. BARNETT AND EVAN D. BERNICK, THE ORIGINAL MEANING OF THE FOURTEENTH AMENDMENT: IT'S LETTER AND SPIRIT (2021); ILAN WURMAN, THE SECOND FOUNDING: AN INTRODUCTION TO THE FOURTEENTH AMENDMENT (2020); ERIC FONER, THE SECOND FOUNDING: HOW THE CIVIL WAR AND RECONSTRUCTION REMADE THE CONSTITUTION (2019). *Compare* Thurgood Marshall, *The Continuing Challenge of the 14th Amendment,* 1968 Wis. L. Rev. 979 (1968) (taking a positive view of the 14th Amendment) *with* Gene R. Nichol Jr., *Examination of Congressional Powers under §5 of the 14th Amendment*, 52 Notre Dame Law. 175 (1976) (taking a negative view of the 14th Amendment).

60 Britton-Purdy et al., *supra* note 8, at 1807-08.

61 *Id*. at 1808.

62 *Id.* at 1809.

speech with commerce."[63] The focus on increased protection of commercial speech and, hand in hand with that, the use of money in funding and paying for speech, is important for the LPE Project's synthesis. This is a means by which these opinions "exacerbate an increasingly oligarchic political economy" and "economic power from political disruption."[64] That is, where the Court's 14th Amendment jurisprudence prevents state actors from remedying inequality, the Court's 1st Amendment jurisprudence is said to allow private actors to ossify that inequality.

The third is more driven by theory and principle. The final plank in the LPE founders' explanation for the public-law side of the "synthesis" moves its focus from the Courts to the role of public choice in shaping academic and administrative attitudes towards public policy. It notes "a growing public-law skepticism toward political judgments about distribution and economic ordering, based on the conviction that these judgments are likely to enforce and entrench the kinds of 'capture' that James Buchanan's 'political economy' emphasized."[65] In the paper's telling, the growth of public-choice ideas led "influential academics to argue that the only appropriate response was a move to market-mediated technocracy, in the form of cost-benefit analysis."[66]

The collective result of these three phenomena — jurisprudential shifts in the 14th and 1st Amendments and the rise in public choice — was, according to the LPE account, a shift in the legal academy in the 1980s and 1990s. The means by which legal academics traditionally seek to exert influence — shaping the opinions and decisions of judges and policy makers — had closed themselves to efforts to directly influence political outcomes. "[D]ebates in mainstream legal scholarship [therefore] migrated to make questions of political economy hard to ask because they were seemingly already settled both theoretically and practically."[67]

D. *LPE's Normative Prescriptions*

Beyond the Twentieth-Century Synthesis offers a unique telling of the development of the law during the latter part of the twentieth century. The details of this telling will be critiqued in Part III. But what is the punchline? What are the LPE project's normative prescriptions?

This question stands as important on its own. But it is particularly important to the extent that LPE is a successor to the earlier critical legal

63 *Id.*

64 *Id.* at 1810.

65 Britton-Purdy et al., *supra* note 8, at 1811 (citing Jedediah Purdy, *Beyond the Bosses' Constitution: The First Amendment and Class Entrenchment*, 118 COLUM. L. REV. 2161, 2172–75 (2018)).

66 *Id.* at 1811–12 (citing Cass R. Sunstein, *The Cost-Benefit State* 4 (Coase-Sander Inst. for L. & Econ., Working Paper No. 39, 1996)).

67 *Id.* at 1812.

studies movement. One of the most trenchant critiques of critical legal students – along with many of its cognate critical fields – was its lack of compelling normative prescriptions for how the law ought to operate differently.[68]

The authors of the *Twentieth-Century Synthesis* share an understanding of the need to center LPE around a call to action – though this understanding does not emanate as a response to the question that challenged earlier critical movements. Rather, the authors frame LPE's normative commitments in a more urgent, if less analytically-grounded, need to respond to their perceived problems identified in the *Twentieth-Century Synthesis*. They write:

> We must replace the Twentieth-Century Synthesis with a different framework for legal thought. At the core of this reconstruction is a renewed commitment to questions of political economy. With others, we have thus begun to practice a scholarship of 'law and political economy' (LPE), rooted in a shared set of insights, concerns, and commitments.[69]

This is a rhetorically interesting move. Unlike earlier critical movements, LPE has chosen an affirmative name for itself. It is not merely about criticizing existing legal norms – or, though critical of the asserted twentieth century synthesis opposing it – its name is the mantle that it wears. The alternative it proposes is that of "law and political economy." *Beyond the Twentieth-Century Synthesis* offers the following explanation of what it means by "law and political economy":

> The term 'political economy' is historically variable and contested. We do not mean the 'political economy' analysis of institutions and policies as practiced in mainstream economics departments, which turns on the application of rational-choice models to government actors or institutions. Rather, we intend the older and more foundational usage familiar to nineteenth-century audiences, which persisted in traditions of 'radical' political economy until a few decades ago. *This* political economy investigates the relation of politics to the economy, understanding that the economy is always already political in both its origins and its consequences.[70]

Of course, if we want to speak of the "older and more foundational usage" of political economy "familiar to nineteenth-century audiences," we are talking of the political economy of Adam Smith. Smith's understanding of political economy will be discussed in Part IV.A. It seems unlikely that the LPE project intends its founding document as an appeal for greater study to *The Wealth of Nations* or *The Theory of Moral Sentiments* – though, as I argue in Part IV generally I expect there is much common ground to be found between critical theorists, LPE, and law & economic scholars. Perusing the results of a Google N-Grams search of "political economy" through the 19th

68 *See, e.g.*, Michael Fischl, *The Question that Killed Critical Legal Studies*, 17 LAW & SOC. INQUIRY 779 (1992).

69 Britton-Purdy et al., *supra* note 8, at 1791–92.

70 *Id.* at 1792.

century yields results such as David Ricardo, John Baptiste Say, and Stanley Jevons.[71] Of course, there were other more "radical" political economists in the 19th Century (certainly Karl Marx comes to mind). But mainstream political economists in the 19th century were the successors to Adam Smith and forebears to the modern "economics" that the LPE project scorns. It is curious to wear "political economy" as a mantle when the key idea the LPE means to invoke from its description of political derives from the adjective "'radical" prepended to it. Radical 19th century views may have power today – but no help comes from coopting a mainstream term as a talisman.

Perhaps more important, the passage quoted above does little to explain the *Twentieth-Century Synthesis* normative commitments, other perhaps than being 'radical' in its thought. The authors do suggest that it was important to "Legal Realists, in their battle against laissez-faire ideology," it "requires attentiveness to the ways in which economic and political power are inextricably intertwined with racialized and gendered inequity and subordination."[72]

The authors promise to "explore these questions in Part II" of the *Beyond the Twentieth-Century Synthesis*. But this exploration only raises more questions. Critique of the LPE characterization of the private-law and public-law sides of its "synthesis" is undertaken in Part III. The questions of interest to LPE, and to the extent they suggest a normative position that LPE takes in its analysis, LPE's normative position as presented in *Beyond the Twentieth-Century Synthesis* is characterized below.

In Part II.A, "From Efficiency to Power," *Beyond the Twentieth-Century Synthesis* asks: "What would it mean to take power once more as a central unit of analysis in law? In the broadest sense, when we teach a canonical case or encounter a legal problem, we might ask quite simply, who has power here, who should have power, and why?"[73] This question is presented in opposition to the assertion that

> The Twentieth-Century Synthesis held that such power was unimportant, either by redirecting attention from it or by denying that power was stratified or structured in ways that matter. By refocusing scholarship on questions structured by transaction costs and externalities, law-and-economics analysis placed questions of distribution and coercion outside the lamplight of methodology. It thus neglected the actual social world comprised of highly disparate resource allocations that are themselves products of background legal rules.[74]

My best understanding of this discussion as a normative prescription is that it is a rejection of ideas such as placing an initial assignment of liability

---

[71] Usage of "Political Economy" from 1800 to 1900, GOOGLE NGRAM VIEWER, https://books.google.com/ngrams/graph?content=political+economy&year_start=1800&year_end=1900 & corpus=26&smoothing=3 (last visited Sep. 27, 2022).

[72] Britton-Purdy et al., *supra* note 8, at 1792.

[73] *Id*. at 1820.

[74] *Id.* at 1819.

on the least-cost avoider in torts cases and, instead assigning liability to whatever party is in the best position to bear the cost of an accident, regardless of their fault in causing the accident or ability to avoid it.

In Part II.B, "From Neutrality to Equality," *Beyond the Twentieth-Century Synthesis* asserts that the Twentieth-Century synthesis's focus on neutrality "has reinforced a very non-neutral drift toward elite control of government."[75] In response, it argues that we should "develop a normative theory of bargaining that centers a substantive ideal of freedom, rather than relying upon the formal idea of uncoerced agreement,"[76] that "moves from the setting of interpersonal bargaining to structural questions,"[77] and that has "a concrete focus on what kinds of equality we want law to generate."[78]

And in Part II.C, "From Antipolitics to Democracy," *Beyond the Twentieth-Century Synthesis* turns to public choice economics, arguing that "in the Twentieth-Century Synthesis, governance – even in its more progressive, reform-oriented aspects – came to be viewed as an antipolitical operation." This followed from public choice ideas, "which portrayed government as inherently prone to capture and corruption."[79] This is not an unfair characterization. As discussed in Part III.B, below, public choice economics raises trenchant descriptive criticisms of the fundamental operation of government that gives pause – and should give pause – to the idea that it can, in the general case, be used to support the public interest (or, indeed, that there even is a thing that can be identified as "the public interest").

Ultimately, I believe that LPE's real bone to pick is with public choice economics – and the most fundamental critique of LPE is based in its failure to engage with public choice economics. As noted in *Beyond the Twentieth-Century Synthesis,* "a comprehensive answer is beyond the scope" of the article. The authors do "identify a number of critical questions and already-emerging frontiers of debate"[80]: the importance of "strengthening existing institutions of electoral democracy"[81]; that the "law should shape the economy to support the institutions and capacities that uphold the equality and efficacy of democratic citizens"[82]; and that "scholarship should consider what moral images of social and political order are implied in a given legal patterning."[83]

---

75 *Id*. at 1824.

76 *Id*.

77 *Id*. at 1825.

78 Usage of "Political Economy" from 1800 to 1900, *supra* note 71.

79 Britton-Purdy et al., *supra* note 8, at 1828.

80 *Id*. at 1829.

81 *Id*. Britton-Purdy explains this to mean "a direct challenge to the ways that antipolitics constrains even the most seamlessly majoritarian of politics", through "commitment to voting rights, overcoming gerrymandering, defending campaign-finance laws, and ultimately challenging the antimajoritarian features of the American constitutional scheme, notably the Electoral College and the Senate." *Id*.

82 *Id*. at 1830.

83 *Id*. at 1832.

These are all central themes in modern (and not-so-modern) public choice scholarship. The general thrust of LPE's normative prescription is: *the world is unjust and we should use government to do something about it*. But, as discussion in Part III.B, public choice explains, in descriptive, not normative, terms why using the government to do something about it is fraught. Public choice is not antipolitic, it is realpolitik. The LPE's conflation of public choice's descriptive analysis of politics as a normative prescription for a politics ultimately renders LPE's normative prescriptions as merely oppositional. They amount to a call to use the machinery of government to impose a preferred set of policies upon the world, without engaging with the risks of such an approach – namely that the policies imposed upon the world may be those preferred by someone else, someone antagonistic to one's own preferred policies.

In other words, LPE's normative call is that "lawyers of the world unite" – but this call forgets that the "first thing those with power do is kill all the lawyers."

II. A REBUKE OF LPE'S CLAIMED "SYNTHESIS"

I have tried above to offer a descriptive account of the LPE "synthesis" that describes its understanding of the status quo against which the field opposingly defines itself. This understanding is seriously flawed. It is incomplete and inaccurate in parts, at times borders on incoherent, and is fundamentally self-contradictory. The discussion below develops this critique of LPE and the "synthesis" upon which it is founded.

While this discussion is critical of LPE and its foundations, there is an uncomfortable truth that needs to be kept front-and-center: While the LPE synthesis's critiques do not stand informed scrutiny, they nonetheless represent widespread understandings. This is most true of how most legal thinkers understand law and economics and its priors: law and economics scholars would generally reject the characterization of law and economics as myopically focused on efficiency and wealthy maximization while having a callous disregard for distributional questions. To the contrary, in my own experience, most law and economics scholars care and think deeply about distributional questions – a topic taken up further in *infra* Part II.A.

Most legal thinkers, be they academics, judges, or policymakers, are not law and economics scholars – and to a very large extent the LPE account reflects the widespread understanding of the descriptive priors and normative goals that law and economics brings to bear on legal analysis. This presents a twofold danger to law and economics. First, it robs the field of its sophistication and ongoing value as a tool for meaningful legal analysis. It makes the field easy to dismiss, unattractive to up and coming researchers, and threatens the contributions that the field has made previously. And second, as we see with LPE, it gives energy to critical fields that would define themselves in opposition to a false but dislikable straw man.

A. *LPE Gets L&E Wrong*

The LPE account of law and economics, which it places at the heart of contemporary private law jurisprudence, is at once commonplace and bizarrely unrecognizable. It is commonplace in characterizing law and economics as focused on efficiency and wealth maximization while putting to the side distributional effects. I will return to what it gets wrong on this front in a moment. But we should focus first on where the LPE account of law and economics is unrecognizable bizarre. *Beyond the Twentieth-Century Synthesis* asserts that

> The agent of law reform imagined here [by law and economics] was not the people but the technician: the judge, economist, or bureaucrat who would calculate hypothetical consumer and producer surplus to order law and policy to serve the aims of wealth maximization.[84]

Anyone who has read Coase's *The Problem of Social Costs* will be simply dumbfounded by this assertion. In *The Problem of Social Cost*,[85] Coase's central subject was the implausibility of the Pigouvian model of using taxes and subsidies to address social costs such as externalities. The Pigouvian model truly was that of the technician: bureaucrats calculating the delta between private benefits and costs and social costs and benefits and using taxes or subsidies to bring private conduct into alignment with the socially optimal level of conduct.

Puzzling over this Pigouvian model of technocratic social design, Coase worries that it "would require a detailed knowledge of individual preferences and I am unable to imagine how the data needed for such a taxation system could be assembled."[86] The Pigouvian model, he concludes, suffers "from basic defects in the current approach to problems of welfare economics."[87] Chief among these defects is that there is no reason to believe that those attempting to calculate such a system of taxes and subsidies know the socially optimal level of any activity – without which, they cannot calculate the social costs of private activity that need to be internalized. Hence, of course, the title of Coase's paper, *The Problem of Social Cost*.

Coase's paper is the ur-text of law and economics. Its entire point is that the model of technicians calculating hypothetical consumer and producer surpluses is a fraught endeavor – and that rather that rely on that model, a better approach is to remove technicians from how we order private conduct and instead to rely on allocating burdens and designing institutions to minimize transaction costs.

---

84 Britton-Purdy et al., *supra* note 8, at 1800.

85 R.H. Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1 (1960).

86 *Id*. at 42.

87 *Id*.

After Coase, perhaps the greatest popularizers of law and economics as a legal methodology are judges Richard Posner and Guido Calabresi (along with his then-student Doug Melamed). Looking at their contributions to the field bring us back to the first point that the LPE characterization gets wrong: its over-emphasis of efficiency as a normative goal.

Judge Posner is perhaps the single greatest figure in the development of law and economics as a field. His book *Economic Analysis of Law*[88] is a tour de force, applying principles of microeconomic analysis to nearly every area of the law. But in reading Posner's work, it is essential to distinguish between his positive and normative claims. This is perhaps confounding to some, because the claims are seemingly similar: descriptively, Posner argues that the common law, in fact, tends to be efficient; normatively, he argues that it ought to be efficient.

It is important to separate these claims for two reasons. First, the success of law and economics as a field, and the impact that it has had on the law, result principally from Posner's positive analysis. The majesty of Posner's *Economic Analysis of the Law* is that he forcefully demonstrates that most judges, in most cases, decide cases in ways that result in efficient outcomes (generally by placing burdens on the least cost avoider). Thus, and entirely independent of any effects that law and economics may have had on legal analysis, the best way for lawyers to practice law – to advise clients, predict judicial outcomes, and likely to influence those outcomes – is to approach legal questions from a microeconomic perspective.

This foundational aspect of law and economics is wholly absent from the LPE synthesis. Under the LPE telling, efficiency is the purpose, not the result, of economic analysis.

Of course, for many, law and economics also carries efficiency as its normative goal (and sometimes conflating the idea of efficiency with that of wealth maximization). Here it is notable that both Posner and Calabresi (who popularized the idea of allocating the burden of avoiding harm on the least cost avoider), among the most important advocates for efficiency as normative goal, are judges, responsible for deciding cases. The LPE critique would have us think of them as would-be technicians: Pigouvian bureaucrats calculating social costs and designing legal rules to maximize total welfare. Those familiar with their work, however, know that nothing could be further from the truth. Rather than Pigouvian, they are Coasean. One of the virtues of the Coasean approach is that it minimizes the role of the decisionmaker as technician. The question for the judge is *who is in the best place to have avoided this harm* – not *what are the private and social costs and benefits as between those involved in this incident and how do I decide this case in order to achieve a socially optimal division of those costs and benefits*.

A brief note is due here regarding the role of distribution in economic analysis of the law. The LPE critique notes, correctly, that law and economics

88 RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW (9th ed. 2014).

tends to discount considerations of distributive efficiency, focusing instead on maximizing allocative efficiency and relying on expressly distributive mechanisms such as taxes to address any distributional inequities.[89] In a survey of major law and economics textbooks, all but one devote substantial attention early on to these issues.[90] Interestingly, the same cannot be said for casebooks commonly used to teach private law subjects to first year law students – an issue that undoubtedly drives much misunderstanding about the normative values of law and economics held by faculty and students alike.

There is a deeper irony in the misunderstandings of law and economics presented in the LPE synthesis, which will carry us into the discussion of the LPE characterization of public law. LPE views law and economics and private law as shaped by law and economics as antidemocratic – a reformation of the law shaped not by the people but by technocratic judges, economists, and bureaucrats.[91] Not only is this characterization wrong – the driving value of law and economics rather being the exact opposite – but the very purpose of the LPE Project is to place legal academics squarely in the role of the accused technocrats.

B. *LPE Gets Political Economy and Public Choice Tragically Wrong.*

The LPE critique of public law – and of public choice, in particular – is circuitous, circular, and self-contradictory. If anything, it demonstrates the importance of public choice generally, and presents a compelling public choice critique of the LPE project.

As discussed in Part I.C, the *Twentieth-Century Synthesis* identifies three "moves" in its discussion of public law institutions. The first two focus on changing First and Fourteenth Amendment jurisprudence over the second half of the of the twentieth century, facilitated (in the LPE telling) by conservative judges that took disproportionate control over the federal judiciary in that period.[92] The third move was the "growing public-law skepticism toward political judgments about distribution and economic ordering" based in ideas originating from public choice scholars such as James Buchanan – ideas such as rent extraction, political capture, and interest group theory.[93]

It is helpful to start with public choice in considering this nexus of arguments. But what is "public choice" theory? Traditional economics

89 Britton-Purdy et al., *supra* note 8, at 1798 ("[L]egal entitlements should always be designed to maximize efficiency, shunting issues of distribution elsewhere (commonly, to the tax code)").

90 MAXWELL STEARNS & TODD J. ZYWICKI, PUBLIC CHOICE CONCEPTS AND APPLICATIONS IN LAW (2009). Perhaps surprisingly, the one book that I found that does not address distributional issues expressly is Stearns & Zywicki – both of whom I know understand these concerns very well. I would posit that they do not address the issue squarely because it is so obvious an issue to them.

91 Britton-Purdy, *supra* note 8.

92 *Id.* at 1807.

93 *Id.* at 1809.

focuses on how private actors make economic decisions – that is, it is the study of "private choice." Public choice is the field of economics that studies how public institution, such as governments, regulators, and other bureaucratic entities, make decisions. Many public choice scholars, in fact, are thought of, or think of themselves, as belonging to the field of political economy.[94] Unlike law and economics, a field with which most scholars are familiar with at least by name, public choice economics is far less well known and has not been as well received by legal scholars – even if many of its ideas have been influential. As the noted in the *Twentieth-Century Synthesis*, "[m]any legal scholars objected to this turn" towards public choice ideas in the 1980s and 1990s – a confounding factor for LPE's assertion that this field played an assertedly formative role in any grand synthesis of twentieth-century legal thought.[95]

What are the core ideas of public choice economics? James Buchanan famously described the field as "politics without romance."[96] The field uses traditional tools of economics to understand how the individual actors that make up public institutions respond to incentives, and how those institutions themselves operate. This strips away the "romance" of politics – the naïve and idealistic understanding of government and the individuals that make up government institutions as purely beneficent, public-minded, actors unaffected by the incentives that affect private actors.

There are at least three core ideas strongly associated with public choice. This first is Arrow's "impossibility theorem," which challenges the very possibility of voting mechanisms as a tool for making collective decisions.[97] Arrow was not the first to show this result – that every voting system is subject to manipulation and that no voting system can guarantee democratically-representative outcomes – but he was the first to show it in a mathematically formalized proof. A second core idea of public choice is that legislatures are made up of individual actors, who respond to individual incentives. Those individuals work to maximize various functions that are separate from, and often conflict with the interests of, those whom they purportedly represent. For instance, public officials often work to maximize their resources, autonomy, or power – and to use those to maintain their public position, not for the benefit of the individuals who put them in that position.[98] And perhaps the best well known of public choice's core ideas is that of interest group theory – roughly the idea that smaller interest groups will generally be better able to lobby and "capture" the favor of government

94 As the LPE founders awkwardly acknowledge in their reference to James Buchanan's "political economy," Nobel Laureate James Buchanan's work is considered among the foundation of public choice scholarship, though he framed his work as political economy.

95 Britton-Purdy et al., *supra* note 8, at 1812.

96 James M. Buchanan, *What is Public Choice Theory?* 32 IMPRIMIS 1, 5 (Mar. 2003).

97 KENNETH ARROW, SOCIAL CHOICE AND INDIVIDUAL VALUES 60 (Eric Maskin ed., Yale Univ. Press 2012).

98 James Q. Wilson, *The Rise of the Bureaucratic State*, 41 NAT'L INT. 77, 81 (Fall 1975).

actors than more diffuse groups, and especially than the most diffuse "interest group" of general voters.[99]

Collectively, these ideas challenge the very idea of a government and of regulation that work for the public interest. They suggest that there are no reliable mechanisms to measure the "public interest" in the first instance. They caution that government institutions will be made of up individuals who use those institutions for their own purposes – even if some individuals are drawn to government and public service out of a commitment to public service, those drawn to public institutions for private gain will tend to acquire greater resources and over time have greater electoral success. And public choice tells us that private interests will generally be able to petition and capture public institutions for their private gain.

As Buchanan says, this is not a romantic view of politics.[100] The power of public choice is that it explains in descriptive terms why this is also an inevitable truth of politics. It needs to be emphasized that this first stage in the development of public choice is a descriptive effort to understand how public institutions work – and it has widespread acceptance. But people like their romance and have an attachment to the idea of government. Public choice finds its critics in its normative prescriptions, which generally call for limitations on the scope of government. If capture and rent extraction are inherent to the public enterprise, the only way to protect against them is the limit to scope of the public enterprise. Either limit it to functions least susceptible to those abuses or keep government small enough that its operation will be difficult to shield from public scrutiny.

The balance of the public-law critique offered in the *Twentieth-Century Synthesis* vindicates the public choice view of politics and public law institutions. The basic argument offered by the authors of the *Twentieth-Century Synthesis* boils down to "conservatives captured public institutions to shape the law in ways that we don't like." This may be an uncharitable characterization – but it is an uncharitable characterization of an uncharitable characterization.[101] Truth be told, many L&E scholars are sympathetic to

---

99 *See* MANCUR OLSON, THE LOGIC OF COLLECTIVE ACTION 53 (1971); Gary Becker, *Competition Among Pressure Groups*, 98 Q. J. OF ECON., 371, 385–394 (1983).

100 Buchanan, *supra* note 96, at 5.

101 For a thoughtful discussion of modern interlocutors' tendency to approach one another uncharitably in debates around public law, see Emad H. Atiq & Jud Mathews recent article, *The Uncertain Foundations of Public Law Theory*, CORNELL J. L. & PUB. POL'Y, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4032904 (forthcoming) (noting "that successful truth-seeking enterprises, like the natural sciences, tend to withhold judgment on matters that are contested among good-faith reasoners; and, moreover, that there is considerable ambivalence among contemporary philosophers of law, who by all accounts are reasoning in good faith, towards the standard views in The rejection by the authors of the *Twentieth-Century Synthesis* of good-faith accounts of changing 1st & 14th Amendment views ignores the substantive disagreements that define these contingent debates. Significant numbers of Americans are more concerned about the power of government and preservation of their individual liberty than the power of the government to intervene in the lives of individuals to assert the

many of the issues that animate the LPE project, including many not the focus of the *Twentieth-Century Synthesis*. This proposition is explored further in Part III. It is not unusual for public choice scholars to lament the seemingly inescapable nihilism that public choice has for the political process. But that results from an understanding that the political process is likely not an effective way to achieve those goals, not from a lack of belief in the importance of those goals. Indeed, this seeming nihilism comes from an understanding that any effort to use the political process to achieve these goals would likely only make them worse as the political machinery is captured and turned to the benefit of narrow interests.

It is telling that the account of public law offered in *The Twentieth-Century Synthesis* seems to understand this challenge. In one of the more revealing passages that comes just after the account of changes in First Amendment jurisprudence over the second half of the twentieth century, the authors note that

> in keeping with the law-and-political-economy premise that state action and economic power are always mutually intertwined, it is key to appreciate that the result of these decisions is not to segregate state power from economic power but to exacerbate an increasingly oligarchic political economy in which private power is readily translated into influence over public decisions.[102]

The LPE project, in other words, appears to import the lessons of public choice into its very foundations, identifying the "idea" that "state action and economic power are always mutually intertwined" as a premise of LPE.[103] This is literally just a restatement of the public choice premise that public actors are influenced by private incentives. And the consequence identified by the authors, that "private power is readily translated into influence over public decisions" may as well be torn from the pages of a public choice paper.[104]

The lesson that LPE takes, however, is exactly the opposite of that taken by public choice scholars. Where public choice scholars see public institutions as dangerous tools readily captured by private interests, LPE sees them as opportunities that can be captured to serve their own interests. Perhaps they are right. Perhaps law devolves to mere power. Perhaps over the course of the twentieth century sources of conservative power were able to capture public institutions to impose their preferred policies on society.

---

rights of others – this, indeed, is a fundamental source of disagreement within the American polity. As Atiq & Mathews argue, "If practical conclusions must be drawn on the basis of the contested proposition, then it must be with explicit recognition that the justification for one's practical choice is reasonably contestable. Doubt of this kind makes a difference to how scholars and judges should argue with one another," yet such an approach is seemingly overtly eschewed by the *Twentieth-Century Synthesis*. *Id*. at 55.

102 Britton-Purdy et al., *supra* note 8, at 1810–11.

103 *Id.* at 1810.

104 *Id.* at 1811.

And perhaps now the best approach for those who disagree with those policies is to wrest back power over those public institutions to use them to impose a different set of policies preferred by a different set of oligarchs.

I do not believe this is the right lesson to learn – I do not welcome the return to the pre-Hobbesian war of all against all and believe that there is a role for public institutions in providing a baseline of public service and moderating the discourse within and policies adopted by a society.[105] I will return to these ideas in Part III to argue that part of a better approach is to forego the war of all against all – or at least of LPE against L&E – and consider the possibility of a synthesis between these competing methodological camps.

III. A MORE SYMPATHETIC SYNTHESIS

The discussion so far has largely focused on the LPE Project's understanding of law and economics, both in the private law setting and in the public law setting in the form of public choice. It would not be overdramatic to say that a goal of the "Law and Political Economy" project is to erase "law and economics" from the law. The methodological constraints imposed by economic analysis and counter-indicated by its conclusions are anathema to LPE's own normative commitments.

But while the discussion above is critical of the LPE movement, my purpose in this paper is not only to criticize but also to offer a different synthesis, one between LPE and L&E. Putting its attacks on the role of law and economics to the side, one would be too hasty to dismiss LPE as a whole, and especially to dismiss the substantive concerns animating its adherents. Indeed, many of the concerns raised by its adherents are familiar to scholars of other stripes, including many law and economics scholars.

LPE's founders are open about their desire to erase economics from the law, sharing expressly their view that "We must replace the [law-and-economics based] Twentieth-Century Synthesis with a different framework for legal thought."[106] They continue, saying that "at the core of this reconstruction is a renewed commitment to questions of political economy."[107] In explaining what this means, they explain:

> The term "political economy" is historically variable and contested. We do not mean the "political economy" analysis of institutions and policies as practiced in mainstream economics departments, which turns on the application of rational-choice models to government actors or institutions. Rather, we intend the older and more foundational usage familiar to nineteenth-century audiences, which persisted in traditions of "radical" political economy until a few decades ago. This political economy investigates the

105 THOMAS HOBBES, LEVIATHAN 519 (A.R. Waller ed., Cambridge Univ. Press 1904) (1651).

106 Britton-Purdy et al., *supra* note 8, at 1791-92.

107 *Id*. at 1792.

> relation of politics to the economy, understanding that the economy is always already political in both its origins and its consequences.[108]

The discussion in this part turns that reconstruction around, arguing that any study of political economy is necessarily anchored in the values of economic thinking. To do so, it looks at three thinkers who assuredly must be considered political economists (unless the LPE movement means to erase or redefine both law & economics *and* political economy): Adam Smith, the original political economist; Elinor Ostrom, one of the leading political economists of the 20th century; and Derrick Bell. Not only are these three individuals unquestionably engaged in political economy, but their work inhabits a place of interest, both substantive and methodological, to both L&E and LPE.

A. *A Different Synthesis*

But I start by echoing a more recent appeal for synthesis between these fields, quoting an extended passage from a 2003 article by Professors Devon Carbado and Mitu Gulati.[109] In this passage, they discuss the relationship between law and economics and critical race theory – though it generalizes to critical legal studies and LPE.

Legal academics often perceive law and economics (L&E) and critical race theory (CRT) as oppositional discourses. Part of this has to do with the currency of the following caricatures:

> L&E is the methodological means by which conservative law professors advance their ideological interests. The approach is status-quo-oriented and indifferent (if not hostile) to the concerns of people of color and the poor. . . . The political effect of L&E is to entrench and obfuscate racial and class hierarchies.
>
> CRT is the methodological means by which radical faculty of color (and especially black faculty) advance their ideological interests. The approach is invested in finding discrimination and characterizes even the most progressive institutional practices as racist. . . . The political effect of CRT is preferential treatment and social welfare programs for people of color — particularly black people.
>
> It would be neither difficult nor interesting to disprove either caricature. Yet both have considerable intellectual and institutional purchase, so much so that they have helped to balkanize L&E and CRT scholarship. . . .
>
> Both sides are at fault. . . .

---

108 *Id.*

109 Devon W. Carbado & Mitu Gulati, *The Law and Economics of Critical Race Theory*, 112 YALE L.J. 1757 (2003).

> A CRT/L&E engagement helps to cure some of the deficiencies in both fields. For example, CRT's notion of race as a social construction can help L&E scholars move to a dynamic conception of race, and L&E's focus on the incentive effects of legal and institutional (norm-based) constraints can help CRT scholars analyze the ways in which the pressures and constraints of the workplace shape both employer and employee behavior. In short, a CRT/L&E joint venture could advance our thinking about how, in the shadow of law, workplace structures and norms affect racial identity — and vice versa.
>
> The argument for a collaboration between economics and CRT (and feminist theory and gay and lesbian legal studies) was made with force in a 1996 essay by Ed Rubin. Rubin argued that the common critical approach to institutional analysis shared by L&E and CRT — both fields reject claims about the neutrality and objectivity of legal rules, albeit for different reasons — would, if combined, produce [] an exciting new methodology for legal inquiry [that could serve as a] unifying discourse in legal academia.[110]

I could not agree more with Carbado and Gulati. It is true that both L&E and LPE (as well as other critical fields) rarely understand and readily misrepresent each other. Yet both are complementary, as sources of questions and analysis. And both have uniquely sharp foci on institutions that make them powerful and relevant in ways that most other fields of legal analysis are not.

If either field is to demonstrate its legitimacy, it must move beyond base characterizations such as offered by LPE in *Beyond the Twentieth-Century Synthesis* and acknowledge the contributions of the other.[111] This is particularly the case given the extent to which the fields actually do concern themselves with similar topics. As discussed previously, both fields share substantial concern with how private interests capture public institutions.[112] This and other overlapping themes are explored further in the discussion below.

I would hasten to note here that while I reject LPE's characterization of L&E as "big C" Conservative – or, indeed, as having any necessary political valence – one must also acknowledge that many conservatives are drawn to law and economics largely because it does tend to confirm their political views. On the other hand, the fields' critics often reject the field and its methodologies for the converse reason: that they believe it is antagonistic to their political views. The materials below draw from the work of three scholars with whom I expect most political proponents of law and economics are not more than passingly familiar (and two of whom for which the same can likely be said of proponents of LPE). That is an unfortunate state of affairs for both camps.

---

110 Carbado & Gulati, *supra* note 109, at 1757–58, 1761 (emphasis in original) (citing to Edward L. Rubin, *The New Legal Process, the Synthesis of Discourse, and the Mircroanalysis of Institutions*, 109 HARV. L. REV. 1393 (1996)).

111 *See* Atiq & Mathews, supra note 101 and accompanying text.

112 *Supra*, Parts I.C and II.B.

B. *Adam Smith*

The way in which LPE invokes "political economy" above should carry a sense of humor to many law and economics scholars. Characterizing their meaning as drawing from the "older and more foundational usage familiar to nineteenth-century audiences"[113] gives their usage an air of mystery, as though the secret books of political economy hold truths that would lead any adherent of the twentieth-century synthesis to rethink their fundamental beliefs. This is humorous, of course, because the founding father of "political economy" was Adam Smith, whose ideas are in no way foreign to the law and economics scholars from whose ideas the LPE so clearly seeks to distance itself.

The discussion below considers some of Smith's core ideas, as are relevant to the intersection of political economy and law and economics. And without doubt, many of Smith's concerns sound familiar to today's readers. Consider, nearly 250 years ago, in Book 5 of *An Inquiry into the Nature and Causes of the Wealth of Nations*, Adam Smith explained that "Civil government, in so far as it is instituted for the security of property, is in reality instituted for the defense of the rich against the poor, or of those who have some property against those who have none at all."[114] It seems, perhaps, that the twentieth-century synthesis, with its effects of fomenting and preserving class-based inequality, was alive and well in the eighteenth century, as well as in the centuries prior.

Of course, Smith, the political economist, was not an advocate for this state of affairs. To the contrary, in *Wealth of Nations*, Smith was describing what had been the case throughout human history, tracing the development of human society's transition from hunters and gathers to shepherds and then to agriculture – and he was arguing that human society was entering a new, fourth, age in which it needed to develop a more equal form of civil government.[115] Most notably, Smith (and his contemporaries) were arguing that a contemporary civil government had to embrace ideas of separation of powers and equality before the law.[116] It is instructive to look at Smith's rebuke of his century's own synthesis and how it relates to our current political economy.

---

113 Britton-Purdy, *supra* note 8, at 1792.

114 ADAM SMITH, AN INQUIRY INTO THE NATURE AND CAUSES OF THE WEALTH OF NATIONS 207 (Edwin Cannan, ed., Methuen & Co. 1904) (1776) [hereinafter WEALTH OF NATIONS].

115 *See id.* at 214.

116 Smith states:

> But upon the impartial administration of justice depends the liberty of every individual, the sense which he has of his own security. In order to make every individual feel himself perfectly secure in the possession of every right which belongs to him, it is not only necessary that the judicial should be separated from the executive power, but that it should be rendered as much as possible independent of that power.

Id. at 214.

1. Smith on the Immorality of Slavery

It is useful to start with some discussion of Smith's views on slavery, which helps to situate Smith's work alongside contemporary critical discussions. Slavery was not a primary subject of Smith's written works; it was, however, a practice that that he comments on in most of them, consistently to express hit strong opposition to the practice. His opposition to slavery was rooted most directly in his moral philosophy, which viewed all people as fundamentally equal and explored morality in terms of one's ability to identify with and express sympathy for those who are different. In his view:

> [t]he difference of natural talents in different men is, in reality, much less than we are aware of . . . . The difference between the most dissimilar characters, between a philosopher and a common street porter, for example, seems to arise not so much from nature as from habit, custom, and education.[117]

He concludes this passage by arguing that it is "the vanity of the philosopher [that he] is willing to acknowledge scarce any resemblance" between himself and those in lesser occupations.[118]

Taking this view from the opposite direction, not only does Smith view slavery as unjust, but those who practice it as barbarically immoral. In his *Theory of Moral Sentiments*, Smith develops the ideas of the "impartial spectator" for evaluating the morality of conduct

> We conceive ourselves as acting in the presence of a person[, an impartial spectator], who has no particular relation, either to ourselves, or to those whose interests are affected by our conduct . . . . It is only by consulting him that we can see whatever relates to ourselves in its proper shape and dimensions.[119]

Upon this standard, Smith clearly views slavery as a barbaric practice and those who practice it as engaging in an unjust enterprise. As he explains,

> There is not a negro from the coast of Africa who does not, in this respect, possess a degree of magnanimity which the soul of his sordid master is too often scarce capable of conceiving. Fortune never exerted more cruelly her empire over mankind than when she subjected those nations of heroes to the refuse of the [jails] of Europe, to wretches [European slave owners] who possess the virtues neither of the countries which they [the enslaved] come from . . . .[120]

---

117 *Id.* at 17.

118 *Id.* at 18.

119 Dugald Stewart, *Introduction* to ADAM SMITH, THEORY OF MORAL SENTIMENTS (Stewert ed., London, Henry G. Bohn 1853) (1759).

120 ADAM SMITH, THEORY OF MORAL SENTIMENTS 299–300 (Stewert ed., London, Henry G. Bohn 1853) (1759).

Indeed, it is notable that Smith's sympathies extend to America's indigenous people. He writes of Columbus and the Spanish backing of his voyages that "the council of Castile determined to take possession of countries of which the inhabitants were plainly incapable of defending themselves. The pious purpose of converting them to Christianity sanctified the injustice of the project."[121] He later broadened his critique to include the English colonies:

> Folly and injustice seem to have been the principles which presided over and directed the first project of establishing those colonies; the folly of hunting after gold and silver mines, and the injustice of coveting the possession of a country whose harmless natives, far from having ever injured the people of Europe, had received the first adventurers with every mark of kindness and hospitality.[122]

Despite his own views on slavery, Smith's moral critiques of the institution are a secondary focus in his work. Rather, his attacks on slavery are more often framed in the pragmatic language of his exploration of capitalism and commercial society: he argues that slavery is economically inefficient. This marks perhaps the greatest divide between LPE and both Smith's political economy and contemporary law and economics: Smith's pragmatic patience as compared to the impatience of LPE scholars; L&E's focus on the limitations of the political process as compared to LPE's willingness to wrest political control to channel the political process to bend it to their preferred outcomes.

Smith takes this approach largely because he views the abolition of slavery as unlikely. He notes that it has been common to most early societies, explaining that "Slavery takes place in all societies at their beginning, and proceeds from that tyrannic disposition which may almost be said to be natural to mankind . . . . It is indeed all-most impossible that is should ever be totally or generally abolished."[123] Practically, he can't imagine that the British monarchy or aristocracy would abolish it, noting that

> the persons who make all the laws in that country are persons who have slaves themselves. These will never make any laws mitigating their usage; whatever laws are made with regard to slaves are intended to strengthen the authority of the masters and reduce the slaves to a more absolute subjection.[124]

---

121 WEALTH OF NATIONS, *supra* note 114, at 63.

122 *Id.* at 90.

123 ADAM SMITH, LECTURES ON JUSTICE, POLICE, REVENUE, AND ARMS 96 (Edwin Canaan ed., 1896) (1763).

124 ADAM SMITH, LECTURES ON JURISPRUDENCE 181 (R.L. Meek et al. eds., LibertyPress 1982) (1766) [hereinafter LECTURES ON JURISPRUDENCE]. *See also id.* at 187 ("To abolish slavery therefore would be to deprive the far greater part of the subjects, and the nobles in particular, of the chief and most valuable part of their substance. This they would never submit to, and a general insurrection would ensue.").

Nor, did he believe, that we could "imagine the temper of the Christian religion is necessarily contrary to slavery."[125]

Smith's arguments against slavery therefore took on a more instrumental character, addressed more to those who might own slaves in their individual enterprises than to those who might abolish the entire institution. His best-known argument is that slavery is economically inefficient. Thus, in *Wealth of Nations* he argues that workers who are allowed to own some of the fruits of their labor "have a plain interest that the whole produce should be as great as possible, in order that their own proportion may be so."[126] But "[a] slave, on the contrary, who can acquire nothing but his maintenance, consults his own ease by making the land produce as little as possible."[127] Similarly, in his *Lectures on Jurisprudence*, Smith adds that not only do slaves have less incentive to work hard than free men, but also that they have less incentive to develop new, more productive, technologies.[128]

Smith also offers a more complex pragmatic argument against slavery, which is unfortunately too involved to offer a brief treatment of here. In summary, however, he argues that two things will happen as a slave-owning nation grows in wealth. First, there will be growing inequality between the slaves and their owners. For Smith, this is a moral inequality: the owners will be further and further removed from the experience of their slaves. As a result, they will sympathize with them less and therefore treat them with greater cruelty. And second, as the nation grows increasingly wealthy, it will rely on an increasingly large population of slaves. This, Smith pragmatically cautions, carries with it the risk of inevitable uprising. It is also remarkable to note the extent to which Smith presents an account that is sympathetic to the Marxist dialectic of class.

2. A Note on the "Dismal Science"

Of course, Smith is best known as an economist — he is arguably the most significant figure in the study of economics — though given his interest in the legal institutions that make up a political economy, one might proclaim him as the first law and economics scholar. It is therefore unsurprising that he has some connection to the field known as the "dismal science." More surprising is that that moniker was given to the field because Smith and other like-minded economists were staunchly opposed to slavery.

---

125 *Id.* at 191.

126 WEALTH OF NATIONS, *supra* note 114, at 365.

127 *Id.*

128 LECTURES ON JURISPRUDENCE, *supra* note 124, at 526 ("It is impossible that [commerce] can be so well carried on by slaves as by freemen, because they can have no motive to labour but the dread of punishment, and can never invent any machine for facilitating their business.").

The name "the dismal science" was given to economics in 1849 by Thomas Carlyle, a British historian perhaps best known for his "Great Man theory" – the view that "The history of the world is but the biography of great men."[129] One minor, but historically important, aspect of Carlyle's theory was that many people – including Blacks and most other historically enslaved people – were hereditarily inferior to other men.[130] It was therefore in their interest to be indentured (such as through slavery) to those most superior men for their own well-being.[131] Unsurprisingly, and relating this discussion back to earlier parts of this paper, Carlyle's views were influential in the early eugenics movement and popular among the American Progressives in the early 1900s.[132]

Contrary to Smith's skepticism that slavery could ever be abolished in England, it in fact was in 1833. This was in large part due to a movement among the generation of British moral and political philosophers most influenced by Smith and his contemporaries. (Though the terms on which slavery was ended echoed Smith's skepticism: to accomplish abolition, the British government had to pay off slave owners. In effect, the government bought every slave in the country to free them, paying British slave owners a sum so significant that the British government took out a loan that was not fully paid off until 2015.)

In 1849, Carlyle published a phenomenally racist satirical letter, *Occasional Discourse on the Negro Question*, in which he argued for the re-establishment of slavery. In this letter, Carlyle takes aim at, among other targets, "Exeter Hall" (the meeting place of a leading anti-slavery coalition) and economists. He writes:

> Truly, my philanthropic friends, Exeter Hall philanthropy is wonderful; and the social science -- not a "gay science," but a rueful – which finds the secret of this universe in "supply and demand," and reduces the duty of human governors to that of letting men alone, is also wonderful. Not a "gay science," I should say, like some we have heard of; no, a dreary, desolate and, indeed, quite abject and distressing one; what we might call, by way of eminence, the dismal science.[133]

The "social science" to which Carlyle refers, "which finds the secret of this universe in supply and demand," is economics.[134] And his criticism of it

---

129 THOMAS CARLYLE, ON HEROES, HERO-WORSHIP AND THE HEROIC IN HISTORY 1 (Fraser ed., 1843).

130 THOMAS CARLYLE, OCCASIONAL DISCOURSE ON THE NEGRO QUESTION 533 (Fraser ed., 1849) (referring to the British as "worthier" than their West Indies slaves).

131 *Id.* at 534–35 (discussing how "European heroism" made savage lands "arable" and keep the natives in a "comfortably idle condition.").

132 *See* Jeffrey A. Tucker, *The Prehistory of the Alt-Right*, Fee Stories (Mar. 8, 2017), https://fee.org/articles/the-prehistory-of-the-alt-right ("Hitler's biographers agree that the words of Carlyle were the last he requested to be read to him before he died.").

133 CARLYLE, *supra* note 130, at 530.

134 *Id.* at 530-531.

is the Smithian prescription that an economy of free men, in which "human governors … let[] men alone," is preferable to one in which those "human governors" keep men as slaves.[135]

3. Smith's views on civil government

Adam Smith's best-known work is *Wealth of Nations*, which is most often thought of as an ode to wealth, and "the bible of capitalism." In truth, it was no such thing. As Smith described the book later in his life, it was a "very violent attack … upon the whole commercial system of Great Britain."[136] Recall that the entire title of the book was *An Inquiry into the Nature and Causes of the Wealth of Nations*. His purpose is to discern why the United Kingdom had become so prosperous a country in a historically short period of time when the rest of the world known to him had remained in a poor agrarian state.

Smith's answer to this question, famously, was that the advent of capitalism and the division of labor had led to the wealth of the nation. But in providing this answer he argued that the abuses of the crown and more general lack of justice in the British government, which were unfortunate hallmarks of earlier ages of society, were untenable in a capitalist society. In particularly, he argued for a separation of powers between the crown and judiciary.[137]

Smith's statement about the purpose of civil government that "in so far as it is instituted for the security of property, [it] is in reality instituted for the defense of the rich against the poor"[138] requires situating his statement within his understanding of the development of society.

In Smith's narrative, human society naturally progresses through several stages of development. Historically, we start as hunters and gathers, then progress to shepherds, and from there to agriculture.[139] Each of these stages has different needs for government and property. Notably, over this course of development, there is a transition from nomadic to tribal, and then to nation-state, conflict.[140] The prosperity of society in these stages of development, to the extent there is any, is very much tied to the ability of its leaders to wage war and provide for the needs of their followers. The governments, therefore, are tied to the wealth and strength of the leaders.[141] In the earlier stages of development, strong leaders effectively barter their strength at war for wealth from their followers. As those leaders transitioned

135 *Id.* at 531.

136 *Adam Smith on Consumption as the Only End and Purpose of Production*, ONLINE LIBERTY FUND, https://oll.libertyfund.org/quote/adam-smith-consumption-only-purpose-production.

137 WEALTH OF NATIONS, *supra* note 114, at 210.

138 *Id.* at 207.

139 *Id.*

140 *Id.* at 205.

141 *Id.* at 207.

from chieftains to kings, accumulating wealth along the way, this bargain changed. Because all the wealth of the land was the king's, citizens were permitted to occupy in land in exchange for working it for the king. In Britain, this ultimately gave rise to the system of nobility described at the beginning of this paper, in which the nobility were effectively the King's deputies, given control over the King's territory in order the manage the peasants; and the peasants were allowed to occupy the land in exchange for working it for the nobility.

Smith's starting observation in *Wealth of Nations* is that Britain was entering a new, post-agrarian, stage: the stage of commerce.[142] He was writing in the early period of the Industrial Revolution, a period of rapid innovation and industrialization that was reshaping society. Only decades prior, there was little commerce in society – the most valuable exchange that most could make was to exchange their labor in the fields for a promise of protection by the King's sheriffs and army.[143] This began to change dramatically over the course of the eighteenth century as new technology rapidly allowed individuals to produce significantly more food through significantly less agricultural labor. Per capita GDP (a measure of the amount of wealth created by each person on the country) in Britain grew by less than 1 percent between 1700 and 1750 – that is, it effectively did not grow at all. It grew by 1.6 percent between 1750 and 1760. Between 1760 and 1770, it grew by an incredible 42 percent (that is *not* a typo), and again by 42 percent between 1770 and 1780.[144]

This dramatic growth in individual prosperity was happening around Smith in real time. It was a period during which the average person went from living on the cusp of starvation to what would then be thought of as middle class. This gave rise to an important change in society: most people went from having to spend all their time working in order to produce enough food to stay alive to having both some amount of free time and some amount of wealth they could spend on leisure. In such a society, citizens want for more than merely the protection of the King's army and sheriffs – indeed, they want more than King or government would ever be able to produce. This is what Smith meant by the transition from an agricultural state to a commercial one. In an agricultural society, the wealth of the nation is kept in its land and limited to the agricultural products that that land can produce. In a commercial society, the wealth is created by its people – indeed, it is its people – and is limited by their imagination and their ability to engage with one another.

---

142 WEALTH OF NATIONS, *supra* note 114, at 161.

143 Guy I. Seidman et al., *The Origins of Accountability: Everything I Know About the Sovereign's Immunity, I Learned from King Henry III*, 49 ST. LOUIS U.L.J. 393, 421 (2005).

144 Stephen Broadberry et al., *British Economic Growth, 1300–1850: Some Preliminary Estimates*, presented at at Economic History Society Annual Conference 2010 at the University of Durham, 26-28 Mar. 2010 (unpublished).

4. Smith's critique of civil government

Smith's critique of civil government as being "in reality instituted for the defense of the rich against the poor"[145] needs to be understood in the context of this transition to a commercial society. In every prior stage of human society, the government is naturally – not rightly or wrongly – structured to protect the property of the chief property-holder. In England, that meant in the interest of the crown or nobility. While nominally the common-law legal system would decide disputes as between the parties, in practice, it was always beholden to the King.

As Smith explains immediately following his statement about the historic purpose of civil government:

> The judicial authority of such a sovereign, however, far from being a cause of expence, was for a long time a source of revenue to him. The persons who applied to him for justice were always willing to pay for it, and a present never failed to accompany a petition. After the authority of the sovereign, too, was thoroughly established, the person found guilty, over and above the satisfaction which he was obliged to make to the party, was likewise forced to pay an amercement to the sovereign. He had given trouble, he had disturbed, he had broke the peace of his lord the king, and for those offences an amercement was thought due.[146]

Such a system was never tolerable to a political philosopher like Smith, who was writing in the spirit of Hobbes, Locke, and Montesquieu. Indeed, in the pages following this quote, Smith goes on to discuss various ways in which a judicial system could be financed, tracing how tying the self-interest of a court to its ability to raise revenue or its need to appease the whims of a ruling political class could systematically distort justice.[147]

This system was particularly intolerable to Smith as we transitioned from an agricultural society to a commercial society.[148] In an agricultural society, the vast majority of peasants, who made up the vast majority of the population, did not have much, if any property for the law to protect – and nobility would settle most disputes by more civilized (and less public) ways than turning to the courts. But in a commercial society, most people were rapidly acquiring property and accumulating wealth. The role of the judicial system was therefore increasingly important, and the interests that it needed to protect more representative of the peasant than the crown.

Smith concludes this portion of his discussion by arguing for a separation of power between the executive and judiciary, arguing that

> When the judicial is united to the executive power, it is scarce possible that justice should not frequently be sacrificed to what is vulgarly called polities. The persons entrusted

145 WEALTH OF NATIONS, *supra* note 114, at 207.

146 *Id.*

147 *See id.* at 210.

148 *Id.* at 161.

> with the great interests of the state [that is, the king] may, even without any corrupt views, sometimes imagine it necessary to sacrifice to those interests the rights of a private man. But upon the impartial administration of justice depends the liberty of every individual, the sense which he has of his own security. In order to make every individual feel himself perfectly secure in the possession of every right which belongs to him, it is not only necessary that the judicial should be separated from the executive power, but that it should be rendered as much as possible independent of that power.[149]

5. Smith's Political Economy of the Common Law

A separation of powers is essential for what Smith at one point called the "tolerable administration of justice."[150] But what is that law being enforced by this tolerable administration of justice? It is important here to note that Smith's primary concern was the judge-made common law, not the statutory law created by legislation. Indeed, Smith's argument for a separation of powers between the judiciary and executive was of course an idea borrowed from Montesquieu – but he omits the legislature from this need for separation.[151]

This is because Smith's primary concern was the protection of property, the laws governing which – in the thinking of Smith and his contemporaries – were largely within the purview of the judiciary to discover through the common law process. In this sense, he was largely influenced by his friend David Hume. In his work on justice and injustice, Hume spoke of "the three fundamental laws of nature, that of the stability of possession, of its transference by consent, and of the performance of promises,"[152] arguing that "It is on the strict observance of those three laws that the peace and security of human society entirely depend; nor is there any possibility of establishing a good correspondence among men, where these are neglected."[153] Broadly, these three "laws of nature" correspond to the three familiar core common law subjects: tort (stability of possession), property (transference by consent), and contact (performance of promises). Collectively, they are the core subjects of what contemporary legal scholars refer to as "private law."

In other words, Smith's political economy recognized the same private-law/public-law divide criticized by the LPE Founders as a "twentieth-century synthesis" – and for much the same reasons the twentieth-century L&E and public choice scholars have argued for that divide.

---

149 *Id.* at 214.

150 *Adam Smith on the Need for Peace, Easy Taxes, and a Tolerable Administration of Justice*, ONLINE LIBERTY FUND, https://oll.libertyfund.org/quote/adam-smith-on-the-need-for-peace-easy-taxes-and-a-tolerable-administration-of-justice-1755 (last visited Oct. 24, 2022).

151 WEALTH OF NATIONS, *supra* note 114, at 213.

152 DAVID HUME, A TREATISE OF HUMAN NATURE book 3, pt. 2, § 6 (1740), Hume Texts Online, https://davidhume.org/texts/t/3/2/6#:~:text=Hume%20Texts%20Online%20SEC%20T.%20VI.%20Some%20farther,by%20consent%2C%20and%20of%20the%20performance%20of%20promises.

153 *Id.*

C. *Elinor Ostrom*

The next case study is Elinor Ostrom. Among other things, Ostrom was a Nobel laureate, one of the leading figures in New Institutional Economics, a past president of the Public Choice Society, and winner of the Frank E. Seidman Distinguished Award for Political Economy.

Ostrom was a highly celebrated twentieth-century political economist. One might therefore expect her to epitomize the opposite of "the older and more foundational usage [of political economy] familiar to nineteenth-century audiences"[154] that the LPE project means to invoke. But this seems unlikely. Her work and values – like Adam Smith's – seem surprisingly in line with the goals of the LPE project.

Ostrom spoke about both when she won the Nobel Prize:

> [After a few years in my position as an Assistant Professor at Indiana University,] I was able to offer a year-long graduate seminar in 1969–1970 on the theories related to urban government and the measurement of public goods and services. … The fall semester gave us an opportunity to review the extensive literature on urban governance and service delivery. There were two dominant approaches—metropolitan reform and public economy. As we began to unpack the theory underlying these approaches, we found both posited that the size of governmental units in a metropolitan area affected the output, efficiency, distribution of costs to beneficiaries, citizen participation, and responsibility of public officials, but the direction of the posited relationship differed.
>
> …
>
> Advocates of the metropolitan reform approach assumed that size of governmental units would always be positive for all types of goods and services. Scholars using a political economy approach assumed that size of governmental units would be positive or negative depending on the type of public good or service. …
>
> We did have a "case" nearby that made our effort to understand these diverse approaches highly relevant. …[T]here were three independent, small police departments serving neighborhoods immediately adjacent to socioeconomically very similar neighborhoods being served by the much larger Indianapolis City Police Department. That gave us a natural experiment. Our study measured the performance of policing in these six urban neighborhoods using survey research and building on recent rigorous methodological studies.
>
> …
>
> Our study generated some surprising findings—at least to those scholars who presumed that larger urban governments would always produce superior public services. In contrast to their neighbors served by the Indianapolis City Police Department, households in the three communities served by their own smaller police departments faced a lower victimization rate, were more likely to call upon the police if they were victimized, received higher levels of policy follow-up, and evaluated the performance of their police department more positively.

154 *LPE Manifesto*, *supra* note 9.

…

In the fall of 1970, I worked with an outstanding group of black students to conduct a study in two poor, independent black communities that we matched to three similar communities being served by the Chicago Police Department. At the time of our study, the two small communities had just a few police officers. Their wages were low and their official cars were frequently out of service because their budgets were so limited. The Chicago Police Department had a force of >12,500 men who were paid relatively high wages. We estimated that expenditures for police service in the three Chicago neighborhoods were 14 times the amounts spent on policing in the small communities. But despite the huge difference in spending, we found that in general the citizens living in the small cities received the same or higher levels of services compared to the residents in Chicago. Although victimization rates were similar, citizens living in the small independent communities were less likely to stay home due to fear of crime, and they agreed with statements that their local police treated all citizens equally according to the law, looked out for the needs of the average citizen, and did not take bribes. The findings from these initial studies were consistent with the political economy theory.

…

We then conducted a much larger field study in the St. Louis metropolitan area, which was served by two large departments … . We found a strong and significant positive relationship between size of police department and per capita costs of police services in a neighborhood as well as the percentage of households that had been victimized. Negative relationships existed between size of department and the percentage of households that were assisted who indicated that police response was rapid, rated the job of the police as outstanding, or evaluated their police to be honest.

Replications of our empirical work were undertaken in Grand Rapids, Michigan, and in the Nashville–Davidson County area of Tennessee. In this full set of studies, no one found a single case where a large centralized police department was consistently able to outperform smaller departments serving similar neighborhoods across multiple indicators.

Thus, we provided very strong support for the posited relationships of the political economy approach. For policing, increasing the size of governmental units consistently had a negative impact on the level of output generated as well as on efficiency of service provision. Why?

In our efforts to understand why smaller police departments so consistently outperformed their better trained and better financed larger neighbors, we developed the concept of "coproduction" of public services. This involves a mixing of the productive efforts of consumer/citizens and of their official producers. It is feasible for governments to produce highways and other physical infrastructure without the active involvement of citizens, but we observed citizens and their officials working together more effectively in the small- to medium-sized communities we studied, and this collaboration has important effects on policing. In the smaller communities, citizens take a more active role in monitoring their neighborhoods, notifying the police rapidly when suspicious activities occur or when victimized. Systematically observing officers on duty in their patrol cars demonstrated that officers in the smaller departments also had greater knowledge about the areas they served. Thus, not only were there diseconomies of scale in the formal production of services in urban areas—as posited by the political economy approach—but human services could not be effectively produced by official agencies

> alone. Citizens are an important coproducer. If they are treated as unimportant and irrelevant, they reduce their efforts substantially.[155]

In this excerpt, Ostrom describes two things. First, she describes her research into the effects of police department size on the quality of that department. Her research found that the quality of a department's policing was inversely proportional to that department's size. This result was contrary to the most common understanding then, (as is largely still true today) that larger, more sophisticated, better resourced police departments would be better police departments.

Second, she offered a way to account for this mismatch, a theory that she calls "coproduction."[156] The basic idea of coproduction, applied in this context, is that "policing" isn't a good or service that is simply procured or provided by the government.[157] You don't go to the store to buy "some policing." Rather, good policing is something that is produced jointly by the community in need of the policing service and the police offering that service. The further removed the police are from the on-the-ground experience, needs, and concerns of a local community, the less ably the police will be able to serve that community. And, though Ostrom doesn't say this, the more removed the police are from a local community, the more they will service the needs and expectations of some other, remote community.

Building on this work, Ostrom went on to study how communities and institutions can govern themselves. A basic theme that runs throughout her work is that communities are generally able to develop rules to govern themselves and that those rules often outperform more formal rules. Thus, farmers are likely to develop customs for how to manage shared pastures that are more effective than if a legislature enacts rules; users of water from an irrigation system are able to develop rules that ensure all users have access to water from the system, without the need for governmental involvement; and outside of extreme circumstances local, community policing is more effective than more typical, government-designed, approaches policing and criminal justice.

As with Smith's critique of civil government and the political economy of the common law, Ostrom's work draws attention to the tension between public law institutions and private individuals. Perhaps even more than Smith, Ostrom finds potential efficiencies in private ordering, especially compared to the inefficiencies of public ordering. What Ostrom's work, in

---

155 Elinor Ostrom, *A Long Polycentric Journey*, 13 Annu. Rev. Polit. Sci., 1-23 (2010).

156 Elinor Ostrom, *Crossing the Great Divide: Coproduction, Synergy, and Development*, 24 WORLD DEV. 1073 (1996).

157 *Id*. at 1079 ("After studying police services in metropolitan areas, however, we had not found a single instance where a large, centralized police department was able to provide better direct service, more equitably delivered, or at a lower cost to neighborhoods inside the central city when these were carefully matched to similar neighborhoods located in surrounding jurisdictions. . . . We were dealing with a public-private industry rather than with the bureaucratic apparatus of a single government.").

effect, says is that we don't need a system of formal, governmentally created and backed rules in order for communities to govern themselves. And, in fact, informal (or non-governmental) rules that a community develops to govern itself will generally outperform the more formal rules that a government-based process would have created.

This, once again, echoes the "twentieth-century synthesis" that LPE would have us reject. But it finds support for this synthesis in terms quite different than those that LPE identifies. Indeed, Ostrom's work more often echoes views and causes that are more likely to hold political valence with proponents of LPE than the conservative proponents of L&E that the LPE movement criticizes. From police reform to environmental regulation, Ostrom's work was ground-breaking and remains relevant. Yet its own foundation is complementary to, and would not survive LPE's uncritical rejection of, the "twentieth-century synthesis."

D. Derrick Bell

The final case study that I consider is Derrick Bell. Professor Bell is not an economist or a law and economics scholar. He is probably best known as one of the founding voices of Critical Race Theory and for his book *Faces at the Bottom of the Well*. But I, as a law and economics scholar, nonetheless find his work compelling and see echoes of it in law and economics. Indeed, as argued below, contemporaneous work in law and economics was exploring similar themes and reaching similar conclusions as his work, albeit not in the context of race. This similarity is perhaps most notable in his idea of interest convergence, which bears notable resemblance to the "Baptists and bootleggers" idea in public choice.[158]

In Derrick Bell's *Faces at the Bottom of the Well*, Bell questions "whether [advances in civil rights] established realistic goals or only token symbols that merely placated blacks and helped white Americans alleviate the shame and guilt of a hateful, oppressive past."[159]

According to Bell, gains made in civil rights for blacks in the 1950s and 1960s were offset by "the courts and the political, social, and economic structure."[160] He blames this on the lasting legacy of racism in the country that continued to his day (and, likewise, continues to our day).[161] Bell says that the jurists and others in the legal community, for their part, "rationalize

158 *See* Derrick A. Bell, Jr., Brown v. Board of Education *and the Interest-Convergence Dilemma*, 93 HARV. L. REV. 518 (1980) (introducing the concept of interest convergence).

159 Stephanie Goodman, *Faces at the Bottom of the Well: The Permanence of Racism*, 28 HARV. C.R.-C.L. L. REV. 244, 245 (1993) (reviewing Derrick A. Bell, Jr, FACES AT THE BOTTOM OF THE WELL: THE PERMANENCE OF RACISM (1992)).

160 *Id.* at 244.

161 *Id.* at 245–46 (". . . a society in which racism is permanent and whites can find justification for the subordination of and discrimination against blacks").

[discriminatory] laws through formal doctrine, though they fail to remedy the reality of discrimination below the surface."[162]

His study of race led Bell to the idea of interest convergence, which states that advancement in blacks' rights only happens when it is simultaneously advantageous for whites.[163] For example, Bell asserts that the iconic *Brown v. Board of Education* decision in 1954 was not made to achieve racial equality, but rather as a propaganda effort during the Cold War.[164]

Most law and economics scholars – and public choice scholars in particular – will recognize this concept immediately. It is a specific example of the "bootleggers and Baptists" idea popularized by Bruce Yandle.[165] Both Bell and Yandle, writing in the early 1980s, identified this "strange bedfellows" aspect to public policy:[166] enacting policy often requires "strange bedfellows" coalitions, and in such coalitions, the interests of the minority group are often secondary to those of the majority group. In Yandle's example, Prohibition was brought about by the convergence in interests between teetotaling Baptists, who opposed alcohol on moralistic grounds, and bootleggers, who knew making alcohol illegal would create a vibrant black market in which they would be key players.[167]

Upon learning of interest convergence, I expect most law and economic scholars – especially those educated in public choice – would recognize the concept as intuitively correct, interesting, and of a kind with mainstream economic thought. Yet I also expect that most law and economics scholars have not heard of interest convergence, especially because it comes from the L&E-antagonistic field of critical race theory. But there ought to be no antagonism here.

There is a tragic realism in Bell's writings – one that, again, echoes themes from public choice. Bell argues that to combat interest convergence, blacks should have "Racial Realism" and realize that their subordination is permanent.[168] Bell believes that, rather than a form of despair, this will free blacks to think of new racial strategies that can bring fulfillment and triumph.[169] For example, in *Faces at the Bottom of the Well*, Bell discusses a hypothetical Racial Preference Licensing Act, in which Congress allows

---

162 *Id.* at 249.

163 Derrick A. Bell Jr., *School Litigation Strategies for the 1970's: New Phases in the Continuing Quest for Quality Schools*, 1970 WIS. L. REV. 257, 267 (1970).

164 Derrick A. Bell, Jr., *Reconstruction's Racial Realities*, 23 RUTGERS L.J. 261, 266 (1992).

165 Bruce Yandle, *Bootleggers and Baptists: The Education of a Regulatory Economist*, 7 REGULATION 12, 13 (1983).

166 Bell published Brown v. Board of Education *and the Interest-Convergence Dilemma*, *supra* note 158, in 1980.

167 *See* Yandle, *supra* note 165, at 13.

168 Derrick Bell, *Racial Realism*, 24 CONN. L. REV. 363, 373–74 (1992); *see also* DERRICK BELL, FACES AT THE BOTTOM OF THE WELL: THE PERMANENCE OF RACISM 12 [hereinafter BELL, FACES] ("Black people will never gain full equality in this country.").

169 Bell, *supra* note 168, at 373–74.

business owners to discriminate against black patrons as long as they pay a tax that supports Black schools, housing, businesses, and the like.[170] This works, according to Bell, not to achieve racial equality, but to "declare[] racial justice in the marketplace by balancing the rights of African-Americans to fair treatment with the needs of whites to 'prefer' certain customers, employees, and contractees."[171] In other words, Bell is making an economic argument for racial equality: a business can make the choice to discriminate, but at a high financial cost to them. The market may better mitigate discrimination because business owners may stop discrimination to avoid paying the tax. "Racial Realism" has been criticized for not providing a clear way to reform civil rights strategies.[172]

Here Bell's style of thinking clearly resembles that of law and economics scholars – he is literally calling for a market mechanism for allocating racist attitudes and collecting Pigouvian taxes to compensate for those attitudes.

But there is a more subtle and important aspect to Bell's tragic realism. As discussed in Part III.B, students of public choice often lament its descriptive nihilism of the political process. As a field, public choice is notable for documenting the myriad pathologies of the political process – the ways in which it is captured by special interests and fails to protect the interests of the majority of any polity – without prescribing solutions to these problems. Its strongest prescription is to limit the threat of these pathologies by limiting the scope of power of government. Just as critical legal theorists were criticized for failing to offer solutions to the myriad criticisms they leveraged against legal institutions, it is not entirely unfair to criticize public choice scholars for not offering solutions for the criticisms they leverage against government.

I say "not entirely unfair" because it is, in some ways, an unfair comparison. Public choice scholars offer a descriptive account of how the political process works and offer a responsive prescription: limit the extent to which we entrust important decisions to this flawed institution and rely instead on other public institutions that are more insulated from capture by special interests. LPE scholars take the opposite approach, arguing that government's frailties present an opportunity to wrest its power for those scholars' own normative goals. Here, Bell's perspective seems more in line with that of the public choice scholars. He does not give in to the nihilism of despair. His "racial realism" does not aspire to wrest control of the system or to change it. Rather, it recognizes the permanence and limitations of the system and uses those limitations to address the permanence as best as possible.

---

170 BELL, FACES, *supra* note 168, at 47–48 (1992).

171 Karen L. Ross, *Combatting Racism: Would Repealing Title VII Bring Equality to All?*, 21 SETON HALL LEGIS. J. 141, 162–63 (1997) (citing to BELL, FACES, *supra* note 168, at 47).

172 *See* Willie Abrams, *A Reply to Derrick Bell's Racial Realism*, 24 CONN. L. REV. 517, 517–18 (1992).

One cannot argue in good faith that either Bell's normative prescription or those of public choice scholars are satisfactory. They are not. But that is the nature of realism. It would be better to live in a world without discrimination, in which government works fairly and equally for all – a world without transaction costs. But the response that the LPE project offers to this dissatisfaction is even less satisfactory. The perspectives of Bell, public choice scholars, and law and economics scholars are, above all else, based in realism. To reject them in favor of a misguided sense of political economy that calls for capturing the means of legal production and turning the machines of government to work to the ends of a preferred class is nothing more than to reject the modern state – law and all – and to return society to a pre-Hobbesian war of all against all.

CONCLUSION

The Law and Political Economy project makes bold claims, both in its characterization of the nature of contemporary law and in its own normative prescriptions responding to that characterization. This article has considered both this characterization and prescriptions and found them lacking. This is tragic. LPE has already demonstrated that it has real power as a movement and, consequentially, has real potential to do damage to the legal institutions best positioned to address the very real concerns animating much of the LPE agenda – and to emphasize this point, there are substantive, real, important societal concerns animating the LPE agenda. The even greater tragedy, however, is the extent to which LPE defines itself oppositionally to the legal mainstream – and especially to economic analysis. The more important synthesis, I argue, is between mainstream legal analysis, notably law & economics, and public choice economics, and LPE. These fields can learn from each other and would be stronger were they to do so.

The characterization of contemporary law is presented as a grand "synthesis" of private- and public-law institutions that developed over the latter half of the twentieth century, focusing primarily on the effects of law and economics (on the private side) and public choice economics (on the public side) in shaping contemporary understandings of the law. On both the private- and public-law side of the ledger, these developments are presented as normatively suspect and in need of a "radical" normative response. But the proponents of the *Twentieth-Century Synthesis* fail to engage substantively with the purely descriptive reasons that law & economics and public choice economics were so powerful in shaping the course of the law. Indeed, the authors' presentation makes basic mistakes on both accounts that suggest, at best, a lack of understanding of the law & economics critique of earlier legal institutions and an absolute failure to engage with the substantive critiques of public choice economics.

This latter point is especially notable because the *Beyond the Twentieth-Century Synthesis* presents a truly fascinating study in public choice. On one

side of the coin, many of the public-law critiques raised by the authors are quite apt. The rise of the conservative legal movement and its effectiveness in capturing various public institutions is a straight-up public choice story. And on the other side of the coin, some of the ideas the authors promise to explore (even if they get only a very superficial treatment in *Beyond the Twentieth-Century Synthesis*) could be pulled straight from a public-choice textbook or contemporary public choice scholarship. How and whether we can structure democratic institutions to be insulated from capture and sensitive to concentrations of power are central questions of public choice economics.

But rather than embrace and learn from public choice economics as a cognate field, LPE instead chooses to castigate it as a source of contemporary concerns and, ignoring its lessons, to lean forcefully into institutional capture as a source of power. The general thrust of LPE's normative prescription is: *the world is unjust, and we should use government to do something about it*. This amounts to a call to use the machinery of government to impose a preferred set of policies upon the world without engaging with the risks of such an approach – namely that the policies imposed upon the world may be those preferred by someone else, someone antagonistic to one's own preferred policies.

In other words, LPE's normative call is that "lawyers of the world unite" – but this call forgets both that not all lawyers of the world agree on all things, and that "first thing those with power do is kill all the lawyers."

The irony of LPE is that political economists, from Adam Smith (arguably the founder of political economy as a field) through today, have long shared many of the concerns to which the LPE project draws attention. Today these thinkers are often thought of as aligned with fields like law & economics and public choice – and rightly so. But their concerns highlight the overlap between LPE and the fields that it castigates. Indeed, not only do these fields share many of the same concerns, but also many of the same methods and conclusions. Rather than demonize, mischaracterize, and ostracize mainstream legal analysis, LPE would be better served to work in dialogue with them – and those fields would benefit from better understanding and taking the concerns of LPE seriously.

HOW CRITICAL THEORY FUNDAMENTALLY CHALLENGES TRADITIONAL INQUIRY IN SOCIAL SCIENCE[1]

Richard Lowery
McCombs School of Business
University of Texas at Austin

Carlos Carvalho
McCombs School of Business
University of Texas at Austin

INTRODUCTION

Academic analysis outside of fields where controlled experiments are relatively easy has always fallen short of the ideals of the scientific method. Much of the intellectual history of such fields centers around continued attempts to bring practices more in line with these ideals. These efforts range widely, from improved statistical analysis designed to allow scientific information to be extracted from nonexperimental data to actual attempts in fields like economics to introduce experimental methods. While debates about how to proceed in this path have always been contentious, until recently there has been a general consensus in much of social science that making fields more scientific wherever possible was the right way to proceed.

A challenge to this consensus has developed over recent decades and now dominates much of the thinking on social issues in universities. Critical theory, colloquially known as "wokeness," explicitly rejects the scientific method in favor of two fundamental propositions: (1) knowledge is gained through personal experience, and the quality of that knowledge is determined almost exclusively by which identity groups that person belongs to and (2) the measure of the value and validity of knowledge is determined by its usefulness in bringing greater power to certain groups at the expense of other

[1] The authors thank the Law & Economics Center (LEC) project on Capitalism and the Rule of Law (CapLaw) at George Mason University for support of this research. The opinions expressed here are the authors' alone and do not represent the official position of the Law & Economics Center, George Mason University, or the University of Texas at Austin. The authors would like to thank the participants in the Law & Economics Center's June, 2021 Research Roundtable on Capitalism and the Rule of Law for their helpful comments. Any comments or suggestions are welcome and may be emailed to richard.lowery@mccombs.utexas.edu.

groups, not by the ability of the knowledge to make predictions about real world phenomena.[2]

These two approaches to learning about society and advising those who must design policies that affect social outcomes clearly conflict in essential ways. This conflict of visions remained somewhat latent within universities; the natural siloing of disciplines led to divergent approaches, with economics and related fields continuing to pursue the elusive goal of reaching the status of a true science while other social sciences, and particularly the new and growing "area studies" fields, pursued the critical theory path. Now, these barriers are breaking down, and the critical theory method appears poised to become the dominant approach in a much wider range of fields, likely including economics and allied disciplines.

The most serious issues that will arise as critical theory approaches grow further in previously scientific or aspirationally scientific fields will include the narrowing of free discourse within these fields. Critical theory approaches to "academic" analysis, somewhat ironically, show a distinct hostility toward criticism and a rejection of the idea of empirical falsifiability. Specifically, disagreeing with maxims of critical theory, especially in the case of critical race theory, is itself considered proof of the assertions of critical theory, as described most succinct with the famous phrase "The very heartbeat of racism is denial."[3] Further, evidence cannot be used to overcome "knowledge" obtained by other means, and the idea of objective evidence itself is largely rejected. Moreover, disagreement or failure to enthusiastically embrace certain ideas is considered a form of violence, and punishment for deviations from these ideas is considered legitimate in the new academic environment.[4]

Furthermore, evidence for the increasing relevance of these ideas in academic fields beyond the usual ethnic and gender studies areas has become hard to dispute, with recent surveys documenting a marked hostility toward ideas that dissent from this framework among university faculty.[5] Even in "hard sciences" the replacement of objective inquiry with pre-determined narratives based on critical-theory-based ideas of oppression appears to be moving ahead apace; for example, a recent editorial from the board of *Nature* embraced such ideas as a priority going forward.[6] That is, a major journal in natural science has embraced an explicitly non-falsifiable belief system as a main governing principle for its future editorial process, along with a commitment to work to push the scientific community in that direction. Thus,

---

[2] Helen Pluckrose & James Lindsay, CYNICAL THEORIES: HOW ACTIVIST SCHOLARSHIP MADE EVERYTHING ABOUT RACE, GENDER, AND IDENTITY – AND WHY THIS HARMS EVERYBODY (1st ed. 2020).

[3] Ibram X. Kendi, *The Heartbeat of Racism is Denial*, N.Y. TIMES Jan. 13, 2018.

[4] Eric Kaufmann, *Academic Freedom in Crisis: Punishment, Political Discrimination, and Self-Censorship*, 57, C.S.P.I. 2021.

[5] Adekoya et al, *Protecting Viewpoint Diversity*, POLICY EXCHANGE, 8, 2020.

[6] Board of Nature, *Tackling Systemic Racism Requires the System of Science to Change* 593 NATURE 313 (2021).

researchers in the social sciences and other disciplines must come to accept that we now work in a very different environment than that in which we expected to operate. Specifically, one must now be very careful to avoid presenting information that deviates from accepted orthodoxy, and one must be prepared for an audience that cannot be persuaded of certain facts regardless of the available information.

Given the speed of the shift in the scientific paradigm, we believe it is valuable to provide a formal conceptual framework for how academic inquiry will proceed in the future, along with a practical analysis of the sorts of tools and methods that will need to be developed and applied going forward. The fundamental challenge will be to continue to provide some relevant information that may be used in decision-making in the face of stark and ever-changing constraints on what can be expressed. For example, few would have believed even ten years ago that certain fields would effectively prohibit the acknowledgment of a fundamental, socially important distinction between males and females, but now the irrelevance of biological sex is considered a required view not only within certain departments but as a matter of institutional policy in certain academic environments.[7] We attempt to provide this framework by presenting formal models of the new intellectual environment, evidence that such an environment may be entering into economics and related disciplines, and tools to potentially operate successfully in this environment.

Our paper proceeds as follows. First, we introduce a formal model of research design in a "woke" environment. We consider a simplified version of the now-ubiquitous Bayesian persuasion models, where a researcher designs an experiment to influence the beliefs of those who receive the information contained in the research.[8] The model captures the two most important elements of the modern Critical Theory or Critical Social Justice approach to research: First, a subset of the consumers of research, the woke, have degenerate priors over what statements may be true. Second, if the woke are forced to confront evidence that the truth is actually something for which they give zero prior probability, the woke respond by "cancelling" the researcher. That is, instead of updating priors, which is impossible when placing zero probability on certain states that turn out to be true, the woke attack the agent who is responsible for providing the new information, giving him a large, negative payoff.

Our model shows how the presence of a woke receiver can influence the choice of research design. Fear of cancellation can discourage researchers fromconductingvaluable experiments. Similarly, the presence of extreme

---

[7] John Morgan, *British University Apologizes for Disinviting Academics Over Gender Views*, Times Higher Education (May 21, 2021), https://www.timeshighereducation.com/news/essex-apologisesacademics-disinvited-over-gender-views.

[8] Dirk Bergemann & Stephen Morris, *Information Design, Bayesian Persuasion, and Bayes Coorrelated Equilibrium,* (Am. Econ. Rev.: Papers & Proceedings 2016, 106(5): 586-591), http://dx.doi.org/10.1257/aer.p20161046.

priors can distort which experiments will prove most useful under the political constraints imposed by wokeness. We characterize how these distortions change as the share of the population that is woke changes. Further, we use the model to investigate how attitudes toward "free speech" will change as the proportion of the population that is woke changes. Notably, the model can be interpreted as showing that the woke will express support for unconstrained research design (i.e. free speech) when they are a relatively small share of the population, moving toward hostility when they become more dominant.[9] Our model thus may help explain the changing attitudes toward free expression among academics and activists.

We next home in more specifically on the problem facing researchers in the developing Critical Social Justice environment. Motivated by the tragic story of Nikolai Vavilov, who stood against Troffim Lysenko at great personal cost,[10] we argue that the fundamental challenge for a researcher in social science in particular going forward will be to develop research designs that allow valid, useful information to flow while avoiding challenging potentially completely incorrect beliefs of the woke. A simple, formal example of the challenges presented by such constraints illustrates both the necessity and utility of developing such methods. This example introduces the concept of *political* and *auxiliary* variables, where political variables are those whose acceptable values are predetermined by the intellectual and political environment, while auxiliary variables are those variables whose values researchers have broad latitude to evaluate.[11]

Armed with this conceptual framework built on our two models, we then describe through examples how Critical Theory approaches operate and how it will affect, in practice, research in economics. Our first motivating example is a paper in political science which was withdrawn in the face of death threats.[12] We then proceed to evaluate two papers in economics that provide an potentially interesting contrast, to show how the Critical Theory approach could be, and possibly already is, applied in economics; papers with similar methodologies may be subject to very different levels of criticism based on the extent to which the results align with the conclusions from alternative "knowledges" as determined by Critical Theory.[13] Finally, we consider how existing empirical techniques may be modified to implement the constraints we have discussed, allowing at least some useful work on controversial topics to proceed.

---

9 *See also* Chanda Prescod-Weinstein, *Making Black Women Scientists Under White Empiricism: The Raclialization of Epistemology in Physics,* 45 SIGNS 421, 421 (2019) (Asserting, among other claims, that "white empiricism. . . harms the people who are othered.").

10 *See generally* Peter Pringle, The Murder of Nikolai Vavilov: The Story of Stalin's Persecution of One of the Greatest Scientists of the Twentieth Century (2008).

11 *See id.*

12 Bruce Gilley, *The Case for Colonialism,* Third World Q., April 17, 2018, at 167.

13 Roland G. Fryer Jr, *An Empirical Analysis of Racial Differences in Police Use of Force: A Response,* 128 J.P.E. 128 (10) 4003-4007; Pascaline Dupas et al., Seminar Dynamics Collective, Gender and Dynamics of Economics Seminars (2021).

Overall, our paper makes an initial attempt to formalize the current problem facing or soon to face researchers in all social science fields. We attempt to provide some direction on how to operate under the constraints of this new environment in a way that still permits some learning about relevant determinants of social outcomes. Further work in these areas is absolutely essential. We plan, for future work, to systematically analyze the constraints facing economists in particular, and how these constraints may have evolved over time.

Much more work will also be needed to understand the methods that can and will be used to engage in research under political constraints, which we also plan to pursue. Further, we must develop an understanding of how to consume research in this environment. Researchers must scrupulously avoid generating results that violate political constraints, and as such observing a series of papers "confirming" political orthodoxy may provide no information about reality at all; given the number of possible specifications, interpretations, and research designs available to address a question, we will always see such a flow of apparently confirming work, while contrarian work will either not be completed or will be hidden from view. Consumers of research must learn to update based on the flow of confirming information relative to the expected flow of confirming information if orthodox beliefs are in fact true, a daunting Bayesian updating task.

I. MODEL

We now introduce our model. We adopt the sender-receiver framework from Bayesian persuasion research. In these models, there is a *researcher* who designs an experiment.[14] The experiment then generates a signal for the *receiver*, who uses the signal to update his beliefs about the true state of the world. Crucially, and in contrast to the "cheap talk" models that follow, the researcher cannot arbitrarily select a signal. Instead, the researcher ex ante commits to a research design that sends a particular signal as a function of the true underlying state, which is not observed by any players in the game. So, for example, the receiver may be interested in the temperature of some liquid, and the researcher would develop a temperature strip that turned a certain color if the temperature exceeded some level determined by the researcher. The researcher cannot lie about the temperature, but he can influence the posterior beliefs of the receiver by choosing the temperature threshold. These models are growing in popularity in economics because they capture the ability of a research designer to influence the beliefs, and therefore policy choices, of other agents without the technical difficulties associated with cheap talk models. These models also separate "persuasion" from deception in a useful manner. For our purposes, we adopt these models

[14] *See generally* Emir Kamenica & Matthew Gentzkow, *Bayesian Persuasion*, 101 AM. ECON. REV. 2590-2615 (2011).

because they can capture the decision process of an honest researcher seeking to positively influence policy in the face of consumers of research who have arbitrary, and potentially quite extreme, beliefs. The model also allows for a natural extension to formally introduce the concept of "cancellation" into the research environment, as described below.

There are three states ($\{A, B, C\}$) and three possible policies ($\{Y, N, HN\}$). The researcher (sender) needs to ex ante design an experiment to reveal additional information about the state, after which the polity (receiver) will decide on the policy. All agents have the same objectives, but there are two possible types of receiver. The receiver may be normal, or the receiver may be *woke*. The normal receiver and the sender put prior probability $\frac{1}{2}$ on state A and $\frac{1}{4}$ on each of states B and C. The woke receiver has a degenerate prior, where $P(A) = 0$, and the prior is uniform over the remaining two states. Crucially, the woke receiver engages in *cancellation* if forced to divide by zero; that is, should the receiver design an experiment that, with positive probability, rules out both B and C, the sender receives a payoff of $-\infty$ in that state. We confine the sender to at most a two piece partition to reflect limitations on the specificity of the research to be completed.

The payoff matrix, common to all players, is:

	A	B	C
Y	1	0	0
N	0	1 + a	1 + b
HN	d	c	d

where

$$a < 1$$
$$b > 1$$
$$c < 2$$
$$c > 1 + a$$
$$d < 0$$
$$|d| > |c|.$$

We first consider the two information policies that do not risk cancellation; $\{\{A, B\}, \{C\}\}$ and $\{\{A, C\}, \{B\}\}$. Under either of these policies, the woke receiver never has to confront information that he cannot accept, so the sender is safe from cancellation. Comparing the payoffs from these two possible research designs reveals that, when the receiver is known to be normal, the sender prefers $\{\{A, B\}, \{C\}\}$. To see this, observe first that $\{\{A, B\}, \{C\}\}$ implements $\{Y, N\}$:

$$\underbrace{\frac{2}{3}+0}_{\text{Payoff for Y after \{A,B\}}} > \underbrace{\frac{1}{3}+\frac{a}{3}}_{\text{Payoff for N after \{A,B\}}}$$

which clearly holds for $a < 1$, while $\{\{A, C\}, \{B\}\}$ implements $\{N, HN\}$:

$$\underbrace{\frac{2}{3}+0}_{\text{Payoff for Y after \{A,C\}}} < \underbrace{\frac{1}{3}+\frac{b}{3}}_{\text{Payoff for N after \{A,C\}}}$$

which holds since $b > 1$. We can then calculate the expected payoff for each research design:

$$\{A, B\}, \{C\}: \frac{1}{2} + 0 + \frac{1}{4} + \frac{b}{4}$$

$$\{A, C\}, \{B\}: 0 + \frac{c}{4} + \frac{1}{4} + \frac{b}{4}$$

and thus $\{\{A, B\}, \{C\}\}$ is preferred since $c < 2$.

We can now consider whether the risk of cancellation inhibits the flow of information in this scenario. If there is no risk that the sender would face a woke receiver, the sender would have the option of providing the partition $\{\{A\}, \{B, C\}\}$. Note that it is immediate that the $\{B, C\}$ message implements N since $|d| > |c|$, and the $\{A\}$ message implements Y. Thus, the $\{\{A\}, \{B, C\}\}$ design is preferred to the $\{\{A, B\}, \{C\}\}$ if

$$\frac{1}{2} + \frac{1}{4} + \frac{a}{4} + \frac{1}{4} + \frac{b}{4} > \frac{1}{2} + 0 + \frac{1}{4} + \frac{b}{4}$$

which clearly holds since $a > 0$.

Since we already know that $\{\{A, B\}, \{C\}\}$ is preferred to $\{\{A, C\}, \{B\}\}$, we can conclude that in the absence of the woke receiver the $\{\{A\}, \{B, C\}\}$ design is preferred. However, as long as there is a positive probability of a woke receiver, the threat of cancellation makes such a research design unacceptably risky. Thus, we have captured the direct effect of wokeness on the effectiveness of research; researchers will tilt their designs to avoid revealing information that will upset a woke audience and will as such prefer to do research that generates ambiguous rather than clear answers.

A second, more subtle and insidious, effect also manifests itself in this example. To explore this other effect, let us designate the probability of facing a woke receiver as w. As long as $w > 0$, the only feasible experiments to use are $\{\{A, B\}, \{C\}\}$ and $\{\{A, C\}, \{B\}\}$. The first experiment is superior when the receiver is normal. But, if the receiver is woke the first experiment implements $\{HN, N\}$ since the woke receiver refuses to consider the

possibility of A. Thus, we can determine the threshold for the probability of a woke receiver under which the researcher must further distort his experiment in order to avoid a bad outcome from the misinterpretation of the experiment by the woke receiver. This fear of a bad outcome is different than the fear of cancellation; the researcher has insulated himself against cancellation, but the dogmatic prior of the woke mean that the lesson they take from the research might result in what the researcher views as a very bad decision since the woke are unwilling to consider the full range of possible states. This distortion occurs if

$$w\left(\frac{d}{2}+\frac{c}{4}+\frac{1}{4}+\frac{b}{4}\right)+(1-w)\left(\frac{3}{4}+\frac{b}{4}\right)<\frac{c}{4}+\frac{1}{4}+\frac{b}{4}$$

which gives $w>\frac{2-c}{2-c-2d}\in(0,1)$. Here, the risk of providing a reasonably useful experiment (i.e. $\{\{A,B\},\{C\}\}$) is that, should the receiver be woke, he will interpret the first outcome as being B for sure, implementing a policy that is disastrous in the more likely case where the true state is A. Thus, in addition to avoiding the most valuable experiment available out of fear of direct repercussions, the sender further distorts the available information out of fear of the endogenous response of the woke receiver. Thus, the presence of the woke receiver creates two distortions in the ability of the sender to communicate with the normal receiver; for a sufficiently high risk of a woke audience or a sufficiently bad outcome when the woke are in charge, we end up with the sender effectively acting as if his policy preferences do not align with the normal receiver; the sender in effect acts as if he is trying to Bayesian persuade the receiver to implement the policy N in as many states as possible.

In this somewhat extreme case we have explored, the entire information exercise is of limited use for the normal receiver when w is above the noted threshold; the normal receiver will implement the same policy that he would if given no information $\frac{3}{4}$ of the time. Such an outcome is not a general result; it is straightforward to tweak the game to have two very similar policies but where the limited information available in the presence of the woke receiver may still be useful in a larger number of states when the receiver is normal. For example, we could augment the policy choice set with another policy that is similar to N, as described below, with $\delta<\varepsilon$:

	A	B	C
Y	1	0	0
N	0	1 + a	1 + b
HN	d	c	d
NU	$0+\varepsilon$	$1+a-\delta$	$1+b-2\varepsilon+\delta$

For sufficiently small $\varepsilon > 0$ and $\delta > 0$, none of the main calculations above change when comparing N or NU to the other policies for the purposes of designing information. If, however, $w > \frac{2-c}{2-c-2d} \in (0,1)$ so that the sender chooses the design $\{\{A, C\}, \{B\}\}$, the normal receiver prefers HN when the state is B, but now prefers NU over N when the state could be A or C. Thus, we have found a scenario where the risk of the presence of a woke receiver leads to a costly distortion of the information environment, but where careful information design can still convey some useful information to the normal receiver in all states. In this particular example, the prejudices of the woke receiver makes it impossible to get him to implement even the marginally more efficient policy; here again since he refuses to consider the possibility that state A may have occurred he sticks with N even though the cost of NU in the C state is lower than the benefit of NU in the A state.

A. *Giving The Woke Their Due*

We have so far proceeded under the assumption that the woke receiver has a distorted view of the world. In the interest of objectivity, we should consider the possibility that the woke may, in fact, be correct about the prior probabilities of the different states. Thus, we now take an agnostic approach to who is correct about the prior and investigate the implications of the potential presence of a woke receiver on the value of research. In this section, we formally consider the prior probability placed on A to be slightly positive and take that probability to zero. This section will also highlight the crucial difference between introducing a woke receiver and a model that simply has heterogenous priors among potential receivers. We revert to the original version of the game here.

We will now calculate the subjective value of the experiments available when the presence of a woke receiver is taken into account versus when it is not. We consider two cases: First, consider the case where there is only a small probability of a woke receiver, $w < \frac{2-c}{2-c-2d}$. We will first compare the value from the perspective of the receiver if he is able to ignore the wokeness constraint even in the presence of the woke receiver, and we will compare it to the value of the optimal information design taking into account the wokeness constraint. Unsurprisingly, the sender perceives the wokeness constraint as costly.

If the sender ignores the woke constraint, then he uses $\{\{A\}, \{B, C\}\}$. This is the optimal policy to follow because it was shown to be the preferred information design in the absence of a woke receiver, and with the infinitesimal prior by the woke receiver it generates identical behavior on the part of the woke receiver. Thus, the payoff perceived by the sender for using this partition is simply

$$1 + \frac{a+b}{4}$$

since this design implements Y for A and N for B and C. Taking into account the wokeness constraint, where $\{A\}$ cannot be allowed to occur, the sender instead uses $\{\{A,B\},\{C\}\}$ since we are in the low w regime. In this case, the actions will be different depending on whether the receiver is woke or normal, giving a payoff of

$$\frac{3}{4} + \frac{b}{4} + w\left(\frac{d}{2} + \frac{c}{4} - \frac{1}{4}\right).$$

So, subtracting this second value from the first, we see that from the sender's perspective the value of being able to ignore the woke constraint, even in the presence of the woke receiver, is

$$\frac{1}{4} + \frac{a}{4} + \frac{w}{4} - w\left(\frac{2d+c}{4}\right)$$

where the last term is positive since $d < -c$. So, perhaps unsurprisingly, the sender would prefer to ignore the constraints imposed by wokeness when there are sufficiently few woke receivers.

We now consider the issue from the perspective of the woke. If the sender is not constrained by the cancellation risk, the woke perceives the payoff as

$$1 + \frac{a+b}{2};$$

The woke is certain that the true state will be either B or C, thus the design $\{\{A\},\{B,C\}\}$ implements policy N with probability 1, with states B and C being equally likely. If, however, the sender takes into account the cancellation risk and uses $\{\{A,B\},\{C\}\}$, the woke perceives an expected payoff of

$$w\left(\frac{c}{2} + \frac{1}{2} + \frac{b}{2}\right) + (1-w)\left(\frac{1}{2}0 + \frac{1}{2} + \frac{b}{2}\right)$$

which gives

$$\frac{1}{2} + \frac{b}{2} + \frac{cw}{2}.$$

So, the woke in fact prefers for the sender to ignore the cancellation constraint if

$$w < \frac{1+a}{c},$$

but this always holds in this case since $w < \frac{2-c}{2-c-2d}$ and $\frac{2-c}{2-c-2d} < \frac{1+a}{c}$, as we now show:

$$\frac{1+a}{c} \overset{?}{>} \frac{2-c}{2-c-2d}$$

$$1 - \frac{c-(1+a)}{c} \overset{?}{>} 1 - \frac{-2d}{2-c-2d}$$

$$\frac{-2d}{2-c-2d} \overset{?}{>} \frac{c-(1+a)}{c}$$

$$-2dc \overset{?}{>} c(2-c-2d) - (1+a)(2-c-2d)$$

$$0 \overset{?}{>} c(2-c) - (1+a)(2-c-2d)$$

$$(1+a)(2-c-2d) \overset{?}{>} c(2-c)$$

$$(1+a)(2-c) - 2d(1+a) \overset{?}{>} c(2-c)$$

$$(2-c)(1+a-c) \overset{?}{>} 2d(1+a)$$

$$(2-c)(c-(1+a)) \overset{?}{<} -2d(1+a).$$

But, this last line is clearly valid because $-2d > 2 > 2-c$ since $-d > c > 1+a > 1$ and $c-(1+a) < 1+a$ since $(c-1)-a < 1$ from the fact that $c < 2$ and $a > 0$, while $1+a > 1$.

Notably, all players in the game, the normal sender, the normal receiver, and the woke receiver, would profess a preference for, in effect, free speech. Each player prefers the $\{\{A\},\{B,C\}\}$ experiment. The sender and the normal receiver prefer this experiment because it allows the normal receiver to distinguish between the states where Y and N are good policies, and it forces the woke receiver to confront the information that he largely refuses to believe, thus implementing what they believe to be the right policies. In contrast, the woke receiver values the $\{\{A\},\{B,C\}\}$ policy because it will convince the normal receiver of the truth of the woke sender's priors, thus preventing the implementation of the Y policy. Note this distinction between a model with a woke receiver and a model that simply has heterogeneous priors. In the latter model, $\{\{A\},\{B,C\}\}$ would be used, and this decision would be perceived as optimal by all players in the game. It is only the

specific characteristic of the woke receiver that he "cancels" the sender if the information does not comport with his prior, which leads to the distortion. Thus, in a sense, the presence of a small number of woke receivers is self-defeating; their presence will lead the researcher to tilt information design away from any setup that could confirm their view of the world due to fear on the part of the researcher.

This result will not always hold. Consider now the case where $w > \frac{2-c}{2-c-2d}$ such that the sender facing the woke constraint will chose design $\{\{A, C\}, \{B\}\}$. As before, the payoff to the sender for the unconstrained design of $\{\{A\}, \{B, C\}\}$ is

$$1 + \frac{a+b}{4}$$

but now the payoff for the constrained design is simply

$$\frac{c}{4} + \frac{1}{4} + \frac{b}{4};$$

both types of receiver chose policies N for $\{A, C\}$ and HN for B. Thus, again unsurprisingly, the sender would prefer not to face the cancellation constraint, since $c < 2$.

From the perspective of the woke receiver, however, matters are different. Under the $\{\{A\}, \{B, C\}\}$ design, the payoff is again

$$1 + \frac{a+b}{2}.$$

Under the $\{\{A, C\}, \{B\}\}$ design, both the woke and normal receiver chose N for $\{A, C\}$ and HN for B. Thus, from the woke perspective, the payoff to this design is simply

$$\frac{1}{2} + \frac{b+c}{2}.$$

Thus, the woke prefer the $\{\{A, C\}, \{B\}\}$ design since $c > 1 + a$; this design allows the woke to distinguish between B and C, which he views as the only useful piece of information, and at the same time this design persuades the normal receiver to implement the N policy when B does not occur. Thus, what is perceived as information destruction resulting from the presence of the woke is viewed by the woke as inducing better policy outcomes. Thus, in this case, where the woke have grown to be a sufficiently high proportion of the receivers, they will now view the cancellation threat as a positive good.

This model thus catches some of the evolution of the attitude towards free inquiry that we have observed in recent times. When the woke were not dominant, they played lip service to the value of free inquiry but reacted with hostility to any results that did not comport with their preconceived views, thus catching many well-meaning but naive researchers by surprise. Now, as they have become dominant, any pretense of favoring free inquiry has been dropped and they are up front with the demands that research conform to their predetermined conclusions.

II. DISCUSSION

The above model is quite abstract, but it captures a number of important considerations facing practical empirical economists going forward. Researchers distort the design of their research to avoid angering woke academics and activists, and thus the most efficient research projects are not completed. Research design may be further distorted by the need to avoid misinterpretation of the results by woke receivers, whose distorted priors may lead to extremely poor outcomes when faced with new information. Note that our baseline setup implies, in effect, a correlation between distorted priors and the tendency to engage in "cancellation" behavior when faced with information that the receiver is unwilling to believe. This correlation drives the nature of the distortions in research design. In principle, agents with more reasonable or flatter priors could instead be the ones to attack researchers when information comes in that leads to updating away from, say, the mode of the prior distribution. We view the correlation between cancellation activity and extreme priors as a reasonable way to capture what we observe; it is somewhat difficult to think of cases where people who took an agnostic stance toward an issue reacted with extreme hostility to information supporting one side or the other of the issue, whereas there has been a proliferation of such behavior among those with strong prior beliefs.[15]

In this environment, we note that the researcher still has some choice as to what "experiments" to run, and thus there remains a role for research in revealing *some* valuable information. Dogmatic adherence to established orthodoxy can be very costly, but a researcher does not do anyone much good if he runs afoul of orthodox ideas. A famous case from botany provides an excellent example of this challenge. Troffim Lysenko developed a pseudo-scientific theory of plant changes that rejected Mendelian genetics, and his theory became favored within the Soviet Union under Stalin. Research by Lysenko and his followers continually confirmed his theories, basically through outright fabrication, which is admittedly outside the scope of our model. Nikolai Vavilov, by contrast, continued to produce legitimate

[15] *See generally* David Acevedo, *Tracking Cancel Culture In American Higher Education*, National Association of Scholars (2021), https://www.nas.org/storage/app/media/New%20Documents/ academic-cancellations-updated-may-6-2021.pdf.

botanical research which contradicted Lysenko's theories. Since Vavilov ended up dying in prison for his resistance to Lysenko, his research ended up being just as ineffective at preventing famine in the Soviet Union as Lysenko's was. What was needed at the time was research designs that would have allowed Soviet scientists to continue to accept Lysenkoism while also learning something relevant about how to maintain crop yields in this constrained environment. In the context of our model, Vavilov consistently ran the $\{\{A\},\{B,C\}\}$ experiment to disprove Lysenkoism, when he needed to be running $\{\{A,C\},\{B\}\}$ to provide some useful policy advice while not directly challenging Lysenko.[16]Ample similar stories are available with respect to the Great Leap Forward and the Cultural Revolution in Maoist China.[17]

As a practical matter, how can researchers achieve such a goal, protecting their ability to produce research while providing at least marginally useful information? In the next section, we consider methodological challenges in the context of a less, but still largely, abstract model of research under woke constraints. First, we consider the possibility of using a pre-registration system to work around these constraints, and we argue that this simple approach is unlikely to address the issues we have identified in our model. We then take an extremely simple example of how woke constraints may impede research and consider what practices might actually work to achieve the goal of providing useful information under such constraints.

III. EMPIRICAL ANALYSIS UNDER CONSTRAINTS: MODEL

A. *Pre-Registration*

The analysis in section I subsection A suggests a highly tempting approach that would appear to entirely solve the woke constraint when the woke population is small. Ex ante, all players agree on the optimal design of the experiment, and therefore in principle "pre-registration" of design should solve the problem of potentially confronting a woke receiver. This solution is appealing because such pre-registration is broadly accepted as best-practices for other reasons relating to the correct implementation of the scientific method. A journal could have a policy of requiring research designs to be submitted up front and published prior to the analysis, with an agreement that withdrawals of papers, misconduct investigations (in coordination with universities) and related penalties could only be applied based on the up front research design and would only be entertained prior to

16 Jan Witkowski, *Stalin's War On Genetic Science*, 454 NATURE 577-79 (2008) (discussing Pringle's The Murder of Nikolai Vavilov).

17 *See generally* JOSHUA EISENMAN, RED CHINA'S GREEN REVOLUTION: TECHNOLOGICAL INNOVATION, INSTITUTIONAL CHANGE, AND ECONOMIC DEVELOPMENT UNDER THE COMMUNE (2018).

the production of the actual results. Under our low *w* regime, such a plan would be fully supported by all parties, woke or not. That is, the confidence the woke have in their beliefs would lead them to agree to such constraints since they would expect an appropriate design to confirm their beliefs.

Such a procedure may indeed have a place, but details of the Critical Theory approach to knowledge mean it is not the panacea that the model may appear to imply. If the objective of the exercise is to convince a well-meaning but naive observer to accept results and refrain from penalizing the researcher, such pre-registration may be valuable. If the objective is to prevent an attack by the woke or to potentially persuade the woke, the process is likely doomed to failure.

To see why, it is necessary to understand that the Critical Theory approach to knowledge does not admit such precommitments to researh design. Knowledge is gained through lived experience and the validity of knowledge is determined by the identity of the individual not by the appropriateness of the empirical procedure.[18] Further, the ethical system underpinning Critical Theory asserts that the only core moral imperative is to struggle by any means to bring power to pre-designated marginalized groups.[19] Thus, if the woke agreed to accept the outcome of a pre-designed research program, and that program then delivered conclusions that did not support the pre-determined truth from the lived experience of a marginalized woke person, then the woke would reject the agreed upon design after the fact and proceed with an attack upon the researcher, all while feeling fully morally justified.

Thus, such a procedure would only serve to potentially reveal the bad-faith actions of woke receivers; the ex ante agreement would potentially protect a researcher when adjudication of complaints is completed by a well-meaning administrator. Empirically, this situation seems rare[20] and in fact in many cases university administrators appear to tolerate or encourage mobs motivated by pre-established conclusions.[21] In light of these issues, we will need to consider other alternatives to pre-registration in order to provide researchers the opportunity to continue to produce useful research under constraints, including constraints enforced by mob violence.

---

18 *See* Sandra Harding, *Rethinking Standpoint Epistemology: What is "Strong Objectivity?*, 36 THE CENTENNIAL REV. 437, 443 (1992); *see also supra* note 1 and accompanying text.

19 *See* Kimberle Crenshaw, *Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color*, 43 STAN. L. REV. 1241, 1258 (1991).

20 Peter Beinart, Editorial, *A Violent Attack On Free Speech At Middlebury*, THE ATLANTIC (Mar. 6, 2017), https://www.theatlantic.com/politics/archive/2017/03/ middlebury-free-speech-violence/518667/.

21 *See* Bret Weinstein, Opinion, *The Campus Mob Came For Me—And You, Professor, Could Be Next*, WALL ST. J., (Mar. 30, 2017), https://www.wsj.com/articles/ the-campus-mob-came-for-meand-you-professor-could-be-next-1496187482.

B. *Political and Auxiliary Coefficients*

We will start by introducing the concepts of political and auxiliary coefficients, which will play a central role in the remainder of our analysis. These are somewhat novel concepts from the perspective of statistics and econometrics, so in this section we define exactly what we mean by these concepts and give an extremely simple example of the issues a researcher faces when addressing a research design that includes both types of coefficients.

A *political* coefficient is a coefficient which has constraints on its value which are determined by some subset of the audience for the research. Crucially, these constraints do not need to correspond to objective reality, and as such standard data analysis practices are not guaranteed to lead to estimates that satisfy such constraints, even for correctly specified models under asymptotic analysis. The most natural example of such a constraint would be in the analysis of discrimination. In the current environment in many fields, and likely in the future in all fields, any coefficient that clearly measures discrimination against what are described as marginalized groups must find evidence of discrimination. Violations of such constraints will often lead to serious personal consequences.[22]

An *auxiliary* coefficient is a coefficient that is of interest for policy purposes but is not yet constrained by the political environment. These coefficients will be the main object of interest for us; the fundamental question at play is how to learn and communicate as much as possible about auxiliary coefficients while avoiding violating constraints on political variables. An example of such coefficient would be the effect of blood pressure on heart disease. Obesity is a clear risk factor for heart disease,[23] but a researcher in an environment where Fat Studies has achieved institutional power cannot find a link between obesity and heart disease but could still identify a link between blood pressure and heart disease.[24] This example also shows the subtlety of the political constraints; in certain environments it may be possible to express the link between blood pressure and heart disease as long as obesity is not discussed. In more severely constrained environments such a discussion may be impossible because of the link between obesity and blood pressure. Researchers will need to become experts in parsing out the exact nature of the political constraints they face; one of our goals here is to develop a coherent framework to facilitate the study and understanding of these constraints so that researchers can continue to provide some value to society wherever possible.

---

22 *See* Salvatore Carbone et al., *Obesity Paradox In Cardiovascular Disease: Where Do We Stand?*, 15 VASCULAR HEALTH & RISK MGMT., 89, 89 (2019).

23 *Id.*

24 Marilyn Wann, FAT STUDIES: AN INVITATION TO REVOLUTION. THE FAT STUDIES READER, xi (2009); Anna Mollow, *Disability Studies Gets Fat*, 30 HYPATIA 199–216 (2015).

We now consider a very simple, abstract example of where one coefficient is political while another is auxiliary, and what issues may arise when trying to communicate information about the auxiliary coefficient. Consider a *researcher* who observes $N > 3$ draws from a normal distribution with mean μ and variance σ^2. The mean μ is a political variable such that any analysis must find that $\mu > \bar{\mu}$, while σ is an auxiliary variable that is relevant for determining optimal policy. The researcher can communicate information to the *reader*, but he cannot reveal the actual data observed; we can think of this as a technological constraint; while publishing a series of draws from a normal distribution and relying on the reader to compute the average himself is not actually hard, our simple example is intended to capture effects that may be at play when the required data analysis is more complex than simple averaging. Further, in some cases the researcher may be prohibited from actually disclosing the data, or disclosing the data itself may be considered a political act. More generally, the primary point of empirical research is to project data onto a space that can be easily communicated to some audience. For simplicity, then, we assume that the researcher is able to communicate by reporting a point in R^n, $n \leq 3$. Informally, we allow the researcher to report some estimate of the mean and the standard deviation, along with a single, one dimensional, piece of auxiliary information.

This informal interpretation maps to the formal structure of the communication game because there are multiple types of readers. An *unsophisticated* reader automatically interprets the first element of the reported vector as the mean and the second element as the standard deviation, ignoring the third element, if applicable. Thus, when communicating when the reader may be unsophisticated, the game is not one of cheap talk, as long as the researcher's objective function depends, directly or indirectly, on the beliefs of the unsophisticated reader. We can thus proceed by describing the researcher as reporting a mean and a standard deviation, even though to a non-unsophisticated (i.e. *sophisticated*) reader these quantities do not necessarily represent such estimates.

Putting aside this formality, we now focus on a fundamental challenge of scientific communication under political constraints. Suppose the researcher seeks to communicate the best estimate of the standard deviation of the distribution in question. He must report a mean of at least $\bar{\mu}$. Suppose the mean of his sample is in fact $\hat{\mu} < \bar{\mu}$. Then, he may find it optimal to report a mean of $\bar{\mu}$.[25] He then has effectively two options to report as an estimate of σ^2: $\hat{\sigma}^2 \equiv \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \hat{\mu})^2$ or $\bar{\hat{\sigma}}^2 \equiv \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{\mu})^2$.[26] The estimate

[25] Clearly the optimal report depends on other features of the environment, but for now we take this simply strategy as given. It will be optimal, for example, if the unsophisticated reader is the dominant audience and the loss function of the researcher is increasing in the reader's error.

[26] It is possible that the second estimator should have $\frac{1}{N}$ instead of $\frac{1}{N-1}$ since there is no degree of freedom issue, as would be the case with a known mean. We neglect this subtlety as it does not seem to be a solvable question and the $\frac{1}{N-1}$ choice slightly simplifies some calculations.

$\hat{\sigma}^2$ has the advantage of reporting the best estimate of the variance, but it contains a contradiction; by using the computed mean instead of the assumed mean, the researcher is acknowledging the dishonesty of the μ report; there is no scenario in which the reported μ and the reported σ estimates can be outcomes of a legitimate estimation procedure. The $\bar{\hat{\sigma}}$ estimator, on the other hand, is exactly what would be obtained if the $\bar{\mu}$ constraint were a valid physical or economic, rather than political, constraint.

The researcher faces both an ethical and practical dilemma. Ethically, should the researcher concede the political constraint while trying to report the auxiliary information as accurately as possible, when such an action will mislead a reader about the nature of the data-generating process since the researcher is reporting his own results falsely? Or, should the researcher, when forced to use a constraint, simply acknowledge that constraint and proceed as if it is a legitimate constraint? In the second case, a sophisticated reader will be unable to invert the reported μ and σ to obtain the true best estimate of σ; even if the researcher signals that $\mu < \bar{\mu}$ by reporting exactly $\bar{\mu}$, this is insufficient information to learn the true best estimate of σ.

An alternative approach is to exploit the third dimension of the signal that we have provided the researcher. A researcher can report the politically required mean, the correctly estimated variance, $\hat{\sigma}$, and a scale factor of how much the variance was scaled from $\bar{\hat{\sigma}}$. We will define that scale factor as γ. This correction factor will need to be named and explained away in some clever and reasonably opaque way such that the sophisticated, normal (i.e. non-woke) reader will be able to understand what is occurring without having the researcher run afoul of the woke reader. A challenge here is that $\hat{\sigma}$ and γ can be inverted to learn $\hat{\mu}$. Specifically, since

$$\gamma = \frac{\sum_{i=1}^{N} (x_i - \hat{\mu})^2}{\sum_{i=1}^{N} (x_i - \bar{\mu})}$$

we obtain the quadratic equation

$$\bar{\hat{\sigma}} = \frac{1}{\gamma}\bar{\hat{\sigma}} + 2\frac{N}{N-1}\hat{\mu}\bar{\mu} - 2N\hat{\mu}^2 - \frac{1}{N-1}\hat{\mu}^2. \quad (1)$$

Revealing γ thus leads to at most two candidates for $\hat{\mu}$, and in certain cases only one of those candidate values will be real and fall below $\hat{\mu}$.

Sufficient transparency in methodology, at least in this case, undermines the opacity needed to avoid violating political constraints. If an audience consisted of sophisticated, normal readers and unsophisticated woke receivers, the three dimensional signal is ideal; cancellation is avoided due to the dogmatic interpretation of the report, and both the best variance and best mean estimate are effectively communicated. More realistically, the audience will be a mixture of sophisticated and unsophisticated and woke and normal receivers, with representation from each quadrant. Thus, the researcher must trade off the risk of cancellation from revealing too much

information that goes against the woke prior, versus failing to convey sufficient useful information to normal receivers. So, for example, here the researcher could add noise to the report of γ, or report something above $\hat{\mu}$ when the political constraint binds, thus breaking the invertability of equation 1. For example, the researcher, instead of reporting γ, could simply report that an adjustment was made to provide a more accurate estimate of σ given the errors in data observed by the researcher. Thus, the report is that $\gamma \neq 1$, which conveys that the estimates have been adjusted to report the true $\hat{\sigma}$, but the information needed to recover $\hat{\mu}$ is not provided. Such obfuscation may be sufficient to avoid cancellation, as the woke would have to believe that any data leading to $\hat{\mu} < \bar{\mu}$ must in fact be erroneous. While this may seem implausible, specialists in areas dominated by woke researchers have demonstrated a significant tendency to accept implausible claims about data in order to confirm their prior beliefs.[27]

A key point in this example, which will hold in practice in more complex settings, is that the political and auxiliary coefficients will be entangled in ways that complicate the appropriate method for communicating research results. In cases where the computation of the political and auxiliary coefficients are independent, the problem is simple, but such a situation is seldom encountered. In the mean and variance case, the constraint on the mean will necessarily inflate the variance calculation when the constraint binds, assuming the constraint is taken as true. Similarly, the degree of adjustment needed for the variance estimate reveals information about the true mean. Related issues will arise in virtually any methodology that relies on analysis of correlations among a set of potential explanatory variables. We will now consider several of those approaches, how they can be modified to address political constraints, and the resulting issues in interpretation and researcher safety when using such methods. But first, given this conceptual framework, we will describe real world cases which illustrate the constraints and responses to these constraints.

IV. INTRODUCTION OF CRITICAL RACE THEORY APPROACHES INTO EMPIRICAL ECONOMICS

While it is beyond debate at this point that certain fields have ceased to develop or maintain actual intellectual standards in favor of a purely

[27] *See* Helen Wilson, *Human Reactions To Rape Culture And Queer Performativity At Urban Dog Parks In Portland*, 27 GENDER, PLACE & CULTURE 307 (2020) (retracted).

activistm approach to academic inquiry,[28] economics and allied fields are still relatively early in this process.[29]

Fortunately, we can observe how the process unfolds by considering examples from relatively closely aligned disciplines and then see how the situation will likely play out in economics. A recent, relatively well documented case from political science is particularly useful.[30] Gilley also evaluated the potential benefits from historical colonial episodes and called for a more data-driven approach to determining the effects of colonial institutions:

> The failure of anti-colonial critique to come to terms with the objective benefits and subjective legitimacy of colonialism points to a third and deeper failure: it was never intended to be "true" in the sense of being a scientific claim justified through shared standards of inquiry that was liable to falsification.[31]

This approach and the general openness to the possibility that colonial institutions benefited local populations led to the following situation, in the words of the editor of *Academic Questions*:

It provoked enormous controversy and generated two separate petitions signed by thousands of academics demanding that it be retracted, that *TWQ*

---

[28] *See, e.g.*, *id.*; Richard Baldwin, *Who Are They To Judge? Overcoming Anthropometry Through Fat Bodybuilding*, 7 (2018) [hereinafter Baldwin, 2018] (retracted); M. Smith, *Going In Through The Back Door: Challenging Straight Male Homohysteria, Transhysteria, And Transphobia Through Receptive Penetrative Sex Toy Use*, 22 <<Sexuality & Culture>> 1542 (2018) (retracted); Richard Baldwin, When The Joke Is On You: A Feminist Perspective On How Positionality Influences Satire (2018) (accepted in <<Hypatia J. Feminist Phil.>>, then retracted) (on file with author); Carol Miller, Moon Meetings And The Meaning Of Sisterhood: A Poetic Portrayal Of Lived Feminist Spirituality (2018) (accepted in <<J. Poetry Therapy>>, then retracted) (on file with authors); Maria Gonzalez & Lisa A. Jones, Our Struggle Is My Struggle: Solidarity Feminism As An Intersectional Reply To Neoliberal And Choice Feminism (2018) (accepted in <<Affilia: J. Women & Soc. Work>>, then retracted) (on file with author); Maria Gonzalez, The Progressive Stack: An Intersectional Feminist Approach To Pedagogy (2018) (third round reject and resubmit in <<Hypatia J. Feminist Phil.>>, then rejected after hoax revealed) (on file with author); Stephanie Moore, Super-Frankenstein And The Masculine Imaginary: Feminist Epistemology And Superintelligent Artificial Intelligence Safety Research (2018) (revise and resubmit in <<Feminist Theory>>, then rejected after hoax revealed); Richard Baldwin & Brandon Williams, Agency As An Elephant Test For Feminist Porn: Impacts On Male Explicit And Implicit Associations About Women In Society By Immersive Pornography Consumption (2018) (in review but rejected by <<Porn Studies>> after hoax revealed) (on file with author); James Lindsay, Peter Boghossian, & Helen Pluckrose, Project Fact Sheet (2018) (on file with authors).

[29] Papers cited *supra* note 27 are examples from the "Grievance Studies Affair," where non-specialist academics were, after very limited study, able to have seven papers in various critical social justice fields accepted over the course of a single year, with another 3 papers at some stage of the revise/reject and resubmit process. Part of the document summarizing the project is included in the appendix with the permission of the authors. Readers who question the unequivocal statement about failure to maintain intellectual standards are encouraged to follow the link to the full document.

[30] Bruce Gilley, *The Case For Colonialism*, 31 ACADEMIC QUESTIONS 167 (2018) [hereinafter *The Case*]; *see also* Bruce Gilley, *The Case For Colonialism*, THIRD WORLD Q., Sept. 2017.

[31] *Gilley*, *The Case*, *supra* note 29, at 173.

apologize, and that the editor or editors responsible for its publication be dismissed. Fifteen members of the journal's thirty-four-member editorial board also resigned in protest. Publisher Taylor and Francis issued a detailed explanation of the peer review process that the article had undergone, countering accusations of "poorly executed pseudo-'scholarship,'" in the words of one of the petitions. But serious threats of violence against the editor led the journal to withdraw the article, both in print and online. Gilley was also personally and professionally attacked and received death threats.[32]

Thus, a paper that simply reviewed evidence and called for a more scientific approach to a question was removed at the behest of academics. The conclusion that colonialism was always and everywhere a negative for local populations is, in the parlance of the previous section, a political variable, and Gilley's error, while trying to put forward an intellectual framework that could lead to improved governance in low-income countries, was to note the possibility of that variable not corresponding to the assumed value. Had he concluded, in spite of evidence, that colonialism always and everywhere harmed local populations, he could have potentially proceeded with an analysis that showed the benefits of certain colonial-based institutions when reintroduced in modern forms (the auxiliary variable).

Political science in many ways is the closest social science to economics; researchers use data and attempts at logical arguments to understand interactions among people. Thus, the fact that conclusions are now in at least some cases pre-determined in this field, with threats of violence used to enforce those conclusions, should raise concerns among economists.

Here, we consider two papers, by Fryer Jr. and Dupas et al., to describe how this situation may move into economics.[33] The first paper is the controversial police-shooting paper that appears to establish that, in the police department covered in the paper, there is effectively no evidence of a racial bias in police shootings.[34] In an uncharacteristically weak criticism of the paper, Durlauf and Heckman argue, among other things, that a selection problem undercuts the argument in the paper since officers may be engaging in racially motivated stops.[35] But, Fryer Jr. convincingly refutes this selection claim by pointing out, as was obvious from the original paper or from basic common sense, that the overwhelming majority of police shootings are a result of a 911 call for service, where selection concerns are notably weaker.[36]

---

32 *Id.* at 167.

33 Roland G. Fryer Jr., *An Empirical Analysis Of Racial Differences In Police Use Of Force*, 127 J. POL. ECON. 1210 (2019); Dupas et al., *supra* note 12.

34 Freyer Jr., *supra* note 32, at 1210-16.

35 Steven M. Durlauf & James J. Heckman, *An Empirical Analysis of Racial Differences in Police Use of Force: A Comment*, 128 J. POL. ECON. 3998, 3999-4000 (2020).

36 Roland G. Fryer Jr., *An Empirical Analysis Of Racial Differences In Police Use Of Force: A Response*, 128 J. POL. ECON. 4003, 4003-04, 4006 (2020).

One wonders why respected researchers would make such a weak criticism and why a respected journal would publish it. Would such weak criticisms be treated with the same level of importance if the results had shown the politically more acceptable effect? This question is hard to answer, but certainly we should be concerned. While all research should be subject to criticism, under a reasonable model of information transmission we may end up systematically biasing research in favor of politically predetermined conclusions if different standards and thresholds for criticism are used based on the political acceptability of the conclusion. Gilley's 2018 article is an extreme example of this effect; if anyone presenting evidence that colonialism has benefits faces deaths threats and has his work publicly attacked by academics based on its conclusions, not methods, and eventually has the work withdrawn, we know ex ante that basically all research on colonialism will show that it is an unalloyed negative.

At best, readers will understand this situation and simply ignore research on this question since it will all say one thing regardless of the truth, thus undermining our ability to learn, one way or the other, about a question potentially affecting the well-being of some of the most disadvantaged and impoverished people on earth. At worst, readers will not understand that the answer is predetermined, and will be persuaded to update their beliefs in favor of assumed opinion since they observe only those analyses with politically acceptable results. It is perhaps too much to ask for readers of academic literature, particularly those who consume such literature for the purpose of making policy, to do the proper Bayesian updating conditioning on the likelihood that research results are being filtered by a murderous, angry mob. There are certainly historical examples of serious consequences for those speaking out against incorrect but popular beliefs, and a consequent failure of the general population to correctly update on the true likelihood of the underlying beliefs given the very few people willing to speak against consensus.[37]

Are we at this point in economics? The ideal experiment is unfortunately effectively impossible. To definitively answer the question, we would require a sample of paired papers, using the same methodology to research the same questions, but reaching different conclusions. We could then evaluate the number, intensity, and placement of critiques of the work and determine if the version that confirms politically acceptable ideas is subject to fewer critiques than the version that questions orthodox academic opinion. Creating such an experiment would be extraordinarily difficult and would likely generate ethical concerns. But, we can make some natural comparisons. The recent paper on seminar environments, Dupas et al. provides a natural comparison to Fryer Jr. Further, since Dupas et al. is still a relatively early stage paper, we can make potentially testable conjectures

[37] *See* SARAH GILMAN, THE SALEM WITCH TRIALS 11-14, 23-27, 41-42 (2017).

about its likely reception which may be informative about how far economics is along the path we describe.

Briefly, Dupas et al. argues that the fact that women are asked more questions in economics seminars, and those questions are coded as more harsh by the volunteer cadre of secret seminar spies recruiting by the authors, proves that economics is biased against women.[38] Putting aside the question of why the main measure of hostility, the number of questions, is a negative sign rather than a positive sign of how a seminar is received,[39] there is a potential selection problem here that should be obvious to any working economist. The sample of seminars includes a large number of job talks, and many universities have explicit or implicit policies that a certain share of job talks be given by women. Even regular seminars must now exhibit greater gender balance, where balance is in terms of fraction of women in the general population, rather than as a fraction of economists at a certain level. These policies appear to bind in many places, though of course no one systematically documents such things for fear of career consequences. If there is no bias against women in academia and thus the probability of a man and a woman being invited conditional on quality is the same, then any intervention to bring in more women than would otherwise have been brought in will lead to, on average, lower quality seminars from women. This effect will be particularly strong in the job talk sample, where there is a clear desire to bring in more women, well above the proportion of women actually in the job pool. Thus, even if the distribution of quality of male and female job candidates are the same, if the fraction of Ph.D. graduates who are women is below the target fraction for job talks,[40] the quality of the female pool will be lower. The paper does not do much to address this potentially important source of unobserved heterogeneity, instead attempting to argue that they have proxies for quality.[41] Notably, their identification strategy introduces elements of the critical theory approach; sexism is assumed to drive all differences, and under this assumption the paper is well identified. If sexism does not drive differences, the paper cannot establish the presence of sexism, but since sexism is assumed, irrefutably, to occur as in the critical theory approach, the identification is valid from a critical theory perspective. Further, raising questions as to whether sexism may not be present would be denial, and thus further evidence of sexism. Notably, one of the authors, Wolfers, was a leader of a campaign to eliminate an editor from the *Journal of Political Economy* for perceived racial insensitivity, suggesting strongly that the paper is positioned in this critical theory space intentionally.[42]

---

38 Dupas et al., *supra* note 12, at 1, 19, 27.

39 *Id.* at 1, 15-17.

40 *Id.* at 9 n.4, table 1.

41 *See* Dupas et al, *supra* note 12, at 17.

42 Gregg Re, *Paul Krugman, Professors Seek Top Economist's Removal From Influential Job for Criticizing Black Lives Matter*, FOX NEWS, https://www.foxnews.com/politics/paul-krugman-professors-

Will people in the economics profession be willing to speak publicly about this gap in the paper, as they were willing to do so when Fryer found results that went against what people wanted to believe?[43] The response that unfolds will provide some evidence as to how much these methods have taken over in controversial areas of economics. Those few remaining areas of social inquiry within academia that may seek to maintain some intellectual standards must be very conscious of these concerns. A veneer of free inquiry, where a subset of authors are willing and permitted to produce research that reaches conclusions that deviate from the political consensus, is insufficient as long as research with politically unpopular conclusions is held to higher standards than research that confirms ideas that are politically popular within academia.

Maintaining a paradigm of holding politically unpopular evidence to a higher standard than politically popular evidence is always scientifically dubious. It isless damaging, however, when there is likely a correlation between political popularity and truth. Thus, it could be argued defensibly that certain long-held customs should be given the benefit of the doubt, and evidence rejecting such customs could perhaps legitimately be treated with a higher degree of skepticism than other evidence.

For example, no successful modern society exists without a strong taboo on cannibalism. Even historically, cannibalism seems to have been practiced in relatively few societies and did not generally survive into modern times as a cultural practice outside of isolated groups.[44] If a researcher were to present claims that cannibalism is in fact a beneficial activity, it is understandable that such claims might be subject to greater scrutiny than claims demonstrating harm from such activity. Scientific objectivity should, perhaps, be balanced against accumulated cultural knowledge.

This tradeoff is not at play in the current environment. The political constraints in question do not arise from any process of accumulated cultural knowledge, but instead grow out of a specific intellectual tradition that explicitly seeks to exert control over institutions in order to promote ideas that would not otherwise achieve broad acceptance.[45] Of possibly greater concern, this movement is an offshoot of an movement that arose explicitly in response to the failure of the ideology of its eventual adherents, leading to a philosophy of "radical skepticism."[46] Roughly, it appears that the beliefs

---

demand-top-economist-lose-his-job-for-criticizing-black-lives-matter (June 22, 2020) ("University of Michigan professor Justin Wolfers and dozens of other academics are leading a massive effort to have a senior faculty member at the University of Chicago removed from his position at the world's preeminent economics journal. . . .").

43 *See* Steven N. Durlauf & James J. Heckman, Comment, *An Empirical Analysis of Racial Differences in Police Use of Force: A Comment*, 128 J. POL. ECON. 3998 (2020).

44 CAROLE A. TRAVIS-HENIKOFF, DINNER WITH A CANNIBAL: THE COMPLETE HISTORY OF MANKIND'S OLDEST TABOO (2008)

45 *See* HERBERT MARCUSE, COUNTER-REVOLUTION AND REVOLT (Beacon Press 2010); ANTONIO GRAMSCI, SELECTIONS FROM THE PRISON NOTEBOOKS (2007).

46 *See supra* note 1 and accompanying text.

underlying much of what is currently used to determine which ideas are acceptable in an academic setting arose from a belief that objective knowledge was impossible, which in turn arose from some sort of unwillingness to accept the empirical failure of a particular revolutionary agenda; the conclusion was then that objective knowledge was impossible, which then led to a dogmatic acceptance of unfalsifiable tenants of "social justice." While the logic for this process seems strained at best, it does appear to have happened, leading to a large set of ideas that have no empirical support and do not represent any accumulated cultural knowledge but now receive almost complete deference in academia.

Further, this critical theory approach rejects the idea of evidence and logic in favor of alternative ways of knowing[47] with a particular emphasis on "standpoint epistemology."[48] Thus, unlike long-established cultural practices and ideas, which can often contain important insights and promote valuable heuristics even in the absence of well understood, empirical foundations, and which in many cases did develop in part through empirical observations and logical processes, "woke" ideas are entirely the construct of institutionally powerful but empirically empty thinkers.[49] Thus, ideas like "systemic racism" and the non-existence of a sex binary deserve no deference in the process of scientific inquiry. In our current environment, however, they receive far more such deference than defensible cultural traditions. But, we must take the environment for what it is rather than what would generate the best environment for academic inquiry. As such, we now move on to consider possible approaches to continuing to pursue empirical inquiry under these constraints.

V. PRACTICAL APPROACHES TO EMPIRICAL ANALYSIS UNDER CRITICAL THEORY CONSTRAINTS

The narrowing of acceptable discourse is likely to have serious repercussions on the use of data to answer important social questions. Still, for at least some time there may be space for empirical analysis to reveal some underlying truths about socially relevant variables, even under the constraints imposed by the expansion of critical theory. We will investigate several approaches to data analysis within this framework. For clarity, we will take a formal approach to defining estimators; in reality, our proposed approaches will probably be implemented in an informal and hidden way. Still, going forward it is important to understand the characteristics of the

47 Allison B. Wolf, *"Tell Me How That Makes You Feel": Philosophy's Reason/Emotion Divide and Epistemic Pushback in Philosophy Classrooms*, 32 HYPATIA 893, 897 (2017).

48 Sandra Harding, *Rethinking Standpoint Epistemology: What is "Strong Objectivity?"*, 36 CENTENNIAL REV. 437 (1992).

49 JOSEPH HENRICH, THE SECRET OF OUR SUCCESS: HOW CULTURE IS DRIVING HUMAN EVOLUTION, DOMESTICATING OUR SPECIES, AND MAKING US SMARTER (2016).

estimators that will effectively be used, and as such we consider the formal exercise to be illuminating.

First, we describe what we believe will be the objective of empirical research going forward. Roughly in order of importance, the objective of empirical work will be:

1. Deliver superficial results in line with political requirements
2. Make sure a sophisticated analyst cannot clearly infer true bad results from the output of the analysis
3. Get good estimates of not-yet-political variables.

Below, we provide a discussion of two of the most natural approaches to dealing with political constraints, the constrained estimator approach and the IV approach, and we discuss further issues related to implementing political constraints.

A. *Constrained OLS/NLLS*

Constrained OLS is a well-established empirical technique used when physical or economic constraints provide prior information to the researcher that can be used to limit the possible values of coefficients. When the model is well-specified, the constrained OLS estimator has the same asymptotic properties as unconstrained OLS. In small samples, constrained OLS can provide lower variance estimates of unconstrained variables; the simplest example would be univariate regression with a constant. If the constant is known to be zero, the estimate of the single coefficient will be improved in small samples by exploiting this information.

In our setting, the naive application of constrained OLS to implement political constraints has the opposite effect; by imposing a constraint that is not based on the underlying data generating process, estimates of auxiliary coefficients will exhibit higher errors in small samples and will in fact cease to be consistent.

In this scenario, a researcher has several options. One option is to run unconstrained OLS and report all auxiliary coefficients as they are, changing only the politically constrained coefficients. Such an approach is probably unwise as it cannot be replicated without revealing that the empirical method is fundamentally dishonest. While in the relatively near future the idea of replicability is likely to be dropped from the remaining social sciences where it still has some influence, for the time being this approach would create potential issues for the researcher.

We propose instead that a researcher can privately run unconstrained OLS and then engage in a sort of meta-search for the algorithmic method which will, under the political constraints, deliver values for the auxiliary coefficients that are closest to the unconstrained estimates. Such a procedure would be facilitated by estimating a constrained nonlinear least squares model, even when a linear model appears most appropriate. Nonlinear models can generate more degrees of freedom in the search parameters and

starting values, thus potentially allowing the researcher to more precisely target the known, correct auxiliary parameter value. Carefully adjusting tolerances and starting points can potentially lead to estimates of auxiliary parameters that closely correspond to the unconstrained coefficients in a way that can be replicated by sharing code. Such an approach certainly violates best practices for such estimation procedures but does not deviate a great deal from currently accepted methodology. A second advantage of such a procedure is that a nonlinear model can be built to provide locally accurate estimates of auxiliary variables, even if the political constraints distort the overall estimate.

Figure 1 provides an example of this approach. Data are simulated from a linear model, and then a linear model is fit to the data. Unsurprisingly, the model fit is quite good. But, we posit a political constraint, where the constant must be equal to zero. The fit of the linear model with this constraint, shown as the dashed line, is poor. Thus, imposing the political constraint interferes with the model fit for the effect of the x variable. We then consider nonlinear models of the form $y = \beta x^{\frac{1}{z}} + \varepsilon$, $z \in N$. Specifically, the top right panel fits the model

$$\underset{z \in N}{\operatorname{argmin}} \int_0^1 (\hat{\alpha} + \hat{\beta} x - \hat{\beta}_{NLLS_z} x^{\frac{1}{z}})^2,$$

where $\hat{\alpha}$ and $\hat{\beta}$ are the coefficients for the correctly specified linear model and $\hat{\beta}_{NLLS_z}$ is the nonlinear least squares coefficient associated with the model $y = \beta x^{\frac{1}{z}} + \varepsilon$. Thus, we maintain the political fiction that the constant term is zero while achieving a better overall fit. But, in many cases, the overall fit will not be what is of interest. The relevant policy question may depend on, for example, the relationship between the y and x variables over a subset of the x space. The lower two panels show the relevance of this, solving instead

$$\underset{z \in N}{\operatorname{argmin}} \int_0^{0.2} (\hat{\alpha} + \hat{\beta} x - \hat{\beta}_{NLLS_z} x^{\frac{1}{z}})^2$$

and

$$\underset{z \in N}{\operatorname{argmin}} \int_{0.8}^1 (\hat{\alpha} + \hat{\beta} x - \hat{\beta}_{NLLS_z} x^{\frac{1}{z}})^2.$$

Thus, depending on which region is most relevant to the policy question, the appropriate specification will be either, in this example, $\beta x^{\frac{1}{5}}$ for the entire region, $\beta x^{\frac{1}{7}}$ for the low-x value region, and $\beta x^{\frac{1}{3}}$ for the high-x value region.

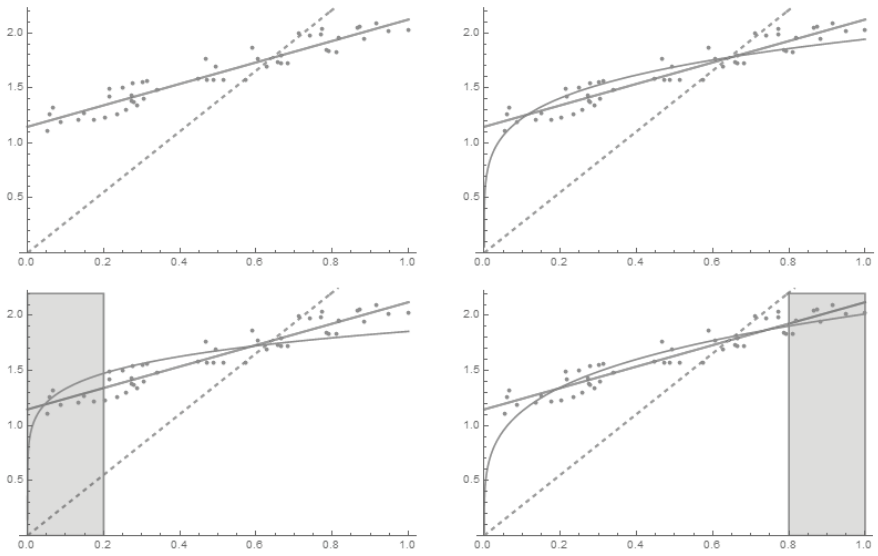


Figure 1

An example of how a nonlinear estimator can be selected to fit different regions of a constrained linear estimator. The top left panel show the (correct) linear specification, and the linear specification without a constaint; that is, where the political variable is the constant and that constant must be zero. The top right panel shows the lowest mean square error for a nonlinear fit over the entire region plotted when using a nonlinear specification of the form $y = \beta x^{\frac{1}{z}} + \varepsilon, z \in N$. The bottom two panels do the same, except fitting over the shaded region of $(0, .2)$ and $(.8,1)$ respectively.

B. *Intentionally Omitted Variables*

An alternative approach when faced with politically constrained coefficients is to simply ignore them by omitting the data on the variable in question. Such an approach, of course, leads to the classic omitted variables problem, which for well understood reasons potentially leads to inconsistent and often quite misleading estimates of other variables; omitting a variable that plays a significant role in causing the outcome in question will lead to too much effect being assigned to variables that correlate with that variable.

Ignoring the variable raises issues of internal consistency of the model. If it is assumed that a variable has a particular effect, then on the surface it may make sense to leave it out of the regression specification since we already know from, for example, the lived experience of a relevant party, then there is no reason to investigate the effect of that variable. From the

perspective of an econometrician this argument of course does not work; all variables that effect the outcome in question should be included in the specification; leaving out a variable will only be innocuous when that variable is uncorrelated with all of the other variables.

This subtlety itself may provide a means to provide information to normal readers while avoiding cancellation by woke readers. Making such a statement as justification for omitting a variable will effectively signal to the reader that the researcher is operating under political constraints, thus potentially triggering a more subtle interpretation of the results. At the same time, if wokeness is correlated with lack of econometric understanding, such a message may in fact ingratiate the researcher to the woke reader. This strategy, while likely valuable in certain cases, does carry risks; the inverse relationship between wokeness and econometric understanding is not deterministic, so caution must be used.[50]

If the researcher determines that omitting the political variable is the appropriate way to proceed, several options are available. As long as there is some other variable that could plausibly influence the outcome but which is, in fact, not available to the researcher, the researcher can proceed with an instrumental variables approach. Alternatively, the researcher could invent some rationale to argue that some variable should be included but cannot be, and then proceed with instrumental variables anyway. In either case, instrumental variables will deliver a consistent estimate of the effect of the auxiliary variable as long as the instrument satisfies the exclusion restrictions; the reason for the omission of a variable is not relevant to the properties of the estimator. A difficulty here, however, is that anything correlated with the auxiliary variable of interest may also tend to be correlated with the intentionally omitted variable, and thus may be seen as a proxy for the omitted variable, risking cancellation.

The practical approach to instrumental variables that we see currently helps to suggest an approach to overcome this issue. The set of potential instruments in many cases is quite large, and therefore a researcher has the opportunity to pick from a large set of potential specifications. In our case, the researcher can carefully select instruments that are not subject to the critique that they proxy for the political variable. Such a selection approach likely removes the most plausible instruments, and thus the ones most likely to work. In contrast, however, to the true omitted variables problem, the researcher in our setting can know the correct parameter value since he has access to the omitted variable. Thus, even if the researcher cannot find a strong instrument that delivers good small sample properties, he can continually search through weak instruments until he finds on that delivers a coefficient estimate as close as possible to the known coefficient. Such a procedure mimics what likely already happens under publication constraints (which needs to be analyzed though a full meta-analysis of papers on

[50] *See generally* Dupas et. al, *supra* note 12 and accompany text.

controversial topics), but in this case is arguably more ethical as the objective is to manipulate the specification to deliver a known, correct result. Of course, it will be quite difficult to know whether the researcher has honestly searched for the specification that delivers the correct coefficient estimate or simply mined specifications for significance, but that is not particularly far from the current situation anyway. Robustness testing becomes more difficult, however, and thus there are some limits to this approach.

C. *Dynamic Political Constraints*

An important feature of the emerging regime is that, while dissent from acceptable views is forbidden, what is acceptable is not stable. Over time, an increasing set of variables is likely to be constrained by the elect, while the nature of the constraint may also change. In some cases, constraints may actually flip; recent experience with Covid policy provides examples of this effect: masking was initially discouraged and those who supported masking were considered foolish and selfish for taking PPE away from front line health care workers; then belief in masking became mandatory, and those who dissented were considered practically murderers. This change, crucially, came without new high quality evidence becoming available. Another example would be travel bans: initially, banning international travel from affected areas was viewed as xenophobic, and thus no researcher could safely argue that travel bans were effective. Later, restricting mobility far beyond an international travel ban became the only acceptable policy solution. Again, these changes were abrupt but not related to existing or new evidence about the efficacy of such bans.

As such, empirical work that attempts to uncover truths about those variables that are still politically unconstrained must be structured so that it can be updated over time to adhere to the changing political requirements for the constrained variables. Notably, getting the best estimates of these not-yet-political variables becomes a dynamic process. A researcher has access to estimates from before a political constraint is applied, and in order to get the best estimate of the non-political variables this information should be used. Further, if the constraint has flipped, the analysis under the two different constraints can be informative about the true value of the non-political variables.

D. *Relationship to Fairness Algorithms*

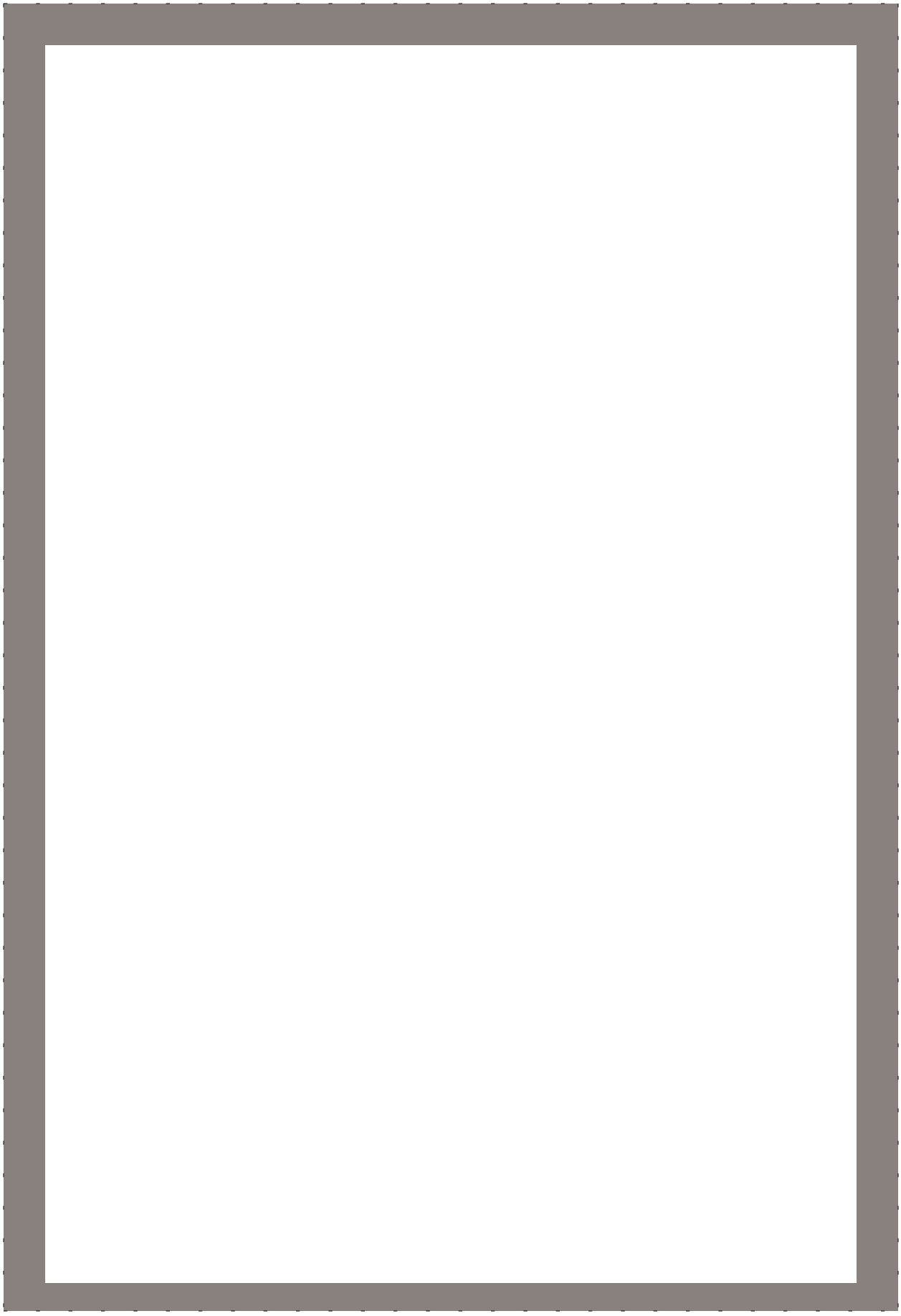
Much of our discussion relates to the growing field in statistics which seeks to develop "fair" algorithms. These methods are designed to implement political assumptions regardless of what actual data suggests. So, for example, if Group A and Group B differ on loan default rates and also differ in underlying characteristics that predict loan default rates, the algorithms are

designed to remove dependence on such underlying characteristics in order to in effect predict identical rates of defaults in the two groups. Our challenge is greater than theirs; whereas the work in that area is consciously avoiding using valid information, our objective is instead to use data in a way that makes false things appear true or at least hides true things, while at the same time still recovering valid information on other dimensions. Formally, some of our proposed methods align with the methodologies used in this other area, but the interpretation and the optimal method to select will differ.

CONCLUSION

We have developed a framework to understand changing norms in social science, as objective empirical analysis loses favor to Critical Theory methods. As activism based on pre-determined conclusions replaces inquiry as the primary function of the university, the space for the honest researcher to contribute in useful ways to important but controversial issues shrinks. Still, we believe honest research must still have a role to play going forward, and we attempt to provide such a researcher with a conceptual framework to understand the new environment and some tools to operate in this environment.

Still, the conclusion of our paper is bleak. Even if researchers are able to successfully navigate the new environment, the set of questions that they can address will be greatly constrained. Perhaps ironically, many of these questions that are or will become off-limits are those whose answers are most needed to improve the lives of the least privileged populations in the US and beyond, populations whom Critical Theorists claim to be working to help. The only alternatives to accepting these constraints appears to be aggressive external pressure to protect individual, rather than collective, academic freedom at universities, as is being attempted in the United Kingdom, or to create new academic units or institutions which are built specifically to permit and encourage dissenting thought. Existing institutions have failed to protect free inquiry, and only such radical changes imposed from outside the academy seem likely to have any chance of staving off the dystopian academic environment we have described.



44682-gme_17-4 Sheet No. 88 Side B 11/29/2022 09:39:42