

The Empirical Study of Legitimate Authority: Normative Guidance for Positive Analysis*

Sanford C. Gordon

Gregory A. Huber

Abstract

We employ key concepts in the normative study of legitimate authority to place the empirical analysis of legitimacy on firmer analytical foundations. Our critical review of empirical research on support for courts, regimes generally, and international organizations highlights the slippage between normative and positive approaches, while simultaneously drawing attention to problems of measurement and critical inferential problems rooted in limitations of research design. We then describe a simple theoretical model that formalizes these considerations. The model reveals conditions under which it is possible to isolate the effect of an authority's legitimacy on citizen behavior net of extrinsic compliance motivations as well as environments in which examination of the antecedents of legitimate authority is most likely to be fruitful.

Keywords: Legitimacy, authority, compliance, obedience, consent, justification, causal attribution, causal mechanisms, mediation

1. Introduction

In this paper we examine three related questions of broad interest across the social sciences. First, is the *legitimacy* of institutions and individuals in positions of authority a meaningful, analytically distinct empirical phenomenon? Second, if it is, is a rigorous, causally identified empirical analysis

*We thank Eric Dickson for his essential contributions to other work in our collaborative research agenda, and gratefully acknowledge the helpful suggestions of Leonie Huddy, Macartan Humphreys, and Margaret Levi. Jennifer Wu provided superlative research assistance. This research is supported in part by NSF grants 1556047 and 1555494.

of its origins and consequences possible? And third, what can empirical scholars interested in the study of legitimacy learn from the rich normative literature on the subject?

Since the seminal writings of Max Weber in the early 20th century, the idea of legitimacy has emerged as a core concept in the positive analysis of human interaction across the social sciences. What might be called the “legitimacy hypothesis” – that beliefs about the legitimacy of authorities and institutions are essential to long term cooperation and compliance – enjoys prominence in numerous fields of relevance to political scientists. In the domestic context, this is particularly the case in the analysis of courts and law enforcement; in the comparative context, in discussions of state-building, counterinsurgency, and pro- and anti-regime sentiment and activity; and in the international context, in examinations of norms and cooperation with supranational organizations. Generally speaking, a common view that emerges across these subject areas is that material incentives alone are inadequate to facilitate social cooperation. At the heart of the legitimacy hypothesis is the idea that citizens may voluntarily comply for reasons *over and above* the likelihood of reward or punishment, but tied instead to the characteristics or behavior of authorities and institutions.¹

We examine the normative political theory literature on legitimacy, and argue that it provides useful interpretive guidance on how to put the positive literature on firmer analytical foundations. Two issues are of particular note: the close association of legitimacy with the *justifiability* of an authority’s actions and the question of whether citizens, as moral agents, have an *obligation* to comply with authority. The normative literature’s focus on these two issues in particular is helpful to us in constructing a definition of legitimate authority as an empirical phenomenon. Specifically, obligation implies a motivation to comply apart from extrinsic, material motivations; while justification implicates citizens’ beliefs about authorities and institutions. Understanding the structure of legitimacy in terms of motivations and beliefs allows us, as positive analysts, to draw on game-theoretic intuitions and equilibrium logic. Specifically, we can think of legitimacy as a feature of an equilibrium in which citizens’ intrinsic motivations are enhanced by those beliefs about authorities, and the actions of governing institutions and their incumbents are consistent with those

1. See, for example, Tom R. Tyler, “The Psychology of Legitimacy: A Relational Perspective on Voluntary Deference to Authorities,” *Personality and Social Psychology Review* 1, no. 4 (1997): 323–345.

beliefs.² (Throughout, we use the term intrinsic motivations to refer to internally generated reasons for behavior [e.g., moral values] and the term extrinsic motivations to refer to external pressures for behavior [e.g., financial rewards, the threat of coercion, etc.])

We want to make it clear at the outset that we do not mean to engage in definitional trickery, merely redefining legitimacy in such a way as to find fault with the extant empirical literature. Instead our aim is to isolate the features of explanations for deference to authorities specific to the idea of their legitimacy that make them analytically distinct from other categories of explanation – particularly those based on extrinsic material rewards and punishments. It is our belief that definitional clarity facilitates improved empirical work and will also allow that empirical work to speak more clearly to the concerns undergirding the normative scholarship.

Next, we review the empirical political science research on legitimacy appearing in selected key journals since 1970 in three subject areas: courts, regimes in general, and international organizations. Our discussion of these literatures suggests the existence of three fundamental challenges to studying legitimacy empirically. First, the term “legitimacy” as defined in the literature tends to conflate a number of analytically distinct factors affecting attitudes about, and cooperation and/or compliance with, authorities and institutions. Attempts to measure beliefs about legitimacy frequently are confounded by this same theoretical ambiguity. Second, legitimacy is often studied in institutional environments where its behavioral consequences are likely to be obscured. And third, studies of legitimate authority are, with a few exceptions, insufficiently attentive to issues of causal inference and identification.

Having explored the empirical literature, we then elaborate on the challenges to inference while providing some guidelines to researchers about research designs with which scholars can uncover legitimating effects of institutions and authority actions, should those effects exist. Our analysis points to situations in which the identification of “institutional” and “personal” legitimacy is and is not feasible; to conditions under which material factors are likely to yield upward vs. downward

2. For an approach similarly grounded in the importance of obligation, see Margaret Levi, Audrey Sacks, and Tom Tyler, “Conceptualizing Legitimacy, Measuring Legitimizing Beliefs,” *American Behavioral Scientist* 53, no. 3 (2009): 354–375.

bias in estimates of legitimacy effects; and to conditions under which variation in perceptions of an institution's legitimacy are most likely to yield observable behavioral implications.

2. The Research Environment

A. Insights from Normative Theory

Though a full account of the development of the concept of legitimacy in normative political theory and philosophy is beyond the scope of the current inquiry,³ several key points of discussion stand out as relevant to our treatment of legitimacy as an empirical phenomenon.

The first of these is the problem of justification. Raz, for example, notes that a key feature of authority (e.g., of the state) is the right to issue *exclusionary reasons* to perform actions, that is, reasons that categorically override or exclude other reasons (for acting differently).⁴ Insofar as the coercive component of authority may conflict with the autonomy of the individual citizen – itself a primary good – the state must have a compelling rationale to exercise this authority. In a similar vein, Habermas states that legitimacy “means that there are good arguments for a political order’s claim to be recognized as right and just,” referring to an order’s “worthiness to be recognized.”⁵ Anscombe associates legitimate authority explicitly with the administration of justice.⁶ And Beetham explicitly ties justification to legitimacy, arguing that in a given power relationship an authority gains its legitimacy from the fact that it can be justified in terms of the beliefs of its participants.⁷ Overall, justification is therefore linked to being able to express reasons, often based in values, for acting as an authority.

What constitutes adequate justification is itself likely to be contingent on the historical peculiarities and moral traditions of the polity in question, a point noted by Habermas in his discussion of the historical evolution of the legitimation activities of different regimes. From the perspec-

3. For a helpful overview, see Morris Zelditch Jr., “Theories of Legitimacy,” in *The Psychology of Legitimacy: Emerging Perspectives on Ideology, Justice, and Intergroup Relations*, ed. John T. Jost and Brenda Major (New York: Cambridge University Press, 2001), 33–53.

4. Joseph Raz, *The Authority of Law*, 2nd (Oxford: Oxford University Press, 2009).

5. Jürgen Habermas, *Communication and the Evolution of Society* (Boston, MA: Beacon Press, 1979).

6. G.E.M. Anscombe, *The Collected Philosophical Papers of G.E.M. Anscombe: III. Ethics, Religion, and Politics* (Oxford: Basil Blackwell, 1981).

7. David Beetham, *The Legitimation of Power* (New York: Palgrave Macmillan, 1991).

tive of liberal political theory, however, justification entails an alignment of government's exercise of power with commonly (or universally) accepted principles. This is the position articulated by Rawls, who associates the legitimate exercise of state power with "a constitution the essentials of which all citizens as free and equal may reasonably be expected to endorse in the light of principles and ideals acceptable to their common human reason."⁸ Relatedly, Nagel associates legitimate liberal governance with being impersonal or impartial.⁹

The second key point of discussion in this literature is the relationship between consent on the one hand and the obligation to obey on the other. Consent, of course, forms the basis of all contractarian theories of governance. A substantial range of opinion exists, however, concerning the duty of the citizen to comply with the edicts of a duly constituted authority. At one extreme lies the Hobbesian view, which holds that the social contract between sovereign and citizen entails an absolute duty of the citizen to comply with the sovereign's commands so long as the commands do not threaten the subject's inviolable right to (physical) self-preservation. At the other extreme, philosophical anarchists articulate a view of individual autonomy as sacrosanct, to the point of rejecting any notion of a citizen owing a debt of obedience to an authority (although citizens are certainly entitled to give weight to the instructions of authorities as part of a constellation of reasons).¹⁰ Dworkin offers a middle ground perspective, endeavoring to articulate an account of legitimacy grounded in the obligation of the citizen to the community rather than to the state itself.¹¹

The linkages among legitimacy, obligation, and consent suggest a behavioral consequence to more legitimate authority, because greater legitimacy will translate into consent and (at least weakly) a tendency to obey or cooperate with an authority. For example, Beetham notes that legitimacy provides a "moral grounds" for cooperation and obedience that exists in a complex set of motivations for compliance that may include both immediate self-interest as well as moral or

8. John Rawls, *Political Liberalism* (New York: Columbia University Press, 1993).

9. Thomas Nagel, "Moral Conflict and Political Legitimacy," *Philosophy & Public Affairs* 16, no. 3 (1987): 215–240.

10. See, especially, Robert Paul Wolff, *In Defense of Anarchism* (New York: Harper & Row, 1970).

11. Ronald Dworkin, *Law's Empire* (Cambridge, MA: Belknap Press, 1986).

normative reasons.¹² Importantly, this means that legitimacy may affect behavior, and also that consent alone is not evidence of legitimacy because there are many other reasons for cooperation. Similarly, Hampton distinguishes between what she refers to as “convention consent” and “endorsement consent.”¹³ The former is behavior that is de facto supportive of, or at the very least not undermining of, the regime. By contrast, endorsement consent is behavior that conveys active approval of the regime. Endorsement consent, she writes, “is a *decision* to support [the regime] because of one’s determination that it is a good thing to support. By giving this form of consent, the subject conveys her respect for the state, her loyalty to it, her identification with it, and her trust in it.” Consent can therefore vary in degree with legitimacy, and all consent is not evidence of legitimacy.

What are the valuable insights from this literature for empirical scholars, and what are some pitfalls to avoid? The first concerns the notion of obligation. Irrespective of where one comes down on the duty to obey an authority, it is clear that normative theorists are articulating a notion of motivation to comply distinct from extrinsic motivations. When we speak of obligation, then, we are speaking of *intrinsic* motivations to cooperate with or support an authority. And, similarly, cooperation alone is not evidence of that intrinsic motivation. Most bluntly, this means that empirical researchers need to isolate intrinsic motivations for behavior in order to instantiate the concept of legitimacy as articulated by normative theorists.

The second insight concerns justification. An authority’s actions implicate the beliefs of agents about the fairness or justice of institutions or individuals, and it is only fair and just institutions that are “worthy” (to use Habermas’ term) of deference. That is, perceptions of fairness (among other values) enhance or diminish the intrinsic motivation to comply.

Together, obligation and justification suggest a normative core to legitimacy in the motivation of citizens to comply with authority. Critically, and borrowing from the language of game theory (a tool we employ in greater depth below), this normative core is an equilibrium: the beliefs associated with justification, and the actions associated with obligation, must be mutually consis-

12. Beetham, *The Legitimation of Power*, 26.

13. Jean Hampton, *Political Philosophy* (New York: Westview Press, 1997), ch. 3.

tent. If the authority acts in an unjustifiable way, obligation is undermined. And if citizens feel no obligation, there is little incentive for an authority to justify its actions. The normative literature accordingly points us in the direction of an equilibrium definition of legitimacy potentially amenable to empirical analysis:

Legitimacy is a feature of an equilibrium in which (a) subjects' favorable beliefs about an authority enhance their intrinsic motivations to comply with its directives and (b) the authority's actions are consistent with those beliefs.

Before explicating this notion of legitimacy and the attendant problems for isolating its causal effects, we offer a caveat: any behavioral model that aims at prediction rather than prescription must depart from the normative language of absolute (or near-absolute) rights and obligations and consider the concepts outlined above in terms of marginality and conditionality. Rather than speak of a general obligation to comply with an authority (or absolute autonomy from such an obligation), we will speak instead of legitimation "on the margin": conditions in which the enhanced intrinsic motivations associated with a legitimate authority incrementally affect the likelihood of compliance, *all things considered*. Given this treatment, the legitimacy of the authority is just one of many factors affecting compliance. Others include extrinsic reasons (the apprehension of punishment or reward), informal norms in the polity, or altruism, a perspective found in some normative work (e.g., Beetham). Given individual-level heterogeneity in other non-legitimacy related beliefs that affect compliance (or probabilistic responsiveness to legitimacy), it is therefore also likely that the behavioral consequences of an authority's legitimacy are likely to vary across individuals (or over time). This approach, which makes the empirical study of legitimacy feasible, is also, of course, what makes the empirical identification of legitimacy so vexing.

B. Extant Empirical Research

We examined empirical research on legitimacy in six major political science journals (*the American Political Science Review*, *the American Journal of Political Science*, *the Journal of Politics*, *Comparative Political Studies*, *International Organization*, and *World Politics*) for the period 1970-

2016.¹⁴ These papers can be divided roughly into three sets: research on public support for courts; research on regime legitimacy, and research on the legitimacy of supranational authorities. Before reviewing this research, we briefly examine its roots in the seminal work of Max Weber and David Easton.

Weberian and Eastonian Roots. The literature referencing legitimate authority as an empirically distinct and important phenomenon is voluminous. From a political science perspective, its roots can be traced to two critical sources. Most familiar is the treatment of legitimate forms of domination by Weber, who defines legitimacy as a belief in, or orientation to, the validity of a social order.¹⁵ For Weber, the idea of legitimacy is itself closely connected to the apparent need of an authority to justify itself. Critically, for Weber the belief in the legitimacy of an authority motivates compliance with authority, *over and above* other motivations, among which Weber includes custom, personal advantage/material motives, and “purely affectual or ideal motives of solidarity.”¹⁶ And empirically, evidence of the legitimacy of a system is manifested in “the probability that to a relevant degree the appropriate attitudes will exist and the corresponding practical conduct ensue.”¹⁷

This Weberian notion of legitimacy is expanded upon in Easton’s seminal *Systems Analysis of Political Life*.¹⁸ Like Weber, Easton ties the notion of legitimacy to compliance rooted in the beliefs of citizens about the normative appropriateness of a regime’s actions: “the most stable support [for a regime] will derive from the conviction on the part of the member that it is right and proper for him to accept and obey the authorities and to abide by the requirements of the regime”.¹⁹ There are two points of elaboration that are critical for understanding the importance of Easton’s contribution to subsequent political science research on legitimacy. The first is that Easton views legitimacy as

14. Specifically, a research assistant obtained all articles using the word legitimacy from JSTOR (for all included journals except CPS) or SAGE (for CPS) for this period. From this initial set, all articles where legitimacy was either an outcome or explanatory concept were retained. We read this initial set of 64 articles and discuss here those we deemed most relevant.

15. [1922] Max Weber, *Economy and Society* (Berkeley, CA: University of California Press, 1978).

16. *Ibid.*, 213.

17. *Ibid.*, 214.

18. David Easton, *A Systems Analysis of Political Life* (New York: Wiley, 1965).

19. *Ibid.*, 278.

a potential determinant of a broader concept that he terms *diffuse support*, defined as “a reservoir of favorable attitudes or good will that helps members to accept or tolerate outputs to which they are opposed or the effects of which they see as damaging to their wants.”²⁰ The second point is that while Easton views belief in the legitimacy of a regime as a “major source of diffuse support in most systems we know about,” he is quite clear in stating that systems can enjoy diffuse support even in the absence of legitimacy.

This last point is critical because it suggests that subjects’ beliefs in the legitimacy of a regime are not a necessary condition for diffuse support of that regime.²¹ Accepting Easton’s premise, then, we cannot infer the existence of subjects’ beliefs in the legitimacy of an institution or authority from evidence of diffuse support. Put bluntly, there are many reasons apart from legitimacy that individuals support existing political arrangements. Inadequately accounting for these other reasons is a key deficiency in the extant political science literature on the subject.

On the Legitimacy of Courts. In Federalist #78, Hamilton famously argued that the federal courts would have “neither force nor will, but merely judgment.” It is unsurprising, therefore, that scholars have suggested that a widespread belief in the legitimacy of the courts can explain how unelected judges maintain support for themselves and the institution of judicial review more generally despite their potentially unpopular decisions and the absence of a direct accountability linkage with the public.

Drawing on earlier work by Caldeira, Caldeira and Gibson construct a battery of survey questions designed to measure diffuse support.²² The battery includes questions about the respondent’s attitudes toward curtailing the power of the Court or even abolishing the institution outright. Gibson, Caldeira, and Spence operationalize “institutional loyalty” toward the Supreme Court explicitly to capture Easton’s “reservoir of good will” account of diffuse support.²³ In addition to

20. Easton, *A Systems Analysis of Political Life*, 273.

21. It need not be a sufficient condition either, but that is not of particular concern at present.

22. Gregory A. Caldeira, “Neither the Purse Nor the Sword: Dynamics of Public Confidence in the Supreme Court,” *American Political Science Review* 80, no. 4 (1986): 1209–1226; Gregory A. Caldeira and James L. Gibson, “The Etiology of Public Support for the Supreme Court,” *American Journal of Political Science* 36, no. 3 (1992): 635–664.

23. James L. Gibson, Gregory A. Caldeira, and Lester Kenyatta Spence, “Measuring Attitudes toward the United States Supreme Court,” *American Journal of Political Science* 47, no. 2 (2003): 354–367.

questions concerning the desirability of limiting the court's power, the survey battery they employ includes questions about trust in the Court, its propensity to favor certain groups over others, and its involvement in politics. In this paper, the authors elide Easton's distinction between diffuse support and legitimacy, even stating that "Easton and many others use 'diffuse support' as a synonym for legitimacy."²⁴

With survey-based measures of support, scholars have also examined how support for courts is shaped by individual factors (e.g., personal ideology and political awareness). Gibson and Caldeira, for example, examine the antecedents of support for the U.S. Supreme Court and the European Court of Justice.²⁵ Other research considers how specific court decisions and their alignment with respondent ideology, along with justifications offered for those decisions, affect support for courts.²⁶ A related literature, building on the idea that diffuse support can shield the courts from the consequences of unpopular decisions, examines the importance of a court's popularity in explaining patterns of conflict between it and legislatures or other bodies.²⁷ This literature therefore

24. Gibson, Caldeira, and Spence, "Measuring Attitudes toward the United States Supreme Court," 356.

25. James L. Gibson and Gregory A. Caldeira, "Confirmation Politics and the Legitimacy of the U.S. Supreme Court: Institutional Loyalty, Positivity Bias, and the Alito Nomination," *American Journal of Political Science* 53, no. 1 (2009): 139–155; James L. Gibson and Gregory A. Caldeira, "The Legitimacy of Transnational Legal Institutions: Compliance, Support, and the European Court of Justice," *American Journal of Political Science* 39, no. 2 (1995): 459–489.

26. For a representative sample, see Brandon L. Bartels and Christopher D. Johnston, "On the Ideological Foundations of Supreme Court Legitimacy in the American Public," *American Journal of Political Science* 57, no. 1 (2013): 184–199; Dino P. Christenson and David M. Glick, "Chief Justice Roberts's Health Care Decision Disrobed: The Microfoundations of the Supreme Court's Legitimacy," *American Journal of Political Science* 59, no. 2 (2015): 403–418; James L. Gibson, "Challenges to the Impartiality of State Supreme Courts: Legitimacy Theory and 'New-Style' Judicial Campaigns," *The American Political Science Review* 102, no. 1 (2008): 59–75; James L. Gibson, "'New-Style' Judicial Campaigns and the Legitimacy of State High Courts," *The Journal of Politics* 71, no. 4 (2009): 1285–1304; James L. Gibson and Michael J. Nelson, "Is the U.S. Supreme Court's Legitimacy Grounded in Performance Satisfaction and Ideology?," *American Journal of Political Science* 59, no. 1 (2015): 162–174; James L. Gibson, Gregory A. Caldeira, and Vanessa A. Baird, "On the Legitimacy of National High Courts," *The American Political Science Review* 92, no. 2 (1998): 343–358; Stephen P. Nicholson and Thomas G. Hansford, "Partisans in Robes: Party Cues and Public Acceptance of Supreme Court Decisions," *American Journal of Political Science* 58, no. 3 (2014): 620–636; James L. Gibson and Gregory A. Caldeira, "Campaign Support, Conflicts of Interest, and Judicial Impartiality: Can Recusals Rescue the Legitimacy of Courts?," *The Journal of Politics* 74, no. 1 (2012): 18–34; James R. Zink, James F. Spriggs, and John T. Scott, "Courting the Public: The Influence of Decision Attributes on Individuals' Views of Court Opinions," *The Journal of Politics* 71, no. 3 (2009): 909–925; James L. Gibson and Gregory A. Caldeira, "Defenders of Democracy? Legitimacy, Popular Acceptance, and the South African Constitutional Court," *The Journal of Politics* 65, no. 1 (2003): 1–30.

27. For example, Tom S. Clark, "The Separation of Powers, Court Curbing, and Judicial Legitimacy," *American Journal of Political Science* 53, no. 4 (2009): 971–989; Gretchen Helmke, "The Origins of Institutional Crises in Latin America," *American Journal of Political Science* 54, no. 3 (2010): 737–750.

similarly equates support for an institution with its legitimacy.

On the Legitimacy of Regimes Generally. A second, related area of research concerns the concept of regime legitimacy. This work, like its counterpart in the courts literature, frequently draws on the Weber/Easton understanding of the bases of support for a regime. An early example of this is found in work by Gurr and McClelland, who define legitimacy as “the extent that a polity is regarded by its members as worthy of support.”²⁸ Research in this tradition tends to compare regime support (1) across political systems, (2) over time, and (3) across individuals (within regimes).

Cross-country studies have found that higher levels of income inequality are associated with reduced trust and lower evaluations of government performance;²⁹ that legitimacy of communist rule in Eastern Europe was based on economic performance, but who also that deference also arose because of absence of credible alternative to the communist regime (arguably a concept independent of legitimacy);³⁰ and that challenges to Soviet and Soviet-proxy rule were more frequent in countries where pre-communist education systems inculcated nationalist conceptions of the right to govern that undercut the legitimacy of communist claims about legitimate authority.³¹ Studies exploiting over-time variation have found that elections followed by changes in control of government enhance legitimacy among those previously excluded from power;³² Others have found that legal reforms in China increased trust in the (national) government;³³ and that the determinants of regime support evolved and in post-Franco Spain.³⁴

28. Ted R. Gurr and Muriel McClelland, *Political Performance: A Twelve Nation Study* (London: Sage, 1971).

29. Christopher J. Anderson and Matthew M. Singer, “The Sensitive Left and the Impervious Right: Multilevel Models and the Politics of Inequality, Ideology, and Legitimacy in Europe,” *Comparative Political Studies* 41, nos. 4-5 (2008): 564–599.

30. Stephen White, “Economic Performance and Communist Legitimacy,” *World Politics* 38, no. 3 (1986): 462–482.

31. Keith Darden and Anna Grzymala-Busse, “The Great Divide: Literacy, Nationalism, and the Communist Collapse,” *World Politics* 59, no. 1 (2006): 83–115.

32. Devra C. Moehler and Staffan I. Lindberg, “Narrowing the Legitimacy Gap: Turnovers as a Cause of Democratic Consolidation,” *The Journal of Politics* 71, no. 4 (2009): 1448–1466. The authors operationalize legitimacy using indices of institutional trust, perceptions of the accountability of political leaders, deference to be governed, support for constitutions, and satisfaction with democratic performance.

33. Xiaobo Lü, “Social Policy and Regime Legitimacy: The Effects of Education Reform in China,” *American Political Science Review* 108, no. 2 (May 2014): 423–437; Susan H. Whiting, “Authoritarian ‘Rule of Law’ and Regime Legitimacy,” *Comparative Political Studies* 50, no. 14 (2017): 1907–1940.

34. Peter McDonough, Samuel H. Barnes, and Antonio López Pina, “The Growth of Democratic Legitimacy In Spain,” *The American Political Science Review* 80, no. 3 (1986): 735–760; Peter McDonough, Samuel H. Barnes, and Antonio López Pina, “The Nature of Political Support and Legitimacy in Spain,” *Comparative Political Studies* 27,

Scholarship focusing on individual-level variation has shown that individual-level reported corruption experience is associated with reduced diffuse support in four Latin American countries;³⁵ that trust in government in the Dominican Republic is correlated with perceptions of economic and political performance rather than democratic values or civic engagement;³⁶ that support for the post-apartheid democracy in South Africa is strongly correlated with race;³⁷ and that among those found guilty of a felony, those who judge the procedures to be fair exhibit more support for law and the broader political system after accounting for their assigned sentence.³⁸

While the research described above focuses mainly on attitudes, one important strand of work focusing on individual-level differences concerns behavior: specifically, which individuals support a regime and comply with its edicts. Mueller and Jukam, for example, show that general dissatisfaction with a governance system (trust, performance, etc., rather than just its particular elected leaders) is correlated with reported anti-system behavior.³⁹ And Giles and Gatlin show that individuals are more likely to exit racially desegregated schools for private schools, a behavior the authors classifies as defiance, if they perceive court intervention to mitigate school segregation as less legitimate, which is operationalized by items assessing beliefs about desegregation moving too quickly and viewing it as undesirable to have courts involved in desegregation policy.⁴⁰

A more recent strand of related research takes an experimental approach: in a laboratory setting, Grossman and Baldassari randomly assign an “authority” who has the ability to punish participants in a public goods game to be either selected at random or by voting from among participants. They find that the elected authority is associated with more frequent contributions and interpret this

no. 3 (1994): 349–380.

35. Mitchell A. Seligson, “The Impact of Corruption on Regime Legitimacy: A Comparative Study of Four Latin American Countries,” *The Journal of Politics* 64, no. 2 (2002): 408–433.

36. Rosario Espinal, Jonathan Hartlyn, and Jana Morgan Kelly, “Performance Still Matters: Explaining Trust in Government in the Dominican Republic,” *Comparative Political Studies* 39, no. 2 (2006): 200–223.

37. James L. Gibson, “The Legacy of Apartheid: Racial Differences in the Legitimacy of Democratic Institutions and Processes in the New South Africa,” *Comparative Political Studies* 36, no. 7 (2003): 772–800.

38. Tom R. Tyler, Jonathan D. Casper, and Bonnie Fisher, “Maintaining Allegiance toward Political Authorities: The Role of Prior Attitudes and the Use of Fair Procedures,” *American Journal of Political Science* 33, no. 3 (1989): 629–652.

39. Edward N. Muller and Thomas O. Jukam, “On the Meaning of Political Support,” *The American Political Science Review* 71, no. 4 (1977): 1561–1595.

40. Michael W. Giles and Douglas S. Gatlin, “Mass-Level Compliance with Public Policy: The Case of School Desegregation,” *The Journal of Politics* 42, no. 3 (1980): 722–746.

effect as arising due to legitimacy after showing that elected and appointed authorities do not behave differently.⁴¹ Olken reports results from a field experiment in Indonesia showing that when local citizens can vote on development projects, they are more knowledgeable, perceive larger benefits from those projects, and are more satisfied than when citizens do not have the opportunity to vote on selected projects.⁴² Further evidence shows relatively minor effects on which projects are chosen, which Olken argues that the process of democratic project selection may enhance the legitimacy of those projects.

Legitimacy of International Organizations. The final area of concern for scholars seeking to document legitimacy perceptions and their effects is that of supranational organizations and institutions.⁴³ This literature is largely qualitative in nature: Grant and Keohand, for example, discuss ways in which supranational authorities, which lack direct ties to those they govern, are accountable and therefore obtain elements of legitimacy.⁴⁴ Barnett discusses the way in which international political actors, and in particular the UN, gain legitimacy, acknowledging the difficulty of distinguishing the strategic logic of supporting the UN from accounts based in the normative legitimacy of the body.⁴⁵ Finnemore considers the difficulty a unipole faces in becoming legitimate – that is, able to garner support from others absent threats of force, and suggests obtaining legitimacy often requires reducing one’s short term power.⁴⁶ Borzel et. al. conduct a quantitative analysis of EU member states’ violations of European law as a function of legitimacy, which they operationalize as citizen support for the rule of law and public support for the EU. They find these factors do not

41. Guy Grossman and Delia Baldassarri, “The Impact of Elections on Cooperation: Evidence from a Lab-in-the-Field Experiment in Uganda,” *American Journal of Political Science* 56, no. 4 (2012): 964–985.

42. Benjamin A. Olken, “Direct Democracy and Local Public Goods: Evidence from a Field Experiment in Indonesia,” *The American Political Science Review* 104, no. 2 (2010): 243–267.

43. Ian Hurd, “Legitimacy and Authority in International Politics,” *International Organization* 53, no. 2 (1999): 379–408.

44. Ruth W. Grant and Robert O. Keohane, “Accountability and Abuses of Power in World Politics,” *The American Political Science Review* 99, no. 1 (2005): 29–43.

45. Michael N. Barnett, “Bringing in the New World Order: Liberalism, Legitimacy, and the United Nations,” *World Politics* 49, no. 4 (1997): 526–551.

46. Martha Finnemore, “Legitimacy, Hypocrisy, and the Social Structure of Unipolarity: Why Being a Unipole Isn’t All It’s Cracked up to Be,” *World Politics* 61, no. 1 (2009): 58–85.

predict violations, while state capacity and material interests do.⁴⁷

C. Observations

Viewed through the lens of our analysis of the normative-theoretic literature on legitimate authority, our brief tour through the recent journal-based empirical literature suggests three preliminary observations.

The first of these concerns the behavioral relevance of diffuse support for, or beliefs in the legitimacy of, specific institutions. Take, for example, the extensive literature on public support for national high courts. In environments where the legal system is largely enshrined, it is generally costless to the individual to criticize or defend it. Moreover, for the typical survey respondent, there is little chance that the absence of support for an institution entails a significant opportunity or motive to disobey it. Thus, the behavioral implications of an absence of legitimacy perceptions may, in many cases, be either unclear or miniscule.

Indeed, research by Caldeira and Gibson on South Africa finds that despite apparently low legitimacy, individuals are willing to defer to legal institutions even in that relatively nascent political system, suggesting that (low) diffuse support need not have substantial behavioral consequences. Below, we discuss the importance for research designs investigating legitimacy to consider cases in which citizens are closer to indifferent between compliance and non-compliance, and where, as a consequence, the behavioral implications of changes or differences in legitimacy perceptions are more likely to arise.

Second, even in situations where the behavioral implications of support for or opposition to institutions are clearer, a more pervasive issue of construct validity comes into focus. As alluded to above, support for a regime or institution is not synonymous with beliefs in the justness or rightness of the regime, its occupants, or its actions. Citizens may (diffusely) support a regime for many reasons: because of its past performance, custom and norms, convenience, or immediate self-interest.⁴⁸

47. Tanja A. Börzel et al., “Obstinate and Inefficient: Why Member States Do Not Comply With European Law,” *Comparative Political Studies* 43, no. 11 (2010): 1363–1390.

48. A related concern involves relevant counterfactuals: specifically, what implicit comparison is a citizen making

Past performance seems particularly salient as an antecedent of Easton's notion of a reservoir of good will: why abandon your support for an institution that regularly produces desirable results when, on occasion, it generates an undesirable one? In this regard, batteries of survey questions that are interpreted as measures of perceptions of legitimacy often implicate attitudes about trust in and satisfaction with an institution, for which a respondent, when answering, might easily consult his or her attitudes about a regime's performance rather than its normative appropriateness. Indeed, part of the reason that legitimacy has been dismissed as a useful theoretical concept by some scholars⁴⁹ is the multiplicity of other, performance-related factors that give rise to support.

Evidence of trust in institutions (by which we mean confidence in their capabilities, decisions, and directives) and norms of compliance with those institutions (by which we mean the tendency of individuals to comply because they perceive doing so as standard, proper, or typical) can also arise from numerous sources aside from perceptions of legitimacy. Citizens may decline to resist an authoritarian regime that delivers substantial GDP growth and threatens to torture subversives: in such cases, a fear of material punishment coupled with possible trust in the regime's ability to deliver economically (or mistrust of the opposition) may confound our efforts to assess perceptions of its legitimacy. A further reason to express caution about the legitimacy-revealing nature of measures of diffuse support is that norms and trust can arise through patterns of equilibrium selection in repeat play. For example, individuals may coordinate on a set of behaviors and beliefs that produce desirable outcomes and involve widespread compliance and regime support. In and of itself, this provides no evidence of a belief in the regime's legitimacy.

A final, closely-related set of concerns, which we develop more fully in a broader discussion in the next section, is about the issues of (1) causal inference and (2) attribution. To begin with, as a

when announcing his or her support for an institution (as against, e.g., a political party)? More broadly, calling into question a political system or institution may be a "cheap talk" survey response, in which individuals are not communicating sincere beliefs or intentions. On the broad problem of "cheerleading" in survey response, see John G. Bullock et al., "Partisan Bias in Factual Beliefs about Politics," *Quarterly Journal of Political Science* 10, no. 4 (2015): 519–578 [in the context of partisan bias]. Such concerns also arise if one attempts to use survey responses to "control for" different explanations for outcomes, particularly if individuals mis-report those behaviors or intentions in ways that are correlated with the treatment or outcome variables of interest.

49. For example, see Theda Skocpol, *States and Social Revolutions: A Comparative Analysis of France, Russia, and China* (Cambridge, MA: Cambridge University Press, 1979); Alan Hyde, "The Concept of Legitimation in the Sociology of Law," *Wisconsin Law Review* 1983 (1983): 379–426.

matter of inference, a key challenge to persuasive causal identification is coming up with a source of variation (in legitimacy) that is unrelated to other causes of the outcome or the outcome itself. We take up next this attribution problem, highlighting how changes in legitimacy often implicate other motivations for compliance. The causal inference issue is particularly thorny in observational analysis because it raises the essential question about what causes differences in perceptions of legitimacy. For example, if one group of individuals perceive their treatment by the courts as less legitimate and report they are less likely to comply with the law, what explains this individual level variation in perceptions and reported behavior? Suppose people who are inclined to disobey (for reasons unrelated to legitimacy) justify their behavior by reporting low legitimacy, or more perniciously, suppose some (omitted from a statistical model) third factor explains both legitimacy perceptions and state support. It is precisely these concerns that have led many scholars to embrace randomized experiments as a tool for persuasive causal identification, because by design they break the correlation between all characteristics of subjects (and therefore outcomes) and their treatment status. At the same time, it is important to note that most extant experiments rely on survey measures of outcomes, which itself returns us to the points raised above about the behavioral relevance and meaning of those responses.

Moreover, suppose we were to empirically detect a relationship between some feature of an institution generally perceived as just or appropriate and either expressed attitudes toward the institution or behavior supportive of it. Could we conclude that the institutional feature under study increased support or compliance by enhancing legitimacy? Even setting aside the issue of non-random assignment of treatment, it is not clear this is feasible absent very strong assumptions. This is because variation in the institutional feature that affects perceptions of its legitimacy may simultaneously affect the performance of, or trust in, the institution.⁵⁰ In the next section, we consider this attribution problem in greater detail and show that it can arise for a number of different reasons. We then provide guidelines that should aid researchers working in the field and hoping to

50. And, in a regression context, if two factors are correlated but only one causes the outcome of interest, regression will tend to give weight to both factors, a pattern that can be exacerbated by (differential) measurement error. This means that efforts to “control for” extrinsic rewards (material incentives) is a complete solution only under specific restrictive conditions.

overcome it.

3. A Positive Model of Legitimate Authority

A. Baseline: The State of Nature

To clarify intuition, we present a very simple model of social interaction. Imagine a stateless polity consisting of two citizens, 1 and 2. Each citizen has an endowment normalized to one, and both simultaneously choose to contribute the endowment to a common pot or keep it for themselves. For each contribution to the common pot, both citizens receive a return of ρ , which lies somewhere between $\frac{1}{2}$ and 1. Absent further elaboration, this structure corresponds to a standard prisoners' dilemma (PD). We modify the PD by adding idiosyncratic shocks to the payoff from contributing for each citizen, labeled w_1 and w_2 , which are drawn from a commonly known probability distribution. The realization of each shock w_i is the private information of citizen i . The payoff matrix, then, is:

		Citizen 2	
		<i>Contribute</i>	<i>Keep</i>
Citizen 1	<i>Contribute</i>	$2\rho + w_1, 2\rho + w_2$	$\rho + w_1, 1 + \rho$
	<i>Keep</i>	$1 + \rho, \rho + w_2$	$1, 1$

Note that the modified game is only a PD for certain draws of w_1 and w_2 ; otherwise, any of the four possible outcomes could be an equilibrium.

B. Enforcement Complementarity

Suppose now that there is an authority with some resources to punish citizens who are caught keeping their endowment. For this section, we treat the authority as a non-strategic actor – in particular, assume that the authority endeavors to punish a citizen if and only he or she observes that citizen to have kept. The authority has two relevant characteristics: capacity and legitimacy. Capacity represents the resources available to the authority to implement punishment. For example, if only one citizen keeps his or her endowment, that citizen is punished with probability c , but if

both keep, each is punished with some probability $c' < c$. Taking our cue from the normative and social psychological literatures, we model legitimacy as an enhancement to the intrinsic benefit to a citizen of contributing, labeled λ_i for each citizen i . For now, we keep legitimacy exogenous, allow it to vary across the two citizens, and assume that each citizen's λ is common knowledge.⁵¹

Solving for the equilibrium strategies of the two citizens (which can be described by cutpoints in w), three insights emerge. First, citizens' choices are strategic complements: that is, the more likely citizen i is to contribute, the higher the expected sanction to j from not contributing, and thus, the stronger the incentive of j to contribute. Interestingly, complementarity is enhanced by an aspect of the authority's *weakness*: the extent to which widespread noncompliance taxes their ability to maintain order. An increase in the state's ability to reliably punish all detected non-compliers reduces the extent of the complementarity.⁵²

Even in the absence of legitimacy concerns, the complementarities between the citizens' strategic choices implies the potential existence of multiple equilibria: if Citizen 1 (2) believes that Citizen 2 (1) will contribute, then she will also believe that the likelihood of punishment for keeping is relatively high, which may drive her to contribute (a "good" equilibrium). But if Citizen 1 (2) believes that Citizen 2 (1) will not contribute, then she will also believe that the likelihood of punishment for keeping is lower, which may drive her to keep (a "bad" equilibrium). The potential existence of multiple equilibria means that complementarities introduce a coordinative aspect into the baseline compliance game.

What, then, is the effect of the authority's legitimacy on the citizens' choices? Most obviously, citizen i is more likely to contribute if he/she views the authority as legitimate (i.e., the probability that i contributes is increasing in λ_i). More interestingly, though, citizen i is more likely to contribute if *the other* citizen views the authority as legitimate (i.e., the probability that i contributes is also increasing in λ_j .) This is owing to the strategic complementarity between the citizens discussed in the previous paragraph: high λ_j increases the likelihood that j contributes,

51. Introducing uncertainty over λ would leave the intuition qualitatively unchanged.

52. Note that this does not mean that enhanced state capacity reduces compliance overall: the dominant, first order effect of an increase in capacity is to increase the probability of compliance.

which increases the expected sanction to i from keeping, which increases the likelihood that i contributes. Even if i does not perceive the authority as legitimate *at all*, the legitimacy of the authority from j 's perspective still enhances i 's compliance for purely extrinsic reasons. More generally, the authority's legitimacy in the eyes of either citizen may dramatically reduce the range of parameter values supporting the "bad" equilibrium.

These effects are roughly equivalent to what Levi terms "quasi-voluntary compliance."⁵³ As an aside, note that legitimate authority is a sufficient but not necessary attribute of Levi's notion of quasi-voluntary compliance. For example, suppose the distributions from which the w_i 's are drawn differ by citizen. If, for example, j is known to be more altruistic on average, this fact will also increase the probability that i contributes, and also be a form of quasi-voluntary compliance. More generally, spillovers from one citizen to another can arise due to both extrinsic and intrinsic (e.g., legitimacy) motivations for the other citizen to comply.

The issue of strategic complementarity implies a violation of the Stable Unit Treatment Value Assumption (SUTVA) that underlies estimation of average treatment effects. Consider a plausibly exogenous increase to legitimacy of the authority in the eyes of one set of citizens that, via strategic complementarity, affects the behavior of another set of citizens who do not recognize that increase. A researcher who ignored this complementarity and compared the behavior of the citizen whose legitimacy perceptions changed to the unaffected citizens would *underestimate* the effect of the change in legitimacy, because it had behavioral effects on both groups. Likewise, a researcher documenting an association between the shock and a widespread increase in compliance (or other form of cooperation with the institution) would be estimating the total effect of the legitimating shock. The total effect would encompass both the direct legitimating effect on the first group of citizens and the purely material effect emerging through the complementarity mechanism. Absent additional assumptions, we would not be in a position to apportion causation across the two mechanisms. Note that an analyst can overcome the SUTVA issue either via research design (e.g., by changing the unit of analysis) or statistically (e.g., via clustering) to estimate a valid estimate of the

53. Margaret Levi, *Of Rule and Revenue* (Berkeley, CA: University of California Press, 1988).

total causal effect. This total effect, while arising due to a change in legitimacy perceptions, does not isolate individual-level changes in legitimacy perceptions (i.e., document the psychological mechanism) necessary to show that any behavioral change arises due to that individual's personal legitimacy perceptions.

To summarize, When authority capacity is limited, strategic complementarities mean that variation in the legitimacy and other motivations for compliance of one citizen affect the incentives of other citizens to comply. These complementarities make isolating which actors' perceptions of legitimacy affect their (versus others') behavior difficult. As such, situations in which such complementarities are minimized will make it easier to isolate the psychological mechanism underlying legitimate authority.

C. Institutional and Personal Legitimacy

We next consider a broader class of inference problems that may emerge even in the absence of strategic complementarities between the citizens.⁵⁴ To do so, we consider what might occur in situations where an authority can take some (possibly costly) action to enhance the *quality* of a governing institution. Before specifying what we mean by quality, consider that there are at least three channels through which this enhancement to quality might cause changes in the behavior of citizens.

1. The quality of the institution may affect the citizens' extrinsic incentives by changing the mapping between citizen action and expected punishment and/or rewards.
- 2a. The quality of the institution may engender greater affinity toward it, and hence toward the act of compliance itself. We refer to this as the *institutional legitimacy* channel.
- 2b. The authority's action can affect the citizens' beliefs about the authority's type (i.e., worthiness), which may in turn affect the intrinsic benefit of compliance with the authority. We refer to this as the *personal legitimacy* channel.

54. This section builds on Eric S. Dickson, Sanford C. Gordon, and Gregory A. Huber, "Identifying Legitimacy: Experimental Evidence on Compliance with Authority" (Typescript, New York University, 2017)

There are many ways to operationalize quality, but taking our cue from the extant literature, we proceed from the premise that the perceived fairness of the institution plays a role in enhancing its legitimacy in the eyes of citizens.⁵⁵ We follow Leventhal, who draws a useful distinction between distributive and procedural fairness, and focus for the moment on the latter, as it highlights the identification issues most clearly.⁵⁶ One critical aspect of procedural fairness is accuracy: the extent to which an institution accurately assigns sanction to non-compliers while exculpating compliers.

We can permute our modified PD from above as follows. First, we shut down the complementarities described in the preceding section by making the probability of punishment independent of the number of contributors. This will have the effect of making the relevant strategic interaction the one between the authority and the individual citizen, rather than the interaction between the citizens. Second, we suppose the authority is a strategic player. Prior to the contribution stage, the authority chooses an observable level of effort e , which affects the accuracy/quality of the institution q . We assume that accuracy is strictly increasing in effort. Authorities are characterized by a type, β , which scales the extent to which she cares about procedural fairness *primitively*, that is, independent of its effects on contributions. Letting Y denote the number of contributions to the common pot, the authority seeks to maximize $\rho Y + \beta q - e$. Finally, we modify the citizen's intrinsic benefit by assuming that it is responsive to three factors: the idiosyncratic term w_i ; institutional legitimacy from the perspective of i , $\lambda_i(q)$; and personal legitimacy, captured by citizen's i 's posterior belief concerning the extent to which the authority cares primitively about fairness, $\hat{\beta}$. Citizens value the extrinsic and intrinsic benefits from contributing, as well as the extrinsic cost of being punished (which, given inaccurate institutions, may happen even if the citizen contributes)

We now move to a consideration of what makes identifying personal or institutional legitimacy separately from extrinsic incentives so challenging.

55. Tom R. Tyler, "Procedural Justice, Legitimacy, and the Effective Rule of Law," *Crime and Justice* 30 (2003): 283–357; Jason Sunshine and Tom R. Tyler, "The Role of Procedural Justice and Legitimacy in Shaping Public Support for Policing," *Law and Society Review* 37, no. 3 (2003): 513–548.

56. Gerald S. Leventhal, "What Should Be Done with Equity Theory?," in *Social Exchange: Advances in Theory and Research*, ed. Kenneth J. Gergen, Martin S. Greenberg, and Richard H. Willis (Boston, MA: Springer, 1980), 27–55.

Identifying Institutional Legitimacy. Holding constant the authority's effort level, consider the effect of a change in institutional quality on the behavior of the citizen. There are two channels through which a change in institutional quality can affect the probability the citizen contributes. The first is the extrinsic channel: by improving the mapping between anti-social behavior and sanctioning, an increase in quality corresponds to an increase in marginal deterrence. The second is the intrinsic, institutional legitimacy channel: an increase in the enforcement accuracy enhances the citizen's positive affect toward the institution and thus the tendency toward compliance. These two channels are captured in the portion of the diagram in Figure 1 in the inner dashed rectangle labeled "Institutional."

FIGURE 1 ABOUT HERE

The multiplicity of channels of influence complicates efforts to isolate the legitimating effect of an increase in institutional quality empirically. Note that randomization of the institutional environment does not alleviate the problem. While it would permit a valid causal estimate of the total effect of institutional quality on citizen behavior, it would not permit the analyst to isolate the legitimating effect of institutional quality. This observation, of course, does not imply that the legitimacy channel is inoperative; indeed, a standard observation concerning compliance with legal norms is that citizens often comply even in the absence of strong material incentives to do so. That said, any observed association between institutional quality and citizen behavior cannot be attributed solely to the legitimacy or extrinsic motivation channels absent additional assumptions.

What is the overall takeaway from these considerations? To estimate the legitimating effects of institutions, it is advisable to consider institutions that have minimal effects on the material incentives of individuals to comply. In many cases, this can present a challenge for the researcher. For example, consider the recommendation of the community policing movement that police legitimacy will be improved if police officers are residents of the neighborhoods they patrol. The residence of the officers may seem materially irrelevant. But police officers who live in proximity to their beats are more likely to possess dense knowledge about the residents of those neighborhoods.

This may affect the accuracy with which they assign sanctions, and thus the material incentives of citizens.

Personal Legitimacy. To examine the personal legitimacy effect, note that in a *separating equilibrium* to the game, an authority with higher β exerts greater effort, and citizens infer the authority's type from the observed effort level and act accordingly. Can we identify the legitimating effect of the authority's effort choice from the observed relationship between the authority's action and the citizen's response?

Not in the current context, as it turns out. The causal effect of the authority's effort choice propagates through three distinct channels. The first is the personal legitimacy channel: by informing the citizen of the authority's type, the authority's effort affects the intrinsic motivations of the citizen, which in turn affects behavior. Second, the authority's effort affects the character of the institution, which in turn affects the extrinsic motivations of the citizen. Finally, the effect of the authority on the institution also affects the intrinsic motivations of the citizen through the institutional legitimacy channel. Figure 1 depicts these channels graphically.

Clearly, in light of this complex casual pathway, one cannot simply regress the citizen's behavior on the authority's action and claim to isolate a legitimacy effect. And controlling for the institution is not feasible, as it is, in this model, deterministically tied to the authority's effort choice. It would seem, then, that identification requires severing the deterministic linkage between the authority's choice and the institution. This, however, introduces two additional subtleties in light of the posited causal mechanism.

First, the researcher might be tempted to randomize the authority's effort choice, or find an instrumental variable that affects effort but neither institutional quality nor citizen behavior. The challenge here is that if effort is known to be randomly or quasi-randomly assigned, the citizen should draw no inferences about the authority's type from the observed level of effort.

Second, suppose the researcher decides instead to randomize the mediating variable q (institutional quality). A necessary and sufficient condition for a separating equilibrium (in which higher types take costlier actions) is that type and effort are complements. This is true by construction

above, given the relationship between quality/accuracy and the authority's effort level. If accuracy is completely randomized, however, then the separating equilibrium will collapse: higher types will gain nothing from their effort, and all types will choose no effort. In this situation, it is no longer feasible to detect personal legitimacy, and further, if accuracy shapes both material incentives and intrinsic motivations (via the institutional legitimacy channel) then those effects remain intertwined as we discussed above. Dickson, Gordon, and Huber describe sufficient conditions for identifying personal legitimacy effects given these issues.⁵⁷ Essentially, what is required is that the authority's level of effort induces *overlapping lotteries* over levels of accuracy that are orderable by stochastic dominance. Under these conditions, the separating equilibrium will be preserved, and the personal legitimacy effect can be identified from the relationship between the authority's and citizen's behavior, controlling for the realized level of quality.⁵⁸

That paper instantiates this strategy in a laboratory experiment as follows: if the authority takes a costly action, with 50% probability the accuracy is "high" and with 50% probability it is medium; if the authority foregoes the costly action, with 50% probability the accuracy is "medium" and with 50% probability it is low. In this design, the high type still has an incentive to take the costly action (thereby revealing her type), but one can obtain medium accuracy regardless of whether the costly action was taken. The paper then compare citizen behavior when the authority took the action with that behavior when she did not, restricting attention to cases in which the realized accuracy level was medium.

In summary, identifying a causal relationship between the action of an authority and response of a citizen that can accurately be described as "legitimacy" or "legitimation" is difficult owing to the relationship between the authority's legitimating action and the institutional environment of the subject, which activates extrinsic motivations as well as institutional legitimacy considerations. This is a particular kind of causal mediation problem, but for the reasons detailed above, the

57. Dickson, Gordon, and Huber, "Identifying Legitimacy: Experimental Evidence on Compliance with Authority."

58. For a related identification strategy – termed "selective trials" – involving probabilistic treatment assignment following a choice by the subject in an experiment, see Sylvain Chassang, Gerard Padró I Miquel, and Erik Snowberg, "Selective Trials: A Principal-Agent Approach to Randomized Controlled Experiments," *American Economic Review* 102, no. 4 (June 2012): 1279–1309.

answer is not as simple as “randomize the mediator.” Looking for situations in which the institutional environment is related probabilistically to the authority’s actions is one way to identify the personal legitimacy effect. Another would be to look for a valid instrumental variable that, along with the authority’s action, affected the institutional environment without separately affecting the authority’s effort choice or citizen behavior. While each strategy provides purchase to identify the personal legitimacy effect, these strategies do not seem feasible in the realm of institutional legitimacy (unless the institution itself has no effect on material incentives).

4. Other Considerations

A. The Case for Indifference

An important research design challenge is identifying cases in which it is possible to ascertain whether a theoretically relevant construct operates as hypothesized. Note that this is analytically distinct from asking whether a concept is relevant or operative in a given domain, but is instead about whether one could detect the importance of a given concept in a given situation. Put differently, we wish to look for instances in which a marginal change in some theoretically relevant concept would produce an observable behavioral effect.

The general question of “where to look” is particularly important in the study of legitimacy because, as we articulated above, a belief in legitimacy is just one among many reasons individuals may comply with an authority of institution. For example, even in the version of the model described above that assumes away complementarities, an individual’s decision to comply is characterized by a cutpoint w^* , and compliance occurs if and only if the sum of idiosyncratic individual-level motivations to comply, material/deterrence motivations, and legitimacy effects exceed this cutpoint. Explicitly modeling compliance in this way makes clear the importance of choosing the right cases for study. Suppose for the moment that the legitimating effect of an institution or authority choice could be isolated from material compliance motivations using one of the strategies described above. Then the average treatment effect of an incremental change in the institution will be proportional to the height of the shock density around the cutpoint. If most

citizens are already inclined to comply (w^* in the left tail of the distribution) or not comply (w^* in the right tail), then even those reforms that yield a large change in the cutpoint may have only negligible behavioral consequences.⁵⁹

What does this mean for research design? We note that many cases where legitimacy is studied are precisely those where compliance seems unlikely to change given posited changes in legitimacy. Take, for example, the question of how Americans respond to Supreme Court decisions they perceive as illegitimate or legal verdicts they perceive as unfair. While prior work provides extensive survey evidence about perceptions of these outcomes, evidence about their behavioral consequences is scant because nearly all individuals will comply with those decisions regardless of marginal perturbations to their legitimacy perceptions for other reasons. In parallel fashion, in regimes where individuals routinely ignore legal verdicts, even substantial increases in legitimacy are unlikely to manifest in detectable behavior changes. Would this mean legitimacy is inconsequential? No, because the lack of a behavioral consequence may arise precisely because individuals are far from indifferent with regard to compliance. Of course, cumulative and large changes in legitimacy may matter for things like Supreme Court decisions, but as a matter of objects of study, the local changes in legitimacy induced by most interventions for these sorts of policies are unlikely to be those with substantial behavioral consequences. This means we may wish to look elsewhere to understand how much and for whom different policies induce behaviorally-relevant changes in legitimacy.

By contrast, in cases where more individuals are close to the compliance threshold, it is much more likely that legitimacy enhancing or degrading policy changes or choices will be consequential, and studying those circumstances will help us to instantiate the behavioral importance of legitimacy (subject to the identification concerns detailed above). Thus, the key takeaway point from this analysis is that if one wants to provide behavioral evidence about legitimacy, one of

59. In environments with complementarities and heterogeneous legitimacy effects, the marginal effect of an increase in the legitimacy perceptions of a subset of citizens can be magnified, via the mechanism described above. For example, thinking through the logic of the game above, suppose both players did not contribute under policy p_1 . If the change in legitimacy associated with a move to policy p_2 induced one player to contribute, then the second player would need a smaller change in legitimacy to contribute than in the absence of complementarities.

two conditions should hold. First, one should look for situations in which, either as a product of researcher intervention or experimental design, individuals are close to indifferent with regard to compliance.

Second, and alternatively, one should focus study on those who are otherwise likely to be close to indifferent with regard to compliance, even if this is only a subset of the larger population affected by a regime. Prior research on regime legitimacy provides suggestive evidence of the fruitfulness of this targeted approach. Among West German college students, it was leftists whose perceptions of the right-leaning German regime's legitimacy that were correlated with self-reports of anti-regime activity.⁶⁰ By contrast, for rightists, for whom the regime's legitimacy was arguably secondary to its ideological attractiveness, no such effects were observed.

B. The Case for Hard Cases

We have, at this point, identified several research strategies for grappling with the deep entanglement between material and legitimacy-based reasons for compliance with an authority. But a deeper understanding of how this entanglement operates in different environments points to another approach to identification: the examination of "hard cases" for legitimacy. These are instances in which, unlike in the cases described above, institutions, policies, and choices that potentially enhance legitimacy *diminish* material compliance motivations. In such situations, the detection of legitimacy effects will be *in spite of*, rather than owing to, the extrinsic motivation channel. Or, to use the language of statistics, estimated legitimacy effects will be downward- rather than upward-biased.

A prominent example of such a case is group- or neighborhood-based profiling in law enforcement: for example, hot spot policing, racial profiling in motor-vehicle and pedestrian stops, and ethno-religious profiling in airport security screening. Suppose a local police department allocates more intense scrutiny to a neighborhood where it anticipates more crime. From the perspective of the police, this may be efficient: the marginal deterrent effect of this reallocation should improve compliance with the law (or detection of lawbreakers). A law-abiding citizen in the heavily

60. Muller and Jukam, "On the Meaning of Political Support."

scrutinized neighborhood may view the allocation as illegitimate: from her perspective, it may be motivated by bias rather than efficiency. Suppose a law enforcement agency moved toward a more equitable allocation of resources, and we simultaneously observed an increase in compliance in the previously targeted neighborhood. The increase in compliance would not, in this circumstance, be attributable to enhanced material incentives to comply, but rather, to enhanced perceptions of the fairness of policing (increased legitimacy) net of diminished deterrence.⁶¹

5. Conclusion

Legitimacy is a foundational concept in both normative political theory and positive political science and is widely seen as being an important explanation for observed patterns of citizen behavior and institutional support. Indeed, the frequency with which the concept of legitimacy is invoked by academics and policymakers alike as an explanation for the success or failure of policy reveals our prior beliefs about its importance. Despite the frequency with which it is invoked, however, there remains substantial theoretical and empirical imprecision about its meaning, which may account in part for the fact that some scholars are highly skeptical of its practical importance.

In this essay, we have discussed the normative underpinnings of the concept of legitimacy and linked those theoretical concepts to the problem of conducting empirical research that provides informative evidence about its importance. This review highlights several core issues for empirical research that invokes legitimacy, including problems of conceptual clarity, measurement, and causal identification. To more fully elucidate these claims, we present a simple model that helps us to understand the causal and theoretical challenges facing empirical researchers.

In this conclusion, we briefly recount what we see as the key points we have identified based on this review and their implications for future work:

1. The concept of legitimacy, as understood in normative work and foundational empirical

61. This, of course, assumes the availability of measures of compliance independent of the police themselves, a nontrivial issue in itself but one beyond the scope of the current discussion. See Eric S. Dickson, Sanford C. Gordon, and Gregory A. Huber, "Profiling in the Lab: Experimental Evidence from a Novel Subject Population" (Typescript, New York University, 2017).

work, is distinct from extrinsic (material) motivations for consent. Legitimacy should be understood as affecting the intrinsic motivations of citizens to comply because they care about the values an institution or authority embraces or instantiates.

2. In equilibrium, legitimacy obtains when citizens hold beliefs about an authority that enhance these intrinsic motivations to comply, and authorities take (justifiable) actions that are consistent with those beliefs. Legitimacy therefore has to be understood in terms of the interaction between what authorities and citizens do.
3. In moving from normative absolutes to empirical prediction, we must recognize that legitimacy is one motivation among many others for differences in support for an authority or state. Legitimacy is therefore likely to matter on the margin in generating compliance.
 - (a) It is important not to conflate evidence for diffuse support with evidence for legitimacy. Individuals may value an authority or institution over feasible alternatives for a variety of reasons, but diffuse support (and proxy measures of trust, confidence, and support for political institutions or authorities) is not itself evidence that such diffuse support arises due to legitimacy.
 - (b) If legitimacy is to matter in explaining behavior, it will only be in cases where, *ceteris paribus*, individuals are otherwise close to indifferent about their decision to comply. Research that invokes legitimacy as causally important must therefore anticipate not only that legitimacy considerations are operative, but also that perturbations to those considerations will have behaviorally relevant implications.
4. Because there are many different motivations for compliance, empirical work must grapple with the fact that institutions, choices, and actions by an authority may simultaneously affect the intrinsic and extrinsic motivations for compliance. This has three key implications for existing work.
 - (a) First, at one extreme, if decisions by a governing authority enhance both material and

non-material motivations for compliance (thereby implicating material incentives, institutional legitimacy, and potentially an authority's personal legitimacy), then empirical research will not be able to isolate the causal effect of legitimacy absent strong assumptions. This problem is likely to be exacerbated in situations where state capacity is limited, because effects on one citizen's willingness to comply (whether due to legitimacy or other factors) can indirectly affect other citizens' incentives to comply. Only in cases where legitimacy-enhancing institutions have *no* effect on the other (material) motivations to comply is it possible to isolate institutional legitimacy effects, and only in situations where strategic complementarities among citizens are minimal can an individual citizen's responses to a posited institutional change be isolated as arising due to its effect on perceptions of legitimacy. For researchers beginning empirical work, a primary question of research design is therefore to assess whether factors thought to affect legitimacy have any effect on extrinsic motivations for behavior.

- (b) Second, if factors that improve legitimacy also affect extrinsic rewards, it is precisely hard cases – those in which enhanced legitimacy and improved material incentives have opposite effects on behavior – that provide the most promising opportunities for understanding the consequences of improved institutional legitimacy. This is because it is these situations that estimates of institutional legitimacy effects will be downwardly biased, and so finding evidence for legitimacy in these situations would provide persuasive evidence not only about the existence of, but also the empirical importance of, enhanced legitimacy. These are also the key “real world” situations in which arguments about the tradeoffs between legitimacy and material incentives are the most pressing for public policy.
- (c) Third, simple randomization of an institution cannot solve this problem of attribution. This is true first because even randomized institutions alter both intrinsic (institutional legitimacy) and extrinsic motivations for behavior. It is also true because if it is an authority's choices that communicate her concern for values the citizen cares about

– thereby enhancing her legitimacy – simple randomization makes her choice meaningless or undercuts the reasons for that action altogether. We specify the conditions under which an authority’s choice to pursue a superior institution can, through conditional randomization with overlapping support in realized institutions, create situations where estimating a personal legitimacy effect is not conflated by material concerns (or institutional legitimacy).

Stepping back somewhat, our view is that legitimacy remains an important area for political science research. With renewed attention to a conceptually delimited and rigorously defined definition of legitimacy, many of the complaints about the theoretical imprecisions surrounding claims of legitimacy may be addressed. At the same time, this more circumscribed understanding of legitimacy itself reveals the thorny empirical challenges that make isolating legitimacy so hard.

All of this begs the question of “why bother?” That some reform “works,” whether for legitimacy or other reasons, may be all that we care about. We close with two observations. First, as academic researchers, we believe one of our key roles is to isolate theoretical mechanisms. Prediction without understanding, particularly given the frequency with which legitimacy is invoked causally, therefore risks misunderstanding. This misunderstanding may itself be undesirable, but if it in turn leads to bad policy advice, the consequences may also be deleterious from a public welfare perspective. Nor is this problem limited to empirical research. Many important normative arguments draw on inferences and intuitions we make about “what causes what,” and as such, making sure we understand whether legitimacy is operative helps us understand whether individuals are in fact motivated by these concerns.

Second, there are many situations in which policymakers confront choices about which (costly) reform to implement. While it is often the case that enhanced legitimacy and improved material motivations for compliance move hand in hand, they do not always do so and sometimes interventions that affect one causal pathway more than another are differentially costly. Fully isolating and understanding the empirical consequences of legitimacy for compliance allows us to better make predictions and recommendations in situations in which the contours of policy involve choices

along these lines.

Biographical

Sanford C. Gordon is Professor of Politics and Associated Professor of Law at New York University. He received his B.A. from Cornell and his Ph.D. from Princeton. Prior to his arrival at NYU, Professor Gordon taught at the Ohio State University from 1999-2002; and in 2005-2006, was a fellow-in-residence at the Center for the Study of Democratic Politics at Princeton. His research has appeared in the *American Political Science Review*, the *American Journal of Political Science*, the *Journal of Politics*, *Public Choice*, the *Quarterly Journal of Political Science*, *Political Analysis*, the *Journal of Law, Economics, and Organization*, and the *New England Journal of Medicine*, among other places.

Gregory A. Huber is the Forst Family Professor of Political Science at Yale University. He earned his Ph.D. from Princeton and currently serves as the associate director of the Center for the Study of American Politics and director of Institution for Social and Policy Studies Behavioral Research Lab at Yale. He is also an associate editor of the *Quarterly Journal of Political Science*. He is the author of more than 50 papers on various topics in American Politics.

Figure 1: Graphical Representation of Identification Issues in the Study of Institutional and Personal Legitimacy. Oval nodes represent observables, and rectangular nodes unobservables.

