The Limits of Evidence-based Regulation: The Case of Anti-bribery Law

Kevin E. Davis[*]

September 2020

Abstract

Evidence-based regulation is a term of art which refers to the process of making decisions about regulation based on evidence generated through systematic research. There is increasing pressure to treat evidence-based regulation as a global best practice, including from US political interests hoping to tame the regulatory state, the OECD, international trade agreements, and academics. However, there are certain conditions under which evidence-based regulation is likely to be a less appealing method of decision-making than the alternative, namely, relying on judgment. Those conditions are: it is difficult to collect data, on either interventions or outcomes; accurate causal inferences are difficult to draw; there is little warrant for believing that the same causal relationships will apply in a new context; or, the decision-makers in question lack the capacity to undertake one of these tasks. These conditions are likely to be present in complex, transnational, decentralized and dynamic forms of business regulation. The global anti-bribery regime is an illustrative case.

CONTENTS

INTRODUCTION

Regulatory decisions ought to be based on the best available evidence, meaning evidence obtained through systematic research. This proposition may seem too obviously valid to debate, equivalent to insisting that regulation be rational. That would explain why evidence-based regulation, a term of art which refers to the process of making regulation based primarily on evidence generated through systematic research, is now widely touted as a global best practice.

At first glance, the only justification for failure to adopt evidence-based regulation is irrationality, faith over reason. This Article explores a different justification for resisting evidence-based regulation: the evidence that it prioritizes might be unhelpful, at least when compared to the alternative basis for decision-making commonly referred to as "judgment," a broad term which encompasses a mix of experience, unsystematic inquiry and theory. The argument here has two main points. First, there are specific conditions under which the balance is likely to be tilted against the evidence-based approach and in favor of reliance on judgment. Second, those conditions are likely to arise in complex, transnational, decentralized and dynamic regulatory regimes. Under these conditions, regulation should be regarded as more of an art than a science.

An exploration of the limitations of evidence-based regulation is timely because there is now pressure to expand its domain, both in terms of subject matter and geography, in a way that might extend to important forms of transnational regulation. Within the US, social scientists eager to demonstrate the value of newly developed research tools such as randomized controlled trials ("RCTs"), and political interests eager to tame the regulatory state, have sought to make principles of evidence-based regulation obligatory for a broad range of regulatory agencies. Meanwhile at the international level, the Organisation for Economic Co-operation and Development ("OECD") regularly encourages its member states to embrace evidence-based regulation. More recently, in signing the Comprehensive and Progressive Agreement for Trans-Pacific Trade Partnership ("CPTTP"), which replaced the Trans-Pacific Partnership ("TPP") after it was rejected by President Trump, a diverse group of Pacific Rim countries representing almost 500 million people explicitly committed themselves to a regulatory approach that very much resembles U.S.-style evidence-based regulation. All of these forces are likely to receive a considerable boost from the award of the 2019 Nobel Prize in Economics to three development economists for a body of work

that advocates an evidence-based approach to interventions aimed at alleviating poverty. The stage is set for a discussion of the implications of these developments for the crucial questions of whether and when evidence-based regulation should be adopted.

This Article provides a framework for analyzing the merits of the evidence-based approach to regulatory decision-making. For the purposes of this discussion, "evidence" means information about causal relationships between past regulatory interventions and outcomes that is derived from systematic research. In other words, evidence means evidence of effectiveness based on systematic analysis. Evidence-based regulation involves four stages: systematically collecting data, drawing inferences about the causal roles of particular regulatory interventions, concluding that those interventions will play the same causal roles in the present context, and implementing the relevant intervention in the new context. Here we will focus on the decision-making that occurs in the first three stages.

There are four basic situations in which the evidence-based approach might be unhelpful: it is difficult to collect data, on either interventions or outcomes; accurate causal inferences are difficult to draw; there is little warrant for believing that the same causal relationships will apply in a new context; or, the decision-makers in question lack the capacity to undertake one of these tasks. These problems all have been discussed extensively in the literature on comparative law, empirical legal studies and law and development.[1] The literature has placed less emphasis on the fact that all four of these potential problems are particularly likely to arise in complex, decentralized and dynamic forms of business regulation. In these settings it is common to find that: regulations are implemented by multiple agencies; those agencies have overlapping jurisdiction in relation to certain actors or activities; there are interactions between the targeted actors, such as when they compete with or emulate one another; and, the agencies are heterogeneous, e.g. because they have different abilities to process information or operate in very different environments. In

---

[1] For surveys of the literature see, for example, Mathias Siems, COMPARATIVE LAW 2ND ED. 27-9 *et passim* (2018) (discussing difficulty of generalizing about the effects of law reform); Holger Spamann, *Empirical Comparative Law*, 11 ANN. REV. L. & SOC. SCI. 131 (2015) (discussing problems of data collection and causal inference); Green and Thorley, *Field Experimentation and the Study of Law and Policy*, 10 ANN. REV. L. & SOC. SCI. 53 (2014) (discussing problems of causal inference); Kevin E. Davis and Michael J. Trebilcock, *The Relationship Between Law and Development: Optimists versus Skeptics*, 56 AM. J. COMP. L. 895, 917-938 (2008) (surveying criticisms of "deterministic" theories which hold that legal institutions might play similar roles in different contexts and discussing why reformers might lack the capacity to identify appropriate reforms).

these circumstances the value of systematic research on the effectiveness of regulatory interventions is likely to be low relative to the value of alternative bases for regulatory decision-making.

Many regimes that regulate transnational business activity fit the profile of regimes in which the value of systematic research is suspect. Examples include the regimes concerned with antitrust/competition, bribery, environmental law, food safety, fraud, intellectual property, money laundering, privacy, securities, terrorist finance, and human trafficking. All of these regimes govern large swathes of business activity. This means that they frequently cover activities that transcend national borders, since a great deal of modern business activity involves the movement of goods through global supply chains or international flows of capital, labor or information. Most countries respect traditional limits on extra-territorial assertions of jurisdiction, and supranational regulatory agencies are scarce. Accordingly, transnational business activity is, if it is regulated at all, likely to be regulated by multiple agencies. And since many forms of transnational business activity are global in scope, the applicable regimes cover agencies and regulated actors that span a wide range of social, economic, political and physical conditions.

This Article uses a case study of anti-bribery law to illustrate these claims about the limits of evidence-based regulation. At its core, anti-bribery law is concerned with discouraging individuals and firms from paying bribes to public officials. A prominent feature of the modern anti-bribery regime—which comprises both anti-bribery laws and the agencies that implement them—is that it regulates bribery of not only domestic but also foreign public officials, meaning officials who wield power in countries other than the country of the enforcing state.[2] This aspect

---

[2] *See* generally, Kevin E. Davis, BETWEEN IMPUNITY AND IMPERIALISM: THE REGULATION OF TRANSNATIONAL BRIBERY (2019) (analyzing the anti-bribery regime). The use of the term "regime" here is slightly idiosyncratic. International relations scholars use the concept of an international regime to refer to an ensemble of international rules and processes focused on a single issue. A sufficiently large set of overlapping regimes is called a regime complex. *See* generally Karen J. Alter and Kal Raustiala, *The Rise of International Regime Complexity*, 14 ANN. REV. L. & SOC. SCI. 329 (2018). To the extent the anti-bribery regime is grounded in international agreements it qualifies as an international regime or regime complex. However, as defined here, the anti-bribery regime also includes norms that emanate from and bind private actors as opposed to only states and interstate organizations. Accordingly, the anti-bribery regime is probably better understood as an example of what Abbot called as a "transnational" regime or regime complex, or what Halliday and Shaffer call a "transnational legal order." *See* Kenneth W. Abbott, The Transnational Regime Complex for Climate Change, 30 ENV. & PLAN. C: GOV. & POL. 571 (2012): 571-590; Terence C. Halliday and Gregory Shaffer, Transnational Legal Orders, in TRANSNATIONAL LEGAL ORDERS, ed. Terence C. Halliday and Gregory Shaffer (2015), 7-21.

of the regime allows the US to, for example, sanction a multinational company which pays bribes to a public official in Nigeria.

Several features of the anti-bribery regime make it a 'best case' in which to validate the claim that systematic research can be unhelpful, or correlatively, a worst case for the claim that systematic research is always helpful. To begin with, the anti-bribery regime is global in scope and the portion that concerns transnational bribery has, to date, mainly been concerned with transactions between multinational firms based in OECD countries and public officials of relatively poor countries. By design, there is a great deal of overlapping jurisdiction, meaning that many instances of misconduct can be sanctioned by multiple enforcement agencies. The actors subject to the regime include many multinational firms so many of them have significant economic and social interactions. For all of these reasons, this is a case in which the low-value research theory suggests that evidence-based regulation is likely to be especially unattractive. However, the insights provided by this case have broader application, some of which have already been developed by scholars who have analyzed fields such as financial regulation[3] and intellectual analyzing the value of evidence-based regulation that applies across jurisdictions and substantive domains.

There are other grounds for concern about evidence-based regulation besides doubts about the value of the evidence upon which it relies. For starters, evidence of effectiveness only is valuable among people who agree on how to define effectiveness. Since effectiveness is a measure of progress towards achievement of an objective, this in turn implies a reasonable amount of consensus about the objectives of regulation. That may be a heroic assumption.[4] Proponents of evidence-based regulation also presume that regulators are not only able but willing to use evidence to enhance their effectiveness. This ignores the possibility that regulators will pursue

---

[3] *See* John C. Coates IV, *Cost-Benefit Analysis of Financial Regulation: Case Studies and Implications*, 124 YALE L.J. 882 (2015) (identifying difficulty of drawing causal inferences as one of several obstacles to cost-benefit analysis of financial regulation).

[4] Christopher Ansell & Robert Geyer, '*Pragmatic complexity' a new foundation for moving beyond 'evidence-based policy making'?,* 38 POLICY STUD. 149 (2017) (suggesting a pragmatic alternative to evidence-based policy in contexts with diverse values); Wayne Parsons, *From Muddling Through to Muddling Up – Evidence Based Policy Making and the Modernisation of British Government*, 17 PUB. POLICY & ADMIN. 43, 54 (2002) (criticizing evidence-based policy for neglecting "people, power and politics" and that "knowledge is utterly embedded in power and value contexts and relationships").

objectives that serve their own interests or the interests of groups with which they are affiliated.[5] Yet another objection to evidence-based regulation is that effectiveness is not the only criterion against which regulation ought to be evaluated. There are compelling reasons to believe that the design and operation of legal institutions ought to be evaluated according to non-instrumental criteria such as legitimacy and respect for due process.[6]

For the sake of analytic clarity this Article will largely ignore these additional concerns about evidence-based regulation. Specifically, the analysis that follows will assume that in the relevant context there is a consensus about the objectives of regulation, regulators are dedicated to pursuing those objectives, and it is accepted that regulators ought to be judged solely in terms of their success in achieving those objectives. These assumptions are unrealistic; they effectively rule out the possibility that divergent values or interests—in short, politics—will influence either regulatory practices or how they are evaluated. The advantage of this analytic approach is that it makes it possible to focus on the influence of a specific set of factors on regulatory effectiveness, namely, the quality of information that regulators rely upon and their capacity to process it.

Part I of the Article describes evidence-based regulation. Part II discusses the limits of evidence-based regulation, presenting several reasons why evidence of the effectiveness of legal regulation might be of limited value in regulatory decision-making. Part III uses a case study of the anti-bribery regime to provide a more concrete illustration. Part IV discusses two courses of action suggested by the analysis in the preceding parts: 1) enhance the value of the evidence used in evidence-based regulation and 2) place greater weight on judgment in regulatory decision-making. This part concludes with a few comments on the prospects for extending the analysis to incorporate political factors.

---

[5] Parsons, *supra* note 4.
[6] Davis, *supra*, chapter 5.

I. WHAT IS EVIDENCE-BASED REGULATION?

A. EVIDENCE-BASED POLICY

Evidence-based regulation is a subset of the broader practice of evidence-based policymaking. Evidence-based policymaking is a term of art which refers to a process of making public policy decisions on the basis of the best available evidence, where the "best" evidence is presumed to include evidence produced by systematic research.[7] Reduced to its essential elements, the process requires policymakers to ask two questions. 'Has this policy intervention worked somewhere else?' and 'Will it work here?' In other words: Did the policy intervention, in combination with other factors, contribute to achieving the outcome of interest?[8] And, will it also play the desired causal role in the present context?[9] This kind of causal analysis can be an important component of a broader analysis of the desirability of an intervention, including a cost-benefit analysis.

The essential feature of evidence-based policymaking is emphasis on using evidence grounded in systematic research to identify the causal role of interventions. The main arguments for giving priority to systematic research as a basis for regulation rest on the fact that the hallmarks of modern research, such as documentation of methods and peer review, facilitate independent assessment and scrutiny of claims' validity.[10] This makes it relatively easy for regulators to determine what weight any given claim deserves, which presumably will enhance the quality of their ultimate decisions. The potential for scrutiny also gives researchers incentives to produce higher quality evidence. In addition, the transparency of systematic research might enhance the legitimacy of decisions that rely upon on it.

Recent interest in evidence-based policy can be traced to the widespread acceptance of evidence-based medicine, an approach to medical practice which demands that clinical decisions integrate individual clinical expertise with the best available evidence derived from systematic

---

[7] Huw Davies, Sandra Nutley, and Peter Smith, *Introducing Evidence-Based Policy and Practices in Public Services* in, Huw Davies, Sandra Nutley, and Peter Smith (eds), WHAT WORKS? THE ROLE OF EVIDENCE IN PUBLIC SECTOR POLICY AND PRACTICE 1-4 (2000) ["WHAT WORKS?"].

[8] Nancy Cartwright and Jeremy Hardie, EVIDENCE-BASED POLICY: A PRACTICAL GUIDE TO DOING IT BETTER, 6 23-36 (2012) (defining the concept of playing a causal role) ["EVIDENCE-BASED POLICY: A PRACTICAL GUIDE"].

[9] EVIDENCE-BASED POLICY: A PRACTICAL GUIDE, 6 (2012).

[10] WHAT WORKS?, *supra*, 6.

research.[11] Some practitioners and proponents of evidence-based medicine and policy endorse explicit hierarchies among types of evidence, with rankings based on the design of the studies that produced the evidence.[12] In the context of medical interventions, at or near the top of the hierarchy are studies that take the form of an RCT. In this kind of study the intervention of interest is applied to randomly selected members of a group drawn from the population of interest so that outcomes for that select group, namely, "the treatment group," can be compared to outcomes experienced by the other members of the group, the "control group." The differences between the mean outcomes of the treatment and control groups are used as an estimate of the causal effect of the intervention. In the medical literature and elsewhere RCTs are regularly referred to as the "gold standard" among sources of evidence of causal effects.[13]

In addition to RCTs, hierarchies of evidence used in evidence-based medicine often assign great value to systematic reviews of multiple studies (which should ideally include RCTs). Accordingly, enormous investments have been made in grading, synthesizing, and disseminating evidence, in the form of systematic reviews and clinical guidelines. In the medical context, many of the systematic reviews are produced by a nonprofit organization called the Cochrane Collaboration.[14]

For present purposes, the key feature of evidence-based policymaking is the prioritization of evidence produced by systematic research, not the fact that there might be a hierarchy among those classes of evidence.[15] Prioritization of systematic research implies downgrading what is often referred to as "judgment," a term which encompasses at least two alternative bases for beliefs about the impact of interventions: 1) theory and 2) unsystematic learning.[16] Theoretically grounded

---

[11] David L. Sackett et al., *Evidence Based Medicine: What It Is and What It Isn't*, 312 BMJ 71 (1996).
[12] M Hassan Murad et al, *New Evidence Pyramid*, 21 BMJ EVIDENCE-BASED MEDICINE 125 (2016); Sandra Nutley, Alison Powell and Huw Davies, WHAT COUNTS AS GOOD EVIDENCE? 10-11 (2013): https://www.alliance4usefulevidence.org/assets/What-Counts-as-Good-Evidence-WEB.pdf.
[13] Angus Deaton and Nancy Cartwright, *Understanding and Misunderstanding Randomized Controlled Trials*, SOC. SCI. & MED. (in press).
[14] http://www.cochrane.org/about-us/our-products-and-services
[15] Some scholars and practitioners question the value of hierarchies of evidence, particularly in relation to policymaking. See generally, Justin O. Parkhurst and Sudeepa Abeysinghe, 'Good' Evidence for Improved Policy Making: From Hierarchies to Appropriateness, 30 SOCIAL EPISTEMOLOGY 665 (2013); Sandra Nutley et al, WHAT COUNTS AS GOOD EVIDENCE?, *supra,* 10-15; UK Department For International Development, HOW TO NOTE: ASSESSING THE STRENGTH OF EVIDENCE, 7-8 (2013): https://www.alnap.org/system/files/content/resource/files/main/dfid-htn-strength-of-evidence.pdf.
[16] Nutley et al., *supra*, 6.

beliefs are generated by deduction from premises that do not purport to be empirical claims about the impact of the relevant intervention, although they might be based on research. Meanwhile, beliefs based on unsystematic learning are grounded in personal experience as well as unsystematically acquired information about the experiences or beliefs of others, often combined with theory. To see these distinctions consider, for example, a prosecutor who believes that increased criminal penalties for transnational bribery will cause a reduction in its prevalence. The basis for that belief could be: a statistical analysis of the correlation between variations in penalties and reports of transnational bribery (research); a belief that higher penalties generally deter misconduct, based on both introspection, personal observation, and statistical analyses of the deterrent effects of punishments for crimes other than transnational bribery (theory); or personal conversations with potential payers and recipients of transnational bribes about their likely behavior (unsystematic learning).

Evidence-based policy has a long history in the US. In fact, since the 1960's RCTs have been extremely influential in policymaking concerning welfare and education.[17] Several nonprofit organizations have played an important role in promoting the practice. One is MDRC, a nonprofit organization formed in 1974 and dedicated to research on interventions that affect the poor.[18] Another is the Campbell Collaboration, a nonprofit organization formed in 2000 and modeled on the Cochrane Collaboration, but focused on synthesis of evidence to guide social policy.[19]

In both the US and the UK the evidence-based approach to policymaking resonated with the New Public Management of the 1980s, which emphasized reinvention of public sector service delivery through, among other things, close attention to measurement of performance.[20] In the UK, evidence-based policymaking was embraced explicitly in the late 1990s by the Labour government

---

[17] Judith M. Gueron and Howard Rolston, FIGHTING FOR RELIABLE EVIDENCE (2013) (history of the use of RCTs to study income-maintenance programs in the US); Manzi, *supra*, 181-191 (describing RCTs in welfare and education).
[18] About MDRC, https://www.mdrc.org
[19] "The Campbell Collaboration promotes positive social and economic change through the production and use of systematic reviews and other evidence synthesis for evidence-based policy and practice." Campbell Collaboration Mission Statement, https://www.campbellcollaboration.org/about-campbell/vision-mission-and-principle.html/.
[20] Christopher Hood, *The "New Public Management" in the 1980s: Variations on a Theme*, 20 ACC'TING, ORGS. & SOC. 93 (1995) (describing the New Public Management).

as part of its "modernising government" initiative.[21] In 2016 the US Congress passed legislation establishing the bipartisan Commission on Evidence-Based Policymaking for the purpose of reporting on how to improve the US government's capacity to generate evidence on the viability and effectiveness of government programs and policies.[22]

Evidence-based policymaking appears to enjoy broad support among academics. It has been especially influential in the field of development economics, where the 2019 Nobel Prize in Economics was awarded to Abhijit Banerjee, Esther Duflo and Michael Kremer, three pioneers in the use of RCTs to analyze the effects of interventions aimed at alleviating poverty.[23] More generally, evidence-based policymaking appears to be compatible with a wide range of "experimentalist" approaches to governance.[24] However, some versions of experimentalist governance are agnostic about the value of systematic research. For example, Charles Sabel and his collaborators emphasize that decisions on how to improve the performance of service-providing units should be informed by comparisons with the experiences of their peers but do not specify whether that information has to be obtained from systematic research.[25]

B. EVIDENCE-BASED REGULATION

This Article focuses on evidence-based regulation rather than the broader phenomenon of evidence-based policymaking. Policymaking is a broader concept because it encompasses interventions aimed at the provision or distribution of goods and services, in addition to regulation. Regulation covers only efforts to influence the behavior of firms or individuals.[26] Regulatory

---

[21] Cabinet Office, MODERNISING GOVERNMENT, chapter 2 (1999) ("Government should regard policy making as a continuous, learning process, not as a series of one-off initiatives. We will improve our use of evidence and research so that we understand better the problems we are trying to address.")

[22] THE PROMISE OF EVIDENCE-BASED POLICYMAKING: REPORT OF THE COMMISSION ON EVIDENCE-BASED POLICYMAKING (2017).

[23] Royal Swedish Academy of Sciences, PRESS RELEASE: THE PRICE IN ECONOMIC SCIENCES 2019, https://www.nobelprize.org/prizes/economic-sciences/2019/press-release/ (announcing award of the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2019 to Abhijit Banerjee, Esther Duflo and Michael Kremer).

[24] Deval Desai and Michael Woolcock, *Experimental Justice Reform: Lessons from the World Bank and Beyond*, 11 ANN. REV. L. & SOC. SCI. 155 (2015); Gráinne De Búrca, Robert O. Keohane, and Charles Sabel, *Global Experimentalist Governance*, 44 BRIT. J. POL. SCI. 477 (2014); Charles F. Sabel and Jonathan Zeitlin, eds. Experimentalist Governance in the European Union: Towards a New Architecture (2010); Ian Sanderson, *Evaluation, Policy Learning and Evidence-Based Policy Making*, 80 PUB. ADMIN. 1 (2002); Michael C. Dorf and Charles F. Sabel, *A Constitution of Democratic Experimentalism*, 98 COLUM. L. REV. Columbia 267 (1998).

[25] Sabel and Zeitlin, *supra*.

[26] John Braithwaite, Cary Coglianese, David Levi-Faur, *Can Regulation and Governance Make a Difference?* 1 REG. & GOVERNANCE 1, 3 (2007).

interventions include efforts to curtail socially undesirable behavior, such as pollution, corruption or violence. They also cover efforts to encourage productive behavior, such as investment in reliance on contractual or property rights. Some interventions might be designed both to distribute and to regulate. For example, the allocation of formal legal titles to squatters on public land might be designed both to confer a benefit and to encourage investment. The focus in this Article is on interventions that are primarily regulatory. So, for example, interventions designed to improve the efficiency of civil courts have regulatory effects, but since they often are designed primarily as ways of improving the distribution of dispute resolution services they will not be the focus here. The discussion here also is limited to legal regulation, meaning regulatory interventions implemented by governmental as opposed to private actors.

To keep the analysis tractable, the discussion in this Article will be limited to the use of systematic research to determine the outcomes likely to be caused by regulatory interventions. This kind of analysis is a key step in any "regulatory impact analysis" aimed at identifying the positive and negative outcomes associated with a particular regulatory intervention. These kinds of analyses are, in turn, necessary steps toward making an ultimate decision about whether the outcomes associated with an intervention are, on balance, desirable. That last step often involves cost-benefit analysis, which is a controversial practice in its own right but is ubiquitous in US regulatory practice. The focus here is only on the initial step of identifying the causal role of regulatory interventions. That causal analysis leads to predictions about the outcomes of the relevant intervention which can be used as inputs in cost-benefit analysis.

Regulation is typically a multi-stage process, beginning with the creation of a regulatory norm or standard and ending with enforcement. Proponents of evidence-based regulation advocate for its use in each stage. For example, they encourage reference to studies of the impact of prior interventions on employment outcomes in creating environmental standards. At the enforcement stage, leaders in the policing community endorse evidence-based approaches to policing[27] and the Campbell Foundation's "Crime and Justice" group has published several systematic reviews on police interventions.[28] As for sanctions, many US courts practice evidence-based sentencing,

---

[27] Cynthia Lum and Christopher S. Koper, EVIDENCE-BASED POLICING: TRANSLATING RESEARCH INTO PRACTICE (2017); Lawrence W. Sherman, *The Rise of Evidence-Based Policing: Targeting, Testing, and Tracking*, 42 CRIME & JUSTICE 377 (2013).
[28] Campbell Collaboration, *supra*.

typically by using studies of recidivism to inform sentencing decisions (on the theory that criminal sentencing should be designed, at least in part, to achieve incapacitation).[29]

Advocates of evidence-based regulation can point to several legal instruments for support. Most notably, in the US, a 2011 executive order issued by President Obama directs federal agencies subject to executive oversight to adopt regulations "based on the best available science."[30] That Executive Order also requires the regulatory system to "measure, and seek to improve, the actual results of regulatory requirements."[31] The Trump administration has maintained this particular Obama initiative and has in fact embraced evidence-based regulation as a way to limit and possibly even roll back regulations.[32]

The Obama Executive Order clearly is intended to encourage regulatory agencies to base decisions on evidence from the natural sciences, such as toxicological studies of the effects of pollutants on health. However, there is no obvious reason, to limit the scope of the order to the natural sciences. The language is certainly open to the interpretation that it demands use of relevant research from the social sciences. Moreover, Cass Sunstein, a prominent legal academic who was a leading figure in the Obama administration at the time the order was issued, seems to believe that the order has a broader scope. Shortly after the Executive Order was issued he published a law

---

[29] Sonja B. Starr, *Evidence-Based Sentencing and the Scientific Rationalization of Discrimination*, 66 STAN. L. REV. 803 (2014).

[30] Improving Regulation and Regulatory Review, § 1(a), E.O 13563, January 18, 2011.

[31] Currently this directive does not cover so-called "independent" agencies, a category which includes major sources of business regulation such as the Securities and Exchange Commission, the Federal Reserve, the Commodity Futures Trading Commission and the Federal Communications Commission. There is, however, considerable support for extension of Presidential oversight to these independent agencies. This would presumably entail extension of the requirements to engage in science-based and empirically grounded regulation. See, Jason A. Schwartz and Caroline Cecot, STRENGTHENING REGULATORY REVIEW: RECOMMENDATIONS FOR THE TRUMP ADMINISTRATION FROM FORMER OIRA LEADERS, 10-12 (2016): http://policyintegrity.org/publications/detail/strengthening-regulatory-review

[32] In April 2018, the EPA proposed to adopt a new policy that would require more of the data and models underlying studies used as a basis for certain regulations to be available to the public for validation and analysis. Environmental Protection Agency, Strengthening Transparency in Regulatory Science, 83 FED. REG. 18,768 (April 30, 2018). The proposal was praised by representatives of the chemical and fossil-fuel industries as well as people skeptical of climate change as a way of enhancing the quality and the legitimacy of the EPA's regulations. However, it was attacked by environmentalists and Democrats as an "attack on science" that would drastically constrain the EPA's ability to make regulations designed to protect public health, since the EPA's regulations in this area have traditionally relied on studies which contain medical data that cannot be made public. Lisa Friedman, *E.P.A. to Limit Science Used to Write Public Health Rules*, N.Y. TIMES (Nov. 11, 2019).

review article citing the order as a justification for increased the use of social science evidence in regulation.[33]

There is similar support for evidence-based regulation in international law. Article 2 of the WTO Agreement on the Application of Sanitary and Phytosanitary Measures ("SPS Agreement") requires parties to ensure that any measures taken to protect human, animal or plant life or health are "based on scientific principles" and are "not maintained without sufficient scientific evidence."[34] Pressure to adopt evidence-based regulation in other substantive areas received a boost with the conclusion of the CPTPP. The first "megaregional" agreement, CPTPP covers a remarkably diverse group of countries: 11 countries scattered around the Pacific Rim, including two non-OECD members (Vietnam and Malaysia). The agreement encourages its parties to require regulatory agencies to conduct regulatory impact assessments that "rely on the best reasonably obtainable existing information including relevant scientific, technical, economic or other information."[35] This language is taken verbatim from guidance that the US Office of Management and Budget issues to federal agencies subject to executive oversight and can be interpreted to require evidence-based regulation.[36] However, it is important to note that the CPTPP's reference to "best reasonably obtainable information" is sufficiently flexible to permit reliance on alternatives to systematic research.

More explicit pressure to adopt evidence-based regulation also has come from the OECD, an organization which essentially serves to represent the views of the world's wealthiest states.[37] The OECD is deeply committed to promoting evidence-based policymaking in general and evidence-based regulation in particular. Its official guidance on best practices in regulation is explicitly premised on the notion that governments' decision-making ought to be "evidence-

---

[33] Cass Sunstein, *Empirically Informed Regulation*, 78 U. CHI. L. REV. 1349 (2011) at 1363-1364 (suggesting in article on social science research evidence that continuing empirical research is desirable and citing Executive Order 13563). At the time of writing Sunstein was Administrator, Office of Information and Regulatory Affairs in the White House's Office of Management and Budget. *Id.*, 1349n. Sunstein's time in government is discussed at length in CASS SUNSTEIN, SIMPLER (2013).
[34] *Agreement on the Application of Sanitary and Phytosanitary Measures*, April 15, 1994, *Marrakesh Agreement Establishing the World Trade Organization, Annex IA*, LEGAL INSTRUMENTS – RESULTS URUGUAY ROUND, Vol. 31, Art. 2.2.
[35] *Trans-Pacific Partnership Agreement*, done at Auckland on 4 February 2016 ("the TPP"), Art. 25.5(2)(d), incorporated by reference into the *Comprehensive and Progressive Agreement for Trans-Pacific Partnership*.
[36] Office of Management and Budget, CIRCULAR A-4 (Sept. 17, 2003): https://obamawhitehouse.archives.gov/omb/circulars_a004_a-4/
[37] Buttonwood, *What is the OECD?,* THE ECONOMIST, July 6, 2017.

based."[38] This view applies to all stages of the regulatory process. For instance, the OECD's guidance on best practices in enforcement and inspections says: "Regulatory enforcement and inspections should be evidence-based and measurement-based: deciding what to inspect and how should be grounded on data and evidence, and results should be evaluated regularly."[39] The same document encourages reviewers to "draw on international experience to evaluate the merits of different organizational approaches to address common public policy goals."[40] At the same time, it is important to note that aside from recommending careful analysis of data from previously implemented interventions, the OECD does not define the term "evidence-based" or refer specifically to the idea of prioritizing evidence derived from systematic research.

Evidence-based regulation also has been promoted by legal academics. This was not always true. Over the years many scholars of comparative law have promoted systematic research on experience in other jurisdictions as a tool for identifying "best practices." [41] However, this advice typically was offered with extensive caveats about the difficulties involved in successfully transplanting or translating legal practices from one context to another.[42] Similarly, proponents of experimentalist governance, which applies to legal regulation as well as other forms of public policy, have been careful to leave room for judgement in deciding whether to apply lessons learned from regulatory experiments in new contexts.[43] Recently, however, a handful of legal scholars, led by Sunstein, have come down more clearly in favor of evidence-based regulation.[44] In a prominent article, Abramowicz, Ayres, and Listokin, advocate greater use of RCTs to inform many different kinds of regulatory reforms, ranging from changes to disclosure requirements for issuers of publicly traded securities to bans on discrimination by employers on the basis of sexual

---

[38] OECD, ANNEX TO THE RECOMMENDATION OF THE COUNCIL ON REGULATORY POLICY AND GOVERNANCE, ¶ 1.1 (2012) ("Regulatory policy defines the process by which government, when identifying a policy objective, decides whether to use regulation as a policy instrument, and proceeds to draft and adopt a regulation through evidence-based decision-making.") http://www.oecd.org/governance/regulatory-policy/49990817.pdf. See also, OECD, OECD REGULATORY POLICY OUTLOOK 2015, chapter 4 (2015) (evaluating countries' in terms of whether they use Regulatory Impact Assessment to support evidence-based regulation).
[39] OECD, REGULATORY ENFORCEMENT AND INSPECTIONS, 17 (2014).
[40] Id., 19.
[41] Siems, *supra* note 1, at 4, 27-9, 372 ("Foreign laws can provide models of how well different sets of legal rules work in addressing a particular problem or in pursuing a particular policy"; "Most comparative lawyers take the view that an evaluation about the 'best rules' can, if cautiously made, be part of a comparative analysis").
[42] Id.
[43] Lisa Larrimore Ouellette, *Patent Experimentalism*, 101 VA. L. REV. 65 65, 105-106 (2015) (distinguishing experimentalism from reliance on experiments along several dimensions, including the ability to "elicit local knowledge about heterogenous conditions" and generate correspondingly diverse policy suggestions).
[44] Sunstein, *supra* note 33.

orientation.[45] More recently, there has been a spate of law review articles advocating regulation based on the application of machine learning techniques to big data.[46]

II.THE CHALLENGES OF EVIDENCE-BASED REGULATION

This Part outlines the challenges inherent in evidence-based regulation. The following sections discuss issues that arise at each stage in the process of evidence-based regulation, namely, data collection, causal inference and generalization, as well as how constraints on institutional capacity might affect this process.

A. DATA COLLECTION

A thorough evidence-based analysis of law's impact on behavior requires data on not only legal interventions but also all the non-legal factors that might affect the causal roles of the legal intervention. The range of data required is extensive because it must cover all of the links in the causal chain between law and behavior.[47] That chain typically begins with norms encoded in legislation or caselaw or a treaty or some other authoritative source; in other words, the 'law on the books.' Those norms are then implemented by various institutions, including state agencies like police and prosecutors, as well as private actors who might, for example, incorporate legal norms into standards or codes or contracts. Some people call the norms-as-implemented, the 'law in action.' The impact of these norms and institutions on other actors depends to some extent on how they are perceived. Those perceptions depend on factors such as the state's efforts to publish legal information, the behavior of private media organizations, and the literacy of the population.

---

[45] Michael Abramowicz, Ian Ayres, and Yair Listokin, *Randomizing Law*, 159 U. PENN. L. REV. 929 (2011). See also, Zachary J. Gubler, Experimental Rules, 55 B. C. L. REV. 129 (2014); John O. McGinnis, *Laws for Learning in an Age of Acceleration*, 53 WM. & MARY L. REV. 305 (2011).

[46] See, e.g., Anthony J. Casey and Anthony Niblett, *The Death of Rules and Standards*, 92 IND. L. J. 1401 (2015); Ariel Porat and Lior Jacob Strahilevitz, *Personalizing Default Rules and Disclosure with Big Data*, 112 MICH. L. REV. 1417 (2014); and contributions to the *University of Chicago Law Review Symposium on Personalized Law*, 86 U. CHI. L. REV. 217-609 (2019). This literature ranges beyond evidence-based regulation as defined in this paper. It covers the use of systematic research not only to identify causal propositions but also to identify features of the environment in which law is operating (such as the psychological characteristics of the regulated actors).

[47] *See* generally, Lawrence M. Friedman, IMPACT: HOW LAW AFFECTS BEHAVIOR, 4-5 (2016) (describing causal chain between "legal acts" and behavior); Kevin E. Davis, *Legal Indicators: The Power of Quantitative Measures of Law*, 10 ANN. REV. L. & SOC. SCI. 37 (2014); Yong-Shik Lee, *General Theory of Law and Development*, 50 CORNELL INT'L L. J. 415, 423-428 (2015) (defining different components of law and how they affect development outcomes).

In addition, the ultimate impact of law on behavior depends on how the law, both in reality and as perceived, interacts with environmental factors such as moral attitudes, social structures and economic conditions.

There are two main obstacles to data collection: cost and measurement error. These problems are interrelated as it may be possible, but costly, to reduce measurement error. These problems also affect different types of legal data in different ways.[48]

At first glance, many kinds of legal data appear to be available at relatively low cost from reliable sources. Statutes, regulations, decrees and judicial opinions are often published. The numbers of legal officials and organizational flowcharts are often recorded for administrative purposes. And the practices of licensing and enforcement agencies are often governed by written rules set out in handbooks, manuals, and guidelines.

The principal challenge in collecting data on legal norms is to code or classify them in a way that makes them comparable. Legal norms can be complex: they can be located in multiple places, they typically are multi-dimensional, and they often are ambiguous. This is why private actors typically resort to trained lawyers to help them identify applicable legal norms. Researchers who cannot tap this kind of legal expertise are likely to find it difficult to collect accurate legal data, particularly when they attempt to study legal systems with which they are not personally familiar. The resulting errors can significantly compromise the results of an analysis. In one famous case, Holger Spamman found errors in legal data collected from thirty-three out of the forty-six countries analyzed in a widely-cited study on the impact of protecting shareholder rights. Once the errors were corrected many influential claims based on the original data could not be substantiated.[49]

As for legal institutions, aggregate data on their formal structures and the officials who inhabit them are often available. Typically, however, it is difficult to obtain data on officials' access to technology, how they are allocated across activities, their educational backgrounds, and

---

[48] For in-depth discussion of the challenges associated with collecting legal data and extensive references to the literature, see Siems, *supra* note 1.
[49] Holger Spamann, *The "Antidirector Rights Index" Revisited*, 23 REV. FIN. STUD. 467 (2010).

informal social ties.[50] Yet those additional kinds of data can be critical to understanding legal officials' ability to learn and engage in coordinated action, factors which might in turn be important determinants of institutional effectiveness.

With respect to both enforcement practices and the behavior of regulated actors, the major challenge is that relevant information often is concealed for strategic reasons. Actors want to conceal illicit behavior and regulatory agencies often do not publicize data on enforcement strategies in order to maintain strategic advantages.[51] For these reasons, successful enforcement actions often are publicized but unsuccessful investigations and unsanctioned misconduct are not. This point should not be overstated, however. Some forms of illicit behavior produce observable physical consequences, such as air pollution, water pollution, deforestation, defective construction, street prostitution and certain kinds of drug abuse. The prevalence of these signs of illicit activity also can be used to draw inferences about the prevalence of associated misconduct, such as corruption on the part of officials charged with regulating the activity in question.

Finally, survey data capturing perceptions of legal regulation on the part of both experts and the general population are available but tend not to be comprehensive. The most comprehensive source is the World Justice Project (WJP), which collects data in 113 countries on "the rule of law as experienced by ordinary people."[52] The WJP data come from two sources: (1) a survey of the general population in each country, and (2) a questionnaire sent to "in-country professionals with expertise in civil and commercial law, criminal justice, labor law, and public

---

[50] Gillian K. Hadfield, RULES FOR A FLAT WORLD: WHY HUMANS INVENTED LAW AND HOW TO REINVENT IT FOR A COMPLEX GLOBAL ECONOMY 214 (2017) ("we don't know…very much about how legal resources are allocated across different kinds of people, problems, and policies").

[51] Regulators also may have incentives to produce "policy-driven evidence", which might involve concealment or distortion of data in order to please donors, allies, powerful political actors, or the general public. For a discussion of this phenomenon in the African context see Justin Sandefur and Amanda Glassman, *The Political Economy of Bad Data: Evidence from African Survey and Administrative Statistics*, 51 J. DEV. STUD. 116 (2013). Since this Article is premised on the assumption of good faith on the part regulators this problem will not be discussed further.

[52] *See* World Justice Project, WORLD JUSTICE PROJECT RULE OF LAW INDEX 2017–2018, 156 https://www.worldjusticeproject.org/sites/default/files/documents/WJP_ROLI_2017-18_Online-Edition_0.pdf.

health."[53] The population surveys are distributed in the three largest cities of each country, but the locations of the in-country professionals are not specified.[54]

All of these categories of data tend to be more readily available for wealthier countries, generally because governments and researchers in poorer countries lack the resources, both financial and institutional, to collect and publish good data. Morten Jerven has documented the frailties of economic data collected by under-resourced African statistical agencies.[55] Legal data almost certainly suffer from similar shortcomings. For example, judicial data tend to be more prone to error in countries which have not implemented electronic filing of court documents ("e-filing"). However, poor countries tend not to invest in court automation—as of 2016 fewer than 6% of countries in Africa and the Middle East, Latin America and the Caribbean and South Asia had implemented e-filing.[56]

Data also tend to be much more readily available for national and supranational—as opposed to subnational or extraterritorial—regulation. This bias compromises the value of data in any domain where the effects of law reflect the combined influence of multiple levels of regulation.[57] This is potentially significant because multilevel regulation appears to be widespread. In most countries the legal system operates at multiple levels, e.g. national, state or provincial, and local or municipal. In addition, for many actors, regardless of where they consider themselves to be located, influential regulation also emanates from supranational bodies, such as the institutions of the European Union, the Inter-American Court of Human Rights, the International Monetary Fund, or the World Bank. In certain spheres, including competition law, tax, privacy, corruption,

---

[53] *Id.* at 157.

[54] *Id.* at 157–58 (stating that the population survey is a sample from 1000 people in the "three largest cities of each country" while the in-country professionals are selected through directories and referrals and vetted by the World Justice Project).

[55] Morten Jerven, POOR NUMBERS: HOW WE ARE MISLED BY AFRICAN DEVELOPMENT STATISTICS AND WHAT TO DO ABOUT IT (2013).

[56] Heike Gramckow et al, GOOD PRACTICES FOR COURTS REPORT : HELPFUL ELEMENTS FOR GOOD COURT PERFORMANCE AND THE WORLD BANK'S QUALITY OF JUDICIAL PROCESS INDICATORS - KEY ELEMENTS, LESSONS LEARNED, AND GOOD PRACTICE EXAMPLES, 73 (2016).

[57] *See* generally, Kevin E. Davis, *Data and Decentralization: Measuring the Performance of Legal Institutions in Multilevel Systems of Governance*, 103 MINN. L. REV. 1619 (2018) (discussing inadequacies of existing performance measures for legal institutions in multilevel systems of governance); Ouellette, *supra*,

terrorist finance, and money laundering, extraterritorial regulation by powerful actors such as the United States and the European Union is also very influential.

B. CAUSAL INFERENCE

The "evidence" in evidence-based regulation consists of conclusions about the consequences of a particular regulatory intervention. In other words, evidence means conclusions about the causal role of interventions implemented in the past. One line of attack on evidence-based regulation challenges the validity of these causal claims about past interventions, also known as the "internal validity" of the relevant studies.[58]

In principle, inferences about the causal role of legal regulation can be drawn through close examination of individual outcomes and intermediate causal mechanisms. For example, a case study of an individual firm's behavior might reveal the extent to which its compliance decisions were influenced by the applicable legal regime by documenting the influence of the law on the beliefs and decisions of key actors. This technique, often referred to as "process tracing,"[59] is of limited value when a policy intervention is expected to have heterogeneous or probabilistic effects. This is typically expected to be the case with legal regulation—a classic example is Oliver Wendell Holmes' well-known distinction between the likely effects of law on the "bad man" and the "good man."[60] In addition, in the regulatory context process tracing requires access to data on enforcement strategies or illicit activity. The relevant actors may be reluctant to disclose this kind of information as part of a small-scale study because of the risk that it will be attributed to them.

On account of the limitations of process tracing, a large proportion of the evidence about the causal role of legal regulation is derived from statistical analyses of correlations between, on the one hand, implementation of the regulatory intervention, and on the other hand, the outcome of interest, across multiple units of analysis. The challenge is to justify the inference that the

---

[58] D. T. Campbell, *Factors Relevant to the Validity of Experiments in Social Settings*, 54 PSYCH. BULL. 297 (1957). On the challenges of causal inference in evidence-based policy see Ian Sanderson, *Evaluation, Policy Learning and Evidence-Based Policy Making*, 80 PUBLIC ADMI. 1 (2002). For discussions of the challenges of causal inference in legal settings see Spammann, *supra* note 1; Green and Thorley, *supra* note 1.

[59] On process tracing, see ALEXANDER L. GEORGE & ANDREW BENNETT, CASE STUDIES AND THEORY DEVELOPMENT IN THE SOCIAL SCIENCES 205–32 (2004); David Collier, *Understanding Process Tracing,* 44 PS: POL. SCI. & POL. 823 (2011); James Mahoney, *The Logic of Process Tracing Tests in the Social Sciences*, 41 SOC. METHODS & RES. 570 (2012).

[60] Oliver Wendell Holmes, Jr., *The Path of the Law*, 10 HARV. L. REV. 457 (1897).

correlation is explained by the fact that the intervention has caused the outcome. This means excluding pure coincidence. It also means rejecting the hypothesis that the outcome was caused by one or more other factors that covaried with, or even caused, adoption of the intervention.

A popular way to draw causal inferences is to use econometric techniques to estimate the functional relationship between measures of potential causal factors and outcomes. For example, a study might examine whether there is a linear relationship between the number of police officers, population, income per capita and crime rates.[61] The challenges associated with using econometrics to infer causality are well-known. One broad challenge, which is not specific to legal regulation, is to specify the function that best fits the data, both in terms of the functional form and the parameters it includes. For example, should the function include population or the logarithm of population? Should the unemployment rate be included as a parameter? It is becoming increasingly common to address this task with the assistance of machine learning methods.[62]

The second broad challenge in using econometrics for causal inference is to distinguish correlation from causation. Success involves ruling out the possibility that an unobserved causal factor—say, the abortion rate at the time of birth of the criminal-aged population—explains the correlation. This challenge frequently arises in analyzing the causal role of regulatory interventions. It often is difficult to rule out reverse causality because the political forces that lead to interventions might be caused by negative outcomes. For example, high crime rates might cause expansions of police forces.[63]

In theory, a compelling way to rule out the possibility that an unobserved factor has influenced both the introduction of the policy and outcomes is through an RCT which randomizes implementation of the policy across the units to be studied. In recent years it has become increasingly popular for social scientists to exploit this feature of RCTs to evaluate the causal impact of policies.[64] Consider, for example, how an RCT might be used to gather evidence of

---

[61] See e.g., Steven D. Levitt, *Using Electoral Cycles in Police Hiring to Estimate The Effects of Police on Crime: Reply*, 92 AM. ECON. REV. 1244 (2002); Thomas B. Marvell and Carlisle E. Moody, *Specification Problems, Police Levels, and Crime Rates*, 34 CRIMINOLOGY 609 (1996).

[62] *See* generally, Susan Athey and Guido W. Imbens, *The State of Applied Econometrics: Causality and Policy Evaluation*, 31 J. ECON. PERSP. 3, 22-27 (2017).

[63] Marvell and Moody *supra*; Ouellette, *supra*, 81 (increased research and development might cause rather than be caused by changes in patent laws through lobbying by research-intensive industries).

[64] For a survey of RCTs involving legal institutions see Green and Thorley, *supra* note 1.

whether or not adding police officers plays a causal role in reducing crime.[65] The first step would be to select a group of sites to be studied from the broader population of sites in which we are interested. Suppose we find 100 sites (the study group). From this group we could randomly select 50 sites (the treatment group) to receive 5 additional police officers, while the remaining 50 sites (the control group) receive no special treatment. If sites in the treatment group average, say, 10 per cent less crime than sites in the control group, then we have support for the inference that the additional police caused a 10 per cent the reduction in crime.

There are, however, several well-recognized reasons why this inference may not be valid. To begin, RCTs are valid methods of estimating average treatment effects, but this does not mean that every single RCT yields an accurate measure of the average treatment effect.[66] The estimates produced by RCTs are only accurate on average. It is possible for the treatment effect in a single trial to be positive even if the treatment effect averaged across repeated trials would be zero. For example, perhaps by pure chance the sites in our control group experienced hotter weather and that was the sole cause of their higher crime rates. To obtain a more precise estimate of the average treatment effect we would want to replicate the trial, perhaps several times. Over the course of several trials we would expect the control and treatment sites to experience the same weather—as well as other confounding effects—on average.

The treatment effect observed in an RCT only represents a valid estimate of the average treatment effect if the treatment is the only reason to expect average outcomes to differ between the treatment group and the control group. All bets are off if this condition is violated. Of particular concern are 'spillover effects,' in which the treatment has somehow affected outcomes in the control group.[67] The spillover problem is especially salient for interventions that involve legal regulation. One reason for this is because the subjects of laws often interact with one another and so any intervention that has a meaningful direct effect on one subject is likely to affect others. Second, information about law often is disseminated broadly—though imperfectly—and perceptions of law are an important channel through which law influences behavior.

---

[65] See e.g., Lawrence W. Sherman and David Weisburd, *General Deterrent Effects of Police Patrol in Crime "Hot Spots": A Randomized, Controlled Trial*, 12 JUSTICE QUART. 625 (1995).
[66] Deaton and Cartwright, *supra*.
[67] Green & Thorley, *supra,* 60.

The first kind of spillover arises when members of the control group have significant interactions with members of the treatment group. This can occur even if legal norms formally apply only to firms or individuals in the treatment group. For example, Abramowicz et al propose that firms be randomly exempted from certain requirements under the U.S. securities laws.[68] However, as they acknowledge, this could give firms in the treatment group (the exempted firms) a competitive advantage over firms in the control group. This would tend to bias any observed treatment effect towards over-estimating the actual effect of imposing the treatment on all firms in the study group.[69] A similar kind of bias will arise in our hypothetical RCT if the greater police presence in the treatment sites encourages wrongdoers to shift criminal activities to the control sites.

Legal regulations also are prone to a second kind of spillover: informational spillovers. One spillover of this sort occurs when outcomes in the control group are affected by its members' knowledge that they have been assigned to the control group, or even the mere knowledge that they are participating in a trial. Control group members might, for instance, behave differently because they resent being denied the treatment or because they know they are being observed. To limit this problem researchers who conduct RCTs strive to ensure that the participants are blind, meaning, they are unaware of the group to which they have been assigned. This is why medical researchers offer subjects in their control groups placebos. Ideally, the subjects would not even realize they were participating in a trial, but informed consent requirements generally make this impossible in medical settings. Lawmaking is generally a public affair and so it is difficult to conduct blind trials of policies that involve enactment of laws. It is, however, often feasible in trials that vary the structure and practices of enforcement agencies since those are not always publicized.[70]

It also will be difficult to test the behavioral effects of targeting a randomly-selected group of firms or individuals for more vigorous enforcement if members of the control group are likely to observe and be influenced by enforcement practices or behavior in the study group. For example, suppose that at least one widely accessible media outlet reports on police practices and criminal

---

[68] Abramowicz et al, *supra*, 991-997.
[69] Abramowicz et al, *supra*, 994.
[70] Green & Thorley, *supra,* 60; Abramowicz et al, *supra*, 949-951.

behavior across the entire area covered by our policing RCT. People in the control sites might observe reports about the actions of police in the treatment sites, believe that the practices apply to them, and act accordingly. Alternatively, members of the control group might observe and emulate any increased misconduct in the treatment group.[71]

C. Generalization

Suppose we have conclusive evidence that a particular policy has played a positive causal role in one or more specific contexts. How helpful will this evidence be in predicting the consequences of adopting the same policy in another context? How helpful will it be in predicting policies besides the ones studied? For example, if we add 5 police officers to sites outside the study group will crime fall by 10 per cent? Will the reduction in crime in the original study group persist for the next year? Will adding 10 police instead of 5, either inside or outside of the study group, cause a 20 per cent reduction in crime? This line of questioning challenges what is known as the "external validity" of the relevant studies.[72]

To be clear, even if adding 5 police officers plays exactly the same causal role in other contexts as reported by a study of previous implementation, we will not know the precise effects of implementing the policy in a new site. For one, most techniques of causal inference yield only noisy estimates of causal relationships, to the point where it is considered standard practice to report an estimate of the error along with the estimate itself. In addition, the prior study may provide limited information. For example, RCTs generally are designed only to reveal average treatment effects. The actual treatment effect for any given unit might be different from the average. Even if an RCT tells us that, across multiple sites, the mean effect of adding the officers has been and will be a 10 per cent reduction in crime, this is a far cry from proof that there will be a 10 per cent reduction in crime in any given site, or even that there will be any reduction at all.

These points aside, there is no general reason to believe that a particular legal intervention will play the same causal role in different contexts or timeframes from the one studied, or that

---

[71] This kind of spillover effect might also arise in the trial proposed by Abramowicz et al. Firms exempted from the securities law obligations they identify might engage in more fraud which might, at least according to some theories of criminal behavior, increase the propensity for crime among firms in the treatment group.
[72] Campbell, *supra*.

similar legal interventions will play similar causal roles.[73] The kinds of outcomes influenced by legal regulation tend to be influenced by a large number of different and ever-changing causal factors.[74] There is no particular reason to believe that the mix of factors, or even their cumulative effect, will remain constant across space and time. In fact, studies of complex social systems suggest that they are capable of producing outcomes ranging from stable equilibria to randomness, and that outcomes, or even the class of outcomes (equilibrium, random, etc.), can be sensitive to initial conditions and path dependent.[75]

So, for instance, the principle that determines the causal role of additional police is unlikely to be as simple as 'more police leads to less crime.' It is more likely to be a conditional principle that reflects the presence of additional causal factors, such as 'more police leads to less crime so long as the additional police are honest and make themselves visible to the population (perhaps because the treatment group has not been blinded to the fact that they are participating in a study).' Alternatively, the intervention's contribution to the outcome may be determined by its interaction with other factors, resulting in a causal principle such as, 'more police leads to less crime so long as the population believes that the police exercise legitimate authority.'

If an intervention's causal role depends in part on supporting causal factors then its effects are likely to vary over time as the mix of factors changes. For example, as crime falls police may become complacent, potential victims may become less vigilant,[76] or criminals may become more creative.[77] Any or all of these factors may cause a resurgence of crime beyond the time frame of the original study. Theoretical models that allow for these sorts of influences on crime permit multiple classes of outcomes to appear over time. Depending on the model and initial conditions,

---

[73] For similar discussions of generalizability see Michael Woolcock, *Using Case Studies to Explore the External Validity of 'Complex' Development Interventions*, 19 EVALUATION 229 (2013); Ouellete, *supra*, 100-102 (discussing randomization of innovation policies across countries); Abramowicz et al, *supra*, 951-7. See also, Kevin E. Davis, *Legal Universalism: Persistent Objections*, 60 U. Tor. L.J. 537 (2010) (challenging claim that any given legal institution will play the same causal role in promoting development in different contexts).
[74] Manzi, *supra* (arguing that outcomes in social systems tend to be characterized by high causal density).
[75] Page, *supra*, 25, 28.
[76] Thomas J. Philipson and Richard A. Posner, *The Economic Epidemiology of Crime*, 39 J. L. & ECON. 405 (1996) (model in which potential victims alter their precautions in response to changes in public enforcement).
[77] Joanna Sooknanan, Balswaroop Bhatt and Donna Marie Giselle Comissiong, *Criminals Treated as Predators to Be Harvested: A Two Prey One Predator Model with Group Defense, Prey Migration and Switching*, 4 J. MATH. RES. 92 (2012) (analyzing a model in which both criminals and victims can adapt by switching locations and victims can take precautions).

crime may fluctuate periodically around a fixed point, manifest jumps from one equilibrium or set of periodic fluctuations to another, or fluctuate randomly.[78]

The impact of an intervention also might vary with its scale, in other words, the number of actors it affects. A policing intervention that is effective in deterring crime in a small treatment group may have quite different effects when it is scaled up to the point where criminals cannot escape its effects by moving to another jurisdiction. The large scale intervention might have a sufficiently large effect on criminals' earnings to reduce the number of people who decide to pursue a life of crime. Alternatively, the effect may be to induce criminals to invest in technology, such as firearms or malware, that allows them to fight or evade the police.[79]

Finally, in a complex world there is no reason to assume that similar interventions will lead to similar outcomes. For instance, adding a second squad of 5 police officers need not have the same consequences as adding the first squad. It is tempting to assume that the relationship between police and crime is a simple linear one, so that each increment in the number of police leads to a proportional reduction in crime. However, the relationship may well be non-linear. For example, it might be the case that 'more police leads to less crime, but each additional police officer has a diminishing, and ultimately negative, impact as the population becomes more defiant.'

D. CAPACITY

So far we have discussed reasons why it might be challenging to produce accurate evidence on the effectiveness of regulation. Some decision-makers will find it more difficult than others to overcome those challenges. The burgeoning literature on state capacity shows that there can be significant variations in the ability of public sector institutions' ability to implement policy, both across and within countries.[80] In cases of low capacity, a common source of the problem is lack of resources, including skilled personnel. Gathering and reviewing social scientific evidence requires considerable investments of time on the part of people who are both trained in social science and

---

[78] Philipson and Posner, *supra* (predicting that crime rate will cycle around a fixed point); Sooknanan et al, *supra* (predicting that crime rates may be either stable, unstable or cyclical).
[79] *Cf.* Ouellette, *supra*, 81-82 (experiments on innovation policy are difficult to interpret because the effects will vary depending on how much of the relevant market they affect).
[80] *See* generally, Matt Andrews, Lant Pritchett, and Michael Woolcock, BUILDING STATE CAPABILITY: EVIDENCE, ANALYSIS, ACTION, 20 (2017) (ranking countries in terms of state capability). See also, Lee, *supra*, 453, 465-7 (discussing importance of states' "evaluative and analytical capacity").

familiar with local conditions.[81] Public sector institutions in many poor countries lack access to these sorts of human and financial resources. Consequently, they lack the capacity to implement the evidence-based approach to policymaking.

Lack of state capacity is likely to present a particularly challenging obstacle to evidence-based enforcement. Law enforcement often is a highly decentralized activity that involves discretionary actions on the parts of large numbers of officials. Most countries can muster the small number of skilled personnel required for tasks such as lawmaking, but countries with limited capacity struggle to assemble large numbers of skilled professionals to engage in ongoing tasks such as enforcement.[82]

III.CASE STUDY: REGULATION OF BRIBERY[83]

The preceding Part suggests that the multifaceted challenges associated with evidence-based regulation are likely to be most visible in legal regimes which: regulate illicit behavior; have a broad geographical scope, encompassing both rich and poor countries; apply to a heterogeneous set of actors operating in varied environments who nonetheless interact with one another along multiple dimensions; and, are implemented in a decentralized fashion by a heterogeneous set of institutions with overlapping jurisdiction. Regulation of vehicular traffic may not raise these concerns, but several regimes that regulate business activities do, including those concerning competition, tax, privacy, money laundering and terrorist finance.

The anti-bribery regime is a particularly good example of a regime in which the evidence-based approach is likely to be of limited value. In the first place, the anti-bribery regime regularly provides for overlapping jurisdiction. This is only in part because corrupt transactions frequently cross geographic boundaries. Overlapping jurisdiction is also a deliberate response to the problem of systemic bribery. In some societies bribery is so widespread and pernicious that local anti-

---

[81] For discussion of factors human beings ought to consider in making use of evidence derived from analysis of big data see, Saul Levmore and Frank Fagan, *The Impact of Artificial Intelligence on Rules, Standards, and Judicial Discretion*, 93 S. CAL. L. REV. (forthcoming 2019).

[82] Andrews et al, *supra*, 107-110 (contrasting "policymaking" with "implementation-intensive imposition of obligations" in terms of the required amounts of institutional capability).

[83] Material in this part is adapted from Davis, *supra*.

corruption institutions have been overwhelmed or subverted and so it is helpful to be able to draw upon foreign institutions for support. The anti-bribery regime also covers an especially heterogeneous set of actors—in recent years the focus of the regime has been upon transactions that involve multinational corporations based in OECD countries paying bribes to public officials in poor countries.

The following sections provide an overview of the global anti-bribery regime and then discuss the challenges associated with adopting an evidence-based approach to regulation of bribery.

A. OVERVIEW OF THE ANTI-BRIBERY REGIME

Every country in the world prohibits bribery of its own public officials. These prohibitions are contained in laws enacted by various levels of government. In the United States, for example, criminal prohibitions on bribery are contained in several federal statutes[84] as well as state penal codes.[85] It is not uncommon for those prohibitions to be enforced by several agencies across multiple levels of government. Brazil is a case in point. Prohibitions found in national anti-bribery laws can be investigated either by state or federal police forces, a variety of independent agencies, or federal or state public prosecutors.[86] Those investigations can lead to administrative, civil, or criminal proceedings initiated by either an independent agency or the federal or state public prosecutor.[87] Judicial proceedings can take place in either state or federal courts.[88] In complex cases, it is not uncommon for multiple agencies to be involved, with varying levels of coordination.[89]

---

[84] 18 U.S.C. § 201 (bribery of public officials), 18 U.S.C. § 666 (bribery concerning programs receiving federal funds), 18 U.S.C. § 1951 (extortion statute, also known as the "Hobbs Act", and interpreted to cover bribery).

[85] *See e.g.* New York Penal Law, §§ 200 *et seq*.

[86] *See* Mariana Mota Prado & Lindsey Carson, *Brazilian Anti-Corruption Legislation and Its Enforcement: Potential Lessons for Institutional Design*, 4 J. SELF-GOVERNANCE & MGMT. ECON. 34, 47–49 (2016) (discussing institutional multiplicity in corruption investigations in Brazil).

[87] *See id.* at 49–52.

[88] *See* Lucas Dotto Borges et al., Bribery and Corruption 2018: Brazil, GLOBAL LEGAL INSIGHTS, https://www.globallegalinsights.com/practice-areas/ bribery-and-corruption-laws-and-regulations/brazil (last visited Mar. 27, 2018) (explaining that "[t]he judicial branch (divided into Federal and State jurisdictions according to the nature of the matters being discussed) is formed by first instance courts, appellate courts, and Superior courts").

[89] *See* Prado & Carson, *supra* (discussing the various Brazilian institutions investigating corrupt actions). For a discussion of a particularly complex case, see Kevin E. Davis et al., *Transnational Anticorruption Law in Action: Cases from Argentina and Brazil*, 40 L. & SOC. INQUIRY 664 (2015).

In recent years extraterritorial regulation has become a prominent feature of the anti-bribery regime. This trend began with the enactment of the US Foreign Corrupt Practices Act of 1977 ("FCPA").[90] The most prominent feature of the FCPA is a series of prohibitions, backed by stiff civil and criminal penalties, on payments to foreign public officials in order to assist in "obtaining or retaining business."[91] These are known as the FCPA's anti-bribery provisions. Just as important but somewhat less prominent are the FCPA's books and records provisions, which require firms with securities listed on US exchanges to keep accurate records.[92] These recordkeeping obligations are complemented by a separate obligation to maintain internal controls that ensure the integrity of corporate records.[93]

Although the FCPA is a US statute it applies to many firms with only tenuous connections to the US. To begin with, the current versions of the FCPA's anti-bribery provisions apply to corrupt practices committed anywhere in the world by US citizens or permanent residents, or by corporations that are incorporated or headquartered in the US.[94] In addition, both the anti-bribery provisions and the accounting provisions apply to firms that list their securities on US markets.[95] Finally, the FCPA applies to anyone who violates the statute "while in the territory of the United States."[96] Strictly speaking, foreign individuals or corporations are only caught by the anti-bribery provisions if they act while in US territory or, in the case of foreign issuers, make use of "the mails or any means or instrumentality of interstate commerce." However, US enforcement agencies take the position that even a wire transfer involving the US financial system or an email passing through US servers will satisfy these requirements.[97]

---

[90] Pub. L. No. 95-213, 91 Stat. 1494 (1977) (codified as amended at 15 U.S.C. §§ 78dd-1, 78dd-2, 78dd-3, 78ff, 78m(b), (d)(1), (g)-(h) (2012)), amended by Foreign Corrupt Practices Act Amendment of 1988, Pub. L. No. 100-418, 102 Stat. 1107, 1415 (1988) (codified at §§ 78dd-1 to 78dd-3, 78ff.

[91] 15 U.S.C. §78 dd-1 to 78dd-3.

[92] 15 U.S.C. §78m(b)((2)(A).

[93] 15 U.S.C. §78m(b)(2)(B).

[94] FCPA § 78dd-2.

[95] FCPA §§78dd-1, 78m(b)(2).

[96] FCPA § 78dd-3.

[97] Criminal Division of the U.S. Department of Justice and the Enforcement Division of the U.S. Securities and Exchange Commission, A RESOURCE GUIDE TO THE US. FOREIGN CORRUPT PRACTICES ACT 11 (2012) ("placing a telephone call or sending an e-mail, text message, or fax from, to, or through the United States involves interstate commerce—as does sending a wire transfer from or to a U.S. bank or otherwise using the U.S. banking system"); Information, United States v. ABB Vetco Gray Inc., et al., No. 04-cr-279 ¶¶ 39, 41 (S.D.Tex. Jun. 22, 2004) (In a settled case, asserting that foreign non-issuer violated FCPA while in US territory by causing agents to wire funds to a U.S. bank account while outside the United States and by receiving an email sent from the US to Scotland).

Since the enactment of the FCPA, most countries in the world have signed treaties committing them to prosecute firms for paying bribes to foreign public officials and to help other countries to recover proceeds of corrupt activity from corrupt officials. Those treaties include the extremely influential OECD Convention on Combating Bribery of Foreign Public Officials in International Business Transactions ("OECD Convention")[98] and the broad-ranging United Nations Convention Against Corruption ("UNCAC")[99] as well as several regional agreements.[100] In addition, in 2017 the members of the UN General Assembly resolved to "substantially reduce corruption and bribery in all their forms" as one of the Sustainable Development Goals (Goal 16, Target 16.5).[101] Many countries, along with the major international development banks, also have the power to 'debar' firms that have engaged in corrupt practices, that is to say, refuse to do business with them in the future.[102] Several countries have followed up on these formal commitments with meaningful enforcement actions. Finally, courts and arbitrators around the world are increasingly reluctant to enforce contracts tainted by bribery.[103]

A body known as the Working Group on Bribery in International Business Transactions (the "OECD Working Group") is officially charged with monitoring the implementation and enforcement of the OECD Convention and its related instruments. The Working Group is made up of representatives from the parties to the Convention and is assisted by a Secretariat at OECD headquarters in Paris. Consistent with OECD practice, the OECD Working Group has embraced

---

[98] Nov. 21, 1997, 37 I.L.M.

[99] G.A. Res. 4, U.N. GAOR, 58th Sess., Agenda Item 108, U.N. Doc. A/RES/58/4 (Oct. 31, 2003).

[100] *African Union Convention on Preventing and Combating Corruption*, July 11, 2003, 43 I.L.M. 5 (2004); *Civil Law Convention on Corruption*, Council of Europe, E.T.S. No. 174, November 4, 1999, in force November 1, 2003; *Criminal Law Convention on Corruption*, Council of Europe, January 27, 1999, E.T.S. No. 173; *Inter-American Convention Against Corruption*, March 29, 1996, 35 I.L.M. 724 (1996).

[101] G.A. Res. 71/313 (July 6, 2017).

[102] *See* generally, Anne-Marie Leroy & Frank Fariello, THE WORLD BANK GROUP SANCTIONS PROCESS AND ITS RECENT REFORMS (2012) and Laurence Boisson de Chazournes & Edouard Fromageau, *Balancing the Scales: The World Bank Sanctions Process and Access to Remedies*, 23 EUR. J. INT'L L. 963 (2012).

[103] Metal-Tech Ltd. v. The Republic of Uzbekistan, ICSID Case No. ARB/10/3, Award (Oct. 4, 2013); World Duty Free Company Limited v. the Republic of Kenya, ICSID Case No. Arb/00/7, Award, ¶¶ 138-157, 172 (Sept. 25, 2006); Kevin E. Davis, *Contracts Procured Through Bribery of Public Officials: Zero Tolerance versus Proportional Liability*, 50 N.Y.U. J. INT'L L. & POL. 1261 (2018).

the evidence-based approach to anti-bribery law,[104] and part of its mandate is to accumulate know-how and develop best practices.[105]

There has been relatively little academic discussion of evidence-based approaches to anti-bribery law. However, at least one prominent scholar, Alina Mungiu-Pippidi, has explicitly recommended an evidence-based approach to anti-corruption policy, which includes anti-bribery law.[106] In addition, one of the RCTs recommended by Abramowicz et al would involve exempting randomly selected firms from the internal controls provisions of the FCPA.[107]

B. DATA COLLECTION

1. Data on laws and enforcement practices

Laws on bribery are invariably published and so tend to be widely accessible. Data on enforcement, however, are much less accessible. Official enforcement data are collected in a haphazard fashion. In the United States, for example, data on enforcement of federal bribery laws by federal prosecutors are reasonably accessible but of dubious quality.[108] Meanwhile, data on bribery prosecutions by state and local authorities in the US are not collected in any central location and may not even be collected by the relevant agencies.

Survey data on perceptions of enforcement of anti-corruption laws, which do not necessarily break out anti-bribery laws, are collected by multilateral development banks and various private actors. For example, as part of its Country Policy Institutional Assessment the

---

[104] OECD Working Group on Bribery, ANNUAL REPORT 2006 (2006), 13; Fritz Heimann and Mark Pieth, CONFRONTING CORRUPTION: PAST CONCERNS, PRESENT CHALLENGES, AND FUTURE STRATEGIES (2018), 84-86.

[105] *See e.g.*, Message from Angel Gurria in, OECD Working Group on Bribery, OECD WORKING GROUP ON BRIBERY 2010 ANNUAL REPORT, 3 (2010) ("By continuing the OECD tradition of developing evidence-based public policies and standards like the Anti-Bribery Convention, we will ensure a stronger, cleaner and fairer world economy.")

[106] See, Alina Mungiu-Pippidi, *The time has come for evidence-based anticorruption*, 1 NATURE HUMAN BEHAVIOUR Article no. 0011 (2017).

[107] Abramowicz et al discuss exempting randomly selected firms or industries from the internal controls provision of Sarbanes-Oxley Act, which overlaps significantly with the internal controls provision of the FCPA. See Abramowicz et al, *supra*, 993 (proposing RCT) and Management's Report on Internal Control Over Financial Reporting and Certification of Disclosure in Exchange Act Periodic Reports, Exchange Act Release Nos. 33-8238, 34-47986, Investment Company Act Release No. IC-26068, at pt.II(A) & V(B). (Aug. 14, 2003), http://www.sec.gov/rules/final/33-8238.htm [http://perma.cc/XZN4-GMZ5] (discussing overlap between Sarbanes-Oxley Act and the FCPA).

[108] Adriana S Cordis and Jeffrey Milyo, *Measuring public corruption in the United States: Evidence from administrative records of federal prosecutions*, 18 PUBLIC INTEGRITY 127 (2016) (describing limitations of two sources of official data).

World Bank asks its staff to rate countries on, among other things, "the accountability of the executive and other top officials to effective oversight institutions."[109] Similar data are collected in multiple countries by other multilateral development banks, as well as private organizations such as the Bertelsmann Stiftung, and the World Justice Project.

Data on enforcement of prohibitions on transnational bribery are more readily accessible than data concerning purely domestic misconduct. This is largely on account of the efforts of the OECD Working Group, which has played a critical role in collecting and disseminating information about both laws and enforcement actions. One of the Working Group's main activities is the operation of a peer review system in which each member's performance in implementing the OECD Convention is reviewed by a team of experts from other member states.[110] The country reviews have proceeded in phases. Phase 1 was limited to evaluation of whether the country's legislation complied with the terms of the Convention. Subsequent phases have examined enforcement as well as follow-up on recommendations from previous phases. The reports resulting from these country reviews are all published on the OECD website.[111]

As far as published data on enforcement are concerned, since 2010 the OECD Working Group has collected and published annual data on completed enforcement actions instituted by the parties to the Convention. The 2016 edition of this report included for the first time data on the prevalence of international co-operation in enforcement.[112] The OECD's published data are supplemented by data produced by Transparency International, an international nonprofit organization based in Berlin. Transparency International periodically produces a glossy report on the quality of countries' implementation of the OECD Convention.[113]

---

[109] World Bank, CPIA CRITERIA 2017, 43 (2017).
[110] http://www.oecd.org/daf/anti-bribery/, accessed March 2, 2018.
[111] Id.
[112] OECD Working Group on Bribery, 2016 DATA ON ENFORCEMENT OF THE ANTI-BRIBERY CONVENTION: SPECIAL FOCUS ON INTERNATIONAL CO-OPERATION (2017). Curiously, the data on international co-operation only covered enforcement actions pursued to completion by the US Department of Justice and SEC; they were collected from press releases in which those agencies acknowledged assistance from foreign enforcement agencies. Id, 8.
[113] Gillian Dell and Andrew McDevitt, EXPORTING CORRUPTION: PROGRESS REPORT 2018: ASSESSING ENFROCEMENT OF THE OECD ANTI-BRIBERY CONVENTION (2018); Fritz Heimann, Ádám Földes, Sophia Coles, EXPORTING CORRUPTION, PROGRESS REPORT 2015: ASSESSING ENFORCEMENT OF THE OECD CONVENTION ON COMBATTING FOREIGN BRIBERY (2015).

The OECD Working Group also helps to disseminate data among regulators on a confidential basis. It hosts biannual meetings of law enforcement officials which include reviews of open enforcement actions in a practice known as the *tour de table*.[114] These meetings apparently play an important role in both monitoring and promoting countries' enforcement of prohibitions on foreign bribery, but the proceedings are confidential. Since 2010 UNCAC has used a peer review mechanism similar to the one employed by the OECD Working Group. So far the UN mechanism has been less successful. Fewer than half of the parties (75 out of 162) have allowed full versions of the reports on their first review to be published on the UN website.[115]

Another potentially useful source of enforcement data relating to transnational bribery is the World Bank, which publishes an annual report on the steps it has taken to enforce its rules on fraud and corruption.[116] The report includes information on whether cases have been referred to the national authorities for further investigation and whether the World Bank is aware of any further action by those authorities.[117] These data on referrals are among the few sources of insight into how national enforcement agencies respond when credible allegations of corruption are brought to their attention.

2. Data on outcomes of interventions

Even if comprehensive data were available on anti-bribery interventions, evidence-based decision-making would be greatly hampered by the scarcity of data on the outcomes associated with those interventions, meaning, data on the prevalence and consequences of bribery.[118]

A few ad hoc efforts have been made to collect data on the incidence of bribery through direct observation, for example, by sending observers to ride with truck drivers who might pay

---

[114] OECD Working Group on Bribery, ANNUAL REPORT 2006 (2006), 13; Fritz Heimann and Mark Pieth, CONFRONTING CORRUPTION: PAST CONCERNS, PRESENT CHALLENGES, AND FUTURE STRATEGIES (2018), 84-86.

[115] UN Office of Drugs and Crime, Country Profiles, accessed February 15, 2018, https://www.unodc.org/unodc/en/treaties/CAC/country-profile/index.html

[116] *See e.g.,* World Bank Group, WORLD BANK GROUP SANCTIONS SYSTEM ANNUAL REPORT FY18 (2018).

[117] *Id.*, 63-64.

[118] For a survey of the sources of data on the incidence of corruption see Tina Søreide, CORRUPTION AND CRIMINAL JUSTICE: BRIDGING ECONOMIC AND LEGAL PERSPECTIVES (2016), 64-73.

bribes to police or customs officers,[119] or collecting data directly from firms[120] or public officials.[121] Creative scholars have also developed ways to estimate levels of bribery by looking for gaps or anomalies in data that suggest hidden or illicit behavior.[122] However, most of what we know, or think we know, about the incidence of bribery continues to come from surveys, whether of individuals, firms or experts. Respondents are generally asked to provide information about either their own experiences—for example, "in any of [your inspections or meetings with tax officials in the last year] was a gift or informal payment expected or received"—or their perceptions of other peoples' experiences—for example, "how pervasive is political corruption?"[123]

A wide range of public and private actors field surveys that ask about the incidence of bribery. The UN Office on Drugs and Crime, agencies of national governments and various non-governmental organizations all also collect data on either experiences with or perceptions of bribery.[124] Among surveys conducted across multiple countries, the best known are probably Transparency International's Global Corruption Barometer and Bribe-Payer's Index.[125] In 2017, under the auspices of the project to establish Sustainable Development Goals, the members of the United Nations General Assembly agreed to measure progress toward the goal of "substantially reducing corruption and bribery" with two survey-based indicators:

---

[119] Benjamin A. Olken and Patrick Barron, *The Simple Economics of Extortion: Evidence from Trucking in Aceh*, 117 J. POL. ECON. 417 (2009).

[120] Sandra Sequeira and Simeon Djankov, *Corruption and Firm Behavior: Evidence from African Ports*, 94 J. INT'L ECON. 277 (2014) (data from customs clearing agents); Shawn Cole, and Anh Tran, *Evidence from the Firm: A New Approach to Understanding Corruption*, INTERNATIONAL HANDBOOK ON THE ECONOMICS OF CORRUPTION (2011): 408 (data from a firm).

[121] John McMillan and Pablo Zoido, *How to Subvert Democracy: Montesinos in Peru*, 18 J. ECON. PERSP. 69 (2004) (data from records of a public official).

[122] See generally Sandra Sequeira, *Advances in Measuring Corruption in the Field*, in Danila Serra and Leonard Wantchekon (eds.), NEW ADVANCES IN EXPERIMENTAL RESEARCH ON CORRUPTION (2012), 145-175 (surveying methods of measuring the prevalence of corruption).

[123] Sometimes surveys ask about the experiences of people "like" the respondent. These questions can be interpreted as requesting either of these types of information. For example, one World Bank survey asks, "When establishments like this one do business with the government, what percent of the contract value would be typically paid in informal payments or gifts to secure the contract?" World Bank, WORLD BANK ENTERPRISE SURVEY: MANUFACTURING MODULE (2018), Question J.6, http://www.enterprisesurveys.org/Data. A respondent might reasonably decide to answer based on their own experience, but if they believe their situation is atypical they might report perceptions of other firms' experiences.

[124] See, e.g., United Nations Office on Drugs and Crime, CORRUPTION IN NIGERIA. BRIBERY: PUBLIC EXPERIENCE AND RESPONSE (2017).

[125] See, https://www.transparency.org/cpi and https://www.transparency.org/research/bpi/overview.

16.5.1 Proportion of persons who had at least one contact with a public official and who paid a bribe to a public official, or were asked for a bribe by those public officials, during the previous 12 months.

16.5.2 Proportion of businesses that had at least one contact with a public official and that paid a bribe to a public official, or were asked for a bribe by those public officials during the previous 12 months.[126]

Some, but not all, of these surveys make it possible to measure levels of transnational as opposed to domestic bribery or corruption. Transparency International's Bribe-Payer's Index, which is based on surveys of business executives who have business relationships with foreign firms, focuses exclusively on transnational bribery.[127] In addition, some enterprise surveys cover local subsidiaries of foreign companies and require them to identify themselves in their responses.[128] By contrast, surveys of the general population typically are unhelpful on this front, either because they cover only domestic corruption or they do not distinguish between domestic and transnational varieties.

Investigations conducted by law enforcement agencies are another important source of information about bribery. Increased use of communications technology such as emails and text messages has made details of corrupt transactions much more accessible to enforcement agencies than before. Some, but not all of the information collected by enforcement agencies eventually makes its way into the public domain. A few agencies, like the U.S. Department of Justice, make detailed public disclosures about the cases they pursue.[129] Journalists are another potentially valuable source of information, particularly in countries like Brazil where the popular press

---

[126] G.A. Res. 71/313 (July 6, 2017).

[127] Transparency International, BRIBE PAYERS INDEX 2011, (2011).

[128] See, e.g., Edmund Malesky et al., THE VIETNAM PROVINCIAL COMPETITIVENESS INDEX: MEASURING ECONOMIC GOVERNANCE FOR PRIVATE SECTOR DEVELOPMENT, 2015 FINAL REPORT, (Labour Publishing House, 2016); Nathan Jensen and Edmund J. Malesky, Nonstate Actors and Compliance with International Agreements: An Empirical Analysis of the OECD Anti-Bribery Convention, 72 INT'L ORG. 33 (2018).

[129] Documentation related to criminal enforcement actions is posted at: https://www.justice.gov/criminal-fraud/related-enforcement-actions. The charging documents and settlement agreements typically include detailed statements of facts.

aggressively investigates and reports on corruption cases.[130] Of course, the cases that result in enforcement actions or journalistic reports are not necessarily representative of the broader population of cases of bribery.

There have been only ad hoc efforts to collect data on the consequences of bribery.[131] These generally take the form of academic studies of bribes paid in very specific contexts. For example, a study of bribes paid by customs clearing agents on behalf of South African firms estimated not only the number and size of the payments but also their effects on tariff revenue and the costs firms incurred to avoid a port staffed by particularly corrupt officials.[132]

C. CAUSAL INFERENCE

A policymaker committed to evidence-based anti-bribery regulation would make concerted efforts to study the impact of various interventions on key outcomes. A wide range of interventions have impacts that are sufficiently uncertain to merit further study. These include: allowing firms to raise extortion as a defense, or providing leniency to actors who voluntarily report their misconduct, or increasing the number of investigators assigned to bribery cases. It would be useful to know how any or all of these interventions affect the behavior of potential wrongdoers. For example:

- Do potential wrongdoers try to avoid the application of the regime by cutting ties with countries that enforce their laws vigorously?
- Do organizational actors such as multinational enterprises and states invest in training and internal controls for their employees and agents?
- Do potential wrongdoers reduce their operations in high-risk jurisdictions?
- Do they reduce the number or the value of bribes paid?

It also would be useful to know how the relevant intervention affects potential victims. This would involve answering basic questions such as:

---

[130] Mauro P. Porto, *The Media and Political Accountability*, in Timothy J. Power and Matthew M. Taylor (eds.), CORRUPTION AND DEMOCRACY IN BRAZIL: THE STRUGGLE FOR ACCOUNTABILITY, 103 (2011).
[131] For surveys see Davis, *supra* and Susan Rose-Ackerman and Bonnie J. Palifka, CORRUPTION AND GOVERNMENT: CAUSES, CONSEQUENCES, AND REFORM, 2nd ed. (New York: Cambridge University Press, 2016).
[132] Djankov & Sequeira, *supra*.

- How many victims are there?
- Who are they?
- What harm have they suffered?
- Have they been compensated?

There are also important questions to be asked about how the regime affects actors who are neither perpetrators nor victims. For instance:

- Are citizens of countries whose officials have been bribed aware of foreign enforcement actions? Do they approve of them? How prevalent do they believe bribery to be?
- How are citizens of poor countries affected when firms disinvest to avoid liability under anti-bribery law?

Even if appropriate data were available it would be difficult to determine the role that the anti-bribery regime, or any given component of it, plays in causing these outcomes. A host of non-legal factors can affect attitudes and behavior. For example, suppose a multinational firm ceases to do business in Nigeria. How can we tell if its decision was based on fear of liability under anti-bribery law or a change in tax law or an assessment of trends in the global economy? And even if we know that the anti-bribery regime mattered, which component was the operative factor? How do we isolate the impact of the FCPA as opposed to the UK Bribery Act or Nigerian law, keeping in mind that in some cases it will be unclear whether any or all of these laws apply?

Case studies of specific firms or individuals who have engaged in bribery can shed light on the impact of anti-bribery law on those particular actors. Every enforcement action amounts to a case study of this sort, at least for the people with access to information about the case. These kinds of studies do not, however, tell us much about how law affects the full range of actors who might engage in bribery. Statistical analysis seems essential for this purpose.

There are few official—meaning publicly-sponsored—efforts to conduct statistical analyses of the impact of anti-bribery law. As we have seen, the OECD and UNCAC have formal mechanisms to collect data on levels of enforcement, and public actors have shown increased interest in supplementing private efforts to monitor levels of corruption. There appears to be little

interest, however, in connecting these two types of data to answer even basic questions such as: Has greater enforcement led to reduced corruption?

It is possible that enforcement agencies conduct these kinds of analyses internally. For example, from 2016 to 2017 the US Department of Justice introduced a "pilot program" that offered specified amounts of leniency to firms that engaged in transnational bribery if they self-reported and cooperated with authorities.[133] After roughly 18 months the policy was made permanent, with minor modifications.[134] The Department of Justice claimed to have based its decision on an analysis that included consideration of the number of voluntary disclosures it received before and after adoption of the pilot program.[135] The substantial increase in voluntary disclosures—from 18 to 30—was unlikely to have been caused by any factor other than the pilot program.

Ideally, the Department of Justice also would have analyzed the impact of the pilot program on firms' efforts to prevent foreign bribery, in other words, its deterrent effect. Theory suggests that the promise of leniency can reduce firms' incentives to take preventive action.[136] The level of preventive effort is difficult to observe directly but changes can be estimated by examining variation in the total number of incidents of misconduct detected (with or without voluntary disclosure). In principle, therefore, the Department of Justice could have studied the deterrent effect of the pilot program by analyzing data on the volume of misconduct reported to have occurred before and after the program's introduction. However, the pilot program probably did not run for long enough for those outcomes to be observable within its timeframe.

---

[133] US Department of Justice, Criminal Division Launches New FCPA Pilot Program, Press Release, April 5, 2016: https://www.justice.gov/archives/opa/blog/criminal-division-launches-new-fcpa-pilot-program

[134] US Department of Justice, Deputy Attorney General Rosenstein Delivers Remarks at the 34th International Conference on the Foreign Corrupt Practices Act, November 29, 2017: https://www.justice.gov/opa/speech/deputy-attorney-general-rosenstein-delivers-remarks-34th-international-conference-foreign. The policy is codified in, US Department of Justice, JUSTICE MANUAL, 9-47.120 – FCPA Corporate Enforcement Policy.

[135] US Department of Justice, id. ("In the first year of the Pilot Program, the FCPA Unit received 22 voluntary disclosures, compared to 13 during the previous year. In total, during the year and a half that the Pilot Program was in effect, the FCPA Unit received 30 voluntary disclosures, compared to 18 during the previous 18-month period.")

[136] See Jennifer Arlen, *Corporate Criminal Liability: Theory and Evidence*, in Alon Harel and Keith Hylton (eds.), RESEARCH HANDBOOK ON THE ECONOMICS OF CRIMINAL LAW 144, 185-186 (2012) (explaining how a regime that conditions corporate sanctions on reporting and cooperation must impose a separate "residual" penalty to incentivize prevention).

The few publicly disseminated statistical analyses of the impact of anti-bribery law have been undertaken by professional academics. These studies generally have been inconclusive. They also demonstrate the challenges inherent in assessing the causal role of regulations of this kind. The most ambitious study to date focuses on anti-corruption law rather than just anti-bribery law and was led by Alina Mungiu-Pippidi.[137] She analyzed the impact of several frequently recommended legal interventions, including enactment of freedom of information legislation, establishment of a dedicated anti-corruption agency, and the creation of an ombudsman.[138] In cross-country analyses she found no statistically significant relationship between those interventions and perception-based measures of the prevalence of corruption, at least during the years 1996-2011.[139] Consequently, those interventions do not figure in her version of an evidence-based anti-corruption policy.[140]

Although Mungiu-Pippidi's conclusion is intuitively appealing, there are at least three reasons to question the validity of the results. First, the measure of corruption might be inaccurate. Second, the study makes little or no attempt to account for enforcement or the influence of anti-corruption regulation emanating from foreign countries.[141] Third, and most importantly, the countries in which interventions were implemented were not randomly selected. Consequently, the results are consistent with the possibility that the interventions actually were effective but were implemented mainly in countries that were accurately perceived, either by international or domestic actors, to be on track to experience increased levels of corruption.

There also have been cross-country studies of the impact of regulation of transnational bribery. Studies of the impact of the FCPA prior to the adoption of the OECD Convention produced conflicting results on whether it discouraged US firms from investing in relatively corrupt countries.[142] Later studies showed that relatively corrupt countries experienced lower

---

[137] Alina Mungiu-Pippidi, THE QUEST FOR GOOD GOVERNANCE: HOW SOCIETIES DEVELOP CONTROL OF CORRUPTION, chapter 4 (2015) ["QUEST FOR GOOD GOVERNANCE"]; Alina Mungiu-Pippidi and Ramin Dadašov, *When do anticorruption laws matter? The evidence on public integrity enabling contexts*, 68 CRIME LAW & SOC. CHANGE 387 (2017).

[138] Id.

[139] Id.

[140] Mungiu-Pippidi, *supra*, 2 (focusing judicial independence as the principal legal mechanism for controlling corruption).

[141] QUEST FOR GOOD GOVERNANCE, 109.

[142] Compare James R. Hines, Jr., *Forbidden Payment: Foreign Bribery and American Business After 1977*, NBER Working Paper No. 5266 (1995) (for US firms, corruption is associated with lower rate of growth in foreign direct

levels of foreign direct investment in and exports from countries that had signed the OECD Convention.[143] Consistent with these findings, a recent study found that the enactment of the UK Bribery Act reduced the value of UK-headquartered firms with exposure to countries perceived to be corrupt, increased the value of direct competitors of UK firms headquartered outside the OECD and without connections to the UK, and caused UK firms to reduce their engagement with corrupt countries.[144] Another recent study suggests that the increase in enforcement of the FCPA that occurred shortly after the OECD Convention came into force discouraged firms from OECD countries—not just US firms—from investing in relatively corrupt countries and led to a net decline in aggregate FDI in those countries (relative to aggregate non-corrupt countries).[145]

As for the impact of the law on misconduct, one study using data from Ghana found that firms whose home countries were parties to the OECD Convention were generally less likely to pay or be solicited for bribes.[146]Another study, using data from Vietnam, found that foreign investors in Vietnam whose home countries were parties to the OECD Convention demonstrated lower propensity to pay bribes after the increase in enforcement that accompanied Phase 3 of the Working Groups' review process, which was initiated in 2010.[147] All of these studies appear to assume that firms were subject to liability for foreign bribery only in their home countries. As

---

investment, joint ventures and aircraft exports) and Shang-jin Wei, *How Taxing is Corruption on International Investors*, 82 REV. ECON. & STAT. 1 (2000) (across 12 source countries corruption is associated with a lower stock of foreign direct investment, but the effect is not especially significant for US firms).

[143] Adrian Blundell-Wignall and Caroline Roulet, *Foreign Direct Investment, Corruption and the OECD Anti-Bribery Convention* (2017) https://www.oecd-ilibrary.org/finance-and-investment/foreign-direct-investment-corruption-and-the-oecd-anti-bribery-convention_9cb3690c-en; Anna D'Souza, *The OECD Anti-Bribery Convention: Changing the Currents of Trade*, 97 J. DEV. ECON. 73 (2012); Alvaro Cuervo-Cazurra, *Who Cares About Corruption?* 37 J. INT'L BUS. STUD. 807 (2006).

[144] Stefan Zeume, *Bribes and Firm Value*, 30 REV. FIN. STUD. 1457 (2017).

[145] Hans B. Christensen, Mark Maffett, and Thomas Rauter, *Policeman for the World: US Enforcement of Foreign Corruption Regulation and Corporate Investment Policies* (2019), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3349272. *Cf.* Leah M. Trzcinski, *The Impact of the Foreign Corrupt Practices Act on Emerging Markets: Company Decision-Making in a Regulated World*, 45 N.Y.U. J. INT'L L. & POL. 1201 (2013). In this study of firms subjected to FCPA enforcement actions between 2000 and 2010, Trzcinski found that 70 per cent continued to maintain facilities or offices in countries where misconduct had been identified. She interprets this as evidence that "companies who have been through FCPA enforcement actions are not necessarily pulling out of high-risk countries." However, it is difficult to know whether this is a valid interpretation in the absence of information on the rate of divestment among firms that were not the subjects of FCPA enforcement actions.

[146] Jennifer Spencer and Carolina Gomez, *MNEs and Host Country Corruption*, 32 STRATEGIC MGM'T J. 280 (2011). Spencer and Gomez also examined whether firms that invested in Eastern Europe were less likely to report a need to engage in bribery if their home country was a party to the OECD Convention. They found no evidence to support this hypothesis.

[147] Jensen & Malesky, *supra*.

noted above, this is an important oversimplification of reality because the FCPA and similar laws in other jurisdictions typically apply to foreign as well as domestic firms.

In principle it would be interesting to conduct RCTs designed to determine the causal effect of varying a particular feature of anti-bribery law. For instance, following Abramowicz et al, we could randomly exempt firms from the FCPA's internal controls provisions in order to determine the impact on rates of bribery and compliance costs. As we have already seen, the obstacles to conducting a valid study of this kind of study are daunting. To begin, firms in the control group that compete with firms in the treatment group would be affected by the treatment since it likely would tilt the playing field either in favor of or against firms in the treatment group. It also might be difficult to prevent firms in the control group from being influenced by enforcement efforts targeted at the treatment group. The treatment is likely to increase the overall number of cases of bribery as firms relax their internal controls. This might increase estimates of the prevalence of bribery, among people in both the treatment group and the control group. There are several theoretical models of corrupt behavior which predict that individual decisions about whether to engage in corruption will be influenced by perceptions of the general prevalence of misconduct.[148] Following that logic, the intervention in this case might increase bribery in both the treatment and the control groups. This spillover effect creates a bias against identifying the true causal role of the intervention because comparison of levels of bribery in the treatment and the control groups will understate the true effect on the treatment group. The true effect can only be identified by a comparing the treatment group to a completely unaffected control group.

D. GENERALIZATION

An evidence-based approach to anti-bribery law presumes that interventions which play a particular causal role in one context will play the same role in other contexts. This is consistent with the approach of the OECD Working Group. As noted above, the OECD Working Group fulfills its commitment to evidence-based regulation by contributing to the development of evidence-based best practices in enforcement of transnational anti-bribery law. It tries to

---

[148] Raaj Sah, *Corruption across Countries and Regions: Some Consequences of Local Osmosis*, 31 J. ECON. DYNAMICS & CONTROL 2573 (2007) (formal model in which individuals' decisions to engage in corruption depend on odds of encountering corrupt counterparties); Pranab Bardhan, *Corruption and Development: A Review of the Issues*, 35 J. ECON. LIT. 1320, 1331-1334 (1997) (discussing frequency-dependent equilibria).

accomplish this mainly through sharing of know-how between law enforcement officials from member states.[149] This strategy presumes that evidence about what works in one jurisdiction can be generalized to other jurisdictions.

It seems difficult to generalize from existing studies of the impact of anti-bribery laws. For instance, Alina Mungiu-Pippidi's cross-country analyses of anti-corruption interventions suggest that the effects of most interventions are context-sensitive. Assuming her findings are valid, they show not only that the effects of the interventions are small on average, but also that those effects vary significantly across countries.[150] Mungiu-Pippidi does generalize about the effectiveness of promoting judicial independence, one of the few legal constraints she finds to be consistently effective in controlling corruption.[151] However, since there are so many different ways of promoting judicial independence this hardly counts as a claim about the causal role of a specific intervention.

It also is difficult to generalize from single-country studies like the US Department of Justice's analysis of its pilot program. There is no reason to believe that information about how promises of leniency in exchange for cooperating with investigators affected a large publicly-traded multinational German engineering firm will shed light on how similar promises will affect a medium-sized family-owned Italian textile exporter, or a state-owned Chinese construction company, or a purely domestic construction company in Brazil. In fact, there are good theoretical reasons to expect heterogeneous effects. When managers of widely-held firms cooperate with enforcement officials they may only be prejudicing other managers. By contrast, asking managers of a closely-held firm to cooperate with enforcement officials is likely to involve asking them to implicate themselves. We can also speculate about whether managers of state-owned firms will be willing to implicate other managers who may have powerful political connections. Also relevant are variations in group loyalty and attitudes toward cooperation with the state, both of which might be influenced by factors such as social structures and perceptions of the legitimacy of the state.

---

[149] OECD Working Group, *supra* (evidence-based policies and standards); OECD Working Group, RECOMMENDATION OF THE COUNCIL FOR FURTHER COMBATING BRIBERY OF FOREIGN PUBLIC OFFICIALS IN INTERNATIONAL BUSINESS TRANSACTIONS, ¶XIV (iv) (best practices) (2009).
[150] Mungiu-Pippidi & Dadašov, *supra*, 391-395 (see figures showing considerable variations in country performance).
[151] Mungiu-Pippidi, *supra*, chapter 4; Mungiu-Pippidi & Dadašov, *supra*.

E. Capacity

Constraints on institutional capacity have clearly affected the ability of at least some enforcement agencies to adopt an evidence-based approach to anti-bribery law. In many countries dedicated anti-corruption agencies have been charged with conducting research in order to support evidence-based anti-corruption regulation.[152] This kind of research can be very resource intensive, in terms of human resources, technology, and funds, especially when it involves surveys and quantitative analyses.[153] The literature on point suggests that many anti-corruption agencies have little funding to devote to research, and as a consequence whatever research they produce has little impact on their activities.[154]

IV. IMPROVE OR ABANDON EVIDENCE-BASED REGULATION?

Policymakers have two reasonable responses when the evidence required by evidence-based regulation is unhelpful. The first is to try to improve the evidence, presumably in collaboration with the scholarly community. The second is to reject evidence-based regulation altogether. This Part explores each of these responses in turn.

A. Improve evidence-based regulation

1. Data collection

In sections II.B and III.B we concluded that data collection is particularly likely to be difficult for data on enforcement agencies and outcomes associated with their activities. The solution to this problem is to invest more in data collection. The critical questions are, what sorts of data ought to be collected, and who ought to collect it?

---

[152] Luís de Sousa, *Anti-Corruption Agencies: Between Empowerment and Irrelevance*, 53 Crime Law & Soc. Change 5, 16-17 (2010).

[153] Id.

[154] De Sousa, *supra*, 17. For evidence of under-funding and resource scarcity in anti-corruption agencies see, Jon ST Quah, *Defying Institutional Failure: Learning From the Experiences of Anti-Corruption Agencies in Four Asian Countries*, 53 Crime, Law & Soc. Change 23, 42 (2010) (Thailand); Alan Doig, David Watt and Robert Williams, *Why Do Developing Country Anti-Corruption Commissions Fail to Deal With Corruption? Understanding The Three Dilemmas of Organisational Development, Performance Expectation, and Donor and Government Cycles*, 27 Public Admin. & Dev. 251 (2007).

On the first point, current practice relies heavily on data collected from small samples of agencies and actors. For example, the cross-country data compiled by the World Bank and the World Justice Project are collected from two or three large cities in each country. Sampling is a perfectly reasonable way to estimate the characteristics of a population; however, the more heterogeneous the population, the less well it can be understood by examining a sample of any given size. Many legal institutions seem likely to lie at the relatively heterogeneous end of the scale; literature on "local legal culture" suggests that there are significant differences across courts within countries.[155] So, for example, courts in New York City and Los Angeles (the US cities analyzed by the World Bank's Doing Business Project) are unlikely to be representative of courts in Texas or Wyoming.

This suggests that an important way to enhance the value of data about legal institutions and their performance is to collect more comprehensive data. Ideally, data would be collected for the entire population of existing legal institutions as well as all the people who interact with them. This means that data on every action taken in the course of making and enforcing laws would be recorded. These kinds of legal data are already collected in the most advanced countries in the world; the challenge is to extend this kind of data collection to the rest of the world. In the case of anti-bribery law, this would entail striving to collect at least as much data on enforcement of domestic bribery cases as the members of the OECD Convention collect about transnational bribery.

As for data on outcomes, the case study of anti-bribery law suggests that the availability of data will increase through expansion in the use of information technology and innovations in social scientific research methods. At the same time, it seems reasonable to presume that data on illicit activities never will be easy to obtain.

---

[155] See, e.g. Teresa A. Sullivan, Elizabeth Warren and Jay Lawrence Westbrook, *Persistence of Local Legal Culture: Twenty Years of Evidence from the Federal Bankruptcy Courts*, 17 HARV. J. L. & PUB. POL'Y 801 (1994); David Nelken, *Using the Concept of Legal Culture*, 29 AUSTL. J. LEG. PHIL. 1 (2004) at 12 (describing variation in court delay within Italy). See also, Sandra Gabriel Mayson, and Megan Stevenson, *Misdemeanors by the Numbers* (April 18, 2019). Boston College Law Review, Forthcoming; University of Georgia School of Law Legal Studies Research Paper No. 2019-14; George Mason Legal Studies Research Paper No. LS 19-03. Available at SSRN: https://ssrn.com/abstract=3374571 or http://dx.doi.org/10.2139/ssrn.3374571 (describing variations in treatment of misdemeanors across 8 jurisdictions within the US).

This brings us to the question of who should collect these data. Current practice relies on a combination of governments, international organizations, not-for-profit entities (including academics), and for-profit entities.[156] Gillian Hadfield has recommended that private actors operating in competitive markets be given greater responsibility for delivering certain legal services, and she claims that those actors will have strong incentives to collect and analyze data about their performance.[157] Even if she is generally correct about for-profit actors' incentives to analyze data, they are unlikely to have incentives to collect all types of data—data on the performance of land courts in rural India is unlikely to be a profit center for a legal information company. If the goal is to collect comprehensive data then governmental actors are likely to be needed to fill gaps in coverage left by private actors.[158] Moreover, regardless of who collects the data, public action, in the form of possible legal intervention, is likely to be necessary to discourage fraud.[159]

2. Causal inference

Social scientists appear to be highly motivated to improve techniques for drawing valid causal inferences, if only because of the demand among people interested in marketing new products and developing new medical treatments. As a result, there is little need for interventions designed to stimulate development of new techniques of causal inference. The main challenge is to ensure that policymakers who employ those techniques keep in mind the distinctive challenges associated with drawing causal inferences about the impact of regulatory interventions. For instance, as we have seen, regulatory interventions tend to have more far-reaching and visible effects than many other interventions that social scientists tend to study. As a consequence, techniques suitable for studying non-regulatory interventions may not be ideally suited for studying regulatory interventions. RCTs would be a case in point.

3. Generalizability

---

[156] Davis, *supra*, 1643-1645 (discussing supply of legal performance measures).
[157] Hadfield, *supra*, 215-218.
[158] Davis, *supra*, 1643-1645. For a discussion of how the prospect of use by regulators might affect private actors' incentives to collect data see, Niva Elkin-Koren and Michal S. Gal, *The Chilling Effect of Governance-by-Data on Data Markets*, 86 U. CHI. L. REV. 403 (2019).
[159] Hadfield, *supra*, 217, briefly acknowledges this danger, noting, "there will be bad mixed in with these good results."

There are two main ways to address concerns about the generalizability of research on the impact of regulatory interventions. One is to make the research more generalizable. This will involve studying a broad range of variants on any given intervention in a broad range of environments. This in turn will require either significant increases in the resources devoted to legal research or significant reductions in the cost of conducting that research. Increased investments in collecting comprehensive administrative data might be helpful in this regard. For example, if more enforcement agencies collected data on investigations and prosecutions of public officials who receive bribes then it would be easier to conduct research on the effectiveness of not only the enforcement actions themselves but also other anti-bribery measures, such as leniency programs, whose success is expected to trigger enforcement actions.

A second way to respond to concerns about generalizability is to give up on the idea of generalization and to base decisions primarily on evidence derived from the context in which the intervention is to be implemented. That kind of evidence can be generated by implementing a prototype version of an intervention and then continuously altering its design in response to feedback about its performance. The design of the prototype might be based on theoretical analysis, or upon "evidence" from past research, but without any presumption that either the theory or the evidence is conclusive. This approach to institutional design is associated with a form of experimentalism known as "design science" or "design thinking."[160] The process of experimentation that culminated in the adoption of the US Department of Justice's corporate enforcement policy is a classic example of this approach to regulation. This approach also entails devoting greater resources to research since it demands a new study for each context.

4. Capacity

Regulators' inability to analyze evidence can stem from either limits on their skills or limits on their access to information. The relative importance of these two obstacles depends on the extent to which regulators plan to generate their own evidence and institutional designs as opposed to obtaining them from others. The more that regulators plan to rely on locally-generated evidence

---

[160] See Sabel, *supra*, Ouellette, *supra*. On design thinking and design science see generally, *Special Issue: New Policy Design*, 47:3 POLICY SCIENCES (2014); Alan R. Hevner, Salvatore T. March, Jinsoo Park, and Sudha Ram, *Design Science in Information Systems Research*, 28 MIS QUARTERLY 75 (2004).

and designs, the greater the demands on their skills and the less important it is for them to access external sources of information.

The problem of inadequate skills admits two responses: either enhance the skillset of the regulators or alter the skills required to understand the evidence. The first option requires either training existing officials or employing new officials with the relevant skills. In some situations access to specialized technology, ranging from basic statistical software to state-of-the-art machine learning tools, may also be helpful. The second option requires effort to produce more user-friendly syntheses of relevant research. This may be as simple as translating existing literature into languages spoken by the officials. It also may involve a concerted effort to produce literature reviews of the kind produced by the Campbell Collaboration.

Solutions to the problem of inaccessible information include subsidizing either the publication or the purchase of books and articles. In the poorest countries officials may not even have access to the Internet.

Generally, as our case study of anti-bribery law suggests, resource constraints are the main obstacle to enhancing capacity for evidence-based regulation. Sometimes those constraints reflect absence of political will, but in the case of many poor countries resource scarcity is an incontrovertible fact of life.

B. ABANDON EVIDENCE-BASED REGULATION

The kinds of evidence prioritized by evidence-based regulation may be of limited value. As we have seen, this is most likely to occur when the interactions between legal regulation and desired outcomes are complex and resources are scarce. In these situations regulatory design based on systematic research may be less effective than designs based on the leading alternatives, namely, unsystematic learning and theory. In these settings legal instruments that appear to require evidence-based regulation should be interpreted flexibly to permit reliance on judgment, even when relevant systematic research is available.

Arriving at the decision to abandon systematic research requires a clear-eyed assessment of the limitations of not just systematic research but also of unsystematic learning and theorizing. Choosing between alternatives based on a consideration of the limitations of only one of the

options risks allowing the perfect to be the enemy of the good. Nonetheless, since the limitations of judgement have been documented extensively elsewhere, they need only be discussed briefly here.[161]

There is evidence that decision-makers who rely on judgment sometimes perform remarkably well in comparison to evidence-based algorithms.[162] This especially likely to occur in settings in which the decision-makers are capable of acquiring and applying genuine expertise. These are settings in which the relevant causal relationships are reasonably strong and stable; decision-makers have access to data that allows them to learn those relationships, including data that are difficult to capture in an algorithm; and, decision-makers are willing to apply their expertise consistently.[163]

At the same time, the relevant literature shows that judgment is far from infallible. There is overwhelming evidence that experts who rely on judgment are susceptible to errors and in some contexts perform poorly in comparison to simple evidence-based rules. This is consistent with a broader literature showing that even expert human decision-makers often rely on heuristics that make them susceptible to numerous biases. They tend to make predictions that fit a remembered pattern or a coherent narrative rather than taking into account all of the relevant information. For example, like baseball scouts who ignore obviously relevant data like a player's on-base average and emphasize players' appearance, people often place excessive weight on some variables and insufficient weight on others.[164] They also are influenced by the ways in which questions are framed, anchor their responses around cues, and ignore basic notions of probability.[165] Even when

---

[161] The literature on the limitations of expert judgment is extensive. For a sampling see: DANIEL KAHNEMAN, THINKING, FAST AND SLOW (2011), 222-244; Philip E. Tetlock, EXPERT POLITICAL JUDGMENT: HOW GOOD IS IT? HOW CAN WE KNOW? (2005); Paul Meehl, CLINICAL VS. STATISTICAL PREDICTION: A THEORETICAL ANALYSIS AND A REVIEW OF THE EVIDENCE (1954); Jon Kleinberg et al, *Human Decisions and Machine Predictions*, 133 Q. J. ECON. 237 (2018); Daniel Kahneman, Gary Klein, *Conditions for Intuitive Expertise: A Failure to Disagree*, 64 AMER. PSYCH. 515 (2009); William M. Grove et al, *Clinical Versus Mechanical Prediction: A Meta-Analysis*, 12 PSYCH. ASSESS. 19 (2000). For discussion of the limitations of this literature see, Megan Stevenson, *Assessing Risk Assessment in Action*, 103 MINN. L. REV. 303, 321-27 (2018) and Starr, *supra* note 29, 850-55. See also, Julia Dressel & Hany Farid, *The Accuracy, Fairness, and Limits of Predicting Recidivism*, 4 SCI. ADVANCES 1 (2018).
[162] Kahneman & Klein, *supra* note 161 at 523-24.
[163] Id.
[164] MICHAEL LEWIS, MONEYBALL (2003).
[165] See generally, Kahneman, *supra* note 161 (documenting heuristics and biases and evidence of their existence).

experts apply principles that are well-known to them they apply them inconsistently.[166] Exacerbating these problems, peoples' subjective confidence in their judgments is not always highly correlated with the accuracy of those judgments.[167] This means that people regularly place either too much or too little weight on their predictions of the most likely outcomes.

Just as we should consider ways of improving evidence-based regulation, we also should consider methods of improving judgement. For instance, there is evidence that the performance of experts can be improved with training.[168] It is also worth considering hybrid approaches to regulatory decision-making that involve systematic aggregation of the judgments of experts.[169] Those judgements can be used either on their own or in combination with results from other types of research. These techniques have been used to elicit judgements about the outcomes of physical processes such as lead pollution or climate change.[170] In principle, they could also be used to elicit views on the outcomes of regulatory interventions. So for example, in the anti-bribery context, the OECD Working Group could systematically elicit judgments from prosecutors around the world on, say, the likely effects of granting a specified amount of leniency on the volume of self-reporting by large multinational firms. Another country could use these data, perhaps together with the results of quantitative studies such as the US Department of Justice's review of its pilot program, to inform decisions about whether and how to adopt a leniency program.

Should policymakers take an evidence-based approach to decide whether to take an evidence-based approach to regulation? Again, the answer depends on the quality of the evidence that can be obtained. In principle, it is possible to run an experiment which compares the

---

[166] Natalia Karelaia, and Robin M. Hogarth, *Determinants of Linear Judgment: A Meta-Analysis of Lens Model Studies*, 134 PSYCH. BULL. 404, 420 (2008) (meta-analysis of literature concluding that inconsistency makes linear models based on human decisions perform better than humans themselves).

[167] Kahneman, *supra* note 161, at 199-221.

[168] Welton Chang, et al., *Developing Expert Political Judgment: The Impact of Training and Practice on Judgmental Accuracy in Geopolitical Forecasting Tournaments*, 11 JUDGMENT & DECISION MAKING 509 (2016) (reviewing literature on effect of training on judgment and finding that training improved performance in a geopolitical forecasting competition); Paul Goodwin, Dilek Önkal, and Michael Lawrence, *Improving The Role of Judgment in Economic Forecasting* in Michael P. Clement and David F. Hendry (eds.), THE OXFORD HANDBOOK OF ECONOMIC FORECASTING (2011).

[169] M. Granger Morgan, *Use (And Abuse) Of Expert Elicitation In Support of Decision Making For Public Policy*, 111 PNAS 7176 (2014). These techniques arguably qualify as forms of evidence-based regulation in the sense that they treat systematic research on experts' assessments of unsystematically generated judgements as a form of evidence.

[170] Id.

effectiveness of officials who make evidence-based decisions and those who do not.[171] For example, a random selection of judges responsible for deciding whether to release criminal defendants prior to trial might be given access to an algorithm that predicts risk of recidivism based on the latest scientific research. Measures of recidivism among judges with access to the algorithm could be compared to those of judges who relied solely on judgement. However, this kind of study might not be feasible. Or there may be reasons to believe the results are unreliable, or inapplicable in the present context, including because the decision-maker will misinterpret the evidence. Under any of these conditions it is appropriate to decide whether to adopt an evidence-based approach without giving priority to systematic research.

C. A NOTE ON POLITICS

This Article has focused so far on whether and when the kind of evidence prioritized by evidence-based regulation provides a sound basis for conclusions about the impact of regulation. We have bracketed the question of how accepting or rejecting evidence-based reasoning is likely to affect the distribution of power in society. In other words, we have set aside questions such as: who will be chosen to employ judgment as opposed to evidence-based reasoning, and what values or whose interests are they likely to promote? The answers to these questions will in turn affect assessments of which regulatory approach is most consistent with distributive justice and legitimacy.

It is difficult to say anything useful on these topics because the analysis necessarily will be highly context-sensitive. For instance, in the US federal government, abandonment of the evidence-based approach might mean shifting power from career civil servants to political appointees. In the US and other OECD countries this shift also might mean diminishing the power of economists and political scientists while enhancing the power of lawyers. According to stereotypes, social scientists tend to attach great weight to the products of rigorous quantitative research while lawyers are more likely to rely on judgment. In poorer countries a shift away from evidence-based regulation might entail reducing the role of foreign experts familiar with published literature that synthesizes evidence from a wide range of countries, and giving greater

---

[171] Stevenson, *supra* note 161 attempts to approximate this kind of experiment by comparing outcomes before and after judges in Kentucky were legally required to use a pretrial risk assessment tool in making decisions about pretrial custody.

responsibility to people who draw unsystematically on local experience. The political significance of these kinds of reallocations of decision-making authority will in turn depend on an array of other contextual factors. For example: Are career civil servants representative of the population as a whole or are they drawn from a tiny elite? Are political appointees required to account for their behavior to legislative committees or only to the officials who appointed them? Are economists or lawyers more likely to be concerned about the impact of their decisions on the poor? Are foreign experts more accountable to the governments that sponsor them, and the economic interests those governments serve, or to the local government? Are local decisionmakers prone to be captured or corrupted by local interest groups?

It is impossible to generalize about the magnitude or the significance of these kinds of political effects. However, nothing in this article is intended to deny that they are worth considering.

CONCLUSION

There is pressure from several directions to expand the substantive and geographic scope of evidence-based regulation: academics' enthusiasm for the use of RCTs to evaluate regulatory interventions, political pressure in the US to expand executive oversight of federal agencies, promotion of evidence-based regulation by the OECD, and, the incorporation of the US federal requirements concerning evidence-based regulation in the CPTPP.

There are good reasons to embrace evidence-based regulation and its hallmark, prioritization of systematic research. In many contexts there are good reasons to believe that decisions based primarily on that kind research will be more accurate and legitimate than decisions that give more weight to judgement. At the same time, there are contexts in which the value of evidence-based research is likely to be limited. These are generally contexts in which data are scarce or of poor quality, casual relationships are complex and the agencies charged with implementing the evidence-based approach have limited capacity.

The reality of modern regulation is that much of it applies in precisely the sorts of situations in which the limitations of the evidence-based approach are likely to be most serious. Transnational

business regulation frequently concerns hard-to-observe illicit activities conducted by a wide range of actors and agencies that interact in complex ways, and is implemented by agencies that vary significantly in terms of their resources and sophistication. The global anti-bribery regime is just one of many examples of regimes of this sort. In these settings, feasible types of research on the impact of past interventions will tend to be of limited value in predicting the impact of future interventions. As a result, careful thought is required about whether and how to use research as opposed to judgement in making decisions about regulation, assuming the goal is to maximize regulatory effectiveness. This Article is intended to identify some of the factors that ought to be taken into account in those deliberations.

A final caveat is in order. Ultimate decisions about whether to adopt evidence-based regulation should take into account many factors that have been deliberately ignored in this Article. There are important questions about whether disagreements about the objectives of regulation render evidence-based regulation infeasible. It is also important to consider whether the case for evidence-based regulation is bolstered or undermined by the need to induce regulators to act in good faith and concerns about the legitimacy as opposed to the effectiveness of regulation. These topics are beyond the scope of the present Article and are left for future research.