

**ENGINEERING GLOBAL JUSTICE:
ACHIEVING SUCCESS THROUGH FAILURE ANALYSIS**

by
David Wiens

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Philosophy)
in The University of Michigan
2011

Doctoral Committee:

Professor Elizabeth S. Anderson, Chair
Professor William Roberts Clark
Professor Peter Railton
Associate Professor Mika LaVaque-Manty

© David Wiens 2011
All Rights Reserved

For Barb, Zack, and Cubby

ACKNOWLEDGEMENTS

I have accumulated many debts while completing this dissertation. My most immediate are those to my committee members — Elizabeth Anderson, Mika LaVaque-Manty, Peter Railton, and William Roberts Clark.

Liz's early and obvious enthusiasm for this project assured me that I was headed in a worthwhile direction and gave me the self-confidence needed to pursue it. Her prompt and incisive comments on numerous drafts guided its trajectory, pressing me to new lines of inquiry that were often hidden from my view. Liz has provided the optimal mix of freedom and direction while the project unfolded, facilitating my growth as a scholar and professional.

Mika has graciously listened to me vent my frustrations with mainstream analytic political philosophy and generously shared his experience in dealing with similar frustrations of his own. His guidance has helped me navigate an intermediate space between the sometimes very different worlds of political philosophy and political theory. My work is much better as a result.

Peter has been an important resource throughout my time at Michigan, with respect to both my research and my development as a teacher. His charity and generosity have made me a better scholar and educator.

Bill has gone far beyond the duties of a cognate committee member; he has been a de facto co-chair, mentor, and close friend. Bill and I met in my first year at Michigan, when I enrolled in his international political economy course to satisfy a degree requirement. I

was naïve about what the course involved; I had never encountered statistical analysis or formal models. Bill challenged me to engage the literature on its own terms and, once I showed interest in doing so, accepted the challenge of helping acquire the requisite competencies. Along the way, I began to see political philosophy through the eyes of a political economist. What I saw inspired this dissertation. Were it not for his investment, my research certainly would not have taken the direction it has. But my greatest debt to Bill is a personal one. Not only has he invested in me as an apprentice scholar; he has invested deeply in me as a person. Were it not for his guidance and friendship, I would be neither the professional nor the person I am today. Thanks, Bill.

The philosophy department at Michigan has provided an outstanding environment in which to apprentice as a professional scholar. Thanks to the many faculty, staff, and graduate students who make this a great place to study. Special recognition is owed to several people. Thanks to Linda Shultes for helping me navigate the labyrinth that is the university bureaucracy. Thanks to Molly Mahony for her generosity, shown not only to me, but to my family as well. Thanks to Jim Joyce and Louis Loeb for their service as department chairs; I also thank them both for taking a personal interest in my development as a philosopher and a professional. Thanks to Sarah Buss for her advice and friendship. Early versions of much that appears here were presented at several meetings of the Graduate Student Working Group and Candidacy Seminar. Thanks to the various organizers for the many opportunities; thanks to participants for fruitful discussions. Finally, Nate Charlow, Billy Dunaway, Jason Konek, and Sam Liao made sure that there was more to grad student life than study; Andrew Stevenson (a non-philosopher) was equally important in this regard. Their friendship has been crucial for pushing me to the end.

Without any obligation to do so, the members of Michigan's political science department generously integrated me into the life of their department. Thanks to Rob Franzese,

John Jackson, Skip Lupia, Jim Morrow, and Scott Page for guiding me through an ad hoc methods education. Thanks to members of the Political Theory Workshop and Nameless Internal Seminar for including me and giving me opportunities to present early versions of various bits of this dissertation. Thanks in particular to Zev Berger for insightful comments on at least two of my presentations. With apologies to those I am forgetting, thanks for encouragement, insight, and friendship to James Atkinson, Carolyn Barnes, Sourav Guha, Mihwa Hong, Trevor Johnston, Dan Katz, Jennifer Miller-Gonzalez, Kate Vandenberg Bradley, and Alton Worthington. Special thanks are owed to Cassie Grafström, Bill MacMillan, and Paul Poast—for timely help, but mostly for treating me as one of their own.

Earlier versions of material that appears here have been presented to various audiences. Material from the introduction and conclusion were presented at a global justice and global governance workshop at the Canadian Political Science Association annual conference; parts of chapter 2 were presented at the American Philosophical Association’s annual Eastern Division meeting; chapter 3 was presented at the British Society for Ethical Theory annual meeting; parts of chapters 5 and 6 were presented together at the 2nd annual Dutch Conference on Practical Philosophy; my early thoughts on feasibility (here included as appendix B) were presented at a workshop on nonideal theory and institutional theory at the Canadian Political Science Association annual conference. Many thanks to the participants at these meetings for their comments. In addition to my committee members, thanks to the following people for insightful discussions and correspondence on various parts of this dissertation: Sean Aas, Zev Berger, Sarah Buss, Steve Campbell, Simon Caney, Billy Dunaway, Lisa Ellis, Pablo Gilabert, Bob Goodin, Jonathan Harmon, Jim Johnson, Trevor Johnston, Jason Konek, Holly Lawford-Smith, Dan Little, Skip Lupia, Jim Morrow, Sarah Moss, Dan Peterson, Jennifer Rubenstein, Alex Silk, Dan Singer, Peter Vallentyne, Leif

Wenar, and two anonymous reviewers for *The Journal of Political Philosophy*. A version of chapter 2 is forthcoming in *The Journal of Political Philosophy* under the title “Prescribing Institutions Without Ideal Theory”; thanks to the publishers for permission to reproduce that work here.

When I was six or so years old, my dad wrote a poem, touting me as “David, the scholar, composing his books.” He must have seen something even back then, although I don’t think he (or Mom) expected this. My parents have persistently encouraged me to become my best self and gave me every opportunity to succeed. Their investment made this dissertation possible.

My boys, Zachary and Jacob, never take me too seriously, teaching me to not take myself too seriously. This has been an invaluable gift, given through the hugs and smiles that greet me at the end of every day regardless of how the day has gone.

This dissertation is not only the culmination of years of graduate study, but of ten years of marriage. Barb has been a loyal partner the entire way, working as hard as me — and sometimes even harder — to make sure I finished well. Barb, you are my lighthouse.

TABLE OF CONTENTS

DEDICATION	ii
ACKNOWLEDGEMENTS	iii
LIST OF APPENDICES	ix
CHAPTER	
1. Introduction	1
2. The Importance of Causal Analysis	14
2.1 Causal Analysis	16
2.2 Conventional Modes of Engagement: A Catalog	19
2.2.1 Mere correlations	20
2.2.2 Taxonomy	23
2.2.3 Crude causal information	26
2.3 When Is Causal Analysis Required?	29
2.3.1 Normative political philosophy	30
2.3.2 Applied political philosophy	37
2.3.3 Conceptual political philosophy	40
2.4 Practical Implications	46
3. Institutional Failure Analysis	48
3.1 Architecture and Engineering in Institutional Design	51
3.2 Unpacking The Ideal Guidance Approach: A Case Study	55
3.3 The Failure Analysis Approach to Institutional Design	60
3.3.1 Identifying failure	62
3.3.2 Diagnosing failure	72
3.3.3 Designing to avoid failure	79
3.4 A Fruitful Approach?	83
4. Interlude: From Theory to Application	84

5. Explaining the Resource Curse	87
5.1 Institutions and the Resource Curse	89
5.2 From Policy Choice to Institutional Design	92
5.3 A Model of Political Competition	95
5.4 Resource Revenue and Political Choice	103
5.5 Empirical Implications	107
5.6 Venezuela: A Test Case	111
6. The Resource Curse and the Demands of Justice	120
6.1 Resources and the Demands of Justice	123
6.2 Are There Alternatives to the Resource Privilege?	127
6.2.1 Pogge’s proposed alternative	129
6.2.2 Wenar’s proposed alternative	130
6.3 Assessing the Alternatives	132
6.4 The Need for Exit Options	138
6.5 Where To Now? A Tentative Proposal	143
6.6 A Last Word	150
7. Conclusion: Toward a Pragmatic Statism	152
 APPENDICES	 160
 BIBLIOGRAPHY	 188

LIST OF APPENDICES

APPENDIX

A.	Technical Appendix	161
B.	An Analysis of Feasibility	171
	B.1 A general schema	172
	B.2 Putting the schema to work	178
	B.3 Concluding remark	186

CHAPTER 1

Introduction

Vast numbers of people across the globe suffer acute deprivation. The numbers have become familiar. The number of people living on less than US\$2.00 per day has remained constant at around 2.5 billion since 1981. That number represents more than a third of the current global population.¹ Every year, 12 million children die before their fifth birthday, 4.5 million due to diarrhoea and pneumonia alone;² less than 100,000 of these deaths occur in developed countries.³ 1.2 billion people lack access to an adequate water supply. In 2009 alone, 23.5 million children were orphaned due to AIDS.⁴ From 1998–2007, 216,500 people died from civil conflict; 60,000 people died from violent campaigns perpetrated by a government or other formally organized group.⁵ 37.4 million people are refugees or internally displaced from their homes, many as a result of conflict or systematic human rights violations.⁶

What do we owe these global poor? Specifically, what do the residents of places like Australia, Canada, the European Union, and the United States owe to the residents of places like Afghanistan, Angola, Chad, the Congo, Haiti, Libya, Myanmar, Sudan, Yemen,

1. World Bank, *World Development Indicators*, International Bank for Reconstruction and Development, 2010. Dollar figure reported in US 2005 dollars at purchasing power parity.

2. World Health Organization, *World Health Statistics*, World Health Organization, 2010.

3. UNICEF, *State of the World's Children – Statistical Tables*, United Nations Children's Fund, 2011.

4. Ibid.

5. Human Security Report Project, *Human Security Report*, Simon Fraser University, 2009/2010.

6. UNHCR, *2009 Statistical Yearbook*, United Nations High Commissioner for Refugees, 2010.

and Zimbabwe? My answer assumes two basic commitments:

The cosmopolitan commitment: All human beings enjoy equal status as the fundamental units of moral concern and the interests of each should thus be extended equal concern and respect by all other human beings. Accordingly, national affiliations or state borders are irrelevant for determining the content and scope of our duties of justice.

The institutional commitment: At minimum, cosmopolitan justice requires that we work toward the establishment of international institutional arrangements that are consistent with individuals' status as moral equals and create the conditions under which each individual is capable of living a minimally decent life.

I do not argue for these commitments here. Instead, I begin to explore the implications of these commitments. Specifically, *how should we design the international system to fulfill cosmopolitan moral principles?*

A burgeoning literature has emerged to answer this question. The generality of these two commitments leaves much room for disagreement. Cosmopolitans differ with respect to the moral theory they take to generate cosmopolitan duties — from utilitarianism,⁷ to social contract theories,⁸ to rights-based theories.⁹ Cosmopolitans also differ with respect to the moral objectives they take as fundamental: to promote welfare,¹⁰ to promote individual autonomy,¹¹ to promote capabilities,¹² to fulfill human rights,¹³ or to satisfy basic needs.¹⁴

Finally, cosmopolitans disagree on the depth and breadth of international institutional

7. Peter Singer, "Famine, Affluence, and Morality," *Philosophy & Public Affairs* 1, no. 3 (1972): 229–243.

8. Charles R. Beitz, *Political Theory and International Relations*, 2nd ed. (Princeton: Princeton University Press, 1999); Thomas W. Pogge, *Realizing Rawls* (Ithaca, NY: Cornell University Press, 1989); Darrel Moellendorf, *Cosmopolitan Justice* (Boulder: Westview Press, 2002).

9. Luis Cabrera, *Political Theory of Global Justice: A Cosmopolitan Case for the World State* (New York: Routledge, 2004); David Held, *Democracy and the Global Order: From the Modern State to Cosmopolitan Governance* (Stanford: Stanford University Press, 1995); Henry Shue, *Basic Rights: Subsistence, Affluence, and U.S. Foreign Policy* (Princeton: Princeton University Press, 1980).

10. Singer, "Famine, Affluence, and Morality."

11. Held, *Democracy and the Global Order*.

12. Martha C. Nussbaum, *Women and Human Development* (Cambridge: Cambridge University Press, 2000); Amartya Sen, *Development as Freedom* (New York: Anchor Books, 1999).

13. Thomas W. Pogge, *World Poverty and Human Rights* (Malden, MA: Polity Press, 2002); Shue, *Basic Rights*.

14. Onora O'Neill, *Bounds of Justice* (Cambridge: Cambridge University Press, 2000).

integration required to fulfill cosmopolitan principles. At one end of the spectrum are those who reject the need for international institutional reform or propose only modest *ad hoc* changes to address specific issues.¹⁵ At the other end are those who advocate a centralized global political authority modeled on the modern federal state.¹⁶ In between these extremes, we find advocates for institutional restructuring at varying levels of depth and breadth. Closer to the middle, we find intermediate proposals such as Gillian Brock's suggestion of global taxation schemes, Darrel Moellendorf's proposal to democratize the World Trade Organization, or Allen Buchanan's plan to institutionalize the preventive use of interstate military force.¹⁷ Between this "limited institutional cosmopolitanism" and the world state, we find cosmopolitan democrats, who propose the creation of a network of interconnected and overlapping agencies, with authority being dispersed both vertically (from local bodies through to global bodies) and horizontally (dispersing issue-specific policy-making authority among a plurality of distinct organizations).¹⁸

Underlying these differences are three commonalities. First, regarding their moral objectives, cosmopolitans are ultimately concerned with *human flourishing*, broadly defined. The fundamental cosmopolitan objective is to see individuals everywhere live decent and worthwhile lives.¹⁹ Second, regarding their institutional proposals, nearly all cosmopolitans advocate an institutional order that includes at least some authoritative suprastate institutions. Institutional cosmopolitanism prescribes at least some institutions that have

15. Singer's or Shue's calls for greater foreign aid are examples of the latter; see Singer, "Famine, Affluence, and Morality"; Shue, *Basic Rights*.

16. Cabrera, *Political Theory of Global Justice*.

17. Gillian Brock, *Global Justice: A Cosmopolitan Account* (New York: Oxford University Press, 2008); Darrel Moellendorf, "The World Trade Organization and Egalitarian Justice," *Metaphilosophy* 36, nos. 1/2 (2005): 145–162; Allen Buchanan and Robert O. Keohane, "The Preventive Use of Force: A Cosmopolitan Institutional Proposal," *Ethics & International Affairs* 18, no. 1 (2004): 1–22; Allen Buchanan, "Institutionalizing the Just War," *Philosophy & Public Affairs* 34, no. 1 (2006): 2–38.

18. Held, *Democracy and the Global Order*; cf. Simon Caney, *Justice Beyond Borders: A Global Political Theory* (New York: Oxford University Press, 2005) and Pogge, *World Poverty and Human Rights*, ch. 7. The term "limited institutional cosmopolitanism" comes from Cabrera, *Political Theory of Global Justice*, 2.

19. Cf. Pogge, *World Poverty and Human Rights*, ch. 1.

mandatory global jurisdiction with respect to some issues and are granted authority to constrain the activity of states in pursuit of cosmopolitan objectives.²⁰

Finally, cosmopolitans' arguments for their favored institutional proposal follow the same general schema.

1. The cosmopolitan commitment generates an obligation to realize *the cosmopolitan objective*.
2. *Cosmopolitan institutional arrangements* are required to realize *the cosmopolitan objective*.

Therefore,

3. The cosmopolitan commitment generates an obligation to establish *cosmopolitan institutional arrangements*.

The salient feature of this schema is that it employs *instrumental* reasoning; cosmopolitans prescribe some particular institutional form as required to fulfill cosmopolitan objectives.

Reducing institutional cosmopolitan arguments to the foregoing schema permits us to evaluate these arguments as a class. Naturally, questions arise regarding the causal connection between cosmopolitan institutional forms and the realization of cosmopolitan objectives. Are cosmopolitan institutional forms required to realize cosmopolitan objectives? Could they could even do so effectively? Institutional cosmopolitans present a variety of reasons to support their case: global institutions are required to ensure compliance with cosmopolitan principles; global institutions are required to ensure adequate provision of global public goods; global institutions can reliably prevent states from depriving their citizens.²¹ Intuitively, these support the claim that cosmopolitan institutional arrangements are indeed required to realize any worthwhile cosmopolitan objective.

Premise 2 in the aforementioned argument schema reveals that questions of global institutional design crucially turn on facts about social causal processes. By what causal

20. Cf. Caney, *Justice Beyond Borders*, ch. 5.

21. Cf. *ibid.*, 159–160.

processes have so many people become so severely deprived? By what processes could cosmopolitans' proposed institutional forms be implemented? How would the proposed institutions affect the processes that lead to severe deprivation? Would they alleviate the deprivation or exacerbate it? Given the centrality of these questions to the global institutional design project, it is surprising to discover that global political theorists pay so little attention to the relevant causal processes. We can classify this general neglect more precisely. Corresponding to the preceding questions, we can charge global political theorists with three methodological flaws: they *ignore the causal processes that generate deprivation*; they *fail to rigorously evaluate the processes required to implement their proposals*; and they *fail to anticipate the consequences of implementing their proposals*. As I demonstrate in chapter 2, these flaws are both widespread and consequential. They put extant cosmopolitan institutional design proposals in doubt.

These flaws are not limited to cosmopolitan political theory; they emerge from philosophers' general approach to institutional design. Philosophers' institutional prescriptions fall under one of two broad categories. *Ideal* prescriptions identify institutional principles that would characterize a fully just system of global institutions. *Nonideal* prescriptions specify institutional arrangements that purport to eradicate or at least mitigate various actual injustices. The conventional wisdom is that ideal theory must precede nonideal theory. Even if we are ultimately interested in proposing feasible institutional solutions, our first task is to identify the principles that constitute an ideal institutional structure. We identify these ideal principles by reflecting on the institutional principles that best fulfill our considered moral judgments under ideal social and political conditions, conditions that abstract away from empirical complications, such as humans' motivational and cognitive limitations, acute deprivation, and frequent inter- and intrastate violence. These abstract conditions more or less permit moral considerations to take center stage in decisions about

how to organize our collective life. The principles that emerge from this exercise subsequently guide and constrain our attempts to prescribe institutional solutions for our decidedly nonideal world. I call this conventional approach to institutional design the *ideal guidance approach*.²²

The ideal guidance approach has considerable intuitive appeal if one takes the institutional design objective to be the realization of “a more just state of affairs.” This broad objective suggests the need for a target at which to aim. Since our selection of targets is primarily guided by moral concerns, and since assuming ideal conditions abstracts away from empirical complexity so as to give moral considerations primacy, ideal theory seems to be the place to start. Without ideal theory to define the principles that would constitute institutional arrangements in a fully just state of affairs, we can’t know which institutional arrangements count as “more just” than the current ones. In A. John Simmons’s words, “We can hardly claim to know whether we are on the path to the ideal of justice until we can specify in what that ideal consists. [...] The requirement that nonideal policies be ‘likely to be successful’ requires that we know how to measure success; and that measure makes essential reference to the ultimate target, the ideal of perfect justice.”²³ So the conventional wisdom goes.

Philosophers’ inattention to causal processes arises from this ideal guidance approach to institutional design. Put simply, the ideal guidance approach places emphasis on moral analysis at the expense of social scientific analysis. In terms of the aforementioned argument schema, the ideal guidance approach places emphasis on premise 1 at the expense of premise 2. Not surprisingly, much of the global justice literature focuses on debates between cosmopolitans and their nationalist and statist rivals: Should we endorse cosmopoli-

22. The *locus classicus* for this view is John Rawls, *A Theory of Justice*, 2nd ed. (Cambridge, MA: Harvard University Press, 1999), 7–8, 215–216.

23. A. John Simmons, “Ideal and Nonideal Theory,” *Philosophy & Public Affairs* 38, no. 1 (2010): 5–36, at 34.

tan moral claims? On what grounds should we do so? What moral objectives should we pursue? These are, of course, important issues. What we ought to do depends on how we answer these questions and philosophers have made important progress in answering them. But the above schema shows that progress on institutional design questions depends not only on progress with respect to moral considerations, but on progress with respect to the relevant social scientific issues as well. If we are to prescribe feasible, stable, and effective institutional solutions, we *must* attend to the relevant causal processes.

The preceding recommends development of an alternate methodology for institutional design, one that takes both moral analysis and social scientific analysis seriously. I lay the groundwork for such an alternative in chapter 3. Against the ideal guidance approach, I propose that philosophers adopt an *institutional failure analysis approach*. This alternative takes its design task to be obviating or averting social failures (as opposed to approximating an ideal).

The failure analysis approach promises to fulfill the aim of nonideal theory more effectively than does the ideal guidance approach. Ostensibly, political philosophers are motivated to shift from ideal theorizing to nonideal theorizing by a desire to propose institutions that are (1) capable of bringing about a state of affairs that is more just than the current state while (2) operating in the midst of current injustice. The ideal guidance approach to nonideal theory has failed to deliver proposals that can meet these desiderata, in large part due to its neglect of social science. The central innovation of the failure analysis approach is to give social scientific analysis a central role in our moral evaluation of institutions; it urges philosophers to attend equally to the issues involved in both premises 1 and 2 of the preceding argument schema. Failure analysis thus enables global political theorists to prescribe more effective solutions to injustice because its focus is squarely on understanding the *injustice*, rather than focusing on trying to understand an ideal of jus-

tice. In so doing, failure analysis better fulfills the objective of nonideal political theory, namely, to think about how, in the midst of current injustice, we might bring about social conditions that are more just than our current conditions.

Abstract reflections are insufficient to justify methodological innovation; methodological principles are vindicated through fruitful application. Hence, with the failure analysis framework in hand, the remaining chapters begin the substantive task of proposing institutional reforms to address widespread deprivation. Specifically, I address a phenomenon that development political economists call “the resource curse.” Counterintuitively, many countries with abundant natural resource wealth, such as the Congo, Myanmar, Nigeria, and Venezuela, suffer serious development failures. These “resource-cursed” countries tend to have stagnant economies and authoritarian governments, and often experience protracted violent civil conflict.²⁴ On its face, the resource curse is a moral disaster. Nearly a quarter of the global poor live in resource-cursed countries. They are ruled by dictators who are infamous for their brutality, corruption, and thievery; examples include Sani Abacha, Omar al-Bashir, Mobutu Sese Seko, Teodoro Obiang, and Than Shwe.

Chapter 5 develops a detailed explanation of the resource curse. To do this, I use game theory as a tool for analyzing important aspects of the strategic interaction between a country’s political leaders and its citizens. Building on previous work by (among others) Robert Bates and Charles Tilly, I argue that states undergo successful political and economic development when domestic institutions constrain state leaders to attend to citizens’ interests. This occurs when citizens’ bargaining leverage vis-à-vis state leaders is sufficient to compel the latter to enact policies that advance citizens’ general interests in secure rights and material well-being.²⁵ The resource curse arises when resource revenue

24. For good surveys of the relevant literature, see Michael L. Ross, “The Political Economy of the Resource Curse,” *World Politics* 51 (1999): 297–322; Andrew Rosser, “The Political Economy of the Resource Curse: A Literature Survey,” *IDS Working Paper*, no. 268 (2006).

25. Robert H. Bates, *Prosperity and Violence: The Political Economy of Development*, 2nd ed. (New York

enters a country with domestic institutions that are too weak to constrain the leader to attend to citizens' interests. Moreover, the flow of resource revenue into a country with weak institutions undermines leaders' incentives to carry out positive institutional reforms, thereby perpetuating the resource curse.

Chapter 6 considers the normative implications of this explanation for international institutional design. Extant prescriptions focus exclusively on institutional reforms that target state leaders' use of resource revenue. In light of the explanation offered in chapter 5, there are two problems with these proposals. First, mechanisms that merely limit the flow of resource revenue into resource-cursed countries are insufficient to improve citizens' bargaining position vis-à-vis their leaders.²⁶ In addition to decreasing state leaders' access to resource revenue, the explanation presented in chapter 5 shows that citizens' bargaining leverage depends on having credible "exit threats," opportunities to avoid absorbing the harmful consequences of remaining in their resource-cursed situation. Second, extant prescriptions neglect the extent to which the persistence of institutions that foster the resource curse is endogenous to the curse itself. Curse-engendering institutions are endogenous to the curse if the incentive structure that generates overconsumption and underinvestment, corruption and patronage, unaccountability and repression also inhibits voluntary establishment of the accountability and transparency mechanisms required to avoid these ills. If this endogeneity thesis is true, then it is futile to first identify (e.g.) underinvestment in human capital accumulation or lack of accountability to citizens as key causes of the resource curse and then prescribe that the state invest more in human capital or establish

and London: W. W. Norton & Co., 2009); Charles Tilly, *Coercion, Capital, and European States, AD 990–1992*, Revised ed. (Malden, MA: Blackwell, 1992). Clark, Golder, and Golder present a helpful formalization of the reasoning; see William Roberts Clark, Matt Golder, and Sona Golder, "The Balance of Power Between Citizens and the State: Democratization and the Resource Curse" (University of Michigan and Penn State University, unpublished manuscript, 2008).

26. Cf. Pogge, *World Poverty and Human Rights*; Leif Wenar, "Property Rights and the Resource Curse," *Philosophy & Public Affairs* 36, no. 1 (2008): 2–32.

mechanisms to ensure citizen oversight of resource transactions.²⁷

I highlight the need to incorporate a focus on citizens' exit options into our prescriptions to address the resource curse. The importance of exit options cannot be overstated. State leaders who lose access to resource revenue can still neglect citizens' interests if those citizens have no where else to go or no alternate leadership candidates to turn to. For any prescription to effectively improve citizens' bargaining leverage vis-à-vis their leaders, it must also provide citizens with exit options. Focusing our prescriptive efforts on exit options has a second advantage. Since increasing leaders' reliance on citizens is incredibly difficult so long as resource revenue continues to enter the domestic political economy, providing certain types of exit options might permit resource curse victims to *escape* the curse even where it cannot be overcome.

The prospects for providing resource curse victims with credible exit options appear unfavorable at first glance. Nevertheless, I tentatively propose a strategy that could mitigate the resource curse in a limited number of cases. My strategy has two components. The first aims to increase state leaders' fiscal reliance on citizens by conditioning official development aid and sovereign loan disbursements on achieving a minimum level of fiscal reliance on non-resource tax revenue. Eligible revenue should derive from taxes on income earned domestically by individual citizens and profits earned by businesses outside the resource sector. In implementing this rule, we induce state leaders to make investments that advance citizens' interests in an effort to increase his eligible tax base. Second, we should establish international peer-to-peer lending networks that could facilitate opportunities for citizens to invest their limited resources abroad. This could enable citizens to reduce the amount of their income that is eligible for taxation that meets the rule, thereby providing them with leverage in bargaining with their leaders over taxation. To keep his

27. Cf. Macartan Humphreys, Jeffrey D. Sachs, and Joseph E. Stiglitz, eds., *Escaping the Resource Curse* (New York: Columbia University Press, 2007).

citizens' investments in the country, the leader may have to concede oversight of taxation and public spending to citizens, which would further advance citizens' interests.

Together, chapters 5 and 6 present a resolutely pragmatic approach to nonideal institutional design. My approach eschews high-minded analysis of political ideals in favor of sustained reflection on the causal processes that generate stubborn problems amidst decidedly nonideal conditions. Successful remedies must work within the constraints defined by humans' motivational and cognitive limitations, resource scarcity, and technological and environmental limitations. Only careful diagnostic work can yield insight regarding the boundaries and possibilities that determine available solutions. Political philosophers risk irrelevance when they forgo rigorous diagnosis and proceed directly from moral philosophy to institutional design.

Before I end this introduction, I must address a worry that I am exhorting philosophers to engage in a project that is not at all *philosophical* in nature. If conventional political philosophy is not cut out to provide effective institutional prescriptions, why not enjoin philosophers to abstain from undertaking such work altogether? Why not follow G.A. Cohen in affirming that "the question for political philosophy is not what we should do but what we should think, even when what we should think makes no practical difference"?²⁸ On Cohen's view, there is a fundamental distinction between basic principles and principles of regulation. *Principles of regulation* are the principles that constrain institutional design; they are "device[s] for having certain effects."²⁹ *Basic principles* are statements of our ultimate normative convictions; they are the standard by which we judge the consequences of principles of regulation. Importantly, basic principles are supposed to be *fact-free*; they are true independent of facts about the world. I'm exhorting philosophers to attend to facts when reflecting on principles of regulation. But, Cohen objects,

28. G.A. Cohen, "Facts and Principles," *Philosophy & Public Affairs* 31, no. 3 (2003): 211–245, at 243.

29. Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), 38, quoted by Cohen.

philosophers ought not attend to principles of regulation (and so need not attend to facts). Political philosophy's primary task is to assess basic principles; we need not attend to facts to do this.

I have two replies. First, whether or not philosophers *should* attend to principles of regulation is irrelevant at this point. Philosophers have taken it upon themselves to prescribe principles of regulation and are wading deep into the territory of institutional design. This seems particularly true of the global justice literature. Insofar as this trend is irreversible, it is sensible to be clear about how this sort of inquiry should proceed. The failure analysis approach speaks to this issue.

Second, and more deeply, the objection assumes that we can sensibly separate our assessment of basic principles from our assessment of facts about the social world. This assumption is mistaken. Suppose we concede to Cohen that basic principles of justice do nothing other than set out a framework for morally evaluating states of affairs; they say nothing about what we ought to do given our evaluations. Even still, whether some principle *P* is a sound principle of justice (a sound principle for evaluating states of affairs) depends on the practical requirements of realizing states of affairs that fulfill *P*. This is because our assessment and selection of evaluative principles is always informed by our judgments about which states of affairs could possibly be realized and the consequences of their realization. *P* is only a legitimate evaluative principle if it exhorts us to adopt appropriate evaluative attitudes with respect to the pursuit or failure to pursue states of affairs required by *P*. Failing to pursue states of affairs required by *P* is blameworthy or otherwise regrettable only if those states could possibly be realized or their realization would not have unwelcome consequences. If fulfilling a basic principle requires implementing principles of regulation that would have perverse consequences (according to the basic principle), then we would have reason to revise our basic principle. Contrary to Cohen's

thesis, our judgments regarding basic principles necessarily attend to (causal) facts about our world.

So global political theorists must attend to facts about the world when prescribing international institutional arrangements. But don't they do this already? If not, where are they going wrong? What sort of facts must they attend to? I answer these questions in the next chapter.

CHAPTER 2

The Importance of Causal Analysis

Global political theorists pay scant attention to a particular type of fact, namely, facts about social causal mechanisms. This inattention is consequential. Causal claims are important for assessing the moral quality of institutions, as well as prescribing effective solutions to injustice; for evaluating the extent to which particular agents' relationships to particular social outcomes generate moral duties; even for presenting an adequate analysis of important political concepts, such as power, freedom, or self-determination. Failure to rigorously analyze the relevant causal mechanisms puts these causal claims — and the arguments on which they are based — in doubt.

To preempt any misunderstanding, my claim is not merely that philosophers ought to attend to, and avoid sloppy reasoning about, the facts, insofar as their arguments depend on facts. Although true, that claim is too obvious to be interesting. Nor do I merely claim that philosophers fail to put forward the requisite causal claims; such claims pervade the literature. Instead, I claim that political philosophers must rigorously analyze the causal mechanisms on which their causal claims depend; they fail to do so to the detriment of their arguments.

In the abstract, my thesis might still seem uncontroversial. Yet I show that philosophical practice does not adhere to the implications of this uncontroversial claim. This is why

I do not argue for my thesis in the abstract. To dispel any notion that philosophers can maintain the status quo, my argument consists of two illustrative surveys. The first canvasses global political theorists' typical appeals to empirical information. This is meant to draw the distinction between analysis of causal mechanisms on the one hand and other sorts of empirical analysis on the other; it also illustrates the ways in which philosophical arguments go astray when they are inattentive to causal mechanisms. The second survey divides theorists' claims into three broad types — normative, applied, and conceptual — and takes stock of the ways in which rigorous analysis of causal mechanisms is required for each type of claim. Perhaps most surprisingly, I show that causal analysis is required in at least some straightforward cases of conceptual political philosophy. The central lesson is that analysis of social causal mechanisms must play a more central role in political philosophical argument than is typical. This is not a “simple methodological lesson”; the flaw exposed here is both widespread and consequential. If right, my argument implies that numerous debates in the global political theory literature must be revisited with an eye toward analyzing causal mechanisms.

Two more words of clarification before I continue. First, my claim is not that political philosophers must ground their arguments on *true* causal analyses as opposed to false ones (although true claims are obviously desirable). I recognize that analyses of social causal processes are matters of ongoing debate among social scientists. My claim is rather that philosophers' causal claims must be based on analyses of causal mechanisms *as opposed to other types of empirical analysis*, such as statistical correlations or historical narrative. Analyzing causal mechanisms forces philosophers to lay bare the causal logic presupposed by their arguments to a greater extent than other types of empirical analysis; this, in turn, facilitates more rigorous scrutiny of the argument.

Second, my discussion of particular arguments is critical throughout. Hence, one

might get the impression that this chapter is primarily concerned to make a substantive point about global justice. Lest anyone misunderstand, I stress that my objective throughout is to expose methodological rather than substantive flaws. Of course, methodological flaws have important substantive implications; thus, I often challenge or reject substantive conclusions to make my point about the need for increased attention to causal analysis. However, a rejection of any particular substantive position I take with respect to any particular dispute is not a rejection of my argument here. To reject my thesis, one must show that the sorts of arguments philosophers make frees them from any requirement to attend to causal analysis or that increased attention to causal mechanisms would not significantly alter the substantive conclusions philosophers draw from their arguments.

2.1 Causal Analysis

To fix terms, say that *empirical information* describes the social world—it tells us how the social world *is* and how it came to be that way or how it could come to be some other way, rather than how it *ought to be* or how it ought to come to be some way. This definition admits a variety of types:

- *Data points*: points of fact demonstrable by empirical methods; for example, that 2.6 billion people live on less than US\$2 per day, or that Angola achieved independence in 1975.
- *Correlation*: the demonstration of a statistical relationship between two or more variables; for example, that democratic governance is statistically associated with higher levels of economic development.
- *Narrative*: a descriptive ordering of events according to some organizational principles (typically time, but also topical relevance or causal salience); for example, a chronology of European diplomatic affairs following the Congress of Vienna.
- *Taxonomy (classification)*: a partition of objects (including states of affairs) according to shared characteristics; for example, a classification of countries into regime types.
- *Interpretive*: an account of the meaning or significance (symbolic, cultural, or normative) attributed (or attributable) to particular practices or events; for example,

Clifford Geertz's claim that cockfights are "fundamentally a dramatization of [Balinese] status concerns."¹

- *Causal*: statements regarding a process that generates some empirical regularity or a particular state of affairs.

Let's focus on the last item for a moment (we will return to the others later). The definition of *causal information* I have offered is consistent with the claim that a set of true causal statements constitutes causal information. For example, on this definition, the statement "inclement weather causes a decrease in attendance at outdoor sporting events" counts as causal information. Political philosophers use this sort of information all the time. To more precisely specify my claim that philosophers are inattentive to causal analyses of the relevant phenomena, I need to distinguish *causal information* from *causal analysis*. We do *causal analysis* when we identify the salient components of a process that generates an outcome and specify their interrelationships. This latter part includes specifying how the components interact and how changes in one part of the process affect (the operation of) other components.

This definition has two notable features. First, it is compatible with various conceptions of social explanation, including *holism* (the view that an individual's behavior is explained as a function of her position or functional role within a social system²), *functionalism* (the view that the occurrence of a social causal process that generates outcome *O* is explained by the [beneficial] consequences *O* has for some entity³), and *methodological individualism* (the view that "all social phenomena (their structure and their change) are in principle explicable only in terms of individuals — their properties, goals, and beliefs"⁴).

1. Clifford Geertz, "Deep Play: Notes on the Balinese Cockfight," in *The Interpretation of Cultures* (New York: Basic Books, 1973), 437.

2. Cf. J.W.N. Watkins, "Methodological Individualism: A Reply," *Philosophy of Science* 22, no. 1 (1955): 58–62, at 58.

3. Cf. G.A. Cohen, "Marxism and Functional Explanation," in *Analytical Marxism*, ed. John Roemer (Cambridge: Cambridge University Press, 1986), 223.

4. Jon Elster, "Marxism, Functionalism, and Game Theory: The Case for Methodological Individualism," *Theory and Society* 11, no. 4 (1982): 453–482, at 453.

Second, the emphasis on analyzing a causal story into its component parts and specifying their interrelationships privileges a *mechanistic* view of social explanation. Identifying a (*social*) *causal mechanism* is a matter of “opening up the black box” and identifying the “cogs and wheels” that connect a cause *C* with its effect *E*.⁵ Less metaphorically, a statement of a causal mechanism provides a detailed account of the connection that enables *C* to generate *E*. Continuing with the earlier example, a mechanistic explanation of attendance at outdoor sporting events identifies the components of the causal link between inclement weather and attendance; for example, that the weather changes some individuals’ recreation preferences, or that the weather blocks some travel routes.

Appeal to mechanisms is especially appropriate for social explanation, for several reasons. First, social scientists are reluctant to grant the status of a law to any empirical social regularity. Those that are sometimes called “laws” — the “law” of demand in microeconomics; Duverger’s “law” in political science — are at best what Jon Elster calls “weak laws”: for any change in an explanatory variable, they permit us to predict the *direction* of the change in the dependent variable, but not the *magnitude* of the change.⁶ Second, law-based explanations do not permit us to discriminate between genuine causal connections and spurious correlations. Relatedly, law-based explanations do not tell us the process by which a set of causes generates its outcome. These last two are important because one of the things in which political philosophers are interested is prescribing effective solutions for averting or mitigating morally problematic outcomes. To give us a sense of which

5. The “black box” phrase comes from Jon Elster, *Explaining Social Behavior: More Nuts and Bolts for the Social Sciences* (New York: Cambridge University Press, 2007), ch. 2; the “cogs and wheels” phrase comes from Jon Elster, *Nuts and Bolts for the Social Sciences* (Cambridge: Cambridge University Press, 1989), 3. For more detailed discussion of causal mechanisms and their role in social explanation, see the essays in Peter Hedström and Richard Swedberg, eds., *Social Mechanisms: An Analytical Approach to Social Theory* (New York: Cambridge University Press, 1998).

6. Elster, *Explaining Social Behavior*, 36. We should be careful here with respect to the law of demand; for, contrary to this law, price increases can lead to *increased* consumption of some luxury goods — so-called “Veblen” goods, after the economist Thorstein Veblen. See Thorstein Veblen, *The Theory of the Leisure Class* (New York: Penguin Books, 1994) for his classic statement of the point.

interventions are likely to be effective, we must know *how* an outcome was generated; a mechanistic story grants such insight.

I can now make my thesis more precise. Despite substantial attention to empirical information of all sorts — *including causal information* — global political theorists are often inattentive to causal mechanisms; that is, they often fail to identify the salient components of the causal process generating the outcome(s) of interest and specify these components' interrelationships. This inattention vitiates their arguments. To help further distinguish causal analysis from other types of empirical information, I use the next section to present some examples of philosophers using different types of empirical information when they should be engaging in causal analysis. Along the way, I show how such misuse of empirical information often leads philosophers astray.

2.2 Conventional Modes of Engagement: A Catalog

Philosophers typically treat social science as providing ready-made, “off-the-shelf” answers to their empirical questions. This encourages the practice of simply citing the social scientific results that support the empirical premises one wishes to employ. This type of passive reliance is problematic. To begin with, it is bad practice. Not all social scientific studies are equal in quality. Many have poorly specified or misinterpreted statistical models or specious formal models, which bias the results of the research.⁷ Citing biased results vitiates an argument that presupposes the state of affairs portrayed by the biased results.

Moreover, a body of social scientific literature is not necessarily relevant to a particular

7. Two examples. Brambor, Clark, and Golder document the pervasive misspecification and misinterpretation of interactive statistical models in the top political science journals; see Thomas Brambor, William Roberts Clark, and Matt Golder, “Understanding Interaction Models: Improving Empirical Analyses,” *Political Analysis* 14 (2006): 63–82. Poast shows how using dyadic data to analyze multilateral events — the dominant mode of statistical analysis in international relations scholarship — generates significantly biased statistical estimates; see Paul Poast, “(Mis)Using Dyadic Data to Analyze Multilateral Events,” *Political Analysis* 18, no. 4 (2010): 403–425.

philosophical debate simply because the former addresses the broad topic dealt with in the latter. There are two general pitfalls here. One is that the data used in the cited empirical literature might be poor measures of the phenomena with which the normative literature is concerned.⁸ The other is relying on an empirical literature that speaks to the effect of some intervention on one type of outcome—for example, the effect of foreign aid on economic growth—when the normative debate is concerned with that intervention’s effect on another type of outcome—for example, the effect of foreign aid on poverty reduction. Any argument that succumbs to either of these dangers is thereby discredited. Passive reliance on social scientific results decreases the likelihood that philosophers are sensitive to the relevance of the studies they cite.

More importantly for my purpose here, passive reliance perpetuates a division of labor whereby causal analysis is left for social scientists so that philosophers can “attend to the philosophical issues.” Maintaining this division makes it less likely that philosophers will become attentive to causal explanations of the outcomes that concern them and more likely to rely on other types of empirical information where causal analysis is needed. To illustrate this danger, consider the following examples of misused empirical information.

2.2.1 Mere correlations

Mathias Risse is concerned to defend an account of our “duty to increase the level of prosperity of the global poor.”⁹ Risse claims that “it will be an empirical question of how actually to discharge that duty and any answer to this question must be informed

8. For example, Sanjay G. Reddy and Thomas W. Pogge, “How *Not* to Count the Poor,” in *Debates on the Measurement of Global Poverty*, ed. Sudhir Anand, Paul Segal, and Joseph Stiglitz (Oxford: Oxford University Press, 2009) criticize extant measures of poverty for failing to adequately capture certain normatively salient features of poverty. For rebuttal, see Martin Ravallion, “How Not to Count the Poor? A Reply to Reddy and Pogge,” in Anand, Segal, and Stiglitz, *Global Poverty*.

9. Mathias Risse, “What We Owe to the Global Poor,” *The Journal of Ethics* 9 (2005): 81–117, at 96; cf. *ibid.*, 89.

by our understanding of the sources of prosperity.”¹⁰ So far, so good. But the empirical literature on which Risse relies to make his case is concerned with “what makes *countries* do well.”¹¹ The problem this poses comes out clearly in the following reconstruction of Risse’s argument.

1. We have a duty of assistance to increase *individual prosperity*.
2. The practical content of this duty is given by our understanding of the determinants of prosperity.
3. Domestic institutional quality is the key to increased *country prosperity*.¹²
4. Therefore, our duty of assistance involves providing aid to improve the quality of domestic institutions (where possible).

Put this way, the argument is clearly invalid; it only goes through if the italicized terms match. We can make the argument valid if we add a premise (one which Risse never acknowledges).

1. We have a duty of assistance to increase *individual prosperity*.
2. The practical content of this duty is given by our understanding of the determinants of prosperity.
3. Domestic institutional quality is the key to increased *country prosperity*.
- (*) Increases in *country prosperity* cause increases in *individual prosperity*.
4. Therefore, our duty of assistance involves providing aid to improve the quality of domestic institutions (where possible).

The argument is now valid. But (*) is not obviously true. Although there is some evidence that poverty reduction is correlated with economic growth, this is by no means a settled question.¹³ Nor is it sufficient to show a mere correlation between poverty reduction and

10. Risse, “What We Owe,” 89.

11. *Ibid.*, 84, note 5, emphasis added.

12. On this point, Risse appeals primarily to Dani Rodrik, Arvind Subramanian, and Francesco Trebbi, “Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development,” *Journal of Economic Growth* 9 (2004): 131–165. Rodrik et al’s dependent variable is country income per capita.

13. For evidence to support a correlation between economic growth and poverty reduction, see David Dollar and Aart Kraay, “Growth is Good for the Poor,” *Journal of Economic Growth* 7 (2002): 195–225; for evidence to the contrary, see Martin Ravallion, “Growth, Inequality, and Poverty: Looking Beyond Averages,” *World Development* 29, no. 11 (2001): 1803–1815.

aggregate growth. Whether economic growth leads to poverty reduction depends on the causal foundations of the growth—for example, the sectoral composition of the economy and the number of people employed in growing sectors; the extent to which growth generates demand for skilled labor and, hence, stimulates investment in human capital. In particular, whether aggregate growth reduces poverty depends on how the gains from growth are distributed, and the gains are not always distributed evenly.¹⁴

To illustrate the point, compare Botswana's record with that of Indonesia. From 1986–1993, Botswana averaged 8.9% aggregate growth and 5.8% per capita growth, while reducing the percentage of its population living on less than \$1.25/day from 36% to 31%. In contrast, during roughly the same time period, Indonesia averaged 7.1% aggregate growth and 5.3% per capita growth, while reducing the percentage of its population living on less than \$1.25/day from 63% to 43%.¹⁵ Clearly, poverty reduction depends on more than just economic growth.

Risse's argument is in trouble because it relies on a mere correlation rather than analysis of the relevant causal mechanisms. It is either invalid or it relies on a controversial premise. If the empirical debate on (*) comes down in his favor, fine. But, importantly, this is insufficient to save his argument. In addition to needing to identify the causal mechanisms linking aggregate growth with poverty reduction, Risse needs to identify the causal mechanisms that generate strong domestic institutions, as well as those that link domestic institutional quality with increased prosperity. The claim that we have a duty to provide assistance to countries with weak domestic institutions hinges on the claim that we are *capable* of helping countries strengthen their domestic institutions. Whether this is so depends on the causal mechanisms underlying institutional development. More subtly,

14. World Bank, *World Development Report 2000/2001: Attacking Poverty* (New York: World Bank / Oxford University Press, 2001).

15. Data from World Bank, *World Development Indicators*, 2009 update.

whether we are able to help countries strengthen their domestic institutions *in ways that lead to increased prosperity* depends on the mechanisms that link domestic institutional quality to increased prosperity. Thus, despite initial appearances, Risse’s appeal to correlations instead of causal explanations is inadequate to establish the desired conclusion.

2.2.2 Taxonomy

As an example of taxonomic information standing in where causal analysis is required, consider a familiar dispute between statist — who claim that egalitarian duties of distributive justice are owed only to co-citizens — and cosmopolitans — who argue that such duties extend beyond state borders. Some statist argue that equality arises as a distributive concern only among individuals who live under common institutions that coercively enforce compliance or where compliance is otherwise nonvoluntary. Such institutions require special justification to those who are subject to them because requiring compliance against an individual’s will violates her autonomy. Only egalitarian principles of distributive justice are able to meet this special burden of justification, since those who fare worst under the shared institutional arrangements can reasonably reject nonegalitarian principles as a basis for institutions with which they are required to comply. Since the state is constituted by a coercive or otherwise nonvoluntary institutional apparatus, equality arises as a distributive concern among co-citizens. Moreover, the state is *unique* in being constituted by a coercive or otherwise nonvoluntary institutional apparatus. Statist claim that international institutions consist in voluntary associations for mutual benefit; those who are unhappy with their fate under international institutions can opt out and look elsewhere for a better deal.¹⁶

16. For an argument along these lines that emphasizes the uniquely coercive nature of the state, see Michael Blake, “Distributive Justice, State Coercion, and Autonomy,” *Philosophy & Public Affairs* 30, no. 3 (2001): 257–296; for an argument that the state is uniquely constituted by citizens’ joint authorship of a system of rules, compliance with which is nonvoluntary, see Thomas Nagel, “The Problem of Global Justice,” *Philosophy & Public Affairs* 33, no. 2 (2005): 113–147. See Andrea Sangiovanni, “Global Justice, Reciprocity, and

A common cosmopolitan response to this argument is to marshal empirical information to show that international institutions are also coercive or nonvoluntary in the ways that require egalitarian principles of justice. The argument runs the same as above. International institutions are coercively imposed or compliance with the rules is otherwise nonvoluntary; coercive or otherwise nonvoluntary institutions must meet a stringent standard of justification, namely, that the principles that regulate the constitution of such institutions not be reasonably rejected by those who are subject to the institutions; the global poor can reasonably reject principles that leave them relatively deprived vis-à-vis the global rich; thus, the constitution of international institutions must be regulated by egalitarian principles.¹⁷

Suppose we grant the cosmopolitan's classification scheme: domestic and international institutions are similarly coercive or otherwise nonvoluntary in the relevant ways. Without showing how domestic and international institutions figure into causal accounts of economic and political development, this taxonomic information is insufficient to license the cosmopolitan's conclusion. Suppose domestic institutions mediate the developmental effect of international institutions such as the World Trade Organization. More specifically, suppose countries with strong domestic institutions — institutions that constrain state leaders to be responsive to citizens' demands for economic growth — are able to capture the developmental gains from international trade facilitated by WTO rules, whereas countries with weak institutions fail to benefit or are even harmed by WTO rules.¹⁸ The following two mechanisms can plausibly explain this variation. First, strong institutions constrain leaders to provide good trade policy, whereas weak institutions permit leaders to implement trade policies that enable a small elite to capture rents from trade but do not benefit

the State," *Philosophy & Public Affairs* 35, no. 1 (2007): 3–39 for a discussion of these “coercion-based” arguments and an alternative, “reciprocity-based” argument for statism.

17. See Beitz, *Political Theory and International Relations*, pt. 3; Moellendorf, *Cosmopolitan Justice*, chs. 2 and 3; Pogge, *Realizing Rawls*, ch. 6.

18. The statistical evidence adduced by Rodrik, Subramanian, and Trebbi, “Institutions Rule” suggests such a story.

citizens in general. The other is that strong institutions constrain leaders to invest in economic sectors that are able to benefit from WTO rules, whereas weak institutions permit leaders to fail to invest at all or to invest in sectors that may benefit a small elite despite being unable to benefit from WTO rules. If such mechanisms were in operation, then underdevelopment occurs not because poor states are unable to extract enough resources from rich states, but because poor states' institutions grant leaders discretionary control over the flow of resources into the country and permit them to use these resources without regard for citizens' interests. If this causal analysis were true (I have made no claim that it is), what would be the effect of, for example, increasing poor states' bargaining leverage vis-à-vis rich states in negotiating WTO rules?¹⁹ Without fundamentally altering domestic institutions, restructuring international institutions in this way would permit the leaders of poor states to control a greater share of global resources without having much of an effect on global inequality. Indeed, holding domestic institutional quality fixed, making international institutions more egalitarian could *exacerbate* underdevelopment rather than mitigate it.

It might be the case that domestic institutions are profoundly shaped by international institutions.²⁰ At the limit, we might find that international institutions determine domestic institutional quality and that more egalitarian international institutions produce stronger domestic institutions that hold state leaders more accountable. Under these conditions, the developmental effect of international institutions is such that we would be inclined to say that egalitarian principles apply to international institutions in the first instance, since this subsequently determines whether domestic institutions bring about outcomes that satisfy

19. See Moellendorf, "World Trade Organization" for such a proposal.

20. As Pogge has so frequently argued; see, e.g., Pogge, *World Poverty and Human Rights*. Joshua Cohen is skeptical about the truth of this claim, as am I; see Joshua Cohen, "Philosophy, Social Science, Global Poverty," in *Thomas Pogge and His Critics*, ed. Alison M. Jaggar (Malden, MA: Polity Press, 2010). David Miller is skeptical about its relevance for allocating responsibility for global poverty; see David Miller, *National Responsibility and Global Justice* (New York: Oxford University Press, 2007), 240.

egalitarian principles.

My point here is not to adjudicate this dispute between statist and cosmopolitans; nor do I wish to defend any particular causal story at this point. (I stake my claim in chapter 5.) Instead, my point is that simple taxonomic information is insufficient to demonstrate whether egalitarian principles properly constrain the design of either domestic or international institutions. To draw out the institutional implications of our commitment to the moral equality of individuals, we must analyze the causal consequences of different institutional arrangements.

2.2.3 Crude causal information

Perhaps the most telling case for my argument is one where the argument appeals to causal information but neglects to conduct causal analysis. Thomas Pogge's argument that the international resource privilege harms the global poor is just such a case.²¹ The story he wants to tell is straightforwardly causal. The resource privilege grants the power to effect legally valid resource transfers to any person or group holding a monopoly of effective force in a territory. This privilege generates incentives to acquire and wield power violently. It generates a stream of revenue for oppressive dictators, who can use that revenue to buy guns and political support to keep themselves in power, even against widespread opposition. Even well-intentioned reformist rulers, if they want to stay in office, are compelled to divert resource revenue to the private fortunes of the political elite or to keeping potential rebels at bay. This incentive structure is supposed to explain the widely-noted negative correlation between aggregate economic growth and high levels of resource wealth (as a percentage of GDP). According to Pogge, the resource privilege partly explains this so-called "economic resource curse" and its associated deprivation.

21. The most cogent formulation — the one recounted here — is found in Pogge, *World Poverty and Human Rights*, 113–114, 163–165. I say more about Pogge's discussion of the resource privilege in chapter 6.

Exactly how the resource privilege explains the resource curse is never made clear by Pogge. This is because, despite employing causal claims throughout, Pogge’s argument fails to specify the operative causal mechanisms at crucial points of the argument. As far as I can tell, the argument goes something like this.

1. *The economic resource curse*: Resource abundance causes low (or negative) aggregate economic growth rates.
2. Resource abundance encourages civil conflict and authoritarian rule — resource rents constitute a big prize for anyone who can gain control of the state apparatus and the revenue stream enables dictators to buy enough guns and political support to put down leadership challenges.²²
3. The international resource privilege explains the effect of resource abundance on the incidence of civil conflict and authoritarian rule — the resource privilege incentivizes violent acquisition of power and enables authoritarians to use force to consolidate their rule.
4. Resource-induced civil conflict and authoritarian rule reduce aggregate economic growth.
5. Therefore, from 3 and 4, the resource privilege explains the economic resource curse.

The argument’s failure to identify specific causal mechanisms in 1 and 4 is problematic, not least because Pogge’s insistence that the resource privilege harms the poor is based on his claim that the resource privilege causes the deprivation associated with the economic resource curse.²³ There are causal explanations of the relevant phenomena on offer.²⁴ But Pogge never concerns himself with sketching, or even citing an account of the operative causal mechanisms. Such an account is what Pogge needs to infer 5 from 1–4. Whether

22. As evidence, Pogge cites Paul Collier and Anke Hoeffler, “On Economic Causes of Civil War,” *Oxford Economic Papers* 50 (1998): 563–573; Michael L. Ross et al., “Does Resource Wealth Cause Authoritarian Rule?” (University of California, Los Angeles, unpublished manuscript, 2000); Ricky Lam and Leonard Wantchekon, “Political Dutch Disease” (Yale University, unpublished manuscript, 2002); and Leonard Wantchekon, “Why do Resource Abundant Countries Have Authoritarian Governments?” *Journal of African Finance and Economic Development* 5, no. 2 (2002): 57–77.

23. There’s another problem with the argument, viz., that 3 is entirely speculative. As Joshua Cohen notes, for 3, and thus 5 to hold, it must be true that domestic institutional quality is sufficiently sensitive to changes in the relevant global rules (the resource privilege in this case) (see Cohen, “Philosophy, Social Science, and Global Poverty”). Pogge offers no evidence to support this. *Pace* Pogge, there is good reason to think that changes to the resource privilege would *not* break the resource curse. I discuss this at length in chapter 6.

24. See the discussion in chapter 5.

and how economic stagnation is causally linked to resource wealth and resource-induced civil conflict and authoritarian rule matters for whether the resource privilege — and those of us who are alleged to uphold it — is responsible for generating the severe poverty Pogge wants to attribute to it.²⁵ For 5 to follow from 1–4, the causal logic that links resource wealth with civil conflict and authoritarianism must be of a piece with the causal logic that links resource wealth with reduced economic growth. This is certainly not satisfied if the two causal stories are incompatible, but neither is it satisfied if we merely show that the two causal stories are consistent. To wit, the “lucrative prize and money-for-guns” mechanism that Pogge uses to link resources with civil conflict and authoritarianism is consistent with at least three causal stories linking resources with reduced growth, only the third of which could plausibly connect the resource privilege to the resource curse.²⁶

- Resource exports could increase demand for a country’s currency, causing its foreign exchange rate to rise, thereby making its other exports uncompetitive and diminishing economic diversification, which is important for growth (the so-called “Dutch Disease”).²⁷
- Resource dependent economies might suffer from volatile commodity prices. During boom times, high prices facilitate government spending on expensive “white elephant projects,” which prove politically difficult to cut during subsequent busts. As a result, spending cuts most often target politically vulnerable projects, which results in underinvestment in development.
- Resource revenue might facilitate clientelism. Rather than attract political support by providing the public goods that foster economic development, political leaders can use resource revenues to provide key civilian leaders with private goods in exchange for political support, thereby diminishing investment in productive activity.

To substantiate Pogge’s claim that the resource privilege causes the deprivation associated

25. Of course, even if the resource privilege is not responsible for generating severe poverty, it might still be the case that those of us who uphold the resource privilege harm others by incentivizing civil conflict and authoritarian rule, both of which constitute harms. But concluding this requires getting a firm grip on the causes of the resource curse.

26. For an accessible overview of the following mechanisms, see Humphreys, Sachs, and Stiglitz, *Escaping the Resource Curse*, ch. 1.

27. Pogge often uses “Dutch Disease” and “resource curse” interchangeably, suggesting a failure to understand that Dutch Disease is but one potential explanation for the correlation between low economic growth and resource wealth. If Dutch Disease were the right explanation, it would be hard to see how the resource privilege is responsible for the curse.

with the resource curse, what we need is an explanation that incorporates both stories as part of a single, coherent, unified causal analysis. But Pogge has yet to meet that burden. Until he does, Pogge's argument gives us little reason to accept his claim that the resource privilege harms the poor.

The point of the foregoing survey is to catalog the different ways in which global political theorists typically incorporate empirical information and distinguish these typical (mis)uses from genuine causal analysis. The discussion in each case has also shown that philosophical arguments often go awry as a result of employing these different sorts of empirical information when causal analysis is required. But there remains the question of when philosophers are required to engage causal analyses. Political philosophers carry out multifarious tasks and it might be intuitively plausible that attention to causal mechanisms is important for only a subset of those tasks. The aim of the next section is to deny this claim.

2.3 When Is Causal Analysis Required?

For my purposes here, we can group the tasks undertaken by political philosophers under three general headings. *Normative political philosophy* concerns itself with questions about how our social world should be arranged. This includes: examining fundamental political principles and employing these principles to normatively evaluate social states of affairs; allocating moral responsibility for outcomes among political agents; and discerning the nature and stringency of the rights and duties agents have. *Applied political philosophy* concerns itself with investigating the concrete practical implications of our normative judgments. Thus construed, applied philosophy has a problem-solving impetus. This includes elucidating the practical content of our duties, as well as proposing measures to

resolve, avert, or mitigate morally problematic states of affairs. Finally, *conceptual political philosophy* concerns itself with analyzing and explicating key political concepts, such as autonomy, equality, representation, and self-determination. The question in this section is whether at least some arguments of each type require attention to causal mechanisms.

Notice that I am not here asking whether all particular arguments must attend to causal mechanisms. Nor am I asking whether there are only certain broad topics that require such attention. There is no general answer that can be given a priori to these questions. Whether an argument should attend to causal mechanisms depends on the particular claims it makes. In light of the survey in the last section, I would conjecture that philosophers should engage in causal analysis more often than not, and certainly more often than they currently do. In this section, I provide additional reasons for thinking that causal analysis should play a central role in political philosophical arguments by presenting another series of case studies. Whereas the survey in the last section was organized by type of information misuse, the following survey is organized by type of philosophical claim. Whereas my aim in the last section was to show how different types of empirical information cannot safely substitute for causal analysis, my aim in this section is to show that there is no general type of argument that can safely ignore causal mechanisms.

2.3.1 Normative political philosophy

Normative political philosophy primarily concerns itself with adjudicating among political principles and using these principles to assess our social world. Plausibly, many normative claims need not attend to causal mechanisms — whether individuals or collectives should be taken as the fundamental units of moral concern; whether individuals should be treated with equal respect; whether moral principles apply universally or are context-relative. Yet, normative political philosophy is not wholly unconstrained by

causal considerations. At the very least, whether we ought to pursue some moral objective M — for example, a distribution of goods according to our favored distributive principle, or the protection of some set of rights — at least depends on whether we can, in principle, realize M .²⁸ And whether M is in principle realizable requires knowing something about how agents like us might realize M in a world like ours. Moreover, assessing the desirability of M requires understanding whether realizing M enhances or detracts from our capacity to realize other important goals. For example, if the realization of global equality of opportunity requires the establishment of a world government and the latter would entail unacceptable costs (moral or otherwise), then we would be inclined to reconsider the desirability of global equality of opportunity as an action-guiding moral aim. Finally, determining who has which duties to whom will depend to some extent on causal considerations. Whether A can be justifiably morally condemned for harming B depends on how A 's conduct is causally related to B 's suffering. Whether A can be justifiably obliged to alleviate B 's suffering depends on A 's causal relationship to B 's suffering, as well as both A 's absolute capacity to alleviate B 's suffering — whether A can cause the reduction of B 's suffering — and A 's capacity to alleviate B 's suffering relative to C 's capacity to do so — whether A is better positioned than C to cause the reduction of B 's suffering. (To illustrate this last point, we might find that a dictator is morally condemnable for causing his citizens' suffering and capable of alleviating their suffering, but his unwillingness to do the latter might entail that those who are capable of doing so have a duty to aid the

28. Note that this is not a point about feasibility; it is more general in the following way. Let ϕ denote, e.g., *bring about global equality of opportunity*. If moral theory says we ought to ϕ , ϕ -ing can fail to be feasible at present without invalidating the judgment that we ought to ϕ ; if ϕ is presently infeasible, the judgment implies that we should take steps to overcome the practical barriers to ϕ -ing. However, if ϕ -ing were in principle impossible for some reason, that fact would violate the “ought implies can” principle and invalidate the judgment that we ought to ϕ . Of course, how we distinguish between what is ruled out in principle and what is merely infeasible depends on which features of the actual world we take as parametric. I am inclined to stay “close to the ground” in the sense of taking a greater number of, e.g., individuals' cognitive, rational, motivational, and moral limitations as parametric than some (perhaps most famously, Rawls). This is a complex issue; I discuss it in some detail in appendix B.

suffering citizens.) All this suggests that at least some normative claims must be attentive to causal mechanisms.

To illustrate my point here more fully, consider again the global equality of opportunity example. As Simon Caney construes it, global equality of opportunity (GEO) requires that “persons (of equal ability and motivation) have equal opportunities to attain an equal number of positions of a commensurate standard of living.”²⁹ Caney argues that our commitment to equality of opportunity is grounded on the conviction that it is unfair if the opportunities available to individuals are determined by morally arbitrary characteristics, characteristics that should not be taken as reasons to treat individuals differently. Paradigm examples here include social class or ethnicity. But national or civic identity are morally arbitrary in the same way that class or ethnicity are morally arbitrary. Thus, if it is unfair for individuals’ opportunities to be determined by class or ethnicity, it is also unfair for their opportunities to be determined by national or civic identity. This implies that GEO is a global principle of distributive justice.

Whether we should take GEO (among other principles³⁰) to govern the constitution of international institutions surely depends on the institutional implications of adopting such a principle as regulative.³¹ For a full assessment of GEO, we need to know (at least) which institutional mechanisms are required to realize GEO and the effect that such institutions would likely have on other important values, such as autonomy or self-determination. If the institutional mechanisms required to realize GEO unduly limit or undermine our

29. Simon Caney, “Cosmopolitan Justice and Equalizing Opportunities,” *Metaphilosophy* 32, nos. 1/2 (2001): 113–134, at 120. It’s important that the principle is put in terms of standard of living rather than the value individuals place on the opportunities. This formulation enables Caney to avoid straightforward versions of the objection that there is no cross-cultural way to assign value to opportunities; thus, GEO is not an attainable objective (see Miller, *National Responsibility and Global Justice*, 62–68). Other defences of a global equality of opportunity include Moellendorf, *Cosmopolitan Justice*; Thomas W. Pogge, “An Egalitarian Law of Peoples,” *Philosophy & Public Affairs* 23, no. 3 (1994): 195–224.

30. See Caney, *Justice Beyond Borders*, ch. 3, sec. 6.

31. Institution *I* is “implied” by a principle *P* when *I* is (causally) necessary to bring about realization of the outcome or procedure mandated by *P*.

capacity to pursue other objectives, then the desirability of GEO as a regulative principle is diminished.

Caney's positive cosmopolitan institutional proposal can be summarized by the following recommendations.³²

- Democratically-elected global and regional bodies that exercise authority over states (e.g., an elected assembly at the United Nations).
- Democratization of the World Trade Organization, International Monetary Fund, and World Bank.
- Creation of a new global economic institution to coordinate existing economic institutions such as the IMF and World Bank and insure that these organizations adhere to cosmopolitan principles. (Presumably, this new institution should be composed of elected officials, although Caney does not say.)
- A permanent UN volunteer military force to facilitate swift military response to humanitarian crimes within states.
- Expanding the jurisdiction of the International Court of Justice beyond interstate disputes, as well as making its jurisdiction mandatory.

Why should we adopt such an institutional framework? Caney claims that global institutions employ four types of mechanisms to realize cosmopolitan objectives, GEO in particular.³³

1. Global institutions “*enable agents to act on their just intentions*” by solving collective action problems, allocating responsibilities, and resolving jurisdictional disputes.³⁴
2. Global institutions *thwart agents' attempts to realize their unjust intentions* by enforcing compliance with cosmopolitan principles and allocating duties on a fair basis.
3. Global institutions “*empower vulnerable actors to protect their fundamental interests*” by establishing enforcement mechanisms that they can use to hold powerful actors accountable.³⁵
4. Global institutions “*transform actors' intentions*” by creating a common framework of rules and thereby “socializing” global agents.³⁶

32. See Caney, *Justice Beyond Borders*, 161–162. For further details, see Simon Caney, “Cosmopolitan Justice and Institutional Design: An Egalitarian Liberal Conception of Global Governance,” *Social Theory and Practice* 32, no. 4 (2006): sec. 2.

33. See *ibid.*, sec. 3.

34. *Ibid.*, 742, emphasis added.

35. *Ibid.*, 742.

36. *Ibid.*, 743.

Despite initial appearances, Caney eschews creation of a world state. Instead, he advocates an institutional scheme wherein the political units hold neither supreme nor comprehensive authority; no political unit holds final authority over any issue within a territory and no unit holds authority over all issues relevant to a territory.³⁷

How we should assess a commitment to GEO in light of this institutional proposal? Caney's discussion here — although more attentive to causal considerations than most — is too general to facilitate further assessment. He never really considers the institutional mechanisms required to enable global institutions to operate as he speculates they would. The most basic questions concern the organization of (coercive) power. To have the desired effects — to enable global institutions to solve collective action problems, enforce compliance, or transform agents' preferences — must (coercive) power be centralized at the global level, or can it remain decentralized? Must global institutions have mandatory jurisdiction, or is voluntary cooperation between states sufficient?

We can pose the problem as a dilemma. Consider Caney's second rationale for establishing a cosmopolitan institutional order — to enforce compliance. Either centralized authority is practically required to insure compliance with cosmopolitan principles or it is not. If it is, then Caney is advocating something closer to a world state than he admits and his proposal encounters familiar objections. Indeed, his proposal to establish a permanent global military force to police humanitarian crimes is evidence that he acknowledges some practical pressure to centralize rather than disperse law enforcement authority in the face of coordination problems and jurisdictional disputes. Sovereign states face severe coordination and jurisdictional problems when it comes to policing human rights; Caney's preferred way to overcome these problems is to establish a global agency with supreme

³⁷. Caney, *Justice Beyond Borders*, 163; see *ibid.*, 149–150 for Caney's discussion of supremacy and comprehensiveness.

authority to address humanitarian crimes.³⁸

If centralized authority is not practically required to ensure compliance — there is some evidence that it is not, at least on some issues³⁹ — then the need to secure compliance no longer provides a reason to create global political agencies that reside “above the state,” vested with enforcement authority. Voluntary cooperation among sovereign states should suffice. In fact, Caney’s discussion of compliance concedes this. When discussing available enforcement mechanisms, Caney points to the ability of international organizations to require compliance with certain rules as a condition for membership; this enables organizations to exchange membership benefits for compliance. Caney further appeals to Chayes and Chayes’ “managerial” model of international cooperation, which argues that non-compliance should be viewed as a management problem rather than an enforcement problem. On this model, states are assumed to have a propensity to comply; compliance problems can hence be adequately handled by establishing dispute resolution procedures, improving states’ capacity to comply through technical and financial assistance, and increasing the determinacy of treaty requirements.⁴⁰ Suppose the managerial model is accurate.⁴¹ Then there is no need to establish authoritative suprapstate institutions; voluntary

38. For objections to a world state, see Immanuel Kant, *Toward Perpetual Peace*, in *Practical Philosophy*, ed. and trans. Mary J. Gregor (Cambridge: Cambridge University Press, 1996), 311–352; John Rawls, *The Law of Peoples* (Cambridge, MA: Harvard University Press, 1999) provides a contemporary example. Notice that the objection that a world state would be tyrannical does not require that it hold supreme authority in all issue areas, only that it hold supreme authority over the legitimate use of force; presumably the latter is required to effectively insure compliance with law. But see Cabrera, *Political Theory of Global Justice*, ch. 6 for a defence against the objection that a world state would be tyrannical.

39. See Barbara Koremenos, Charles Lipson, and Duncan Snidal, “The Rational Design of International Institutions,” *International Organization* 55, no. 4 (2001): 761–799, as well as other essays in the same journal issue.

40. Abram Chayes and Antonia H. Chayes, “On Compliance,” *International Organization* 47, no. 2 (1993): 175–205.

41. Downs, Rocke, and Barsoom persuasively argue that when we observe high treaty compliance rates in the absence of coercive enforcement mechanisms, this is because the treaty does not require states to deviate much from what they would have done in the absence of the treaty. In short, high compliance indicates relatively shallow levels of cooperation. When deep cooperation is required — when the gains from defection are high — coercive enforcement mechanisms must be established to secure compliance. See George W. Downs, David M. Rocke, and Peter N. Barsoom, “Is the Good News about Compliance Good News about Cooperation?” *International Organization* 50, no. 3 (1996): 379–406.

cooperation among states is sufficient to handle management problems.

(Notice that we can press similar dilemmas with respect to Caney's first and third mechanisms. Either centralized authority is required to overcome collective action problems or reliably check powerful states, or it is not. If it is, Caney advocates something close a world state, which diminishes his proposal's appeal. If it is not, then overcoming collective action problems or checking states' power no longer provide reasons to establish a cosmopolitan institutional order.)

Let's return to the central point of this section: at least some arguments for normative claims must attend to causal mechanisms. To illustrate this point, I have shown that a prominent argument for adopting global equality of opportunity as a principle of distributive justice is incomplete because it fails to consider the institutional implications of a commitment to GEO. This is important because our judgments about the desirability of the institutions implied by a commitment to GEO inform our judgments about the desirability of endorsing GEO as a regulative principle. In examining Caney's positive institutional proposal, I have not argued that we should reject GEO as a regulative principle. Instead, I have shown that Caney's case for a cosmopolitan institutional scheme — which is presumably motivated by a desire to realize outcomes that conform to GEO, among other principles — neglects important tensions in his own thought because he fails to undertake a causal analysis of the institutional mechanisms on which his case relies. Were he to do so, he might find the appeal of his own institutional proposal diminished, thereby weakening his case for adopting GEO as a principle of distributive justice.

As an important sidepoint, note that my argument here shows that causal analysis is important not only for nonideal theory — that is, theorizing about what to do under conditions of persistent injustice — but also for ideal theory.⁴² Caney is definitely doing

42. I discuss the ideal-nonideal theory distinction at greater length in chapter 3.

ideal theory; he means GEO to be among the regulative principles for a fully just global institutional order. But, as the preceding shows, doing ideal theory does not free us from having to undertake a full institutional analysis, which includes examining the operation of the institutional mechanisms that would be in place under ideal conditions as well as attempting to discern the effects ideal institutions would have (or would be likely to have) under the conditions in which they are intended to operate.

2.3.2 Applied political philosophy

Applied political philosophy concerns itself with the concrete practical implications of our normative political principles, to “put these principles to work in the world,” so to speak. Applied prescriptions aim to tell us which courses of action we have most reason to take in light of our normative judgments about actual states of affairs and our reasoning about how to address social problems. Accordingly, our reasoning about concrete prescriptions is properly constrained by non-moral considerations, including feasibility, stability, and efficiency considerations. Whether we have reason to undertake a particular course of action or implement a particular policy or institutional reform will depend on the causal consequences of that action, policy, or reform; in particular, whether the action, policy, or reform in question is likely to actually bring about the state of affairs we seek to realize. This suggests that attention to causal mechanisms is key for applied work.

To illustrate this point, consider Peter Singer’s work on global poverty.⁴³ For nearly 40 years, Singer has argued that the residents of affluent developed countries have a moral duty to direct a substantial portion of their income to global poverty relief and advocated

43. See, e.g., Singer, “Famine, Affluence, and Morality”; Peter Singer, *One World: The Ethics of Globalization* (New Haven and London: Yale University Press, 2002); Peter Singer, *The Life You Can Save: Acting Now To End World Poverty* (New York: Random House, 2009). Contrary to my thesis here, Singer expresses skepticism regarding the extent to which social sciences like political economy and political sociology “offer a relevant ‘body of knowledge’” for political philosophers (Peter Singer, “Achieving the Best Outcome: Final Rejoinder,” *Ethics & International Affairs* 16, no. 2 (2002): 127–128, at 127).

humanitarian aid as an effective concrete means to discharging that duty. Singer's argument is motivated by his oft-repeated "drowning child" thought experiment: You're on your way to an important meeting wearing a nice suit and shoes; you see a child drowning in a shallow pond; the child's life is vastly more important than your suit and shoes; you ought to wade into the pond and save the child. A key feature of this scenario is that we do not know how the child came to be drowning in the pond in the first place. Such information might be unnecessary when determining what to do in this case because it is immediately obvious that the way to save the child is to pull her out of the pond.⁴⁴ But global poverty is not nearly so easy to figure out. This is where the analogy to the drowning child breaks down. The drowning child scenario is an isolated emergency; most cases of severe deprivation are chronic. Isolated emergencies are relatively easy to solve: if possible, remove the immediate threat.⁴⁵ Chronic cases are embedded in complex causal webs involving not only physical and geographical factors (as in the drowning child case), but also important political, cultural, and economic factors that interact with each other in complicated ways (unlike the drowning child case). Individual cases of severe deprivation are thereby much more difficult to diagnose and solve. If we are to figure out who has duties to relieve poverty, we need to know how people became impoverished. If we are to determine the content of such duties, again, it matters how people became impoverished. This is why causal analysis is key, but Singer offers us no such analysis. Without a diagnosis, we cannot know whether Singer's prescribed solution intervenes at the right place

44. But maybe not. What if the child's trip into the pond resulted in a neck injury? Then we would want to exercise appropriate caution instead of simply yanking the child from the pond. The point here is that, even in Singer's toy example, the specifics about what to do are not as obvious as Singer implies.

45. As Jennifer Rubenstein notes, emergencies often have identifiable "event-like" features, meaning that emergencies are often precipitated by shocks, be they environmental (e.g., an earthquake or flood), economic (e.g., the 1997–98 Asian financial crisis), physical (e.g., armed conflict), or political (e.g., a coup). See Jennifer Rubenstein, "Distribution and Emergency," *The Journal of Political Philosophy* 15, no. 3 (2007): 296–320. This definition doesn't deny that the shocks themselves might have long, complicated causal stories; it simply states that shocks are the proximate causes of emergency outcomes. In light of this, the claim "remove the immediate threat" amounts to saying "mitigate the consequences of the identifiable shock".

in the causal process (if at all), whether it exacerbates poverty, or whether it generates new problems altogether.⁴⁶

Relatedly, Singer never really engages the debate about the efficacy of humanitarian aid programs as a means to relieving global poverty.⁴⁷ His general answer to the efficacy question is two-fold. First, if there are more effective means than humanitarian aid, then we should participate in those too; the more general conclusion is that affluent people should be making significant sacrifices to relieve global poverty. But, second, there is no intuitively plausible reason to think that humanitarian aid is ineffective (even if not the most effective) for relieving poverty, so (other things equal) we should be giving to humanitarian organizations.⁴⁸ Singer does posit that humanitarian aid is ineffective because the allocation of official aid dollars is dominated by irrelevant political considerations.⁴⁹ He then claims that if aid dollars were allocated according to the morally relevant criteria, aid would be effective. Rather than give up on humanitarian aid, we should reform aid allocation to make aid more effective. But Singer overlooks numerous reasons why his proposal might not work and so ignores the evidence either way. One prominent reason is that current official aid practices might actually be effective at achieving their primary aim, which is not to relieve poverty but to advance participants' strategic interests.⁵⁰ If

46. Nevermind the fact that Singer's analogy ignores the thorny issues regarding the victims' role in bringing about their impoverished conditions. As Dale Jamieson notes, many of the victims of the Ethiopian famine of 1983–1985 were involved in a civil war, and many of the Hutus dying in refugee camps in the Eastern Congo in the mid-90s participated in the 1994 genocide of Rwandan Tutsis. See Dale Jamieson, "Duties to the Distant: Aid, Assistance, and Intervention in the Developing World," *Journal of Ethics* 9 (2005): 151–170, at 156. The point is not that we should leave those responsible for bringing about their suffering to their fate, but that we may be unwittingly (and perhaps unjustifiably) taking sides in a much larger issue if we fail to understand the causal process that generated the suffering.

47. For an accessible introduction to the relevant issues, see Jeffrey D. Sachs, *The End of Poverty: Economic Possibilities for Our Time* (New York: Penguin, 2005) on the pro-aid side and William Easterly, *The White Man's Burden: Why the West's Efforts to Aid Have Done So Much Ill and So Little Good* (New York: Penguin, 2006) on the critical side.

48. This reply is expressed most clearly in Singer's later work. See, esp., Peter Singer, "Poverty, Facts, and Political Philosophies: A Reply to 'More Than Charity'," *Ethics and International Affairs* 16, no. 2 (2002): 121–124; Singer, "Achieving the Best Outcome."

49. Singer, *One World*, ch. 5.

50. Cf. Jamieson, "Duties to the Distant," 161. For a model of aid-for-policy transactions that supports

aid was never meant to relieve poverty in the first place, there will be little incentive to reform the system to insure that it does so more effectively. The general problem here is that Singer prescribes an intervention without doing due diligence on the efficacy question, which requires rigorous analysis of the causal effects of foreign aid as well as the ways in which aid causally interacts with successful development trajectories.

The main lesson is this: we generally require doctors to offer a diagnosis of the problem and assess the efficacy of the prescribed intervention before we place any confidence in their prescriptions. Insofar as they are keen to prescribe solutions to moral problems, we should require as much from applied political philosophers too.

2.3.3 Conceptual political philosophy

Conceptual political philosophy specifies the content of political concepts, such as equality, representation, and self-determination; it also specifies the rules for determining which objects fall into the concept's extension. Sometimes this conceptual work is done for its own sake; often, it is interwoven with and undertaken in service of normative or applied objectives. In any case, this conceptual work aims to help us "get a grip" on the social world. Concepts constitute the framework within which we think and talk about the world. Conceptual political philosophy thus aims to construct a framework that facilitates our attempts to understand the world. So conceived, we seek a tight connection between our concepts and the world so as to improve our grasp of the latter; we want our concepts to "carve the world at its joints," so to speak. This suggests that our examination of political concepts would do well to attend to the causal structure of the social world, to be sensitive to the ways in which the things in the world to which our concepts refer are causally related to other things in the world.

this conclusion, see Bruce Bueno de Mesquita and Alastair Smith, "Foreign Aid and Policy Concessions," *Journal of Conflict Resolution* 51, no. 2 (2007): 251–284; Bruce Bueno de Mesquita and Alastair Smith, "A Political Economy of Aid," *International Organization* 63, no. 2 (2009): 309–340.

Consider David Miller's examination of the concept of national responsibility.⁵¹ Miller argues that developed countries and their residents are not required as a matter of global justice to relieve poverty abroad when responsibility for the poverty falls to a subgroup in society or to the society as a whole. Thus, *national responsibility* is key for his argument.

On Miller's view, a group *G* is collectively *outcome responsible* for a state of affairs *S* if the gains or losses associated with *S* properly accrue to a sufficient proportion of the members of *G*.⁵² An individual *I* shares in *G*'s outcome responsibility under one of two conditions. In the *like-minded group* model, *I* shares in *G*'s outcome responsibility for *S* insofar as *I* provides support for a collective activity that leads to *S*. *I* need not causally contribute directly to *S*; for example, *I* need not lob a hand grenade to share in *G*'s outcome responsibility for the razing of a village. *I* need not even be present at the village. All that is required is that the members of *G* "share aims and outlooks... [and] *recognize* their like-mindedness"; further, "individual members act... in the light of the support they are receiving from other members of the group."⁵³ Importantly, *I* can provide such support by merely participating in the community, even if *I* opposes the actions.⁵⁴

In the *collective practice* model, *I* shares in *G*'s outcome responsibility for *S* if *I* benefits from a cooperative practice that results in *S* and that practice treats *I* fairly. This will be true when, for example, *I* is given a fair chance to influence the rules governing the practice or to shape *G*'s decision with respect to the action that led to *S*.

How can *I* avoid sharing in *G*'s outcome responsibility for *S*? For Miller, *I* "must take all reasonable steps to prevent the outcome occurring."⁵⁵ Here, "reasonable" is determined by the expected harm of the outcome and *I*'s expected cost of dissenting. But

51. Miller, *National Responsibility and Global Justice*.

52. See *ibid.*, 87 for further discussion of outcome responsibility.

53. *Ibid.*, 117.

54. See *ibid.*, 118.

55. *Ibid.*, 121.

what Miller makes clear is that mere abstention from the activity or practice in question or simply voicing opposition — even voting against undertaking the action — is not sufficient to avoid sharing in the group’s responsibility.

The final step of the analysis is to show that the concept of “collective responsibility” applies to nations. Miller defines a nation as a group whose members share a common identity and public culture, who recognize special obligations to fellow members, who aspire to be self-determining, and for whom persistence of the group constitutes a good. Nations can be said to act collectively in at least two senses. First, the group’s practices and individual members’ conduct are expressions of their common identity and culture; second, the political decisions will “embody to a greater or lesser extent the articulated beliefs and attitudes of the nation in question.”⁵⁶ To the extent that a nation is self-determining — that is, not ruled by a colonizing power or by an authoritarian dictator — the nation can be said to exhibit features of the like-minded group model. But there are some cases where the like-minded group model applies even if the nation is ruled by an autocrat. This will be the case if an autocratic leader’s policies derive from beliefs and attitudes that are aligned with the public culture at large. Under these circumstances, national responsibility will be shared by individual members, even when they simply voice dissent. To the extent that the individual members benefit from and are treated fairly by collective decision-making procedures or distributive practices, the nation will exhibit features of the cooperative practice model.

How can attention to causal mechanisms improve Miller’s analysis? Let’s focus on Miller’s claim that passive forms of resistance such as abstention or casting a dissenting vote are insufficient to absolve an individual *I* from sharing in a nation *N*’s collective outcome responsibility for its underdevelopment. *I* expresses tacit consent to *N*’s de-

56. Miller, *National Responsibility and Global Justice*, 126.

velopment policy when I fails to take “all reasonable steps” to avert underdevelopment. Whether we should accept this claim must surely depend on I ’s capacity to influence N ’s development policy by taking “all reasonable steps” (beyond mere dissent) to prevent underdevelopment.

One way to put this is in game theoretic terms. Suppose I receives a payoff of U from the status quo policy, which leads to underdevelopment, and a payoff of D from an alternate policy that leads to successful development, with $U < D$. Suppose further that I incurs a cost $C > 0$ from actively opposing the status quo policy, which will be effective with some probability p . Under these conditions, I will only undertake action to oppose N ’s policy if she expects to receive a net payoff from opposing the status quo that is at least as much as the net payoff she expects from not opposing the status quo. In symbols, I opposes the status quo if

$$pD + (1 - p)U - C \geq U,$$

which simplifies to

$$D - \frac{1}{p}C \geq U. \tag{2.1}$$

From (2.1), we see that the likelihood that I actively opposes the status quo policy depends on the probability that I ’s opposition will successfully influence policy in a positive direction and I ’s cost for undertaking such action. Holding D and U fixed, we see that I becomes less likely to actively oppose the status quo as p decreases or C increases.

The foregoing shows that I ’s failure to actively resist N ’s status quo policy need not imply consent of any kind to the policy, tacit or otherwise. Instead, under these conditions, we might more accurately interpret I ’s failure to resist as *resignation* to her lack of capacity to influence policy. To the extent that renouncing communal norms is costly to

I — for example, because *I* is relegated to the margins of the community — *I*'s continuing participation in the community will be a rational response to her circumstances despite her (potentially vehement) disagreement. This complicates the claim that *I* can be held responsible for underdevelopment if the relevant policy was chosen on the basis of beliefs and attitudes aligned with “the public culture.” *I* might prefer to alter the public culture with which the policy is aligned but recognizes her own incapacity to do so. Moreover, such incapacity might be sustained by collective action problems. Suppose it is true that the number of people who are opposed to the status quo policy or the public culture is sufficient to change the policy or culture through concerted collective action. Nevertheless, dissidents might not be able to come together to effect the desired changes because they are not aware that there are a sufficient number of dissidents in their midst,⁵⁷ or even if everyone knows that a sufficient number of dissidents exist, incentives might be structured so as to generate a free rider problem.

Many of the conditions under which individuals will be tempted to resign themselves to their deprived circumstances hold in underdeveloped countries. The global poor are typically unable to hold their political leaders to account, electorally or otherwise, and the government's capacity to suppress dissent makes it costly to oppose government policy. As a result, many of the global poor are often observed to passively accept their lot. If the mechanism at work in the model more or less accurately depicts one of the mechanisms at work in the actual world, then this passive acceptance should not necessarily be interpreted as tacit consent, as Miller is inclined to do, but might signal resignation. Whether we can properly allocate a share of collective outcome responsibility to the global poor for their own deprivation depends in part on the mechanisms that explain their observed passivity

57. Kuran explores the phenomenon of “preference falsification,” where people publicly “falsify” their genuine preferences because publicly revealing the latter is believed to be too costly. See Timur Kuran, *Private Truths, Public Lies: The Social Consequences of Preference Falsification* (Cambridge, MA: Harvard University Press, 1995).

in the face of that deprivation.

Of course, the foregoing considerations are not sufficient to ground any particular claims about an individual's share of her nation's collective responsibility for its underdevelopment, nor any claims about the global affluent's (lack of) responsibility to undertake actions to mitigate underdevelopment, for that matter. We sometimes do require that individuals undertake risky actions against their own rational interest. We must also ask whether the poor individual or the global rich are in a position to stimulate collective action that could overturn the deleterious development policy, and whether the affluent are in a position to help alter development outcomes. So simple game-theoretic reasoning is not enough to show that the individual or the affluent avoid responsibility or condemnation.⁵⁸ But it does show that Miller cannot straightforwardly reason from observed behavior to an analysis of national responsibility, nor to attributions of individual responsibility.

The upshot here is that any specification of the concept of national responsibility must be attentive to the bargaining dynamic from which group decisions arise and individual members' location in that dynamic vis-à-vis the group's power brokers. Failing to attend to the structure of this dynamic means that our specification will be overly reliant on the behavior we observe. But, as my simple model shows, observed behavior often belies important features of the situation. To determine whether *I* is properly held responsible for *N*'s underdevelopment, we must know not only what *I* does under actual conditions, but what *I would do* were conditions different. In game theoretic terms, we need to discern what happens "off the equilibrium path," not just the behavior we observe in equilibrium. Such counterfactual reasoning is constitutive of causal analysis. Since causal mechanisms are supposed to be general causal patterns that explain the causal link between initial conditions and outcomes, engaging in causal analysis prompts us to attend to counterfactual

58. Thanks to Elizabeth Anderson for pressing this point.

scenarios in search of causal mechanisms. It also enables us to make hypotheses about counterfactual scenarios. I take both of these advantages to show how causal analysis can enrich our specification of political concepts such as national responsibility.

2.4 Practical Implications

The general line I have been pushing is that analysis of social causal mechanisms is crucial for a wide range of arguments in global political theory. Our moral analysis and subsequent prescriptions are only as good as our understanding of the causal structure of our social world. We need to understand both the causal process generating the outcomes about which we are concerned and the causal consequences of potential remedies.

The preceding points are likely obvious to many when stated in the abstract. But the foregoing surveys show that political philosophers run afoul of them often enough to warrant explicitly highlighting the importance of causal analysis. Moreover, previous discussions that have sought to motivate more effective interaction between philosophy and the social sciences have located the problems elsewhere—for example, reliance on insufficient data⁵⁹ or appeals to the wrong kind of data.⁶⁰ To be sure, data problems exist. But effective incorporation of empirical information does not merely adduce enough of the right kind of data. Ultimately, effective political philosophy relies on *credible explanations of the data*. Compelling political philosophy must proceed from—be informed and constrained by—compelling social scientific theorizing about causal mechanisms.

This is not to say that political philosophers should become full-time social scientists. We divide the labor for a reason—there are gains to be had from specialization. Normative and conceptual considerations have their rightful place and philosophers are well-placed to

59. Leif Wenar, “What We Owe to Distant Others,” *Politics, Philosophy & Economics* 2, no. 3 (2003): 283–304.

60. Nicole Hassoun, “Making the Case for Foreign Aid,” *Public Affairs Quarterly* 24, no. 1 (2010): 1–20.

give them their due. But morally evaluating states of affairs requires knowing something about their causal antecedents; making effective prescriptions requires understanding the causal logic of the situations we seek to redress. Such insight does not require that philosophers produce novel explanations, but it does require analytically taking apart the causal process and scrutinizing the role played by each of the moving parts, as well as their interconnections. Doing this with adequate rigor requires critically engaging social science *on its own terms*.⁶¹

Several important methodological questions remain. Should a focus on causal mechanisms replace political philosophers' traditional focus on normative issues? If not, how should global political theorists integrate causal analysis with normative analysis? Where does causal analysis fit in a coherent approach to morally progressive institutional design? I address these questions in the next chapter.

61. I stress "on its own terms" because philosophers sometimes lapse into meta-level discussions about the limits of social scientific research methods when what is called for is ground-level criticism of an explanation's capacity to account for the relevant data. For example, when considering the limitations of the institutional view of development he favors, Risse mentions philosophical worries about the capacity of social science to predict the likely consequences of future actions on the basis of observations of the past and the difficulty of applying cross-country statistical evidence to the development policy of specific countries (Risse, "What We Owe"). These are certainly issues, but they are irrelevant given that Risse's objective in the section in question is to pick out one view of development from among several as able to best account for the data. What Risse should be doing here is adjudicating between competing explanations of development instead of raising worries that apply to *any* social scientific explanation. In picking out the institutional view as providing grounds for a particular duty of assistance, Risse presupposes favorable answers to his philosophical worries.

CHAPTER 3

Institutional Failure Analysis

Much recent work in global political theory prescribes ways to design the international institutional system so as to eradicate or at least mitigate severe and systematic deprivation.¹ The conventional wisdom is that such prescriptions are the province of nonideal theory but that ideal theory is required as a guide for nonideal theory. Rawls captures the conventional wisdom well:

Nonideal theory asks how [a] long-term goal might be achieved, or worked toward, usually in gradual steps. It looks for policies and courses of action that are morally permissible and politically possible as well as likely to be effective. So conceived, nonideal theory presupposes that ideal theory is already on hand. *For until the ideal is identified*, at least in outline — and that is all we should expect — *nonideal theory lacks an objective, an aim, by reference to which its queries can be answered.*²

On this view, our first task is to specify fully just *principles of regulation*, those principles that regulate the constitution of fully just institutional arrangements.³ *Nonideal theory* seeks to work out the application of these just principles of regulation to conditions of actual injustice. Call this the *ideal guidance approach*.

1. This chapter is forthcoming in *The Journal of Political Philosophy* under the title “Prescribing Institutions Without Ideal Theory.” It is reproduced here with minor revisions by permission of Blackwell Publishers.

2. Rawls, *The Law of Peoples*, 89f; emphasis added. Cf. Rawls, *A Theory of Justice*, 7–8, 215–216.

3. My use of the term “principle of regulation” is due to Cohen, “Facts and Principles,” 241. For Cohen, principles of regulation are distinguished from *basic principles*, the latter of which are statements of our fundamental normative convictions. Basic principles are the standard by which we evaluate the consequences of different principles of regulation.

The ideal guidance approach uses idealizing assumptions to identify just principles of regulation and then proceeds to apply these moral principles to the practical question of institutional design. Detailed analyses of the causal processes required to implement the the prescribed institutions play a peripheral role at best. Simply put, granting ideal theory priority in our evaluation of institutions emphasizes the moral aspects of institutional design at the expense of causal analysis. This is mistaken. The preceding chapter showed that causal analysis must play a central role in our theorizing about global institutional design. We need a new approach to institutional design that gives causal analysis a central role.

Recent debate on the relationship of ideal theory to nonideal theory is unhelpful for characterizing an appropriate alternative to the conventional wisdom. On the one hand, those who reject the notion that ideal theory is required as a guide for nonideal theory reject ideal theory altogether as “normatively useless” without offering any alternative method for prescribing institutional reforms.⁴ On the other hand, those who defend ideal theory against this charge claim that ideal theory *properly understood* is still useful as a guide for nonideal theory despite the limitations noted by critics, in part because there appears to be no other way to go about prescribing morally progressive institutional reforms.⁵

4. See Raymond Geuss, *Philosophy and Real Politics* (Princeton and Oxford: Princeton University Press, 2008); Colin Farrelly, “Justice in Ideal Theory: A Refutation,” *Political Studies* 55, no. 4 (2007): 844–864; Charles W. Mills, “‘Ideal Theory’ as Ideology,” *Hypatia* 20, no. 3 (2005): 165–184. Amartya Sen is a partial exception here. He joins the others in rejecting ideal theory — or “transcendental” theory, as he calls it — as being neither necessary nor sufficient for nonideal purposes, but goes beyond the others in offering a vaguely developed “comparative” methodological alternative. See Amartya Sen, “What Do We Want From A Theory of Justice?” *Journal of Philosophy* 103, no. 5 (2006): 215–238 and Amartya Sen, *The Idea of Justice* (Cambridge, MA: Harvard University Press, 2009).

5. Ingrid Robeyns, “Ideal Theory in Theory and Practice,” *Social Theory and Practice* 34, no. 3 (2008): 341–362; Simmons, “Ideal and Nonideal Theory”; Zofia Stemplowska, “What’s Ideal About Ideal Theory?” *Social Theory and Practice* 34, no. 3 (2008): 319–340; Adam Swift, “The Value of Philosophy in Nonideal Circumstances,” *Social Theory and Practice* 34, no. 3 (2008): 363–387; Laura Valentini, “On the Apparent Paradox of Ideal Theory,” *Journal of Political Philosophy* 17, no. 3 (2009): 332–355.

Identifying a viable methodological alternative requires that we pay less attention to characterizing the ideal-nonideal distinction and focus squarely on the salient methodological question: *how should we approach the task of prescribing institutional solutions to address actual injustice?* The issue here is identifying the relevant inputs as well as the appropriate procedure for turning those inputs into institutional design prescriptions. Although this issue is clearly related to debate concerning the ideal-nonideal theory distinction, I do not want to engage that debate directly. The main problem here is that “nonideal theory” is ambiguous between three different conceptions: (1) theorizing that identifies intermediate institutional reforms that help us transition from actual institutional arrangements to fully just institutional arrangements; (2) theorizing that identifies institutional arrangements that we should aspire to implement under actual conditions; and (3) theorizing that prescribes feasible institutional solutions to actual injustice.⁶ Given the multifarious philosophical controversies surrounding that distinction, retaining the term “nonideal theory” allows this ambiguity to persist and potentially obscures my central question.⁷ Hence, to maintain focus on the aforementioned methodological question, I will refer to theorizing that prescribes feasible institutional solutions to actual injustice as *clinical institutional theory*, or *clinical theory* for short. To reiterate: the central issue here is how, methodologically speaking, we should approach the task of clinical theory.

6. We might call these (1) transitional theory, (2) nonideal aspirational theory, and (3) clinical theory respectively. Failing to disambiguate “nonideal theory” has had the effect of conflating the various conceptions. So, e.g., it’s not uncommon to find philosophers who think that doing clinical theory amounts to doing transitional theory. Indeed, this is what the ideal guidance approach suggests. My discussion of Allen Buchanan in section 3.2 below illustrates this confusion.

7. Here’s a sample of the controversies I want to avoid: Is the difference between ideal and nonideal theory a matter of function, where the former specifies an ideal of justice and the latter specifies how to proceed in the here and now? (Relatedly, how are we to understand this function of nonideal theory given the aforementioned ambiguity?) Or is the distinction simply a matter of the former employing idealizing assumptions to derive principles of justice whereas the latter assumes a more realistic picture of the world? Related to the latter issue, is the difference a matter of the number of idealizing assumptions or the degree to which the assumptions employed are ideal? Is the difference between the two sorts of theorizing one of kind or degree? Settling these questions is beside the issue here.

Against the ideal guidance approach, I propose that clinical theorists should adopt an *institutional failure analysis approach*, which takes its design task to be, first and foremost, one of obviating or averting social failures (as opposed to approximating an ideal). The key innovation of this approach is to give causal analysis a central role in our evaluation of institutional arrangements. Failure analysis enables clinical theorists to prescribe more effective solutions to injustice because it focuses squarely on understanding the *problem*, rather than focusing on trying to understand an ideal of justice. In so doing, failure analysis better fulfills the objective of clinical theory, namely, to think about how, in the midst of current injustice, we might bring about social conditions that are more just than our current conditions.

3.1 Architecture and Engineering in Institutional Design

Institutions are, in Douglass North's familiar words, "the rules of the game in a society or, . . . the humanly devised constraints that shape human interaction."⁸ More specifically, institutions are sets of (formal or informal) rules that establish roles and stabilize behavioral norms and expectations for occupants of those roles; these norms and expectations subsequently regularize patterns of interaction among individual agents.⁹ Since social outcomes arise from the aggregation of particular interactions, institutions are important for shaping social outcomes and their consequences for individuals' lives. For this reason, institutions are morally significant. Of course, an institution's effect on outcomes is not deterministic. An institutional structure sets the range of possible outcomes and makes

8. Douglass C. North, *Institutions, Institutional Change and Economic Performance* (Cambridge: Cambridge University Press, 1990), 3.

9. Note that the inclusion of informal rules somewhat broadens the definition of an institutional structure beyond the familiar Rawlsian focus on the "basic structure of society." Rawls limited his consideration of institutional structures to the "major social institutions," which comprise "the political constitution and the principal economic and social arrangements" (Rawls, *A Theory of Justice*, 6). Given that informal institutional arrangements can have just as much impact on individuals' lives as formal institutions, I see no reason to restrict our attention to the latter.

some outcomes more probable than others. Thus, when assessing institutional arrangements, we don't only care about the observed outcomes actually realized by institutions, but also the range and likelihood of outcomes they make possible.¹⁰

To the extent that we make choices over alternative sets of rules to regulate interactions, we design institutions. The task of clinical institutional design comprises two distinct sets of problems: architectural problems and engineering problems. To get some traction on this distinction, consider it in reference to a material structure such as a skyscraper or a bridge. The architect's objective is to design a structure that creatively organizes its components using mass, space, form, texture, light, shadow, and so on, in a way that embodies some set of values (often functional, economic, artistic, and aesthetic values).¹¹ The engineer's task is to apply mathematical and scientific principles to solve technical problems involved in the design of the structure within the constraints set by physical, technological, economic, environmental, and ethical considerations.¹² To be sure, the two tasks intersect — there is a discipline called “architectural engineering” — but, on the whole, architects understand their task as more closely aligned with art whereas engineers uniformly understand themselves as doing “applied science.”

Analogously, the architectural problems comprised by institutional design concern the ways in which different configurations of institutional components embody different sets of values. Here the salient values are likely to be functional, economic, and (of primary importance to moral philosophers) moral values. The engineering problems comprised by institutional design concern the application of social scientific principles to the design of

10. Thanks to an anonymous reviewer for *The Journal of Political Philosophy* for bringing this point to my attention.

11. Cf. Francis D.K. Ching, *Architecture: Form, Space, and Order*, 3rd ed. (Hoboken, NJ: Wiley, 2007), ch. 1; Leland M. Roth, *Understanding Architecture: Its Elements, History, and Meaning* (Boulder: Westview Press, 1993), ch. 1; and Simon Unwin, *Analysing Architecture*, 2nd ed. (New York: Routledge, 2003), ch. 1.

12. Cf. Gerhard Pahl et al., *Engineering Design: A Systematic Approach*, 3rd ed., ed. Ken Wallace and Luciënne Blessing (London: Springer, 2007), ch. 1 and Mark T. Holtzapple and W. Dan Reece, *Foundations of Engineering*, 2nd ed. (New York: McGraw-Hill, 2003), ch. 1.

institutions with an eye to making them capable of withstanding the pressures to which they will be subject.

A key feature of the analogy is that the architectural and engineering problematics are mutually constraining as a result of their interdependence. A generic example from the world of residential construction serves to illustrate the point. For any particular house we build, the architect gives us a set of drawings specifying the characteristics of the house — the location of the interior walls, the size of the window and door openings, the pitch and shape of the roof. For the most part, the architect confines herself to specifying the way the house is to *look*. When it comes to specifying the required structural features to make the architectural design work, the drawings are turned over to an engineer. Whereas the architect specifies the size of the openings, an engineer specifies the requirements for the “header” (a beam-like support placed above a window or door opening) to carry the load to be placed above the opening. The architect specifies the pitch and span of the roof, but it is a roof truss engineer who makes sure that the roof trusses are able to span the distance between load bearing points. If it is ever the case that a house can’t be made to work structurally as the architect designed it — for example, the openings are too wide to accommodate adequate load bearing support — the drawings are sent back to the architect for modification. So the engineer specifies the structural limits within which the architect’s design must work. But the engineer doesn’t have a *carte blanche* when it comes to devising structural solutions; these are to be consistent with the architect’s design aim as far as possible. If the architect’s drawings call for large, wide open living spaces, the engineer is not free to unilaterally add interior walls to accommodate a simpler roof truss design. Hence, the architect and engineer each set limits on the other’s set of possible solutions for achieving their objective.

One way to characterize where conventional clinical theorizing has gone awry is to

say that it has ignored the engineering issues comprised by the task of institutional design. Conforming to my earlier distinction, clinical theorizing has largely focused on moral analysis and paid insufficient attention (intentionally or not) to causal analysis. This is not to say that moral analysis is unimportant. Theorists have made some important progress in thinking about the sorts of moral values we want our institutions to embody and how different institutional configurations might embody different values. But to make progress on the problem of designing institutions solutions for unjust conditions, we must make progress on both the architectural problems and the engineering problems. If the analogy I have drawn is apt, then progress on the architectural problem is insufficient. This is because a structure that effectively embodies a chosen set of values but is incapable of withstanding the pressures to which it is subject ceases to embody the chosen values once it ceases to exist as designed.

Focusing almost entirely on the architectural issues has had the unfortunate effect of recommending an institutional design procedure that limits the extent to which clinical theorists pay attention to the relevant engineering problems. Since philosophers have traditionally treated institutional design as a predominantly moral problem, their design projects have typically focused on the application of principles of justice to the design of actual institutions. This is why ideal theory has been thought to guide and constrain clinical theory. The idea is that we want our actual institutions to bring about more just states of affairs. The “more just” here intimates a target ideal according to which the justice of states of affairs can be measured, which suggests that clinical theory must seek ways to close the gap between current states of affairs and ideally just states of affairs. We require ideal theory to characterize a fully just institutional order, which serves as a *regulative ideal*, a guide to insure that our clinical theorizing is aimed in “the right direction.” If a candidate proposal violates one or more of the principles of regulation identified by ideal

theory, or if it does not deliver institutions that more closely approximate the ideal, it is discarded. This ideal guidance approach leads us to judge design proposals according to their fit with the principles that regulate ideal institutional arrangements. Accordingly, we acquire a tendency to ignore the engineering question: *how will the proposed institution fare in the face of the pressures to which it will be subject?* The issue here is not simply that clinical theorists are insufficiently attentive to important feasibility considerations. Feasibility considerations are but one set in the class of engineering considerations, which also includes stability, effectiveness, and efficiency considerations. Moreover, even if we simply pay increased attention to feasibility, the ideal guidance approach leads our consideration of this issue to be circumscribed by the principles of regulation identified by ideal theory. We surely want our designs to be feasible, but our approach to design leads us to be primarily concerned with the extent to which our designs comport with ideal principles of regulation. This leads us to give short shrift to the engineering problems comprised by institutional design and, hence, to causal analysis.

3.2 Unpacking The Ideal Guidance Approach: A Case Study

Allen Buchanan's proposal to reform the international practice of recognition among sovereign states nicely demonstrates the ideal guidance approach in action.¹³ Buchanan's proposal is an especially illuminating case study for two reasons. First, he clearly intends to prescribe a feasible solution to actual injustice.¹⁴ This is important because once we are assured that Buchanan *intends* his proposal to be so, we can ask whether he succeeds in prescribing reforms that are likely to "produce moral improvement in the particular system

13. Allen Buchanan, *Justice, Legitimacy, and Self-Determination: Moral Foundations for International Law* (New York: Oxford University Press, 2004).

14. "[W]e should eschew speculation about what constitutes a *comprehensive* set of ideal substantive institutional principles and concentrate on nonideal theory. . . . [W]e should focus on ascertaining which principles, if implemented, would produce moral improvements in the particular system that now exists" (*ibid.*, 67).

that now exists.”¹⁵ Second, he self-consciously adopts the ideal guidance approach:

The task of ideal theory is to set the most important and most distant moral targets for a better future, the ultimate standards for evaluating current international law. Nonideal theory’s task is to guide our efforts to approach those ultimate targets. . . .¹⁶

This results in clinical theorizing that is deliberately and assiduously constrained by ideal theoretic principles, which enables us to investigate the effect of this approach on Buchanan’s nonideal prescription.¹⁷

The structure of Buchanan’s book reflects his methodological orientation. Part one constitutes Buchanan’s ideal theory. This comprises two basic theses concerning the design of the ideal international system.

1. Institutions — in particular, the international legal system — must be designed to protect and promote *basic* human rights. In Buchanan’s words, basic rights are “rights that, if respected, protect those interests that are most crucial for a [sic] having a good human life.”¹⁸ These include a right to life, a right to physical security, a right against enslavement, a right to the means of subsistence, and a right against systematic racial, ethnic, or sexual discrimination.
2. International law ought to require nation-states to satisfy a minimal constitutional democracy condition.¹⁹

These theses undergird a justice-based conception of political legitimacy: a state exercises political power legitimately only if it respects basic human rights and is minimally

15. Buchanan, *Justice, Legitimacy, and Self-Determination*, 67.

16. *Ibid.*, 60f.

17. Buchanan’s characterization of nonideal theory here clearly resembles what I’ve called transitional theory. Given that I’ve suggested that we should distinguish between transitional and clinical theory, how can I now claim that Buchanan is doing clinical theory? The quote in footnote 14 suggests that Buchanan also conceives of his proposal as part of clinical theory. In light of footnote 6, my point here implies that Buchanan has failed to distinguish between the different conceptions of nonideal theory and, thus, identified the task of clinical theory with the task of transitional theory. Since this is in effect what the ideal guidance approach recommends, this is further evidence that Buchanan has adopted that approach.

18. Buchanan, *Justice, Legitimacy, and Self-Determination*, 129.

19. *Ibid.*, 142–147. The arguments presented to support this second thesis derive it from the first thesis. Thus, (2) is, strictly speaking, a derivative thesis. But, as Buchanan notes, ideal theory comprises not only basic principles of justice that are to be satisfied by any institutional structure, but also “concrete principles that specify the institutional arrangements common to all systems. . . that satisfy the constraints laid down by the most basic principles of justice” (67). Buchanan says that we are largely ignorant about principles of the second type with respect to the international legal system, with “one notable exception”: that nation-states should be minimally constitutionally democratic. Hence, (1) is a basic ideal theoretic principle of the first type, while (2) is a basic ideal theoretic principle of the second type.

democratic.²⁰

In part two, Buchanan applies his ideal theoretic account of political legitimacy to reforming the institutionalized practice of sovereign recognition. The problem posed by the current practice is that it permits human rights-violating states to enjoy all the prerogatives of sovereign states, including “support for their territorial integrity and . . . noninterference in their internal affairs,” as well as the capacity “to participate (in theory as equals) in the basic processes of international law.”²¹ Hence, the current practice prevents us from adequately protecting individuals’ human rights within the borders of rights-violating states. To preempt future rights-violators from using sovereignty as a shield against external interference, Buchanan proposes that international law be reformed to make a *new* polity’s recognition as sovereign conditional upon satisfying four conditions²²:

1. *Internal Justice Condition*: the state must protect (or must not violate) the basic human rights of its citizens.
2. *External Justice Condition*: the state must not violate the basic human rights of citizens of other states.
3. *Nonusurpation Condition*: the state must not come about by usurping a legitimate state.
4. *Minimal Democracy Condition*: the state must be minimally democratic.

These conditions are both necessary and sufficient. Each and every state that meets these conditions must be granted recognitional legitimacy; no state that fails to meet these conditions should be recognized.

To philosophers, Buchanan’s argument looks just fine. He identifies a morally-problematic feature of international law and suggests a solution that at least plausibly addresses the problem. So what is my objection to this approach? It is not that the ideal guidance approach yields the wrong verdict about the (in)justice of the current practice of recognition

20. Buchanan, *Justice, Legitimacy, and Self-Determination*, 234.

21. *Ibid.*, 266.

22. On the first two, see *ibid.*, 269–272; on the third, see 275f; on the fourth, see 278f.

or identifies the wrong reasons for thinking that this practice is unjust. An alternative approach might ultimately arrive at the same verdict. Instead, the problem is that this approach at best arrives at an *incomplete* analysis, not least by emphasizing only a subset of the salient considerations for institutional design. Since international law issues from the activity and resolutions of states (as represented by the relevant government officials),²³ the feasibility and effectiveness of Buchanan's proposal hangs on its sensitivity to a host of considerations concerning both the likelihood and the depth of international cooperation on reforming the practice of sovereign recognition in a way that will drastically limit states' control over their internal affairs. Here I raise several considerations that Buchanan neglects and their implications for his proposal.

For Buchanan's proposal to have any effect on human rights performance, states must be willing to enact a binding resolution with enforcement provisions.²⁴ States are usually willing to bind themselves in this way only if the institutional mechanism is necessary to coordinate their activity to achieve a key policy objective.²⁵ One question, then, is whether states take human rights protection abroad as a foreign policy objective *that overrides competing policy objectives*. This seems implausible given even a cursory examination of the historical record. For example, US foreign policy history is checkered with support

23. I use "state" as shorthand for "the government officials who are taken to represent a polity in international affairs." Consequently, the interests that matter are those of the state officials. Officials' decisions are influenced by their constituents' policy preferences *via* the domestic institutional mechanisms in place for holding state officials accountable. Where such mechanisms are robust, officials' policy decisions largely reflect the interests of the people at large. Where those mechanisms are weak, officials have more latitude in their policy decisions. Cf. Bruce Bueno de Mesquita et al., *The Logic of Political Survival* (Cambridge, MA: MIT Press, 2003); Fiona McGillivray and Alastair Smith, *Punishing the Prince: A Theory of Interstate Relations, Political Institutions, and Leader Change* (Princeton and Oxford: Princeton University Press, 2008). I discuss this in more detail below in chapter 5.

24. See Emilie Hafner-Burton and Kiyoteru Tsutsui, "Justice Lost! The Failure of International Human Rights Law To Matter Where Needed Most," *Journal of Peace Research* 44 (2007): 407–425; Oona A. Hathaway, "Do Human Rights Treaties Make a Difference?" *Yale Law Journal* 111, no. 8 (2002): 1935–2042; Oona A. Hathaway, "Why Do Countries Commit to Human Rights Treaties?" *Journal of Conflict Resolution* 51 (2007): 588–621.

25. Cf. Robert O. Keohane, *After Hegemony: Cooperation and Discord in the World Political Economy* (Princeton and Oxford: Princeton University Press, 2005); Helen V. Milner, *Interests, Institutions, and Information* (Princeton: Princeton University Press, 1997); McGillivray and Smith, *Punishing the Prince*.

for rights-abusing dictators who were otherwise amenable to US foreign policy objectives, as well as operations to overthrow democratically elected governments who were deemed hostile to US interests.²⁶ In addition, rights-abusing states with great natural resource wealth continue to find support despite their human rights record, such as China's (among others') continued support for the Burmese military junta.

Rights-abusing states in particular will have little incentive to sign on to reforms that will effectively prevent them from recognizing new rights-violating states. To comply with the norm would only draw greater public attention to their own rights-violating practices and endanger any perceived legitimacy they will have attained. Although Buchanan's proposal is not supposed to threaten a currently recognized rights-violator's *legal* status, they will certainly fear that the new norm could arouse domestic opposition that would be sufficient to drastically limit their ability to achieve their objectives. And given that some of the most internationally influential states are among the worst rights-abusers or most prominent supporters of rights-abusers (most notably, China and Russia), Buchanan's proposal is likely to meet stiff resistance among this crowd.

Moreover, many states that are relatively rights-respecting at home support rights-violating states abroad for a variety of reasons. Prominent among these reasons is the gains accrued from cooperation with such states. For example, 41% of China's merchandise exports go to the US and EU, while Australia, Canada, and the US all count China as both a top-five source of imports and a top-five destination for exports.²⁷ In addition, many states with significant natural resource wealth turn out to be authoritarian, rights-abusing regimes.²⁸ This means that otherwise rights-respecting states must cooperate with

26. Examples of the former include Fulgencio Batista, Mobutu Sese Seko, Pol Pot, and Saddam Hussein (before the late-1980s). Examples of the latter include Iran in 1953, Guatemala in 1954, Democratic Republic of Congo in 1960, and Nicaragua during the 1980s.

27. See country trade profiles for Australia, Canada, China, and the US, found in the WTO's Statistics Database, <http://stat.wto.org/Home/WSDBHome.aspx?Language=>, accessed July 2008.

28. I discuss this so-called "resource curse" in detail in chapters 5 and 6.

rights-abusing states to meet their substantial natural resource needs.²⁹ Given that similar benefits would accrue to states who cooperate with future rights-abusing regimes seeking recognition, states have little incentive to withhold recognition. Were they to do so, they would forego the substantial gains from cooperating with rights-abusing regimes.

Summing up: Buchanan's proposal neglects several considerations that are important for assessing the feasibility and effectiveness of his proposal. This is because he takes "justice... as the fundamental vantage point from which to evaluate the existing international legal system and to formulate proposals for improving it."³⁰ In other words, the set of considerations to which Buchanan gives adequate attention is restricted by the primacy he gives to his ideal principles of justice. He thereby fails to acknowledge where our interest in actually improving human rights protection might require us to make tradeoffs between our moral ideals on the one hand and feasibility and effectiveness on the other. To the extent that Buchanan's ultimate objective is to prescribe reforms that "would produce moral improvements" in our world, this is a serious blindspot. To overcome this blindspot, our institutional design prescriptions must aim at more than simply approximating ideal principles of justice. We must prescribe solutions that are capable of overcoming social problems as we find them in the actual world.

3.3 The Failure Analysis Approach to Institutional Design

We need an alternate approach to clinical theory that integrates moral analysis with causal analysis. Put another way, we need an approach that pays sufficient attention to

29. Perhaps most important among these is the need for oil. Roughly 36% of the world's oil is produced by unquestionable human rights abusers (Saudi Arabia, Russia, Iran, and China), and of the other top-15 producers, Amnesty International reported serious human rights concerns in seven of them (US, Mexico, Venezuela, Kuwait, UAE, Nigeria, and Iraq). Oil data from "Rank Order – Oil Production," *The CIA World Factbook*, Central Intelligence Agency, 2008. Human rights information from Amnesty International, *Report 08: The State of the World's Human Rights*, London: Amnesty International, 2008.

30. Buchanan, *Justice, Legitimacy, and Self-Determination*, 73.

engineering concerns in addition to the architectural concerns. In this section, I develop an approach to clinical institutional design that draws on insights from an engineering approach to design called “failure analysis,” which takes obviating failure, rather than approximating an ideal, to be the primary design aim. A failure analytic approach to institutional design avoids the shortcomings of the ideal guidance approach by dispensing with the need for ideal principles of regulation to guide the design process. In the place of ideal theory, institutional failure analysis is guided by a process of hypothesis formulation and evaluation. This places the design emphasis on overcoming actual social problems rather than on closing some gap between the actual and an ideal. Importantly, detailed analysis of causal mechanisms plays a central role throughout.

The motivating insight of failure analysis is well-stated by Henry Petroski:

Desire, not necessity, is the mother of invention. New things and the ideas for things come from our dissatisfaction with what there is and from the want of a satisfactory thing for doing what we want done. More precisely, the development of new artifacts and new technologies follows from the failure of existing ones to perform as promised or as well as can be hoped for or imagined. Frustration and disappointment associated with the use of a tool or the performance of a system puts a challenge on the table: Improve the thing. Sometimes, as when a part breaks in two, the focal point for the improvement is obvious. Other times, such as when a complex system runs disappointingly slowly, the way to speed it up may be far from clear. In all cases, however, *the beginnings of a solution lay in isolating the cause of the failure and in focusing on how to avoid, obviate, remove, or circumvent it.*³¹

We can see here a thumbnail sketch of failure analysis as a design process. The process starts first with dissatisfaction, with a sense that some designed artifact does not work as well as we might like. It then proceeds to diagnose the problem: In what does the failure consist? What caused it? Upon analyzing the failure, the designer seeks to design something that will avoid the same fate by improving its capacity to withstand similar pressures, by removing exploitable weaknesses, or by constructing a design that is not

31. Henry Petroski, *Success Through Failure: The Paradox of Design* (Princeton and Oxford: Princeton University Press, 2006), 1, emphasis added.

subject to the same causal mechanisms. Once the new design is complete, the designer tries to anticipate ways in which the new design might fail and, if any potential weaknesses are found, tries to improve the design to avoid these shortcomings.³²

From this sketch, we can isolate three main phases of the failure analysis design process: (1) *identification* of a failure, such as a flawed product or service; (2) *diagnosis*, that is, analysis of the character and cause of the failure; (3) *designing* an artifact to overcome identifiable failures, including potential future failures. The design objective is, quite simply, to create artifacts that avoid failure as far as is feasible. Developing the institutional failure analysis approach involves elaborating on each of the three phases as they pertain to clinical institutional design. I now discuss these in turn.

3.3.1 Identifying failure

An institutional design project is motivated by an initial sense of dissatisfaction with some feature of the social world, whether this be an observed undesirable outcome or social arrangements that impose an undue risk of realizing some undesirable outcome. However, an apparent disanalogy between engineering design and institutional design arises immediately. Failure analysis (in the context of artifacts) presupposes a well-defined design objective, which includes a set of general specifications that an artifact must meet. This well-defined objective makes it easy to identify failure. For example, a bridge should hold its intended load across the length of its span while withstanding environmental pressures, such as wind load or earthquakes. Bridges that fail to meet this design objective, or can only do so by incurring unacceptable costs, are readily identifiable as failures.

In contrast, social and political institutions do not come with well-specified design objectives. Indeed, politics is the process of contesting which ends are to be pursued by

32. Cf. Henry Petroski, *To Engineer is Human: The Role of Failure in Successful Design* (New York: Vintage Books, 1992), 44. A problem arises here for the analogy, viz., that the intended “use” of institutional arrangements is usually much more contested than that of engineered artifacts. I address this difficulty below.

institutions. Consequently, social failures seem identifiable only by reference to some particular perspective. For example, asymmetric bargaining power in trade negotiations is a failure for those who inhabit a weak bargaining position, but the asymmetry rarely disturbs those who benefit from it. In general, this means that social failures are notoriously difficult to identify. The institutional failure analysis approach appears to suffer from an early setback.³³

This apparent difficulty gives life to the ideal guidance approach. How can we know when some feature of the social world counts as a failure unless we know how the social world *should* be constituted? It is natural to think that ideal theory is helpful here. Ideal theory yields a well-defined design objective; it identifies principles that serve as a general specification any institutional order must meet to count as just. With ideal principles of regulation in hand, failure identification becomes a simple matter: institutional arrangements whose principles of regulation fall short of or otherwise differ from the ideal count as a failure. Since ideal theory derives principles of justice from a putatively *impartial* perspective, we need not worry about any particular perspective tainting our judgment. Ideal principles of justice give us the requisite impartial critical edge.

A clarification is in order before responding to this worry. “Ideal” ordinarily connotes something like “that to which we (ought to) aspire.” Paradigmatic examples of ideals in this sense include equality, individual liberty, human flourishing, and national self-determination. With this in the background, my rejection of the ideal guidance approach is apt to be interpreted as the claim that clinical theory ought not appeal to ideals in this ordinary sense. Such an interpretation misunderstands my claim. Ideal theory has come to signify something very particular since Rawls introduced the term. Ideal theory is not “a theory of ideals,” but a way of theorizing about political principles that focuses

33. Thanks to Peter Railton for identifying this disanalogy and to Bill Clark for pressing me on the “politics-as-end-contestation” point.

on specifying the principles of regulation that undergird an ideal institutional structure. These principles are arrived at by reflection on how to best express our moral ideals (in the ordinary sense) under ideal conditions, where “ideal conditions” denotes social and political circumstances that more or less permit moral considerations to take center stage in decisions about how to organize our collective life. For example, Rawls’s difference principle is supposed to be the principle that best expresses our collective commitment to an ideal of society as a system of fair cooperation among citizens conceived as free and equal.³⁴ But the derivation of that principle makes idealizing assumptions to avoid complicating the moral analysis too much. These include, among other things, that society is self-sufficient and closed to transactions with outsiders; that citizens accept and know that others accept a common set of principles; and that citizens fully comply with the demands of the principles of justice.³⁵

The ideal guidance approach subsequently takes the principles specified by ideal theory as regulative when theorizing about the principles that ought to guide institutional design under actual conditions. My rejection of the ideal guidance approach denies that the principles of regulation that express our commitment to our moral ideals under ideal conditions can or should offer any guidance for clinical theorizing. Importantly, this does not entail that clinical theorizing ought to refrain from appealing to moral ideals in the ordinary sense, or *values* as I will call them. In particular, it is open to the failure analyst to appeal to moral values when discussing the (in)justice of any particular social arrangements. To continue the above example, my rejection of the ideal guidance approach denies that Rawls’s difference principle can or should offer any direction for our thinking about the justice of institutions that affect distribution under actual conditions. This is consistent with appealing to the underlying value of society as a system of fair co-

34. Rawls, *A Theory of Justice*, secs. 1, 3, 4.

35. *Ibid.*, 4, 8, 216.

operation among citizens conceived as free and equal when assessing the institutions that are actually responsible for distributive outcomes. This is because the difference principle is not a conceptual truth; we do not arrive at the difference principle by simply analyzing the concepts expressed by the value. The difference principle is a particular expression of that value, which follows from our reflection on that value *given a certain simplified conception of the political world*. To the extent that the derivation of that principle is sensitive to changes in initial conditions, its service as an expression of an important value under different conditions will be in question.³⁶ But this problem affects only particular principles *qua* expressions of values, not the abstract values themselves. Thus, the latter remain available to the failure analyst when assessing institutional arrangements.

I now turn to a development of the failure analysis approach to identifying social failures. My point of departure draws on Sen's distinction between "comparative" and "transcendental" — that is, ideal theoretic — approaches to justice.³⁷ The central question for the comparative approach to justice is "How might we advance justice (or mitigate injustice) in a society?" A comparative assessment focuses on ranking alternative social arrangements with respect to some moral dimension(s). This differs from what Sen calls the "transcendental" approach to justice, which takes as its central question "What constitutes a just society?" Not surprisingly, then, the transcendental approach focuses on identifying the principles that govern ideally just social arrangements.

36. This point clearly echoes the basic idea expressed in the "general theory of second best": an institutional arrangement that is optimal under ideal conditions is unlikely to be so once we deviate from any of those conditions. The optimal arrangements under nonideal conditions are likely to require (perhaps drastic) alterations to the principles of regulation identified by ideal theory; see R.G. Lipsey and Kelvin Lancaster, "The General Theory of Second Best," *The Review of Economic Studies* 24, no. 1 (1957): 11–32. In Bruce Talbot Coram's words, "radical alterations in institutions may be required to accommodate small shifts in initial conditions" (Bruce Talbot Coram, "Second Best Theories and the Implications for Institutional Design," in *The Theory of Institutional Design*, ed. Robert E. Goodin [New York: Cambridge University Press, 1996], 91). For further discussion of the application of the general theory of second best to nonideal political theory, see Robert E. Goodin, "Political Ideals and Political Practice," *British Journal of Political Science* 25, no. 1 (1995): 37–56.

37. Sen, *The Idea of Justice*, introduction and pt. 1.

Although Sen identifies the crux of this distinction as being one of distinct objectives, the key difference between the two approaches is more accurately described as one of method.³⁸ Transcendental theory employs the ideal guidance approach to identify principles to govern our selection of institutional arrangements. The comparative approach identifies these principles by examining and reflecting on our comparative judgments of actual or feasible states of affairs. My development of the failure analysis approach explores this difference in method. In particular, I propose that taking comparative judgments about alternative social conditions as the starting point for clinical theorizing leads to more effective institutional design proposals.

Failure is identified by examining the particulars of the state of affairs that initially motivates the design project and then comparing this state to alternative feasible states of affairs. The contrast cases can be actual or counterfactual; the key is that their realization be feasible. If we take actual cases, we can be confident that the contrast class presents us with alternatives that are in some sense feasible, since they present states of affairs that are already realized. We need to be a little cautious when including counterfactual cases in the contrast class, since the fact that they are not currently realized leaves us uncertain

38. On the one hand, Sen associates the transcendental approach with the objective of identifying perfectly just institutional arrangements that are “themselves [the] manifestations of justice” (Sen, *The Idea of Justice*, 82). On the other hand, he associates the comparative approach with the objective of identifying “institutions that *promote* justice” (82), which requires giving due attention to “the actual societies that would ultimately emerge” from the “actual behaviours of people and their social interactions” (6). This way of putting the distinction is mistaken. Sen’s exemplars of the transcendental approach — Hobbes, Locke, Rousseau, Kant, as well as contemporaries such as Rawls, Dworkin, Gauthier, and Nozick — are not uniformly concerned with identifying perfectly just institutions; Hobbes and Locke come immediately to mind here. Moreover, transcendental institutionalists seem no less concerned with “social realizations” than comparativists. Rawls — Sen’s primary transcendentalist target — was eminently concerned with the social conditions that would be realized under alternative principles of justice. In Rawls’s words, “one conception of justice is preferable to another when *its broader consequences are more desirable*” (Rawls, *A Theory of Justice*, 6, emphasis added). In fact, practitioners of the transcendental approach seem to have the same objective as practitioners of the comparative approach — viz., that of identifying principles that should govern the selection of institutional arrangements that promote justice in our world. These points were made to clear to me by participants in an online reading group on Sen’s book, organized by Blain Neufeld at <http://www.publicreason.net>.

about the extent to which their realization is feasible. However, this should not preclude our making comparisons with counterfactual alternatives. We do not want our sense of which states of affairs are practically possible to be limited by what is actual. The general point is that we should be judicious in selecting our contrast cases, since these are going to determine which conditions to take as problematic and which to take as moral goals.³⁹

A case about which there is reasonable disagreement will work best to illustrate my point here, so say we are dissatisfied with health care provision in the United States. For our set of contrast cases, we might select Canada, Cuba, Mexico, Nigeria, Russia, Sweden, and Tanzania. (I stick with actual cases for simplicity.) We then start by making a rough intuitive ordering of these cases according to the relative justice of their health care provision schemes — for example (from best to worst): Canada, Sweden, United States, Cuba, Mexico, Russia, Tanzania, and Nigeria. This first cut ordering will be relative to some particular interpretation of justice, is likely to focus on some dimensions of health care provision at the expense of others, and is almost certain to be contested. This is fine for now; all we need is some set of orderings to serve as the raw material for the next step of failure identification.

The next step is to justify any particular ordering as capturing morally salient differences between the cases. To do this, one reflects on the considerations motivating any particular ordering and offers reasons for thinking that these considerations are (among) the morally salient ones when it comes to judging health care regimes. Example considerations include performance along objective health benchmarks (life expectancy, infant mortality rates, disease rates, and so on), health care spending efficiency, scope of access to health care, and source of health care provision (public vs. private provision). Sup-

39. More needs to be said about how to select counterfactual cases. Since feasibility is a workhorse concept here, its definition is crucial. To avoid interrupting the flow of this dissertation, I relegate a detailed analysis of feasibility to appendix B.

pose one's ordering is primarily driven by a country's performance according to objective health benchmarks, as the above ranking is,⁴⁰ while another's ranking is primarily driven by access-related considerations. On behalf of the above ranking, one might say that health outcomes are key when judging health care regimes because a society should be primarily concerned with the objective well-being of its citizens and positive health outcomes are important markers of well-being. On behalf of access-related considerations, one might say that citizens, in virtue of their common status as citizens, are entitled to equal treatment when it comes the allocation of health care resources. This is far from a complete characterization of what takes place at this point, but the picture being painted is sufficient to illustrate the point. Once we have a set of first cut orderings, we set about justifying an ordering as authoritative, which leads us to reflect upon the moral values we endorse and our reasons for endorsing them, as well as identifying the principles that best express those values. Moral justification requires us to engage in the process of supporting the moral authority of an ordering with impartial reasons — reasons that do not appeal to any particular person's situation or interests. Such a process will include appeals to abstract moral values. Importantly, this is not the same as identifying the principles of regulation that govern the ideal health care regime (or, the health care regime of the ideal society).

At some point, this moral debate will lead us to some shared judgments (although we are unlikely to arrive at complete consensus). For example, all parties to the debate might agree that health outcomes and access-related considerations are both important, although they might disagree on their reasons for thinking so or the relative weight assigned to each. No matter. This rough agreement still permits us to make judgments of the following sort: improving objective health outcomes in Nigeria and Tanzania constitutes an improvement in health care provision; increasing access to health care provision in the United States

40. The above ranking is according to life expectancy at birth, 2009 estimates. Source: 2009 *CIA World Factbook*.

constitutes an improvement in health care provision. These judgments imply judgments of failure. When we judge that some state of affairs *S* can be improved upon, we are committed to the claim that *S* is not as good as it could be. But not just that. Since our comparative judgments issue from our reflection on the moral values that underpin our judgments, our reasons for taking any ordering as morally authoritative imply that *S* is not as good as it *should* be. In Sen's words, these judgments identify "remediable injustices."⁴¹ On the failure analysis approach, a failure just is a remediable injustice.

One might object that this makes "failure" an overly capacious concept, which undermines its critical edge. Perhaps "failure" should be reserved for social conditions we deem severe injustices requiring urgent remediation. But failure need not be constituted by utter inability to meet design expectations. Instead, a failure is constituted by the presence of a remediable design flaw.⁴² We can comfortably acknowledge that instances of failure will differ along a number of dimensions, including ease of identification, severity, and (moral) urgency. A bridge collapse constitutes a greater failure than an unwieldy water bottle. Similarly, avoidable famine, genocide, total breakdown of the rule of law, and arbitrary detention and torture are more grievous and more urgent than, say, disparities in educational quality or employment opportunities.⁴³ It is true that the generosity of the failure concept will preclude the identification procedure from generating fine-grained distinctions among failures, which could help us set "remediation priorities." But that's not the job of *the identification procedure*. The identification procedure simply seeks to identify

41. Sen, *The Idea of Justice*, vii.

42. For example, Petroski spends a great deal of energy illustrating the failure analysis approach to engineering using the development of presentation pointers as an example. Before laser pointers (the design flaws of which Petroski points out), presenters used long wood pointers — sometimes up to 7 or 8 feet long — that enabled them to point at items on a screen without having to walk in front of the screen. But these pointers became unwieldy and caused some presenters to fatigue. Although not a particularly urgent problem, this, in Petroski's mind, counted as a failure in virtue of these design flaws. See Petroski, *Success Through Failure*, ch. 1.

43. I certainly don't mean to trivialize the injustice of the latter two. It just seems to me that the ability to remediate the latter injustices are luxuries afforded by having eradicated the former.

the members of the set of failures. Once we have identified (some of) the members of this set, we engage in further reflection and debate about our remediation priorities, debate that will and should appeal to considerations generated by our diagnosis of the failure and our anticipation of the effects of various intervention possibilities.

The preceding has exposed an important difference between the ideal guidance and failure analysis approaches. On the former, a social process or outcome is identified as a failure because it diverges from the processes or outcomes that would arise from ideal institutional arrangements identified in the abstract. For example, an actual distribution of some good is a failure because it deviates from the distribution that would result from fully just institutional arrangements. Failure is thus constituted by a gap between the state of affairs produced by the actual institutional order and the state of affairs that would result from an ideal institutional scheme.

On the failure analysis approach, there is no comparison with an ideal institutional order because there is *ex ante* target institutional order. As a result, there is no preexisting blueprint from which actual institutions could diverge. Instead, failures are identified by making comparisons between actual and feasible states of affairs and finding that some actual states are not as good as they could (and should) be. Consequently, we don't need a blueprint of ideal institutional arrangements to tell us which social conditions constitute failures and which institutions should be established to overcome the problems.

In fact, such a blueprint is liable to bias our identification of failures because it pre-judges what counts as a problem and thereby restricts our attention to certain features of the social world. Take Dewey's criticism of *laissez-faire* liberalism (i.e., libertarianism) as an example.⁴⁴ Libertarianism identifies individual liberty with individual economic enterprise more or less unconstrained by government regulation. The concomitant institutional

44. What follows draws from John Dewey, *Liberalism and Social Action* (Amherst, NY: Prometheus Books, 2000), ch. 2, *passim*.

ideal consists of an unregulated market for all goods in which people could have an interest.⁴⁵ When a libertarian assesses the justice of an institutional order, her attention is restricted by her expectation that a just institutional order includes (at least) an unregulated market. The problem, in Dewey's words, is that libertarians have "put forward their ideas as immutable truths good at all times and places." Accordingly, libertarian principles have become reified: "they [hold] that beneficial social change can come about in but one way, the way of private economic enterprise, socially undirected, based upon and resulting in the sanctity of private property." This is due to a lack of "historical sense"; libertarians have been "blinded... to the fact that their own special interpretations of liberty, individuality and intelligence were themselves historically conditioned, and were relevant only to their own time..." Once the sought-after reforms were accomplished, what was once a force for social change thus became a force in favor of the status quo. Consequently, adherence to the *laissez-faire* ideal blinds libertarians to the obstacles to effective liberty brought about by an unregulated market. Since their attention is directed by their favored institutional ideal, these liberals have failed to see the ways in which liberty is restricted by that very ideal.

The general lesson here is this: One does not need to know what the ideal institutional order looks like to be able to identify social failures. The ideal guidance view locates the blueprint at the wrong place, prior to the failure identification phase. But the creation of the blueprint is a *result* of the *design* process. There is no blueprint prior to the design phase, let alone the identification phase.

45. See Nozick, *Anarchy, State, and Utopia*, ch. 7, especially the section titled "How Liberty Upsets Patterns." See also Milton Friedman, *Capitalism and Freedom* (Chicago: University of Chicago Press, 1962).

3.3.2 Diagnosing failure

Once we have identified a social problem, we set about diagnosing that problem. This diagnostic phase incorporates both moral and causal analyses. A return to the architecture/engineering analogy is useful here to get a grip on how the normative and empirical are intertwined. Architectural failures and engineering failures differ in virtue of the distinct design aims of architecture and engineering. Architectural design fails when space is poorly organized, or when component textures, materials, colors, etc. are poorly juxtaposed, or when the structure is a poor fit — functionally, artistically, aesthetically — with the surrounding environment. As examples, consider the structures that often show up on “ugly building” lists, such as Boston City Hall, the Experience Music Project (Seattle), and the Scottish Parliament Building (Edinburgh).⁴⁶ Engineering design fails when structural elements are unable to withstand environmental pressures, or when the object functions poorly or not at all under the conditions for which it was designed, or when the object is unsafe for use. Examples of engineering failures include bridge collapses (Tacoma Narrows Bridge, 1942), building collapses (Hyatt Regency walkway, Kansas City, 1981), nuclear accidents (Chernobyl, 1986), and space accidents (*Columbia*, 2003). From the different lists of examples, it should be clear that architectural and engineering failures have distinct characters. Accordingly, the two require different sorts of diagnoses. When diagnosing architectural failures, we aim to identify the features of the object that deviate from accepted standards or norms of architectural design, to characterize the ways in which the design contravenes design values. On occasion, architecture that was initially criticized is later seen as pushing the field in a positive, innovative direction. Hence, architectural analysis also involves re-evaluating the standards by which we evaluate architecture. In

46. Sources: “The World’s Ugliest Buildings” (www.forbes.com/2002/05/03/0503home.html; edition.cnn.com/2008/WORLD/europe/10/22/ugliest.buildings/; members.virtualltourist.com/vt/t/1c7/).

contrast, when diagnosing engineering failures, we aim to identify the causal processes that generated the failure, to identify the weaknesses in a structure or process and the pressures that were able to exploit that weakness.

Given that institutional design comprises both architectural and engineering problems, our diagnosis of institutional failures requires two types of analysis. Moral analysis aims to identify the features of an institutional structure that deviate from widely-accepted norms and values, but also to re-evaluate those norms and values. Causal analysis aims to identify the causal mechanisms generating the social conditions we seek to alter. In practice, the tasks of moral and causal analysis are rarely separable. Our evaluation and selection of the moral principles by which we assess institutional arrangements will be informed by our causal analysis of current conditions. If, for example, income inequality is an unavoidable feature of collective economic life and we think collective economic life is important (or inevitable), then we might reduce the weight we give to particular egalitarian moral principles when morally assessing institutions. But if income inequality is a result of institutions that unnecessarily and unjustifiably restrict the economic opportunities of an underclass, we might retain those same egalitarian moral principles as important standards for institutional assessment.

Moral considerations also play an important role in identifying the mechanisms to which our causal diagnoses pay attention. For example, commodity price volatility almost certainly plays a causal role in generating the resource curse.⁴⁷ Thus, the independent market decisions of investors and consumers that are responsible in the aggregate for this price volatility are, at least in part, causally responsible for the misery associated with the resource curse. However, we don't typically identify these independent market decisions as causes of the curse and not simply because it is unlikely that we could adequately

47. See Humphreys, Sachs, and Stiglitz, *Escaping the Resource Curse*, ch. 1.

coordinate those decisions to avoid price volatility. In addition, we take it for granted that the market freedom that generates price volatility is a value that we should protect. Consequently, we turn our attention to other causal mechanisms. In the case at hand, we hold the fact of price volatility fixed and turn our attention to the mechanisms that make the economic performance of resource dependent states vulnerable to price shocks. All this is to say that moral and causal analyses are practically inseparable. Nevertheless, it can be useful to think of moral and causal analyses as analytically distinct. With this in mind, I now elaborate on my sketch of the diagnostic task.

Moral diagnosis is prompted by a desire to understand the ways in which a social process or outcome constitutes a *moral* problem. This involves considering which values we want our social life to embody, why these values are important, and which moral principles best express those values. Do we prize equality? What about equality is important? Is individual liberty a key value? How is liberty restricted under current conditions? Is human flourishing an important social goal? Why? How should we understand “flourishing”? Although these questions inquire about abstract moral values and principles, on the failure analysis approach, we do not settle these questions solely in the abstract. Instead, our reflection on the moral principles we want our social life to realize is guided by reflection on actual social conditions.

Consider Mill’s vigorous criticism of Victorian marriage contracts as the sort of moral diagnosis I have in mind.⁴⁸ Mill begins his criticism in earnest by examining “the conditions which the laws of this and all other countries annex to the marriage contract.”⁴⁹ Such contracts left a woman effectively propertyless and thereby without economic security should her husband die or divorce her. Within marriage, a woman had little say over the

48. John Stuart Mill, “The Subjection of Women,” in *The Subjection of Women*, ed. Dale E. Miller (New York: Random House, 2002), esp. chs. 2 and 3.

49. *Ibid.*, 153.

disposal of her husband's property. Without any property of her own, she was financially dependent on her husband and without any credible exit threat should her husband abuse her. The marriage contract precluded any notion of marital rape, leaving a woman as little more than a sex slave. Women had no legal rights over their children; these were granted only to men. What's more, women were compelled to this position of servitude because they were banned from pursuing the means to independence, such as an education or a career. Even as she had no property, the law forbade her to sell her labor. Mill compares the position of women to that of a slave:

I am far from pretending that wives are in general no better treated than slaves; but no slave is a slave to the same lengths, and in so full a sense of the word, as a wife is. Hardly any slave, except one immediately attached to the master's person, is a slave at all hours and all minutes; in general he has, like a soldier, his fixed task, and when it is done, or when he is off duty, he disposes, within certain limits of his own time, and has a family life into which the master rarely intrudes. . . . But it cannot be so with the wife.⁵⁰

The implicit argument here is that the Victorian marriage institution subjected women to a position that was deemed unfit for slaves and that the continued subjection of women to such conditions was inconsistent with earlier judgments against subjecting slaves to such conditions. In other places, Mill compares marriage to political tyranny, implying that the marriage contract subjected women to a position relative to their husbands to which no man would have consented in relation to a political ruler.

Mill's strategy involves enumerating the conditions that result from a particular institution and then exposing the conflict between these conditions and the moral principles we might justifiably endorse upon reflection. This results in both moral criticism of the conditions of marriage and a re-assessment of the convictions that keep those conditions in place. What is striking about Mill's strategy from the perspective of conventional political theory is what he doesn't say. Mill does not argue for a set of ideal principles of justice

50. Mill, "The Subjection of Women," 155.

and then employ them to justify treating women as equals. He does not say “It will be well to commence the discussion of gender discrimination by considering which principles of justice would regulate the constitution of the marriage institution in the ideal society and then apply these principles to the marriage institution as it stands.” Starting with abstract principles of justice permits us to rationalize concrete social conditions, to point to the ways in which the status quo is consistent with the requirements of these principles, which often requires minimizing (or altogether leaving out) inconvenient facts. By starting with concrete conditions, we cannot be let off the hook. We must come face-to-face with the details of our social reality and try to reconcile those details with our convictions. Often times, we cannot.

One might turn around and press this claim against the failure analysis approach, namely, that failure analysis has a conservative bias. After all, on my view, coming face-to-face with social reality only leads to judgments of injustice (and thus a need to prescribe interventions) if we find features of that reality dissatisfying. Accordingly, we might follow G.A. Cohen in claiming that “the question for political philosophy is not what we should do but what we should think, even when what we should think makes no practical difference.”⁵¹ But Cohen’s claim is consistent with my rejection of the ideal guidance approach and gives me a way to reply to this “conservative bias” charge. The failure analysis approach rejects the claim that ideal principles of justice are useful for guiding clinical institutional design. But it need not reject the claim that ideal theory can be useful for guiding *our evaluative attitudes* toward our social reality, including actual institutional arrangements. Indeed, this seems to be the appropriate role for ideal theory within the failure analysis approach. Finding gaps between actual institutional arrangements and ideal institutional arrangements motivates acute dissatisfaction with current arrangements, thereby

51. Cohen, “Facts and Principles,” 243.

initiating the design process. Thus, ideal theory can be important for *motivating* design prescriptions. This is consistent with the claim that the principles it yields ought not *guide* clinical institutional design.⁵²

Importantly, on the failure analysis view, the diagnostic phase involves not only identifying the ways in which current conditions undermine important values, but also *re-evaluating* the standards by which we assess social conditions. Moral principles are adopted in light of particular social conditions. Under conditions of inequality, particular egalitarian principles come to the fore; under conditions of slavery or tyranny, liberty is championed. But social conditions are in continual flux: “old principles [might] not fit contemporary life as it is lived, however well they may have expressed the vital interests of the times in which they arose.”⁵³ The diagnostic phase demands that we reconsider our moral principles in light of our social reality to avoid adopting principles that are ill-suited to current conditions and to prevent principles of justice from becoming reified.

Keeping in mind the purely analytic distinction between moral and causal analysis, our causal diagnosis is an entirely empirical task. We are interested in *explaining* the outcome, not in assessing it according to normative criteria. As noted above (section 2.1), causal analysis involves identifying the salient components of the causal process(es) that generate an outcome and specifying their interrelationships. This latter part includes specifying how the components interact and how changes in one part of the process affect (the operation of) other components.

It is important to note exactly what sort of work causal analysis is doing here. To explain, let’s distinguish between two types of causal analysis. *Speculative causal analysis* specifies the mechanisms by which an institutional prescription could bring about some

52. Thanks to an anonymous reviewer for *The Journal of Political Philosophy* for pressing me to address this issue.

53. John Dewey, *The Public and its Problems* (Athens, OH: Ohio University Press, 1954), 135.

desirable outcome; *explanatory causal analysis* specifies the mechanisms that generate the outcomes we observe.

To illustrate how speculative causal analysis proceeds, consider Simon Caney's discussion of global institutional design. As discussed earlier (section 2.3.1), Caney argues that there should be a layer of suprastate institutions that have authority to constrain the activity of states in pursuit of cosmopolitan objectives. He discusses four mechanisms by which suprastate institutions can realize cosmopolitan objectives.⁵⁴

1. Suprastate institutions can help solve collective action problems.
2. Suprastate institutions can enforce compliance with cosmopolitan principles and hold powerful agents accountable.
3. Suprastate institutions can empower weak agents to protect their interests.
4. Suprastate institutions can transform powerful agents' incentives.

The important thing to note is that Caney does not discuss the mechanisms by which cosmopolitan objectives fail to be realized in the actual world. Put differently, Caney's discussion of mechanisms is not a diagnosis of any social failure; it is a list of the ways in which authoritative suprastate institutions *might* bring about cosmopolitan objectives if implemented. I call this "speculative" causal analysis because it implicitly depends on a particular diagnosis of our failure to realize cosmopolitan objectives. Whether the operation of the specified mechanisms would bring about cosmopolitan objectives depends on the mechanisms operating to inhibit their realization. I find it intuitively plausible that collective action and compliance problems mitigate the realization of cosmopolitan objectives, but this is beside the point. What we need here is to explicitly examine the causal process(es) that inhibit our fulfillment of cosmopolitan objectives. This is what diagnosis amounts to; this is what speculative causal analysis fails to provide.

54. See Caney, "Cosmopolitan Justice and Institutional Design," 742ff.

It is crucial to understand that the diagnostic task requires not just any old discussion of causal mechanisms, but *explanatory* causal analysis in particular. We do not seek to understand a process by which some prescription might bring about the desired outcome, but to understand the process(es) by which realization of our desired outcome is inhibited. Since successful clinical theorizing crucially involves explanatory causal analysis, it is a weakness of the ideal guidance approach that it typically ignores its importance.

3.3.3 Designing to avoid failure

Once we have a working diagnosis, we set ourselves to the design task. This balances design objectives that follow from both the moral and causal analyses. Normatively, the design aim is to prescribe institutional solutions that will bring about social conditions that comport with the moral values we can endorse on the basis of impartial reasons. The design aim that follows from our causal analysis is to prescribe feasible institutional solutions that can intervene effectively at important places in the causal process to improve the outcome.

Again, although we may analytically separate the normative and empirical design aims, they are tightly intertwined in practice. Return again to the architecture/engineering analogy. Architects and engineers mutually constrain the design process. As my construction example in section 3.1 illustrated, the physical limits of different structural materials and their various possible configurations constrain architectural possibilities; architectural standards and values constrain the set of desirable engineering solutions. Similarly, the moral principles we choose to endorse are constrained by the means required to realize those principles, while our assessment of interventions is constrained by the moral cost of implementing those interventions. To illustrate these points, consider the following (perhaps extreme) examples. If the realization of global income equality requires the es-

establishment of a world government and the latter would entail unacceptable costs (moral or otherwise), the foregoing requires that we reconsider our endorsement of global income equality as an attractive moral aim. Similarly, if the best means to consolidating democratic governance requires drastically circumscribing the liberty of dissenters, the foregoing requires that we look for other ways to consolidate democratic reforms.

A key part of the design phase is design evaluation. We are not interested solely in overcoming *identified* failures; we also seek to forestall *identifiable* failures. As Petroski puts it, the design process “may be considered a succession of hypotheses that such and such an arrangement of parts will perform a desired function *without fail*.”⁵⁵ Each institutional design proposal is similarly a hypothesis that the institution as designed will successfully achieve its objective under the conditions in which it will be required to operate. Given the stakes, we can’t accept such hypotheses blindly. Nor is intuition a reliable check. Thus, the last phase of the design process is to evaluate our design hypotheses for potential weaknesses and potentially negative path dependencies. Will the proposal generate morally perverse consequences? Will the institution be exploited by enterprising opportunists? Will it attract sufficient compliance? Will it close off important possibilities for improvement in the future? Should we find weaknesses, we return to the drawing board to find ways to shore them up. If simple fixes are not available, we need to consider overhauling the original proposal. The aim is to establish institutions that can foster and coordinate interactions in a way that, when aggregated, lead to morally improved social conditions that keep open possibilities for future improvement, as well as mitigate or contain the negative consequences of socially destructive interactions.

55. Petroski, *To Engineer is Human*, 44; original emphasis. “Without fail” need not mean “never to be replaced.” The point of emphasis here is the ability to achieve an objective. It is perfectly acceptable to deliberately design something that is temporary and aims at achieving an intermediate goal, so long as the design doesn’t close off the possibility of future progress. Thanks to Elizabeth Anderson for bringing this point to my attention.

The importance of the evaluation phase and the willingness to revise in light of potential weaknesses suggests that clinical theorizing about institutions is a fluid, experimental process. We do not seek to propose an institutional configuration for all time. Nor do we aim to put together a “master plan” that encompasses an entire system of institutions. We are, of course, interested in uncovering the interactions between distinct components of a larger system and avoiding negative interactions as far as possible. But we should not hold out hope for a fully worked out ideal. Our vision is too limited, our knowledge too local. Each proposal is tentative and experimental, aiming at piecemeal, incremental progress.

None of the foregoing shows that the design process need not be guided by target states of affairs. Indeed, when stating the aim of the design phase, I claimed that we had “morally improved social conditions” in view. Doesn’t this suggest the need for ideal theory as a guide to clinical theory? Don’t we need ideal theory to guide our thinking about what constitutes “morally improved social conditions”? To make the objection stronger, note that it need not rely on a view of ideal theory as delivering a singular “best” state of affairs. The easy reply in this case is that ideal theory is not up to the job.⁵⁶ All that is required to get the objection off the ground is that the design phase must be *prospective* in the sense that it requires a more or less well-specified target state of affairs and that ideal theory is required to identify the principles of regulation that govern the targeted arrangements. This is sufficient to vindicate at least a restricted version of the ideal guidance approach.

The fact that clinical theory seeks to bring about a moral improvement of social conditions means that clinical institutional design must aim at *something*, namely, morally improved social conditions. But on the failure analysis approach, the design phase is largely *retrospective* in the sense that our sights are set by looking backward, at the places we have been rather than at the places we would like to go to. We design institutions to

56. See Sen, *The Idea of Justice*, ch. 4.

overcome failure, not to realize an ideal. To be sure, the evaluation phase requires that we try to anticipate potential future failures and we might say that this make failure analytic design prospective. But this is not prospective in the sense used by the ideal guidance approach. More accurately, it is *counterfactually* retrospective. That is, the evaluation phase examines where we would have been had we implemented the original design proposal. Our design aim then is to overcome these counterfactual failures.

Designing to overcome failures has no need for the principles of regulation identified by ideal theory. All we need to know is (1) which possible solutions are feasible; (2) which of the feasible solutions are morally acceptable; and (3) which of the feasible, morally acceptable solutions are likely to effectively intervene at the appropriate place in the causal process generating the failure. None of this makes reference to the principles of regulation identified by ideal theory. One might argue that we need ideal theory to identify the morally acceptable solutions within the feasible set. But there's no reason to employ the overwrought framework of ideal theory to help us here. Attempting to identify ideally just arrangements and the principles that regulate them is liable to distract from the task of identifying solutions. In any case, ordinary moral reasoning is sufficient for identifying which feasible options are morally acceptable. Return to the earlier discussion of the comparative method for identifying failures (section 3.3.1). A similar method is effective here too. We take our current social conditions and compare them to the conditions that would arise were we to implement some particular feasible institutional solution. We then ask ourselves: Do we think the counterfactual conditions are acceptable? Are the counterfactual conditions an improvement upon current conditions? On the basis of which principles do we make these judgments? Can we justifiably endorse these principles upon reflection? This comparative process constitutes nothing other than ordinary moral reasoning. Of course, by "ordinary," I do not mean to suggest that our identification of morally

acceptable options will be easy. My point is that the design process needs nothing like ideal theory to identify principles of regulation to serve as targets for clinical theory. The resources we need to prescribe morally progressive institutional solutions are available without having to take on board the baggage of ideal theory.

3.4 A Fruitful Approach?

I have argued that we should abandon ideal theory in our attempts to prescribe institutional reforms to address actual injustices and adopt a failure analysis approach to clinical institutional design. The primary drawback of the ideal guidance approach is its myopia; it focuses on understanding and applying an ideal of justice at the expense of a detailed understanding of the problem. Failure analysis overcomes this myopia by refocusing our attention on the problem. By integrating causal analysis with moral analysis, a failure analytic approach is more sensitive to the complexities of the problems we wish to address without sacrificing sensitivity to the important moral considerations that rightly constrain our attempts to address our social failures.

The preceding development of the failure analysis approach has remained largely abstract. It remains to be seen how clinical theorizing about particular institutional design problems can be improved by adopting this approach. We would be right to be skeptical of methodological innovations that have not been put to the test, that have not been shown to yield fruit. Hence, the best way to vindicate the preceding discussion is to employ the failure analysis framework in thinking about problems like extreme poverty and inequality, war and military intervention, and gender and racial discrimination. This is the task for the remainder of this essay.

CHAPTER 4

Interlude: From Theory to Application

I have argued for two basic claims thus far:

1. Political philosophers must attend to causal mechanisms.
2. Our institutional design prescriptions must start from diagnoses of social failures.

These claims mutually reinforce each other. A requirement to attend to causal mechanisms puts pressure on philosophers to seek causal explanations of the states of affairs for which they prescribe institutional reforms; a requirement to diagnose social failures puts pressure on philosophers to attend to causal explanations of the failures. These two requirements have been brought together under the failure analysis approach to clinical institutional design.

Any appeal the failure analysis alternative might have remains limited so long as its articulation remains abstract. The ideal guidance approach has gained widespread adherence because it bears an appealing kind of fruit: it yields institutional design prescriptions that reflect normative principles derived from systematic reflection on our deepest moral commitments. I don't deny that this is a significant achievement. Instead, I deny that the prescriptions yielded by the ideal guidance approach can fulfill the promise of non-ideal theorizing, namely, to present institutional remedies that can effectively bring about morally improved states of affairs amidst unjust and otherwise nonideal conditions.

Widespread adoption of the failure analysis methodology is vindicated only if it can fulfill this promise. Of course, whether this is so depends on whether the prescriptions this approach yields are successfully implemented, and implementation is largely out of theorists' hands. But successful implementation starts with proposals that are reasonably expected to produce the desired results. In particular, the proposal must present institutional solutions that will be feasible, efficient, stable, and effective for mitigating injustice. I argue that the failure analysis methodology yields such proposals. This claim is justified only by the successful application of failure analytic principles to concrete problems of institutional design. I begin the task of application in the remaining chapters.

A final methodological word before I proceed. Because failure analysis requires that we start with a detailed diagnosis of the problem at hand, we must select problems that are relatively narrowly described. In other words, failure analysis is not equipped to address amorphous problems such as *global poverty*. Instances of poverty are so varied that any diagnosis that seeks to unify all cases under a single explanation is bound to leave us with institutional prescriptions that are too abstract and general to assess.¹ Instead of tackling global poverty in toto, failure analysis recommends starting with analytically tractable problems — high incidence of easily curable diseases or the corruption that pervades many developing countries' public administration.

The remaining chapters focus on development failures that arise in countries with great natural resource wealth. There is a growing consensus among development political economists that countries that are fiscally dependent on revenue from resources like oil, natural gas, or copper are more likely to be governed by authoritarian regimes and suffer economic stagnation. Many of the countries afflicted by this “resource curse” contain a significant percentage of the global poor; these include Angola, Democratic Republic of

1. This is perhaps another problem with much of the global justice literature; it tries to take on problems that are analytically unmanageable and proposes solutions that are too vague to assess.

Congo, Myanmar, Nigeria, Sudan, Venezuela, and Yemen. Cursed countries comprise more than 300 million poor people, nearly a quarter of the 1.4 billion people who live below the World Bank's US\$1.25/day poverty line — almost half of the global poor if we exclude China and India.² They are often riven by civil conflict, and their residents are harried and left destitute by corrupt leaders who prioritize personal aggrandizement over providing their citizens with basic public goods, such as law enforcement or public infrastructure. Some of the characters involved are infamous for their brutality, corruption, and thievery; names like Sani Abacha, Omar al-Bashir, Mobutu Sese Seko, Teodoro Obiang, and Than Shwe head the list.

The resource curse is a moral disaster from any number of normative perspectives. Victims of the resource curse have little say in the governance of their lives; they suffer from acute welfare deprivation; their basic needs often go unmet; their physical and mental integrity are continually threatened; their control over possessions is insecure; they are subject to arbitrary and predatory rule. Any ordering of states of affairs along any dimension of moral interest (e.g., welfare, autonomy, human rights, needs, and so on) will surely list resource cursed conditions near the bottom. Given this, we can leave aside the failure identification stage; the resource curse is a paradigmatic case of a social failure.

Having identified a failure, we proceed to the diagnostic phase. Part of our task here is to characterize the ways in which the resource curse constitutes a moral problem. Since our normative analysis must focus on the causal processes that actually generate the curse, I start my assessment of the resource curse with a causal diagnosis of the problem. This appears in the next chapter. Chapter 6 draws out the implications of this diagnosis for our obligations to mitigate the privation experienced by resource curse victims.

2. This statistic is lower than it should be since the data for Libya, Myanmar, and Sudan — all populous resource dependent countries with high poverty rates — are missing. Poverty data from UNDP, *Human Development Indicies*, United Nations Development Programme, 2008.

CHAPTER 5

Explaining the Resource Curse

States that are fiscally dependent on exportable point-source natural resources,¹ such as oil, natural gas, or copper, are more likely to be governed by authoritarian regimes or have unstable or stagnant economies.² Recent theoretical and empirical work on this “resource curse” highlights the fundamental importance of domestic institutional quality in explaining why this is the case.³ At bottom, the effect of resource revenue on development is a conditional one: resource dependent countries with weak institutions are resource-cursed, while resource dependent countries with strong domestic institutions can escape

1. On the theoretical importance of point-source vs. diffuse resources, see Philippe Le Billon, “The Political Ecology of War: Natural Resources and Armed Conflicts,” *Political Geography* 20 (2001): 561–584. Snyder and Bhavnani discuss the importance of a similar distinction, that between “nonlootable” and “lootable” resources; see Richard Snyder and Ravi Bhavnani, “Diamonds, Blood, and Taxes: A Revenue-Centered Framework for Explaining Political Order,” *Journal of Conflict Resolution* 49, no. 4 (2005): 563–597.

2. Seminal papers include: Michael L. Ross, “Does Oil Hinder Democracy?” *World Politics* 53 (2001): 325–361; Wantchekon, “Why do Resource Abundant Countries Have Authoritarian Governments?” (resources and regime type); and Jeffrey D. Sachs and Andrew M. Warner, “Natural Resource Abundance and Economic Growth,” *NBER Working Paper Series*, no. 5398 (1995) (resources and economic growth). There is also a literature that claims to show that resource abundance causes higher rates of civil war onset, but this connection is more controversial than the other two; see Christa Brunnschweiler and Erwin Bulte, “Natural Resources and Violent Conflict: Resource Abundance, Dependence, and the Onset of Civil Wars,” *Oxford Economic Papers* 61, no. 4 (2009): 651–674; Paul Collier and Anke Hoeffler, “Greed and Grievance in Civil War,” *Oxford Economic Papers* 56, no. 4 (2004): 563–595; James D. Fearon, “Primary Commodity Exports and Civil War,” *Journal of Conflict Resolution* 49, no. 4 (2005): 483–507; Macartan Humphreys, “Natural Resources, Conflict, and Conflict Resolution: Uncovering the Mechanisms,” *Journal of Conflict Resolution* 49, no. 4 (2005): 508–537.

3. Among others, see Nathan Jensen and Leonard Wantchekon, “Resource Wealth and Political Regimes in Africa,” *Comparative Political Studies* 37, no. 7 (2004): 816–841; Halvor Mehlum, Karl Moene, and Ragnar Torvik, “Institutions and the Resource Curse,” *The Economic Journal* 116, no. 1 (2006): 1–20; James A. Robinson, Ragnar Torvik, and Thierry Verdier, “Political Foundations of the Resource Curse,” *Journal of Development Economics* 79, no. 2 (2006): 447–468; Ross, “Does Oil Hinder Democracy?”

the deleterious effects of resource wealth.

This institutional story gets something fundamentally right: the resource curse is ultimately a story about the extent to which domestic political institutions enable rulers to pursue their own interests at the expense of citizens' rights and well-being. But standard institutional analyses fail to consider the full effects of resource revenue on the institutional environment itself. This is a significant blindspot. If the persistence of weak institutions is endogenous to the resource curse, then the conditional effect of resources on development is more subtle than the standard institutional claim. The more subtle claim is this: if a country does not have firmly entrenched strong institutions *before it becomes fiscally reliant on resource revenue*, then weak institutions will persist and it is vulnerable to the resource curse; if a country has firmly entrenched strong institutions *before it becomes fiscally reliant on resource revenue*, then it can avoid the resource curse.⁴

In this chapter, I present a simple formal model to support the thesis that the persistence of weak institutions in resource-cursed countries is endogenous to the resource curse. Extant theoretical models take the institutional setting as given and then theorize about the effect of resource revenue on governance and economic performance within the given institutional environment. In contrast, the model I present is a more general model of domestic competition over political outcomes. I aim to identify the underlying conditions to which governance and economic performance are sensitive. The model reduces these outcomes to two parameters: rulers' dependence on citizens for support and the credibility of supporters' exit threats. Having shown that development outcomes are sensitive to these parameters, I then theorize about how resource revenue affects them. The model shows that if a ruler does not depend on support from a broad coalition of citizens prior to the inflow of resource revenue, resource revenue undermines any impetus to increase his de-

4. I define weak and strong institutions in the next section.

pendence on their support. However, if a ruler relies on citizens' support prior to resource discovery (and citizens' have credible exit options), then the ruler uses resource revenue to advance citizens' interests.

5.1 Institutions and the Resource Curse

Let's begin with some definitions. Let "ruler" denote the individual or group who retains ultimate authority to implement or block the implementation of policy, as well as the loyal bureaucrats to whom particular tasks are delegated.⁵ We can distinguish between two broad types of domestic institutional structures. Generally, *strong* institutions empower (formally or informally) a broad coalition of citizens to effectively check the formation and implementation of policy; *weak* institutions, in contrast, afford the ruler wide latitude to implement policy without having to consult citizens' interests. Roughly, strong institutions limit a ruler's discretion over policy decisions, while weak institutions *fail to* limit a ruler's discretion over policy decisions. Of particular importance here is the extent of a ruler's discretion over resource revenue. Weak institutions afford rulers wide latitude to use resource revenue as they see fit, whereas strong institutions empower citizens to oversee the use of resource revenue.

Stated generally, the core propositions of the standard institutional story are these:

The economic resource curse: Weak domestic institutions permit a ruler to allocate resource revenue *in ways that inhibit sustainable economic productivity*.

The political resource curse: Weak domestic institutions permit a ruler to allocate resource revenue *in ways that undermine stable executive accountability* to a broad coalition of citizens.

Debate remains over how to specify the italicized phrases. Although political economists disagree about precisely which mechanisms cause the curse, there does seem to

5. Where convenience dictates, I refer to rulers using male pronouns.

be widespread agreement that the curse is a result of various policy failures. This has led to an emphasis on choosing the right policies to avoid the pernicious developmental effects of resource wealth. Initially, development economists tended to focus on various macroeconomic policy failures — lack of sectoral diversification, in particular, lack of investment in internationally competitive manufacturing; overconsumption of resource revenue and failure to smooth spending, thereby increasing budget vulnerability to price volatility; and underinvestment in human capital (e.g., education and health care).⁶ But it was soon noted that macroeconomic policy failure is a *political* outcome. Thus, the resource curse literature has come to focus on the contribution of political factors to the curse. These include unaccountable states with weak bureaucratic structures, which undermine their ability to govern effectively; a political economy rife with corruption and patronage; and lack of transparency in resource sales and revenue spending.⁷ In particular, several papers have focused on the fact that resource wealth decreases the state ruler's need to raise revenue via taxation. This is developmentally deleterious for at least two reasons. Since it is not their money being spent, citizens' incentive and ability to hold rulers fiscally accountable is diminished. The lack of accountability weakens the ties between the rulers and citizens, widening the gap between their respective interests. Investing in development — which is good for citizens — need not be in the interest of rulers. Moreover, the lack of fiscal oversight permits rulers to invest in their personal aggrandizement with impunity — even ignoring the fact that citizens' ability to hold rulers accountable is severely compromised by the latter's capacity to suppress dissent.⁸

6. Richard Auty, ed., *Resource Abundance and Economic Development* (New York: Oxford University Press, 2001); Alan Gelb and Associates, *Oil Windfalls: Blessing or Curse?* (New York: Oxford University Press, 1988); Jeffrey D. Sachs and Andrew M. Warner, "The Curse of Natural Resources," *European Economic Review* 45 (2001): 827–838.

7. Paul Collier, *The Bottom Billion: Why the Poorest Countries are Failing and What Can Be Done About It?* (New York: Oxford University Press, 2007); Jensen and Wantchekon, "Resource Wealth"; Ross, "Does Oil Hinder Democracy?"

8. Cf. Terry Lynn Karl, *The Paradox of Plenty: Oil Booms and Petro-States* (Berkeley and Los Angeles:

This standard institutional explanation gets something fundamentally right: the resource curse is ultimately a story about the incentives generated by the domestic political economy into which resource revenue flows; it is a story about the extent to which domestic political institutions enable rulers to pursue their own interests while disregarding ordinary citizens' rights and well-being. However, standard institutional analyses fail to consider the extent to which the persistence of the weak institutions is endogenous to the resource curse. What if the incentive structure that generates overconsumption and underinvestment, corruption and patronage, unaccountability and repression also inhibits voluntary establishment of the institutional mechanisms required to avoid these ills?

Upon reflection, it is difficult to see how things could be otherwise. Positive institutional reform arises from voluntarily implementing policies that increase rulers' accountability to citizens and thereby secure citizens' rights and improve general economic well-being. If—as the institutional explanation claims—the resource curse arises from an institutional environment that encourages bad policies (i.e., policies that disregard citizens' rights and general well-being), we have little reason to think that voluntary institutional reform is likely under those same institutional conditions.

As noted above, the standard institutional explanation makes a conditional claim: states with weak institutions suffer the curse, while those with strong institutions can avoid the curse. But if the persistence of weak institutions is endogenous to the resource curse, then this conditional claim is too simple. The endogenous institutions thesis implies a more subtle conditional claim:

If a country does not have firmly entrenched strong institutions *before it becomes fiscally reliant on resource revenue*, then its weak institutions will per-

University of California Press, 1997); Hussein Mahdavy, "The Patterns and Problems of Economic Development in Rentier States: The Case of Iran," in *Studies in Economic History of the Middle East*, ed. M. A. Cook (London: Oxford University Press, 1970); Mick Moore, "Political Underdevelopment: What Causes 'Bad Governance'," *Public Management Review* 3, no. 3 (2001): 385–418; Michael L. Ross, "Does Taxation Lead to Representation?" *British Journal of Political Science* 34, no. 2 (2004): 229–249.

sist and it is vulnerable to the resource curse; if a country has firmly entrenched strong institutions *before it becomes fiscally reliant on resource revenue*, then it can avoid the resource curse.

This claim consolidates those kept separate by the distinct economic and political resource curse claims, revealing them to be components of a single story. The economic resource curse strikes when institutions facilitate the misallocation of resource revenue in the domestic political economy; the political resource curse strikes when institutions permit the ruler to allocate revenue in a way that undermines accountability to a broad coalition of citizens. Weak institutions lie at the bottom of both of these phenomena. Institutions that permit rulers to undermine accountability mechanisms facilitate misallocation of resource revenues in the political economy. Hence, the resource curse is embodied by the persistence of weak institutions induced by fiscal reliance on resource revenue.

I make the case for this claim in the remainder of this chapter. This endogeneity claim has important implications for the efficacy of standard prescriptions to address the resource curse. I consider these in the next chapter.

5.2 From Policy Choice to Institutional Design

Extant theoretical models take as given the institutional setting into which resource revenue flows and then consider the effects resource revenue has within the given institutional structure.

Robinson, Torvik, and Verdier model an institutional setting where clientelism is permitted — that is, the ruler is permitted to offer public sector jobs in exchange for political support — and then show that resource booms increase clientelism.⁹ Since the private sector is assumed to be more productive than the public sector, these increases in public sector employment decrease overall economic productivity. Since resource booms improve ex-

9. Robinson, Torvik, and Verdier, “Political Foundations of the Resource Curse.”

traction efficiency and thereby have a positive effect elsewhere in the economy, Robinson et al conclude that institutional obstacles to clientelism play a key role. Institutions that facilitate clientelism make it more likely that resource booms will lower total income, while those that circumscribe clientelism will benefit from resource booms.

Mehlum, Moene, and Torvik distinguish between *grabber friendly* institutions and *producer friendly* institutions — between institutions that encourage rent-seeking activity and those that encourage economically productive activity.¹⁰ To model the extent to which institutions are grabber friendly, they include a term λ that measures producers' average share of the resource rents relative to grabbers' average share; a higher λ means that institutions are producer friendly. Mehlum et al then derive several key results. Not surprisingly, higher quality institutions — more producer friendly institutions — encourage more productive activity than lower quality institutions. Further, as resource revenue increases, institutional quality must also increase to encourage productive activity and discourage rent-seeking behavior. Finally, resource windfalls raise total income when institutional quality is high, but lower total income when institutional quality is low.

Wantchekon models an institutional setting where the state's law enforcement capacity is low, giving the ruler discretionary control over the budget.¹¹ When the risk of armed rebellion is low, the ruler uses resource rents to secure political support rather than invest in economic productivity; when the risk of armed rebellion is high, the ruler invests in repressive technology to secure his rule. Wantchekon attributes an *incumbency advantage* to this control over the budget and uses this incumbency advantage to explain democratic breakdown and autocratic consolidation in resource abundant countries. Wantchekon concludes that institutions that enforce a transparent and rule-oriented distribution of resource rents reduces the ruler's discretionary control over the budget and thereby improve democratic

10. Mehlum, Moene, and Torvik, "Institutions and the Resource Curse."

11. Wantchekon, "Why do Resource Abundant Countries Have Authoritarian Governments?"

stability.¹²

Each of these models treats (changes to) institutional quality as exogenous. They consider how resource revenue induces clientelism, or rent-seeking behavior, or repression within a particular institutional environment. But notice that these types of behavior not only affect development outcomes, but also have a weakening effect on institutions. Clientelistic public sector employment undermines the coherence and competence of the state bureaucracy, which in turn decreases the state's capacity to collect taxes and increases the ruler's discretionary control over the budget. Rent-seeking undermines a rational, principled budgetary process, which facilitates waste and corruption. Repression consolidates a ruler's power and alters the effective channels by which political power can change hands. All this raises a natural question: how does resource revenue affect the institutional environment that facilitates, for example, clientelism, rent-seeking, or repression? Failure to consider this question represents a significant shortcoming of the resource curse literature.

Put simply, extant theory describes the ways in which resource revenue induces policy choices that are deleterious for political and economic development, but it neglects any effect resource revenue might have on the underlying conditions that determine those policy choices. To fill this gap, I present a general model of domestic competition over political outcomes. In the model, groups of domestic constituents use whatever bargaining leverage they have to induce the ruler to choose policies that advance their interests. When constituents lack bargaining leverage, the ruler is free to set policy as he sees fit; when constituents possess bargaining leverage, they are able to constrain the ruler's choices.

My aim is to identify the conditions that determine constituents' bargaining strength. Since relative bargaining strength shapes political choice, the model identifies the underlying strategic conditions to which choices (and, hence, development outcomes) are

12. Cf. Francesco Caselli and Tom Cunningham, "Leader Behaviour and the Natural Resource Curse," *Oxford Economic Papers* 61, no. 4 (2009): 628–650.

sensitive. The model reduces constituents' bargaining strength to two parameters: (1) the extent to which a ruler depends on constituents for (financial, political, military) support to retain political power; and (2) the credibility of supporters' exit threats (roughly, the extent to which they can quit negotiations with the ruler without harming themselves). Having shown that policy selection is sensitive to these parameters, I then theorize about how resource revenue affects them.

One advantage of the model is that it presents a single framework for analyzing the effect of resource revenue on both institutional design and the policy choices made within any particular institutional setting. As I noted earlier, voluntary institutional reforms result from political *choices*. Citizens receive assurance that their ruler will protect their rights and promote their material well-being when institutional mechanisms are established to limit the ruler's discretion over policy. A ruler implements such limitations when he must. To elicit cooperation from citizens, a ruler commits himself to advance citizens' interests by limiting his discretion over policy. Institutional design is here made a choice variable. This permits us to study the effect of resource revenue on institutional design by examining the effect of resource revenue on the conditions that determine political choices more generally.

I present the model informally in the next section, highlighting the conditions that shape political choices. (My reasoning is formalized in appendix A.) The following section theorizes the effects resource revenue has on these conditions and, thus, on political choices.

5.3 A Model of Political Competition

The lesson we learn from the institutional explanation of the resource curse is that the curse is avoided when rulers choose to advance citizens' general interests — when rulers

choose to implement domestic mechanisms that limit their discretion over policy or when rulers choose policies that advance citizens' interests in secure rights and material well-being. The question guiding our examination of the model is this: *Under what conditions do rulers advance citizens' general interests?*

I start with the truism that political choices arise from bargaining between different groups within a polity, identified by their policy goals on a given issue. To gain insight into the conditions under which rulers advance citizens' interests, we must examine the general strategic logic that regulates these bargaining interactions. The literature on institutional development yields two basic insights. The first states that a state ruler's dependence on a reliable source of revenue to consolidate and sustain his monopoly on the use of effective force in a territory leads him to exchange goods for private citizens' provision of revenue. Specifically, rulers provide things like domestic and international security, secure property rights, favorable macroeconomic policies, and political representation in exchange for citizens' tax payments and help with securing loans.¹³ Consolidating one's monopoly on the use of force could be motivated by any number of considerations—the need to secure personal political survival; the need to protect one's assets; the need to secure trade routes; the need to reliably enforce the law. The ruler's motivation is immaterial; he has incentive to seek citizens' resources insofar as he is unable to pay the cost of consolidating

13. This general claim has a long and distinguished pedigree. See, for example, Robert H. Bates and Da-Hsiang Donald Lien, "A Note on Taxation, Development, and Representative Government," *Politics & Society* 14, no. 1 (1985): 53–70; Bates, *Prosperity and Violence*; Margaret Levi, *Of Rule and Revenue* (Berkeley and Los Angeles: University of California Press, 1988); Douglass C. North and Robert Paul Thomas, *The Rise of the Western World: A New Economic History* (Cambridge: Cambridge University Press, 1973); Douglass C. North and Barry Weingast, "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England," *Journal of Economic History* 49, no. 4 (1989): 803–832; Mancur Olson, "Dictatorship, Democracy, and Development," *American Political Science Review* 87, no. 3 (1993): 567–576; Charles Tilly, "War Making and State Making as Organized Crime," in *Bringing the State Back In*, ed. Peter Evans, Dietrich Rueschemeyer, and Theda Skocpol (Cambridge: Cambridge University Press, 1985); Tilly, *Coercion, Capital, and European States*. Many of these sources present case studies—early modern European cases in North and Thomas and Tilly; contemporary African cases in Bates—to illustrate the operative strategic logic.

his monopoly on the use of force from the revenue generated by his personal assets.¹⁴ As Charles Tilly puts it in a study of the origins of the modern European state:

[T]he people who controlled European states and states in the making warred in order to check or overcome their competitors and thus enjoy the advantages of power within a secure or expanding territory. To make more effective war, they attempted to locate more capital. . . . In the long run, the quest inevitably involved them in establishing regular access to capitalists who could supply and arrange credit and in imposing one form of regular taxation or another on the people and activities within their sphere of control. [. . .] Variations in the difficulty of collecting taxes, in the expense of the particular kind of armed force adopted, in the amount of war making required to hold off competitors, and so on resulted in the principal variations in the forms of European states. *It all began with the effort to monopolize the means of violence within a delimited territory adjacent to a power holder's base.*¹⁵

Using “[v]ariations in the difficulty of collecting taxes” and so on, to explain the “principal variations in the forms” of states points to the bargaining dynamic underlying public goods provision. Wherever citizens could make tax collection difficult (e.g., by hiding liquid assets from tax collectors) or wherever internal rivals made consolidating a monopoly on the use of force expensive (e.g., by presenting a credible threat to rebel), the ruler was compelled to make concessions in exchange for cooperation — to provide protection, or to enforce stable property rights, or subject himself to the oversight of citizens’ representatives. Where this was not the case, the ruler could extract from citizens with impunity. We can state the point more generally. Rulers are compelled to limit themselves and provide public goods wherever two conditions are met: first, they depend on the cooperation of some group of citizens to pursue their objectives, whatever those are; second, the citizens on whose cooperation the ruler depends have credible exit options, that is, they can withhold their cooperation without making themselves worse off than they would be were they to cooperate and provide the ruler with revenue.¹⁶

14. The ruler’s objective need not even be the consolidation of his monopoly on the use of force. He has incentive to seek citizens’ resources insofar as he is unable to personally finance the pursuit of his objectives, whatever they are.

15. Tilly, “War Making and State Making,” 172.

16. The reasoning presented here is formalized by Clark, Golder, and Golder, “The Balance of Power.”

The first insight obscures the fact that a ruler's need to respond to the demands of a group of citizens need not produce public goods. Political outcomes depend not only on the bargaining strength of citizens relative to the ruler, but also on the composition of the ruler's support coalition. This leads to the second basic insight: all else equal, the ruler is more likely to provide *public* goods as (1) the number of people on whom the ruler depends for support increases, or (2) the credibility of the supporters' options for backing a leadership rival increases.¹⁷

The logic here is not complicated. Rulers want to retain power. Typically, they require the support of some subset of the population (the "winning coalition") to stay in power; without the loyal cooperation of their supporters, rulers are vulnerable to challenges from leadership rivals. Members of the winning coalition are drawn from a larger subset of the population, the "selectorate." The selectorate is composed of all members of a polity that could potentially become members of the ruler's winning coalition. In a democracy, the selectorate comprises all those who are eligible to vote; in a military junta, powerful military officers; in a monarchy, the nobles. Supporters pledge their allegiance to the leadership candidate who can credibly offer them the best package of benefits. Thus, to retain the loyalty of his supporters, a ruler provides them with a package of benefits that is better than the package they could expect to receive from a leadership rival. When a ruler depends on a small winning coalition, it is most efficient for him to provide his supporters with private goods—monopoly grants, control over tracts of natural resources, public sector jobs, or opportunities for corruption. Since the ruler must retain the support of a

17. See Bueno de Mesquita et al., *The Logic of Political Survival*; for an accessible overview, see Bruce Bueno de Mesquita et al., "Political Competition and Economic Growth," *Journal of Democracy* 12, no. 1 (2001): 58–72. On the importance of the last point, see William Roberts Clark et al., "Why Do Autocrats Overachieve? Political Competition and Material Well-Being in Comparative Perspective" (University of Michigan and Rutgers University, unpublished manuscript, 2010). The last point is qualified in an important way below. For cases illustrating the operation of this logic in post-colonial Africa, see Robert H. Bates, *When Things Fell Apart: State Failure in Late-Century Africa* (New York: Cambridge University Press, 2008).

relatively small group of people, and since the benefits of private goods can be *targeted* (i.e., they accrue only to the holder of the good), private goods provision is a relatively cheap way to provide a high value package of goods to supporters. However, as the size of the winning coalition grows, the ruler's spending must be spread over more supporters and the value of an individual's private goods package diminishes. At some point, the winning coalition becomes large enough that it is more efficient for the ruler to provide *public* goods — security, law enforcement, individual liberties, infrastructure, or investment in human capital (e.g., education or public health)—rather than private goods. Although the benefits of public goods are *nonexcludable* (i.e., they accrue to everyone in the polity if they accrue to anyone) and so cannot be targeted to supporters, there are economies of scale in providing public goods, which enables a ruler to provide his supporters with a higher value package of public goods than the package of private goods that could be provided to each supporter for the same amount of total spending. Thus, a ruler becomes more inclined to provide public goods as the size of his support coalition grows, all else equal.

Similarly, in the absence of reliable sources of non-tax income (e.g., natural resources or foreign aid), a ruler becomes more likely to provide public goods as the credibility of his supporters' threats to back a rival increase, even if the absolute size of the winning coalition is small. More credible threats translate into greater bargaining leverage for supporters, which enables them to demand a greater package of private goods in exchange for their support. To meet the increasing demands of his supporters, a ruler must raise revenue to finance private goods provision. Without reliable sources of non-tax income, the ruler must turn his attention to increasing tax revenue. This means increasing the number of taxpayers or increasing the income of his tax base. A ruler can do both by fostering broad-based economic growth, which is most effectively and efficiently done by

providing public goods rather than private goods. Consequently, absent reliable sources of non-tax revenue, a ruler becomes more inclined to provide public goods — albeit as a means to finance private goods provision — as the credibility of his supporters’ threats to back a rival increases, all else equal.¹⁸

We can summarize the overall picture we get from these two basic insights as follows. Wherever a ruler requires the (political, financial, military) support of a subset of the population to retain power and those supporters gain bargaining leverage from credible exit threats (e.g., asset mobility, private militias, credible leadership rivals), the supporters will be able to extract favorable concessions from the leader. Whether the benefits of those concessions accrue only to individual supporters or to citizens more widely depends on the composition of the ruler’s support coalition — whether he only requires the support of a small group of elites or of a broad coalition of citizens — and the extent to which the leader must stimulate economic growth to finance his provision of goods.

This general picture emphasizes the importance of citizens’ bargaining strength relative to the ruler in shaping political choices. But it neglects the importance of both the *relative* bargaining strength of *distinct* groups of citizens vis-à-vis the ruler and the *relative* value of their loyalty to the ruler. A ruler’s choices depend not only on the bargaining strength of the ruler’s supporters relative to the ruler, but also the bargaining strength of supporters vis-à-vis the ruler *relative to the bargaining strength of other groups of citizens*. Similarly, a ruler’s choices depend not only on the value of supporters’ loyalty to the ruler, but also the value of supporters’ loyalty to the ruler *relative to the value of other groups’ loyalty to the ruler*.¹⁹ Thus, to enhance our understanding of the conditions under which rulers advance citizens’ interests, we must examine the logic of political choice when the

18. See Clark et al., “Why Do Autocrats Overachieve?”

19. Cf. “Not only the ruling class, but *all classes whose resources and activities affected the preparation for war*, left their imprint on European states” (Tilly, *Coercion, Capital, and European States*, 27; my emphasis).

ruler must negotiate with multiple groups, not simply with his supporters.

To incorporate this insight, suppose a ruler must now negotiate with two groups, an elite class and a class of ordinary citizens.²⁰ Suppose further that the ruler provides public goods only in exchange for citizens' cooperation; if he attempts to secure elites' cooperation, he provides private goods.²¹ He keeps everything for himself if he requires neither elites' nor citizens' cooperation. Finally, suppose that public goods provision advances citizens' interests generally, whereas private goods provision advances only the interests of their recipients and detracts from the interests of non-recipients.

An actor has *valuable loyalty* just in case the ruler prefers to concede to her control of the prize over which they are bargaining rather than retain control of the prize for himself. An actor has a *credible exit option* just in case she can quit the negotiation with the ruler without making herself worse off than if she had continued and received the worst possible bargaining outcome. Examples, alluded to above, include the ability to move one's assets to avoid appropriation by the ruler, whether by taxation or theft; the ability to provide personal security and resist the ruler's demands with a private militia; the ability to emigrate; or the ability to back a leadership rival in exchange for a greater package of goods.

The strategic logic that governs the ruler's choice under these circumstances is the same as above. So it remains true that he must make concessions to any group whose support he requires to retain power and who have credible exit threats. Although they remain necessary conditions, citizens' possession of valuable loyalty and credible exit threats are no longer *sufficient* to produce public goods, as they are above. To avoid bad outcomes — outcomes where citizens' interests are harmed — it must also be true that either the elites have no credible exit options or their support is less valuable to the ruler for

20. The following reasoning is presented formally in appendix A.

21. This assumption is required for the model to produce any interesting variation; if the ruler provides public (or private) goods regardless of his supporters' identity, the logic underlying the ruler's choice is obscured.

achieving his objectives than is the citizens'. If elites have credible exit options and their support is more valuable to the ruler than the citizens', then even if citizens have credible exit options and valuable loyalty, the ruler will provide private rather than public goods.

Summing up: Rulers advance citizens' interests when two conditions are met. The first is a disjunctive requirement: either the ruler depends on the support of a broad coalition of citizens and elites have no exit options; or the leader depends *more heavily* on the support of a broad coalition of citizens to achieve his objectives than on the support of a small elite group. Second, the citizens whose support is required have credible exit options. When these conditions are satisfied, state rulers are constrained to advance citizens' general interests.²²

Obviously, the model is schematic in an important sense: it grossly oversimplifies the dynamics of political choice by reducing actors and demands to simple parameter values. There are many ways we could complicate the model: we could make the prize divisible; elites' demands could induce the provision of public goods (as in early modern England or postcolonial Botswana); rulers could provide citizens with private goods (as in Saudi Arabia or Venezuela); we could add civil society groups to the bargaining table or allow individuals to be members in multiple groups; the ruler could seek cooperation from multiple groups in civil society. We haven't even discussed all the ways we could complicate the sequence of moves made by the actors. All these amendments would make the model much more faithful to the complex situations we actually observe.

These amendments would also make the model considerably less tractable and thereby undermine its point. The aim of the model is not to adequately capture the reality of political choice or even to accurately predict actual outcomes. I use the model to trace the strategic logic underlying political choice; the assumptions I make serve this purpose

22. The claim made in this paragraph corresponds to proposition A.8 in the formal appendix.

alone. Actors and demands are reduced to parameters so that we can identify the parameters to which development outcomes are sensitive, as well as the ways in which outcomes are sensitive to the salient parameters. Once we precisely trace the logic, we can better anticipate how changes in the parameters will affect the outcome. For example, using the logic identified by the model, we can expect that if elites' interests were best served by a mix of private and public goods and they had the requisite bargaining leverage, then the ruler would provide a mix of private and public goods, which would produce a mix of development successes and failures. Or suppose the ruler depends (in different ways) on both the elite and citizens and both groups have credible exit options. We could anticipate that the ruler would divide the prize between the elites and citizens to placate them both.

My point here is that the model acquires a certain amount of flexibility in virtue of its simplicity. We can apply the model's insights to a wide and complex range of situations *because* it is so simple.

5.4 Resource Revenue and Political Choice

How does resource wealth affect the logic of political choice? The most straightforward effect is to reduce the ruler's need to "earn" his income by reducing his reliance on tax revenue. In Mick Moore's words, "[s]tate revenue can be considered 'earned' to the extent that the state apparatus has to put in organizational and political effort in working with citizens to get its money."²³ Simply, resource wealth reduces the ruler's need to put in effort to elicit cooperation from citizens to get his money. The importance of this point has been repeatedly stressed in case studies of the Arab gulf states (Bahrain, Kuwait, Qatar,

23. Moore, "Political Underdevelopment," 389.

Saudi Arabia, United Arab Emirates),²⁴ Iran,²⁵ and Venezuela,²⁶ as well as more general studies of the link between taxation and democratization.²⁷

When the ruler no longer depends on citizens as a source of revenue, he no longer needs to negotiate with citizens to secure their financial cooperation. The ruler retains discretion over the budget, which facilitates revenue misallocation throughout the economy.²⁸ The ruler can use resource revenue to finance consolidation of his rule and so fend off leadership challenges without requiring cooperation from citizens (at the polls, in arms, etc.). In terms of the theory sketched above, a ruler flush with resource revenue no longer depends on the citizens' fiscal support to finance his attempts to retain power. Hence, he need not establish institutional mechanisms that grant citizens oversight of the policy formation process, such as representative bodies or a transparent budgetary process. More perniciously, where a ruler does not depend on the electoral support of a broad coalition of citizens to stay in office, he is freed from having to use resource revenue to secure wide electoral support. This permits him to retain more of the revenue for himself with impunity.

Relatedly, resource wealth frees the ruler from the need to foster economic growth to meet demands for private goods. As noted above, when the ruler relies on tax revenue to finance private goods provision, it is possible that the benefits of concessions to a small group of elites accrue to citizens more widely.²⁹ For example, as North and Weingast ar-

24. Lisa Anderson, "The State in the Middle East and North Africa," *Comparative Politics* 20, no. 1 (1987): 1–18; Hazem Beblawi, "The Rentier State in the Arab World," in *The Rentier State: Nation, State and the Integration of the Arab World*, ed. Hazem Beblawi and Giacomo Luciani (London: Croom Helm, 1987); Kiren Aziz Chaudry, *The Price of Wealth: Economies and Institutions in the Middle East* (Ithaca and London: Cornell University Press, 1997).

25. Mahdavy, "Patterns and Problems."

26. Karl, *The Paradox of Plenty*.

27. Ross, "Does Oil Hinder Democracy?"; Ross, "Does Taxation Lead to Representation?" Herb registers tentative disagreement; see Michael Herb, "No Representation without Taxation? Rents, Development, and Democracy," *Comparative Politics* 37, no. 3 (2005): 297–316.

28. Cf. Wantchekon, "Why do Resource Abundant Countries Have Authoritarian Governments?"

29. See Clark et al., "Why Do Autocrats Overachieve?"

gue, if elites interests are served by concessions that foster widespread economic growth, such as secure property rights and the right to oversee taxation, then eventually citizens more generally come to benefit from the ruler's responsiveness to elites.³⁰ But this possibility is diminished by fiscal reliance on resource wealth, which tends to encourage rent-seeking rather than productive activity on the part of elites.³¹ The benefits demanded by elites in the context of resource dependence tend to be private rather than public in nature. In fact, if the absolute value of the resource revenue is large enough, rulers might preempt citizen discontent and foster loyalty by providing private goods to citizens-at-large, such as direct cash transfers (the Arab gulf states) or transfers in kind (Venezuela). Moreover, the availability of resource revenue enables the ruler to meet these demands without needing to raise tax revenue and, thus, without needing to provide productivity-enhancing public goods. Without the need to foster economic growth, the ruler has little incentive to implement a rational budget process or establish a coherent development strategy.

Finally, resource wealth affects the availability of exit options. Oil wells and copper mines can't be moved. Consequently, when the ruler neglects the interests of resource holders, the latter can't force policy concessions by threatening to take their enterprise elsewhere.³² Thus, even if their support is essential for political survival, resource holders are vulnerable to state predation due to a lack of credible exit options.

In general, the state's fiscal reliance on resource wealth generates a curse because it permits a ruler to rule *autonomously*, that is, to rule without having to consult citizens'

30. North and Weingast, "Constitutions and Commitment."

31. See Jean-Marie Baland and Patrick Francois, "Rent-Seeking and Resource Booms," *Journal of Development Economics* 61 (2000): 527–542; Aaron Tornell and Philip R. Lane, "The Voracity Effect," *American Economic Review* 89, no. 1 (1999): 22–46; Ragnar Torvik, "Natural Resources, Rent Seeking, and Welfare," *Journal of Development Economics* 67 (2002): 455–470. See Bates, *When Things Fell Apart* and Karl, *The Paradox of Plenty* for illustrative examples.

32. It has been suggested to me that resource holders could threaten to destroy the resource stock. But this is not a credible threat. Consider the fact that foreign oil companies in Venezuela continued to do business despite losing 78% of their profits to the state by 1970! (ibid., 112.) Even if resource holders are left with little after state appropriation, that's still more than nothing.

interests. The theory here comports with the central findings of the resource curse literature.³³ When a ruler no longer needs citizens' money, he has no incentive to increase his tax base by fostering economic growth. This frees the ruler to use the revenue to pursue his personal aggrandizement or politically motivated "white elephant" projects rather than invest in human capital accumulation or sectoral diversification. Nor does he have any incentive to develop a rational and meritocratic bureaucracy, since he has no need to effectively collect taxes. This increases the ruler's discretionary power in policy-making, while freeing him to use public sector jobs as patronage to buy support. When a ruler no longer counts on citizens' political support, his performance is no longer subject to citizen oversight. This frees the ruler to pursue his own objectives by any means necessary with impunity.

So far, the story is a familiar one. But the model I have presented shows that the resource curse is more pernicious than previously thought. What makes the resource curse so difficult to overcome is the fact that the autonomy afforded a ruler by resource wealth undermines any incentive he might have *to implement institutional reforms that limit his autonomy*. The resource curse is avoided when state rulers are constrained by a strong institutional structure, by institutional mechanisms that insure that rulers use resource revenue to finance initiatives that advance citizens' general interests, and enable citizens to remove them in the case of failure to do so. Historically, the development of strong institutional structures is the result of protracted bargaining between rulers and citizens — typically carried out over decades, even centuries — the need for which arises from rulers' need to obtain citizens' cooperation to fend off leadership challenges. By undermining a ruler's need to negotiate with citizens to consolidate his rule, resource wealth undermines a central impetus for positive institutional change. This is why it makes sense to refer to the

33. See the discussion in section 5.2.

resource curse as a *trap*. Once resource revenue is introduced into a weak institutional environment, it inhibits implementation of the institutional reforms needed to overcome the negative developmental effects of resource wealth. In other words, the weak institutions that engender the resource curse are endogenous to the causal process precipitating the curse.

The central lesson is this: the resource curse occurs not because the incentives generated by the domestic political economy encourage negligent, corrupt, and sometimes violent predators to acquire and wield power to the detriment of citizens, but because the domestic political economy generates incentives for *anyone who ultimately controls resource revenue* to rule in a negligent, corrupt, and sometimes violent fashion. More to the point, the resource curse strikes whenever resource revenue flows into a domestic political economy with a weak institutional structure. Under such conditions, it matters little who holds power or how they acquired it. Once in office, the lack of institutional constraints permits the ruler to rule autonomously, which subsequently permits him to rule without regard for citizens' general interest in secure rights and material well-being.

5.5 Empirical Implications

According to the theory I have presented, the following claim is true:

If a country does not have firmly entrenched strong institutions *before it becomes fiscally reliant on resource revenue*, then its weak institutions will persist and it is vulnerable to the resource curse; if a country has firmly entrenched strong institutions *before it becomes fiscally reliant on resource revenue*, then it can avoid the resource curse.

If this claim is true, what should we observe empirically? Contrary to a recent paper by Haber and Menaldo, the theory does not imply a systematic relationship between fiscal reliance on resource revenue and the quality of governance; an increase in fiscal reliance on resources need not precipitate a decline in institutional quality and a decrease in fiscal

reliance need not precipitate improvements in institutional quality.³⁴ The theory I have presented is consistent with multifarious observable relationships. Following an increase in the state's fiscal reliance on resource revenue, institutional quality could: (1) remain the same; (2) decline; (3) improve temporarily or sporadically but consistently return to lower levels; (4) improve but remain below some threshold level of strength (in the case of weak institutions); (5) perhaps even improve above some threshold level of strength (in the case where strong institutions preceded the increased fiscal reliance). Hence, against Dunning's claim to the contrary, an institutional theory of the resource curse can account for democratic persistence following increases in resource revenue.³⁵ If the theory is consistent with so many observations, is there any way to tell whether the theory is wrong?

The theory does imply clear counterexamples; *it tells us what we should definitely not observe*. If the theory is correct, then if weak institutions precede a state's fiscal reliance on resource revenue, we should not observe stable and consistent improvement in institutional quality in the face of sustained fiscal reliance on resources. Do we observe such counterexamples?

As a start, consider Haber and Menaldo's classification of fiscally reliant states — those whose governments received an average of 5% of their revenues from resources from 1972 to 1999. Suppose we take the Polity regime measure — an index constructed from components measuring the competitiveness of political participation, the openness and competitiveness of executive recruitment, and the constraints on the executive's policy making power — as a good proxy for institutional strength.³⁶ In Haber and Menaldo's analysis, a

34. See Stephen Haber and Victor Menaldo, "Do Natural Resources Fuel Authoritarianism? A Reappraisal of the Resource Curse," *American Political Science Review* 105, no. 1 (2011): 1–26.

35. See Thad Dunning, *Crude Democracy: Natural Resource Wealth and Political Regimes* (New York: Cambridge University Press, 2008), 54.

36. On the Polity measure, see Monty Marshall and Keith Jagers, "Polity IV Project: Political Regime Characteristics and Transitions, 1800–2006" (University of Maryland, unpublished manuscript, 2008). We will have reason to reject this as a measure of institutional strength later.

country is democratic if it scores 7 or above on the Combined Polity 2 measure.³⁷ According to their classification, there are seven countries that “democratized during or after a resource boom”: Botswana, Ecuador, Mexico, Mongolia, Peru, Russia, and Venezuela.³⁸ Of these, Botswana is misclassified using their own coding rules; it received a Polity 2 score of 7 or above from 1969 through 2007, but only became fiscally reliant on diamond revenue during the early 1970s. Of the remaining cases, two of them have been classified as democratic for a comparatively short time — Mexico (since 2000) and Russia (since 2001) — and a third, Peru, has not been stably democratic (1980–1991, 2000–present). As noted earlier, the theory is consistent with temporary or sporadic institutional improvements; it implies that fiscal reliance on resource revenue inhibits *sustained* institutional improvement. Hence, although we should not exclude these cases as potential counterexamples, we should reserve judgment on them until they have been stably democratic for a longer period.

We are now left with three strong candidate counterexamples to the theory: Ecuador, Mongolia, and Venezuela. Of these, Venezuela presents the most striking potential counterexample to my theory. Regarding fiscal reliance, Venezuela has sustained much higher levels of reliance on resource revenues for much longer than either Ecuador or Mongolia. By 1932, oil provided more than 30% of government revenues and its mean fiscal reliance on oil was 53% through 2007. Ecuador, by comparison, started receiving resource revenues in 1972, averaging nearly 29% through 2005. Alternatively, Venezuela’s total resource income averaged \$1693 per capita from 1930 to 2007, dipping slightly below \$1000 per capita only twice since 1946. Mongolia, by comparison, averaged \$242 per capita from 1957 (its first year receiving resource revenue) to 2006, topping \$500 per

37. The original Polity 2 measure ranges from –10 (full autocracy) to 10 (full democracy). Haber and Menaldo normalize it to range from 0 to 100. Hence, in their data, a country is democratic if it scores 85 or above. See Haber and Menaldo, “Do Natural Resources Fuel Authoritarianism?” 4, 5.

38. See *ibid.*, 6.

capita only twice (2005, 2006).³⁹

Regarding institutional quality, Venezuela had been held up as a model for other Latin American countries, hailed for the stability of its democratic institutions from the time it emerged from military rule in 1958 until the political crisis experienced in the late-1990s. By comparison, Mongolia has only been classified as a democracy for a comparatively short time (since 1992), while Ecuador's democratic run lasted roughly 21 years, showing signs of instability beginning in 2000.

In sum, Venezuela has sustained higher levels of fiscal reliance on resource revenue, for a longer time, while sustaining stable democratic institutions for longer than either Ecuador or Mongolia. In addition, Venezuela established democratic institutions after a longer period of relatively high resource dependence (roughly 20 years) than either Ecuador (7 years) or Mongolia (12–15 years). Given all this, the theory implies that if any country should have succumbed to the resource curse, it should have been Venezuela. Venezuela should not have been able to sustain democratic institutions for so long given the depth and duration of its resource dependence prior to and after the emergence of democratic institutions.

Examining the Venezuelan case is thus a good preliminary test for my theory. Specifically, we must probe the Venezuelan case in more detail for evidence that the mechanisms posited by my theory were not at work; in particular, we must examine whether Venezuelan institutions were strong in the sense defined above — namely, that they empowered a broad coalition of citizens to effectively check the ruler's discretion over policy formation. If this turns out to be true, we have good reason to at least modify the theory I have presented above.

39. Statistics derived from Haber and Menaldo's dataset, received from authors. Fiscal reliance data is not available for Mongolia.

5.6 Venezuela: A Test Case

Is Venezuela a genuine counterexample? Was it able to establish strong institutions despite its heavy resource dependence prior to and after democratization? Reflection on Venezuela's institutional history suggests not. Scholars seem to agree that Venezuela's history has been marked by extreme presidentialism. Even during its democratic period, Venezuela's institutions have granted the chief executive wide, and in some cases unchecked, discretion over the formation and implementation of policy. I summarize the key points here.

After achieving independence from Spain in 1811, Venezuela was ruled by decentralized and competing *caudillos*, strongmen who commanded personal militias. State power was finally centralized under military rule in 1899. Military rule lasted until 1958, with a brief period from 1945–1948 — known as the *trienio* — when Venezuela was governed by a popularly elected party, Acción Democrática (AD). When the dictator Marcos Pérez Jiménez was overthrown by a military coup in 1958, the ruling junta invited the major political parties — AD, the Comité de Organización Política Electoral Independiente (COPEI), and the Unión Republicana Democrática (URD) — to form the government. Given a second chance, the political parties established a broad consensus for the direction of the new democracy. The Pact of Punto Fijo, the Minimum Program of Government, and the Worker-Employer Accord were signed prior to elections in 1958 by the interested parties — in addition to the three political parties were FEDECAMARAS and the CTV, the national business lobby and labor union, respectively. Accordingly, Venezuela's democracy has come to be known as a *pacted democracy*; the parameters for future decisions — those concerning development strategy in particular — were set out by a bargain among political and economic elites.⁴⁰

40. For further details on these pacts, see Brian F. Crisp, *Democratic Institutional Design: The Powers and*

These pacts were borne of the demise of the trienio period. According to Brian Crisp, the AD was overthrown in 1948 because it failed to cultivate consensus among competing civil society groups, especially business and labor.⁴¹ Bringing the key interests together to agree on a development strategy was important for consolidating the fledgling democracy. But these pacts also secured elite privilege. In Terry Lynn Karl's words, the pacts "ensure[d] [the parties'] survival by selectively meeting demands while limiting the scope of representation in order to reassure traditional dominant classes that their vital interests will be respected."⁴² It is true that subsequent governments undertook massive social initiatives, creating jobs and providing social services to the poor. But state institutions were designed to benefit the elite disproportionately.⁴³

The key to this institutionalization of privilege was centralizing political power in the executive branch.⁴⁴ Although the Constitution of 1961 did not formally entrench the president's dominance, this became the case in practice. Two key powers were afforded the president. First, the president had the authority to suspend constitutional guarantees and rule by decree. Economic liberties were suspended throughout the period from 1961 to 1999 (when a new constitution was drafted), permitting the president to set economic and financial policy by decree throughout the period. (The legislature had the authority to reinstate suspended guarantees, but never reinstated economic liberties.) Second, the president had the authority to create consultative commissions and the state agencies that compose the decentralized public administration (DPA). This included unrestrained authority to appoint the members of these commissions and agency governing boards.

Typically, the president created a commission to research policy solutions to a particu-

Incentives of Venezuelan Politicians and Interest Groups (Stanford: Stanford University Press, 2000), 25ff and Karl, *The Paradox of Plenty*, 94–101.

41. Crisp, *Democratic Institutional Design*, 23.

42. Karl, *The Paradox of Plenty*, 93.

43. This point is repeatedly stressed by both Crisp and Karl.

44. The following draws primarily from Crisp, *Democratic Institutional Design*, chs. 3–5.

lar problem and then proposed draft legislation on the basis of their recommendations. Commissions were dominated by federal government officials, and representatives of private capital (FEDECAMARAS) and organized labor (the CTV). Although commissions institutionalized consultative relations between the state and civil society, only a narrow segment of civil society was represented. In Crisp's words:

Consultative commissions are used to open the political system to some participants while keeping it closed to others. Bureaucratic institutions are not simply neutral means of making decisions but instead are tools used by the powerful to control and channel conflict.⁴⁵

The state agencies that compose the DPA were the president's key instruments for implementing policy. Each agency was created with funds from the executive (financed by oil revenues) and members of the governing board were appointed by the president. Like the commissions, these were also composed primarily of federal officials and representatives of capital and labor. Most state activity was carried out by these agencies; by 1980, state agencies accounted for 70% of government spending.⁴⁶ Yet agencies' budgets were not subject to congressional oversight.

Although the legislative branch had formal powers to check the executive, the electoral system undermined its efficacy as a constraint on executive decision making. Until 1993, Venezuela was a single-district, closed list proportional representation electoral system. Voters voted for a party list. Parties were allotted seats in the legislature according to the percentage of the popular vote they received. Legislators were chosen sequentially from a list made by the party until the party's allotted seats were filled. The list was closed (voters couldn't alter order of candidates) and blocked (voters couldn't add candidates to the list). Importantly, this meant that party officials were more important for the legislators' political survival than the voters. This permitted the parties to exercise strict discipline

45. Crisp, *Democratic Institutional Design*, 119.

46. *Ibid.*, 68.

over legislators. This crucially limited voters' exit options by relegating legislators to a peripheral role in policy making and granting them little effective oversight over the president. Exit options were further limited by the pacts signed in 1958, which determined that the state's development strategy was to be the same regardless of the party in power. If voters were disappointed with the party in power, they could change their affiliation, but they could not change the country's development strategy.

In short, from 1958 to 1999, Venezuela's institutions were such that the president was usually the only elected official involved in the formation and implementation of policy. The president exercised discretion over appointments to consultative commissions and state agency governing boards, and most state spending was beyond citizen oversight.

The concentration of political power in the executive has only become more extreme since Hugo Chávez took office in 1998. Shortly after taking office, Chávez convened a Constituent Assembly to revise the 1961 constitution, a procedure that went beyond the provisions for constitutional review outlined in the 1961 constitution. Moreover, Chávez was able to implement his preferred procedure for electing Assembly members, which resulted in an Assembly dominated by Chávez supporters.⁴⁷ The new constitution drafted by the Assembly was ratified in a nationwide referendum on December 15, 1999. One week later, the Assembly went beyond the newly ratified constitution and adopted the Decree for a Transitional Regime. It subsequently dissolved all branches of government, including the national Congress, state legislatures, and state governors, appointing new officials in place of those elected during the 1998 popular elections. Crucially, the Assembly dissolved the historically independent Supreme Court of Justice and Electoral Council, creating new bodies and appointing new members in their place.⁴⁸ In effect, Chávez gained complete

47. Allan R. Brewer-Carías, *Dismantling Democracy in Venezuela: The Chávez Authoritarian Experiment* (New York: Cambridge University Press, 2010), 11–12.

48. *Ibid.*, 13–23.

control of the country's key institutions. This control was bolstered in 2007 when Chávez used his decree authority to implement constitutional reforms that had been rejected in a popular referendum earlier that year.⁴⁹ Since 1999, Chávez has used his control of these institutions to harass and exclude opponents. With the creation of the PSUV (Partido Socialista Unido de Venezuela/United Socialist Party of Venezuela) and its assimilation into the state's bureaucratic apparatus, traditional party pluralism has effectively been replaced with single-party government.

It is true that Chávez has been elected to office on three occasions (1998, 2000, 2006) and has survived a recall referendum (2004). Despite this, Freedom House has declared that "Venezuela is not an electoral democracy. While the act of voting is relatively free and the count is fair, the political opposition is forced to operate under extremely difficult conditions, and the separation of power is nearly nonexistent."⁵⁰ In Javier Corrales' words, Chávez has "eliminated the contradiction between autocracy and political competitiveness."⁵¹

This brief review suggests that Venezuela's institutions have been persistently weak; thus, it is not a counterexample to the theory I have presented.

Did fiscal reliance on oil have anything to do with the persistent weakness of Venezuela's institutions? Answering that question gives us an opportunity to probe the mechanisms posited by my theory.

Oil was first discovered at Lake Maracaibo in 1901. Between 1920 and 1925, oil's share of total exports went from 1.9% to 41.6% and hit 91.2% by 1935.⁵² From 1980

49. Brewer-Carías, *Dismantling Democracy in Venezuela*, 381.

50. Freedom House, *Freedom in the World—Venezuela Country Report*, 2010, URL: <http://www.freedomhouse.org/template.cfm?page=363&year=2010>.

51. Javier Corrales, "Hugo Boss," *Foreign Policy*, no. 152 (2006): 33. See also Javier Corrales and Michael Penfold, *Dragon in the Tropics: Hugo Chávez and the Political Economy of Revolution in Venezuela* (Washington, D.C.: Brookings Institution Press, 2011).

52. Karl, *The Paradox of Plenty*, 80.

to 2009, oil's share of total exports remained high, averaging 89%.⁵³ As noted above, oil's share of government revenues has averaged 53% since 1930. From 1964 to 1984, Venezuela sustained its heaviest reliance on oil, with oil providing an average of 66% of government revenues, reaching a peak of 83% in 1974.

Oil's centrality to the Venezuelan economy was secured by the 1943 Hydrocarbons Act. This bill established the "fifty-fifty" principle: the state must obtain at least 50% of production profits.⁵⁴ The state simultaneously moved away from taxing citizens. (Attempts to increase taxes on anything other than oil extraction were persistently rejected until the early 1990s, when the economy was already in severe decline.⁵⁵) This law "established a permanent temptation to cut into the profits of foreign companies as a means of sustaining oil-subsidized activities while avoiding the taxation of domestic groups."⁵⁶

The pacts signed in 1958 further consolidated oil's central role in the economy, with the interested parties agreeing to an oil-based development strategy. Committed to "sowing the petroleum," subsequent administrations persistently milked foreign oil companies for revenue; by 1970, the state retained 78% of profits.⁵⁷ The oil industry was finally nationalized in 1976. Administration after administration used oil revenue to prop up an overvalued currency and keep taxes low. Oil also financed massive social spending programs, permitting the parties to avert popular discontent and foster patronage, a strategy Karl calls "preemptive inclusion."⁵⁸ But most importantly, oil revenues facilitated the creation of hundreds of state agencies, each with their own patrimony. Because agency budgets were not subject to congressional oversight, and because governing board appointments

53. Data from WTO statistics database, URL: <http://stat.wto.org/Home/WSDBHome.aspx?Language=E>, accessed May 2011.

54. Dunning, *Crude Democracy*, 160–161; Karl, *The Paradox of Plenty*, 87.

55. Dunning, *Crude Democracy*, 164.

56. Karl, *The Paradox of Plenty*, 88.

57. *Ibid.*, 112.

58. *Ibid.*, 93, 104.

were discretionary, presidents retained unchecked discretion over the use of agency funds. Thus, oil-financed state agencies “gave public authorities a set of tools with which to shape economic development, subsidize private capital, pay off political allies, and co-opt enemies.”⁵⁹ Huge oil revenues enabled the president to placate potential opposition while autonomously pursuing a development strategy that disproportionately benefited an elite few. This trend continued unabated during the implementation of IMF austerity measures in the early 1990s and has only become exacerbated during Chávez’s tenure.⁶⁰

Addiction to oil revenues also precipitated a severe fiscal crisis in the late 1980s and early 1990s; this ultimately led to the destabilization of pacted democracy, culminating in Chávez’s rise to power.⁶¹ Beginning in 1973, high oil prices fueled a massive increase in government spending, nearly tripling from 1973 to 1974 and then doubling again from 1979 to 1981.⁶² Huge capital-intensive projects received most attention; the government rapidly expanded the state-run steel, aluminum, and petrochemical industries. The newly acquired oil industry also required significant investment, as oil companies had ceased investment and exploration sometime in the late 1960s in anticipation of nationalization. Awash in petrodollars, state enterprises paid little attention to efficiency measures; waste was extensive.⁶³

When oil prices collapsed in the early 1980s, the failure to cultivate broad tax extraction capacity left it without a domestic revenue base. So the government tapped the international credit markets. With deep oil reserves as collateral, loans came fast and easy.

59. Crisp, *Democratic Institutional Design*, 36.

60. See Brewer-Carías, *Dismantling Democracy in Venezuela*; Corrales and Penfold, *Dragon in the Tropics*; Michael Shifter, “In Search of Hugo Chávez,” *Foreign Affairs* 85, no. 3 (2006): 45–59; Kurt Weyland, “Will Chavez Lose His Luster?” *Foreign Affairs* 80, no. 6 (2001): 73–87.

61. “Chávez’s appeal cannot be explained without acknowledging the deep dissatisfaction with the existing political and economic order felt by much of the population in Venezuela” (Shifter, “In Search of Hugo Chávez,” 46).

62. See the table at Karl, *The Paradox of Plenty*, 165.

63. *Ibid.*, 111, 125, 140.

They had the added advantage of enabling the parties to avoid politically unpopular tax reforms. A prolonged period of loose spending in boom times had triggered a voracious demand for more spending; expenditures increased even as revenues decreased.⁶⁴ Oil had cultivated unsustainable expectations; Venezuelan institutions were too weak to rein in the rent-seeking, waste, and corruption. Facing a severe debt crisis, Venezuela was forced by the IMF to implement tough austerity measures in the face of widespread domestic opposition. Following coup attempts in February (led by Hugo Chávez) and November 1992, the incumbent, Carlos Andrés Pérez, was removed from office in August 1993 under charges of embezzling \$17 million of government funds.

There are two key points to extract from this case study. First, Venezuelan rulers have historically exercised vast and often unchecked discretion over the formation and implementation of policy, even throughout its democratic period. Thus, Venezuela's institutional trajectory is consistent with the theory I have presented here; it is not a counterexample. Second, oil dependence helped to consolidate presidential dominance through the democratic period and beyond. The practice of preemptive inclusion is evidence that the parties required citizens' cooperation to some extent; Venezuela's history of mass protests, general strikes, and sporadic civil violence ensured this. But declining reliance on general tax revenue enabled the president to avoid congressional control of the budget. With the creation of fiscally autonomous state agencies, the president gained discretionary control over a significant percentage of government revenues. Inevitably, the president became the target of influence trafficking. With so little oversight over spending, misallocation of petrodollars, waste, and corruption were all but guaranteed.

Adequately testing my theory requires systematically examining additional candidate counterexamples in greater depth. I must put this off for another time.⁶⁵ But the above

64. See the table at Karl, *The Paradox of Plenty*, 165.

65. Preliminary research suggests we can exclude Russia as a counterexample too. M. Steven Fish de-

study reveals an important lesson for future identification of potential counterexamples. The resource curse is not about a lack of *democratic* institutions per se, but a lack of institutions that constrain rulers' discretion over the formation and implementation of policy. Democratic institutions are often strong institutions, as I define them. But, as the Venezuelan case demonstrates, they need not be; formally democratic institutions can be consistent with a largely unconstrained executive in practice. This cuts the other way too; strong institutions need not be democratic. Hence, potential counterexamples are not necessarily those countries that democratize during or following increased reliance on resource revenue. Instead, we must look for countries that appear to have begun with weak institutions prior to becoming resource dependent and subsequently established strong institutions. I leave this for future research.

scribes Russia as a "superpresidentialist" state and argues that resource revenue has sustained weak institutions, facilitating revenue misallocation and corruption. See M. Steven Fish, *Democracy Derailed in Russia: The Failure of Open Politics* (New York: Cambridge University Press, 2005).

CHAPTER 6

The Resource Curse and the Demands of Justice

The resource curse occurs when resource revenue enters a country with weak domestic institutions — institutions that leave state rulers free to form and implement policy without regard for citizens' general interest in secure rights and material well-being. Roughly, rulers are free to neglect citizens' interests when they do not rely on citizens' support to retain political power or citizens cannot credibly threaten the ruler to withhold cooperation. If a country has weak institutions prior to becoming resource dependent, resource revenue undermines rulers' reliance on citizens' support, perpetuating weak institutions.

The claim that the persistence of weak, curse-inducing institutions is endogenous to the resource curse has important implications for an effective prescriptive program to address the curse. The prescriptions that emerge from standard institutional analyses of the resource curse tend to focus on domestic institutional reform. The thought supporting standard prescriptions is a natural one: if the curse arises from domestic institutional failures, then we could encourage policies that avert the resource curse if we improved domestic institutions in the relevant ways. For example, resource-cursed states that exhibit low ruler accountability should establish mechanisms that permit citizens to oversee resource revenue management; states that are vulnerable to price volatility as a result of short-sighted budgeting practices should stabilize expenditures to avoid the negative development ef-

fects of boom-bust cycles; states that overconsume should instead invest in increasing productive capacity in non-resource sectors and in human capital accumulation.¹ This line of reasoning is further encouraged by examples of resource abundant countries who have managed to avoid the curse — Canada and Norway are prime examples. What these countries have in common are strong institutions that encourage government accountability and, consequently, policies that advance citizens’ general interest in secure rights and economic well-being.

This prescriptive program is sustained by assuming away the possibility that continued reliance on resource revenue causes weak institutions to persist. Consider this offering from a recent collection of essays: “[we] assume throughout that both countries and companies can and should do something to more effectively and fairly develop oil resources. We assume in particular that *governments are willing to take sometimes bold and difficult steps to succeed where most states have failed.*”² This is clearly a consequential assumption: “[i]f states are unable or unwilling to take such steps, then the best solution may well be to leave the oil and gas in the ground.”³ In other words, attempts to overcome the resource curse while governments that remain fiscally reliant on resource revenue are *not* willing to take “bold and difficult steps” to limit their discretion over the revenue are unlikely to be successful.

The impetus for assuming that governments are willing to implement positive institutional reforms is understandable: it turns the resource curse into a tractable problem, amenable to technocratic solutions. But this yields prescriptions that ignore the possibility that weak institutions persist so long as resource revenue continues to enter the domestic political economy. If the endogenous institutions thesis is true, then it is futile to first

1. See, among others, Joseph Bell et al., *Natural Resource Charter*, 2010. URL: <http://www.naturalresourcecharter.org>; Humphreys, Sachs, and Stiglitz, *Escaping the Resource Curse*.

2. Humphreys, Sachs, and Stiglitz, *Escaping the Resource Curse*, 14–15, emphasis added.

3. *Ibid.*, 15.

identify, say, underinvestment in human capital accumulation or lack of accountability to citizens as key causes of the resource curse and then prescribe that the state invest more in human capital or establish mechanisms to insure citizen oversight of resource transactions. Such prescriptions ignore the full extent of the problem.

If the standard explanation is correct, then overcoming the resource curse is simply a matter of improving the domestic institutions of resource-cursed countries. But if weak institutions persist so long as the state remains fiscally reliant on resources, voluntary institutional reform is inhibited by the continued flow of resource revenue into the domestic political economy. In fact, my argument in the last chapter suggests that the prospects for general institutional reforms that can directly strengthen the institutions of resource-cursed countries are slim. This is plainly bad news, since strong domestic institutions are the key to overcoming the resource curse. In this final chapter, I try to show that it's not all bad news. Overcoming the resource curse will certainly be much more complicated than previously thought, but there are worthwhile options to explore.

The explanation presented in chapter 5 identifies two necessary conditions for an effective prescription: (1) it must increase state rulers' reliance on civil society groups whose demands induce the provision of public goods, and (2) it must provide citizens with credible exit options. This first condition seems intuitively obvious. Most extant prescriptions, if not explicitly aiming to increase ruler reliance, would in effect do so. But extant prescriptions neglect the second condition. The importance of this condition cannot be overstated. Without also granting citizens' exit options, institutions that merely satisfy the first condition are insufficient for improving citizens' bargaining leverage vis-à-vis their rulers. Rulers who rely on citizens can still neglect their interests if those citizens have no better options. Focusing reform efforts on the second condition has a second advantage. Since increasing rulers' reliance on citizens is incredibly difficult so long as resource revenue

continues to enter the country, providing certain types of exit options — opportunities to emigrate, for example — would permit resource curse victims to escape the resource curse even where it cannot be overcome.

To highlight the importance of providing citizens with exit options, I consider recent proposals to address the resource curse presented by Thomas Pogge and Leif Wenar. Although these proposals would increase ruler reliance if implemented, they encounter numerous problems; not least, they would produce morally undesirable consequences if implemented, and each encounters feasibility challenges. But most importantly here, they fail to provide citizens with exit options and, so, are insufficient to effectively overcome the resource curse.

6.1 Resources and the Demands of Justice

Pogge and Wenar have built upon the standard institutional explanation of the resource curse in their respective moral assessments of the resource curse.⁴ The general lesson both draw from their reviews of the relevant literature is that resource abundant countries are cursed because resource revenue is held by the wrong hands: “The resource curse is not a curse that falls on poor countries because they have abundant resources. [...] The blessing of resources turns into a curse when tyrants and insurgents are allowed to sell off a country’s resources while crushing popular resistance, and to use the proceeds in ways that make the people worse off.”⁵ But both think that the standard explanations neglect an important consideration, namely, the extent to which rules regulating international

4. See Wenar, “Property Rights.” Pogge’s most cogent discussion of the issue is found in Pogge, *World Poverty and Human Rights*, 113–114, 163–165; but cf. Thomas W. Pogge, “‘Assisting’ the Global Poor,” in *Ethics of Assistance: Morality and the Distant Needy*, ed. Deen K. Chatterjee (Cambridge: Cambridge University Press, 2004); Thomas W. Pogge, “Real World Justice,” *The Journal of Ethics* 9 (2005): 29–53; Thomas W. Pogge, “Recognized and Violated by International Law: The Human Rights of the Global Poor,” *Leiden Journal of International Law* 18, no. 4 (2005): 717–745.

5. Wenar, “Property Rights,” 8, 9; cf. Pogge, *World Poverty and Human Rights*, 114. In drawing this conclusion, Pogge and Wenar follow the spirit, if not the letter, of the writings of prominent development economists, such as Paul Collier; see Collier, *The Bottom Billion*, ch. 3.

resource trade causally contribute to the persistence of the failed institutions that permit resource revenue to fall into the wrong hands. The main culprit is what Pogge calls the “international resource privilege,” which grants the power to effect legally valid resource transfers to any person or group holding a monopoly on the use of effective force in the territory containing the resources in question. When it comes to international resource transfers, the resource privilege insures that “might makes right.”⁶ The story is a relatively simple if sparse one. The resource privilege creates incentives for predators to acquire and wield power violently. This explains the proliferation of civil war, as rebel groups are able to sustain themselves by forcibly controlling resource-rich territory and selling off the resources. Once in power, resources provide predatory rulers with a valuable revenue stream, which is used to buy arms and political support to keep themselves in power, even against widespread opposition. Even reform-minded rulers, if they want to stay in office, are compelled to divert resource revenue to the private fortunes of the political elite or to keeping potential rebels at bay. As Wenar notes, such clientelism facilitates widespread corruption and hinders effective development investment, which helps explain economic stagnation.

For both Pogge and Wenar, the fact that resource wealth is a curse constitutes an injustice because the wrong hands come to hold resource revenue as a result of international rules upheld by the rich and powerful to promote their own economic advantage.⁷ But they differ in the details here. For Pogge, the international resource privilege helps pro-

6. Wenar, “Property Rights,” 13.

7. Although it should be noted that Pogge and Wenar differ on the extent to which the resource privilege is *imposed* upon underdeveloped countries by developed ones. Pogge claims that the rich actively impose the resource privilege upon the poor (Pogge, *World Poverty and Human Rights*, 115), whereas Wenar is more circumspect, claiming only that it “persists by custom because powerful global actors have strong interests in maintaining the status quo” (Wenar, “Property Rights,” 15). Wenar’s doubts about any fact of imposition are discussed in more detail in Leif Wenar, “Feasible Reform of International Trade in Resources,” in Jagger, *Pogge and His Critics*, sec. 3.

duce the resource curse and, consequently, the vast misery associated with it.⁸ Thus, our enforcement of the resource privilege is unjust because it violates a duty not to impose foreseeable and unnecessary harm.⁹ For Wenar, the resource privilege is analogous to a rule that grants thieves legitimate title to stolen goods — it effectively treats predators as the legitimate owners of natural resources and permits them to conclude resource transactions without citizens’ consent, thereby violating citizens’ property rights in their country’s natural resources.¹⁰ Hence, our enforcement of the resource privilege is unjust because it violates a duty to respect citizens’ property rights.

To be precise about the nature of Pogge’s and Wenar’s criticisms, note that there are two distinct sites at which we can morally assess the resource curse — transactions and institutions. To illustrate the distinction, consider the difference between assessing instances of theft and assessing a rule that, for arbitrary reasons, prevents members of a racial minority from holding private property. In the former, we assess the extent to which transactions violate principles of justice; in the latter, we assess the extent to which institutions violate principles of justice. Similarly, with respect to the resource curse, we might judge that individual *transactions* that put resource revenue in the hands of authoritarians are unjust because they foreseeably visit grave harm upon ordinary citizens at the hands of predatory dictators (with Pogge), or violate citizens’ property rights (with Wenar). Those who engage in such transactions — the agents of countries and firms that purchase resources or territorial concessions from authoritarians — do so unjustly because their actions violate clear principles of justice — that we ought not impose foreseeable and unnecessary harm

8. Pogge, *World Poverty and Human Rights*, 113–114.

9. Cf. Pogge, “‘Assisting’ the Global Poor”; Thomas W. Pogge, “Severe Poverty as a Violation of Negative Duties,” *Ethics & International Affairs* 19, no. 1 (2005): 55–83.

10. Wenar, “Property Rights,” sec. 3. Wenar cites numerous legal documents — both international and domestic — to support citizens’ claim to rightful ownership of the resources in their territory: Article 1 of the two International Covenants (on Civil and Political Rights and on Economic, Social, and Cultural Rights); Article 21 of the African Charter on Human and Peoples’ Rights; Iraqi and Angolan domestic law; as well as various UN resolutions and declarations.

on others and that we ought to respect property rights.

With respect to *institutions*, we might judge that the resource privilege is unjust because it generates an incentive scheme that foreseeably induces negligent, corrupt, and sometimes violent characters to acquire and exercise power to the detriment of ordinary citizens (with Pogge), or confers *de facto* legitimacy upon resource transactions brokered by predatory dictators, which violates citizens' property rights (with Wenar). Thus, not only do individual resource transactions with authoritarians violate clear principles of justice, but *the rule* that grants legitimate title to authoritarians, and our enforcement of it, is also unjust because *it* violates our duty to not impose foreseeable and unnecessary harm or to respect property rights.

Although Pogge and Wenar undoubtedly lament the unjust transactions that generate the resource curse, their criticisms focus intently on institutions. This is clearly so for Pogge, whose criticism is part of a long-standing program aimed at showing that global poverty results from unjust international institutions; the resource privilege is a recurring example of how the "global institutional order" harms the global poor.¹¹ Although Wenar discusses transactions more than Pogge does, his institutional focus is clear. Wenar's reply to the following objection is telling. The objection concedes that "what we pay to fuel our cars might end up being spent on [dictators'] torture chambers or personal jets," but then argues that "this is the way it is in... a globalized market economy... Some of the money we pay at the pump may go to support tyrants, but that seems just a part of modern life."¹² Wenar's response is that this is based on a "misdirection of attention from the institutional to the natural," and that the resource curse "results from a *defect in the rules that allocate control* over [natural] resources."¹³ To be sure, individual authoritarians and those

11. See the references in footnotes 4 and 9 in this chapter.

12. Wenar, "Property Rights," 8.

13. *Ibid.*, 8, emphasis added.

who transact with them violate citizens' property rights. But this is hardly the end of the story. It need not be this way, after all; international rules need not grant legitimate title to predatory dictators. Accordingly, "[t]he resource curse results from a *failure of institutions*: specifically, a *failure to enforce* property rights."¹⁴ Hence, the resource privilege is unjust as a rule because it fails to prevent unjust transactions.

6.2 Are There Alternatives to the Resource Privilege?

Let's concede both the normative claim — that the resource privilege is unjust — and the causal claim from which it is supposed to follow — that the resource privilege causally contributes to the resource curse. There remains an important question, namely, whether these claims entail a duty to establish an alternate rule in place of the resource privilege, as Pogge and Wenar suppose.¹⁵ This will be so only to the extent that there are admissible alternative institutions or rules that are capable of preventing the resource curse.¹⁶ This is a simple point about "ought implies can": if the rationale for institutional reform is to avert the resource curse, and there are no admissible alternate rules that can effectively avert the curse, then it is not the case that we ought to reform the resource privilege.¹⁷

Say R^* is an *admissible alternate rule* if it meets three conditions:

1. R^* does not violate any moral principle;
2. R^* is not too complicated for the parties who will be subject to R^* to understand and apply; and

14. Wenar, "Property Rights," 9, emphasis added.

15. I use "establish an alternate rule in place of" and "reform" interchangeably throughout, despite their lack of synonymy. (The latter usually connotes piecemeal change, while the former usually connotes wholesale change.)

16. Although it is a slight abuse of terminology, convenience dictates that I group the injustices identified by Pogge and Wenar under the term "resource curse" for the remainder of the paper. The abuse is only slight because the resource curse is clearly causally related to the injustices with which Pogge and Wenar are concerned.

17. Note that this is consistent with the deontological flavour of Wenar's focus on property rights violations. Wenar's rationale for reforming the resource privilege is an instrumental one, viz., to effectively enforce property rights. Insofar as there is no admissible alternate rule that effectively enforces property rights, we have no duty to reform the resource privilege.

3. Implementation of R^* is feasible; at a minimum, we can reasonably expect the relevant parties to enact R^* and establish mechanisms to credibly enforce R^* .

There will certainly be disagreement about how to further specify these conditions, but we need not settle such disputes here.¹⁸ What is important is the more general point that we do not have a duty to reform a rule R if there is no alternative R^* that meets these three conditions and is capable of preventing the injustice in question. More specifically, to justify the claim that we should reform the resource privilege, there must be some alternate rule for regulating international resource transactions that meets the three admissibility conditions and, if implemented, would effectively avert or mitigate the resource curse.

To get a handle on what sort of rule could effectively avert or mitigate the resource curse, recall the central lesson of chapter 5: weak institutions persist — and thereby facilitate the misallocation of resource revenue — so long as resource revenue continues to flow into a domestic political economy with an antecedently weak institutional structure. This is true *regardless of who holds power or how they acquired it*.¹⁹ This suggests that any rule that fails to prevent the flow of resource revenue into countries with weak institutional structures will fail to alter the outcome. Whether we have a duty to reform the resource privilege depends on whether there are any admissible alternatives to the resource privilege that prevent resource revenue from entering domestic political economies with weak institutional structures. Are there any such alternatives? If there are, are they admissible?

Pogge and Wenar have each proposed alternatives that aim to keep resource revenue from entering resource-cursed political economies. Considering their proposals proves to be instructive for thinking more generally about the space of possible alternatives. I briefly

18. I discuss the notion of feasibility at greater length in appendix B.

19. My brief study of Venezuela in the last chapter is a striking case in point. Pogge repeatedly presents Nigeria's Olusegun Obasanjo as an example of a democratically elected chief executive who failed to use resource revenue to implement good policies due to the inability of the domestic institutions to constrain rent seeking behaviour on the part of elites. Oddly, given his sensitivity to the effect of institutions, he seems to not recognize the instruction this example provides for thinking through the relationship of the resource privilege to the resource curse.

outline the proposals before raising some difficulties.

6.2.1 Pogge's proposed alternative

Pogge proposes an “amend-and-adjudicate” strategy (as I will call it). His proposal has two main components. First, fledgling democracies whose economies depend heavily on resource exports should amend their constitutions to include a clause that declares resource transfers executed by future authoritarians illegitimate.

[A] constitutional amendment in which our country declares that only its constitutionally democratic governments may effect legally valid transfers of ownership rights in public property and forbids any of its governments to recognize ownership rights in property acquired from a preceding government that lacked such constitutional legitimacy.²⁰

By putting investors on notice that the resource transfers negotiated with authoritarians will be challenged should democracy be restored, the amendment aims to deter investment in resource extraction under authoritarians. Without more or less stable investment demand, resources become less valuable to would-be authoritarians, thereby reducing the size of the prize awaiting those who successfully acquire political power by unconstitutional means. The anticipated effect is to reduce the number of attempts to acquire power unconstitutionally and thereby consolidate democratic reforms. To ensure that the deterrent effect of the amendment is not undermined by authoritarians who simply suspend the amendment, Pogge proposes the establishment of a “Democracy Panel.” This is “an international panel, composed of reputable, independent jurists living abroad who understand [the country’s] constitution and political system well enough to judge whether some particular group’s acquisition and exercise of political power is or is not constitutionally legitimate.”²¹ This “adjudication” part of Pogge’s proposal is meant to restore investors’ confidence in their

20. Pogge, *World Poverty and Human Rights*, 163. I should note that the amendment is meant to be prospective rather than retrospective. The amendment prohibits any of the country’s *future* constitutionally legitimate governments from recognizing as legitimate the resource transfers executed by *future* unconstitutional governments that precede the former.

21. *Ibid.*, 156.

title to legitimately transferred resources. Unless obtained from a government that is determined illegitimate by the panel, investors have no reason to worry about the legitimacy of their title to the resources. This is supposed to restore the incentives to invest in extraction under democratic regimes while diminishing those same incentives under nondemocratic regimes.

6.2.2 Wenar's proposed alternative

Wenar focuses on establishing mechanisms that respect the property rights of the citizens of resource abundant countries. At a minimum, respecting citizens' property rights requires a government to acquire "some sort of valid consent" from its constituents when transferring resources.²² In light of this, the crux of Wenar's proposal is to prohibit resource transfers that take place without citizens' valid consent. To give valid consent, citizens must have at least minimal civil and political rights. As an authoritative standard for determining whether citizens have the requisite rights, Wenar employs the widely-used *Freedom in the World* report, a rating of countries according to their respect for civil and political rights published by Freedom House.²³ To create "the strongest legal framework," he makes the least controversial assumptions possible: a Freedom House rating of 7 — the lowest possible rating — on either civil liberties or political rights is "decisive indication that no regime can legitimately sell resources from that country."²⁴

With this legal framework in place, Wenar turns to the matter of enforcement. Nationally, he proposes to grant standing to the citizens of a resource-cursed country to bring civil suits against resource corporations in the latter's domestic courts.²⁵ Alternatively, resource

22. Wenar, "Property Rights," 20.

23. For details on the Freedom House ratings, see <http://www.freedomhouse.org>.

24. Wenar, "Property Rights," 25.

25. Wenar cites three cases as precedent: *Doe v. Unocal*, 2003 WL 359787 (9th Cir.); *WIWA v. Royal Dutch Petroleum Co. & Shell Transport and Trading Co.*, WL 319887 (S.D.N.Y. 2002); and *Kensington v. PNP Parnibas*, 05 Ci. 5101 (LAP) (S.D.N.Y. 2005).

corporations could be tried domestically for the crime of receiving stolen goods.²⁶ Successful civil or criminal cases against resource corporations would help enforce citizens' resource rights by deterring the former from dealing with predatory regimes.

To enforce the proposed property rights regime internationally, Wenar prescribes what he calls a “trust-and-tariff” trade policy. The “tariff” part of the proposal requires compliant states to identify defector states — states who host resource corporations that deal with illegitimate governments or states who have direct such dealings of their own — and levy tariffs on imported goods coming from defectors. Insofar as the tariff harms the competitiveness of defector state firms in compliant state markets, the tariff serves as a punishment for acquiring resources from predatory regimes and aims to deter future such acquisitions. Further, we can think of the tariff as a “compliance premium” on defector state imports, which enables compliant state consumers to “wash their hands” of the illegitimate resource transfers. This is where the “trust” part of the proposal comes in. Once an illegitimate regime comes to power, compliant states should establish a “Clean Hands Trust” for the citizens of the resource-cursed state. This trust fund would be filled by the proceeds from the tariffs levied by compliant states against defector state imports until the fund's principal equals the value of the resources illegitimately transferred from the predatory government to defector states. The money in the fund is then held for the citizens of the resource cursed country until the minimum conditions for legitimate resource transfers are met. Once these conditions are met, the money (the fund's principal plus interest) is transferred to the citizens of the resource-cursed country (or, perhaps more accurately, their legitimate representatives).

Wenar uses a well-chosen example to illustrate the operation of the trust-and-tariff mechanism. Sudan has a Freedom House rating of 7. Suppose China acquires $\$n$ of oil

26. Wenar cites the National Stolen Property Act (18 USC 2314) and the extraterritorial jurisdiction of a Federal statute on the receipt of stolen goods (18 USC 662) as support for this litigation strategy.

from Sudan. To enforce the Sudanese people's resource rights, the US should establish a Clean Hands Trust for Sudan and fill it until its principal equals \$*n*. The money put into the fund should be raised from tariffs levied against Chinese imports. The money is held in trust for the Sudanese people until the minimum conditions for legitimate resource transfers are met, at which point, the contents of the fund are transferred to the Sudanese people.

6.3 Assessing the Alternatives

Are these proposed alternatives to the resource privilege admissible? I doubt it, for two sets of reasons. The first concerns practical issues; the second involves normative concerns. I discuss each of these in turn.

The fact that each of these rules would, if effective, drastically limit the availability of legitimately traded resources generates significant distributional consequences, which vitiates their feasibility.²⁷ In the case of Pogge's proposal, a genuinely effective Democracy Panel would affect the availability of resources, the terms of transfer, and the identity of legitimate trading partners. Given that most resources are located in countries that are likely to be ruled "illegitimate" by the panel, powerful actors (actors with significant bargaining power) have incentives either to weaken the terms of the agreement *ex ante* and thereby compromise the effectiveness of the panel or to capture the panel *ex post* and use it to serve their own interests. Consequently, credible decentralized enforcement — essential to keeping resource revenue from entering resource-cursed political economies — would constitute a heroic feat of collective action, a prospect we would be unwise to rely upon. In

27. On the difficulty for cooperation presented by distributional conflicts, see Robert Axelrod and Robert O. Keohane, "Achieving Cooperation under Anarchy: Strategies and Institutions," *World Politics* 38, no. 1 (1985): 226–254; James Morrow, "Modeling the Forms of International Cooperation: Distribution Versus Information," *International Organization* 48, no. 3 (1994): 387–423; James D. Fearon, "Bargaining, Enforcement, and International Cooperation," *International Organization* 52, no. 2 (1998): 269–305.

the absence of credible enforcement mechanisms, corporations would ignore the panel's judgements, as well as the amendment itself, and resource transfers would carry on as usual, leaving the resource curse intact. In any case, the prospects for effective cooperation on the matter appear bleak.²⁸

Wenar's proposal initially looks more promising than Pogge's, but it too has serious limitations stemming from its distributional effects. One problem is the potential for inhibiting international cooperation on other important global issues. While Wenar is correct to note that defector states do not have a legitimate complaint when "antitheft tariffs" are levied against their goods, it's also true that such tariffs must be recognized as legitimate by the World Trade Organization if they are to be upheld in the trade dispute resolution process. Suppose the WTO deems antitheft tariffs legitimate. If, for example, the US levies an antitheft tariff against Chinese goods because China buys Sudanese oil, the WTO rules in favor of the US in any dispute resolution process and any retaliatory trade measures by the Chinese are open to legitimate dispute by the US. Of course, China is still free to retaliate by obstructing cooperation on other important issues, such as security and climate change, not to mention future trade negotiations. Given the importance of China to achieving global cooperation on these issues, this is a high price to pay.

Although this is a serious problem, the main challenge here is not what happens when such a rule is in place. From a feasibility standpoint, the main problem is establishing a rule to recognize antitheft tariffs as legitimate in the first place. Again, such a rule has important distributional implications for the parties to WTO negotiations. As evidenced by the intractability of the Doha Round talks — which have been characterized by a deep

28. Wenar has noted two further limitations. First, that Pogge's proposal requires that a resource abundant country be democratic to enact the amendment, since a nondemocratic government is unlikely to enact such an amendment. Second, the proposal generates anti-democratic incentives. Anticipating a challenge to its title should democracy be restored, resource corporations and their host governments have incentives to help prop up the authoritarian regime from whom the resources were purchased. See Wenar, "Feasible Reform," 15.

divide between developed countries led by the European Union and the US on one side and developing countries led by Brazil, China, India, and South Africa on the other — the interested parties have highly divergent interests. This is enough to make cooperation difficult. Add to this the importance of certain resources as manufacturing inputs and the importance of a dynamic, export-oriented manufacturing sector for economic development, and it's unlikely that all parties to WTO negotiations would agree to implement a rule that legitimizes the use of antitheft tariffs.

We can generalize these particular practical concerns to any rule that (like Pogge's and Wenar's) limits the entry of resource revenue into weak domestic political economies. First, since so much resource wealth is located in countries with weak institutional structures, any such rule will drastically limit the availability of valuable resources, undermining the credibility of any decentralized enforcement mechanism. Further, monitoring and enforcement costs associated with such a rule might well be too high to justify implementing the rule. Second, the distributional conflicts that would arise from any such rule have great potential to inhibit international cooperation, not only on economic matters, but on other important matters as well, such as security, migration, and the environment. Even if such a rule could effectively avert the resource curse, it may well generate new problems whose costs outweigh the benefits of implementing the rule. Third, difficulties associated with judging the extent to which strong domestic institutions are adequately entrenched to avoid the curse and the length of time required to entrench strong institutions makes any such rule burdensome to implement. When is a domestic institutional structure strong enough to avert the resource curse? When is an institutional structure entrenched enough to withstand pressure to revert? These questions have no definitive answers. If we sincerely want to help avert the resource curse, our difficulty answering these may make us hesitant to resume trading with a resource abundant country that has been put on watch.

In cases where our hesitance leads us to withhold resource revenue from countries who are prepared to benefit from resource revenue, the rule unfairly burdens citizens of the resource abundant country by robbing them of important opportunities to improve their lives. (Not to mention potential charges of unjustified paternalism on the part of countries who prevent resource revenue from entering the country.)

What about the normative concerns? An obvious one arises once we acknowledge that Pogge's and Wenar's proposed rules, if enforced, would drastically cut our access to legitimately traded resources. Global economic dependence on resources such as oil, gas, and copper guarantees that a rule that prevents us from trading with resource-cursed countries would have adverse effects on the global economy. This would undoubtedly harm citizens of developed countries through massive job loss and and drastic price increases, but it would also cause severe deprivation in the developing world. Many developing countries rely on resource-intensive industrial development to increase employment and reduce domestic poverty. Further, drastic price increase would preclude countries struggling to industrialize from successfully doing so; an industrial development strategy would be put out of reach for many countries. Ironically, Pogge's and Wenar's proposals to alleviate the resource curse might increase the depth and extent of severe poverty. This consequence is morally unacceptable.

(Instead of a general prohibition on acquiring resources from cursed countries, we might implement *ad hoc* measures to encourage regime change in particular cases. The long trade embargo against apartheid South Africa is a case in point. Even still, such measures must be considered on a case-by-case basis, with sufficient attention given to salient contextual factors that are likely to influence the effectiveness of the action. In light of my argument in chapter 5, this means at a minimum investigating the likely effects of the proposed action on citizens' bargaining position vis-à-vis their ruler. Any use of such

measures must also give full consideration to the moral complexity of such actions.²⁹⁾

Additionally, the resource privilege actually upholds certain important values in international affairs and perhaps should, on balance, be retained. Consider an analogy with Article 2 in the United Nations Charter, which prohibits most types of intervention in the domestic affairs of sovereign states. This norm of nonintervention has certainly permitted some brutal dictators to get away with genocide and other heinous crimes. But, on balance, there are good reasons for such a prohibition: armed intervention often has drastically negative consequences that outweigh the benefits of any particular intervention,³⁰ we want to constrain powerful states' impulse to engage in neo-imperialist "nation-building",³¹ and such prohibitions protect important values such as the right to self-determination and the need to limit the resort to war.³² Similarly, although the resource privilege permits the resource curse, on balance, there might be good reasons to uphold the rule. Suppose the rule included a clause that a vendor government be deemed legitimate by its citizens. This potentially opens the door to serious market disruptions and exacerbates the rent-seeking behavior underlying the curse by encouraging enterprising opposition parties to continually (but without ground) challenge the governing party's legitimacy in an attempt to secure greater control of resource rents. This is highly likely in many resource abundant countries, wherein politics are often dominated by ethnic or regional rivalries. In addition, the

29. For example, Gordon discusses the complexity of economic sanctions; see Joy Gordon, "A Peaceful, Silent, Deadly Remedy: The Ethics of Economic Sanctions," *Ethics & International Affairs* 13, no. 1 (1999): 123–142.

30. Cf. Richard W. Miller, "Respectable Oppressors, Hypocritical Liberators: Morality, Intervention, and Reality," in *Ethics and Foreign Intervention*, ed. Deen K. Chatterjee and Don E. Scheid (Cambridge: Cambridge University Press, 2003); Iris Marion Young, "Violence Against Power: Critical Thoughts on Military Intervention," in Chatterjee and Scheid, *Ethics and Humanitarian Intervention*.

31. Michael Byers and Simon Chesterman, "Changing the Rules About Rules? Unilateral Humanitarian Intervention and The Future of International Law," in *Humanitarian Intervention: Ethical, Legal, and Political Dilemmas*, ed. Jeffrey L. Holzgrefe and Robert O. Keohane (Cambridge: Cambridge University Press, 2003); Simon Chesterman, *Just War or Just Peace? Humanitarian Intervention and International Law* (New York: Oxford University Press, 2001).

32. C.A.J. Coady, "War For Humanity: A Critique," in Chatterjee and Scheid, *Ethics and Humanitarian Intervention*.

current rule makes normative sense against the background commitment to nonintervention articulated in the Charter. Thus, we have good reason to uphold the resource privilege (or something like it) to the extent that we have good reason to uphold a commitment to nonintervention.

It is, of course, impossible to declare all potentially effective alternates to the resource privilege inadmissible, simply because there are many that have yet to be proposed. Hence, the foregoing considerations are not conclusive. But they do strongly suggest that any rule that prevents resource revenue from entering domestic political economies with weak institutions is unlikely to be admissible, for both practical and normative reasons. As I argued earlier, we do not have a duty to reform rules for which there are no effective admissible alternatives. Thus, to the extent that no effective admissible alternates emerge, *we do not have a duty to reform the international resource privilege.*

Of course, this does not imply that we have no duties whatsoever to victims of the resource curse.³³ Let's concede that the resource privilege permits unjust transactions; especially with respect to Wenar's argument concerning property rights, I find this a compelling claim. Then, at the very least, agents of the countries and firms that transact directly with dictatorial "resource thieves" owe compensation to the citizens of those countries, even if the transactions are unavoidable (e.g., because there simply aren't enough legitimate vendors to meet necessary resource demands). Moreover, the resource privilege regularizes a pattern of unjust resource transactions. This plausibly grounds a claim on behalf of those who are harmed by the privilege to receive compensation from those who contribute *indirectly* to the regularization of unjust resource transactions, including, among others, those who enjoy products made with unjustly acquired resources or whose jobs involve using unjustly acquired resources.³⁴ Thus, even if the resource privilege fails to generate

33. I'm grateful to Elizabeth Anderson for comments that led to the following discussion.

34. Relatedly, Barry argues that those who merely suspect their actions have contributed to another's de-

a duty to reform the resource privilege, there are other grounds for, at a minimum, duties of compensation.³⁵

Nothing I have said so far is sufficient to establish that the residents of developed countries have duties of justice to the victims of the resource curse, much less define the scope or stringency of any such duties. Nevertheless, the above considerations are suggestive. If the foregoing is right, an important question arises, namely, how to discharge any duties we have to the victims of the curse. We find guidance on this question by examining the ways in which proposals like Pogge's and Wenar's would be *ineffective* in addition to being inadmissible.

6.4 The Need for Exit Options

Even if we can overcome the concerns presented in the previous section, proposals such as Pogge's and Wenar's would be ineffective for overcoming the resource curse. The theory presented in chapter 5 identifies two necessary conditions for a prescription to effectively address the resource curse: (1) it must increase state rulers' reliance on civil society groups whose demands induce the provision of public goods, and (2) it must provide citizens with credible exit options. At best, Pogge's and Wenar's proposals fulfill the first condition: they would increase rulers' fiscal reliance on a broad coalition of citizens by depriving them of resource revenue. But they — along with most other extant prescriptions³⁶ — do nothing to address citizens' exit options. The importance of this condition

privation should be willing to err in favor of the deprived; see Christian Barry, "Applying the Contribution Principle," *Metaphilosophy* 36, nos. 1/2 (2005): 210–227. Further discussion of this so-called "contribution principle" is found in Christian Barry, "Understanding and Evaluating the Contribution Principle," in *Real World Justice*, ed. Andreas Follesdal and Thomas W. Pogge (Dordrecht: Springer, 2005).

35. A further potential ground for duties of compensation, which I decline to discuss in any detail here, rests on the charge that weak institutional structures and their developmental consequences are a legacy of colonial rule. See, e.g., Daron Acemoglu, Simon Johnson, and James A. Robinson, "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review* 91, no. 5 (2001): 1369–1401.

36. See, e.g., Humphreys, Sachs, and Stiglitz, *Escaping the Resource Curse*.

cannot be overstated. Rulers who rely on citizens can still neglect their interests if those citizens have nowhere else to go or no alternate leadership candidates to turn to. For any prescription to effectively improve citizens' bargaining leverage vis-à-vis their rulers, it must also provide citizens with exit options.

Focusing reform efforts on exit options has a second advantage. Since increasing rulers' reliance on citizens is incredibly difficult so long as resource revenue continues to enter the domestic political economy, providing certain types of exit options might permit resource curse victims to *escape* the curse even where it cannot be overcome. In lieu of prescriptions to overcome the resource curse, we can compensate curse victims by providing them exit options. In fact, given evidence that foreign aid perpetuates rulers' unaccountability to citizens, exit options may be the best form of compensation.³⁷

What is an exit option, and how can we improve resource curse victims' access to them? An *exit option* represents an opportunity to avoid the (deleterious) consequences entailed by remaining in one's current environment. The important feature of an exit option in this context is its expected net benefit (expected benefit less expected costs) relative to one's current situation. In the bargaining literature, an exit option is characterized by the payoff one expects to receive if one abandons the negotiation and no agreement is reached. In the present context, an exit option is characterized by the net benefit one expects to receive if one quits her current situation to avoid the consequences of remaining in that situation. An exit option is *credible* if the net benefit of quitting a situation is greater than

37. Bueno de Mesquita and Smith present evidence that foreign aid is predominantly given to unaccountable governments, as this facilitates the donor's extraction of policy concessions; see Bueno de Mesquita and Smith, "Foreign Aid and Policy Concessions"; Bueno de Mesquita and Smith, "A Political Economy of Aid." Additionally, the theory presented in chapter 5 suggests that *unearned* income is problematic regardless of the source; it need not derive from resource extraction. Foreign aid is just another form of unearned income. Morrison presents further evidence that resource revenue and foreign aid have similarly anti-democratic effects, although via a different mechanism than the one specified here; see Kevin M. Morrison, "Natural Resources, Aid, and Democratization: A Best-Case Scenario," *Public Choice* 131 (2007): 365–386; Kevin M. Morrison, "Oil, Nontax Revenue, and the Redistributive Foundations of Regime Stability," *International Organization* 63 (2009): 107–138.

the net benefit of remaining in that situation. In the particular case of the resource curse, a credible exit option consists in an opportunity to quit a resource-cursed situation that provides greater net benefit than what one expects to receive by remaining in the resource-cursed situation.

What sorts of things could potentially constitute exit options for curse victims? We find some guidance on this by looking to the exit options frequently emphasized in the institutional development literature cited in section 5.3. In the context of early modern Europe (esp. 17th century England), theorists have focused on citizens' ability to hide their income from taxation, as well as the real threat of armed revolution they posed.³⁸ Bueno de Mesquita et al.'s selectorate theory emphasizes the credibility of supporters' threats to shift their support to leadership challengers.³⁹ This suggests we might be able to provide curse victims with credible exit threats by increasing the liquidity of their assets or their ability to move already liquid assets, increasing their ability to credibly threaten armed resistance, or increasing their capacity to credibly threaten to shift their support to leadership challengers.

Unfortunately, each of these faces difficulties that must be overcome if they are to be pursued in the resource curse context. With respect to asset mobility, most victims of the resource curse have few assets to move and the prospects for obtaining liquid assets are slim, since economic stagnation is a component of the curse. Further, as noted above, oil wells and copper mines can't be moved, meaning that resource holders can't exit without leaving their assets behind. But the main problem here is that a person may be able to move their assets out of the ruler's reach without being able to move their body out of reach. Thus, so long as rulers' do not rely on citizens, asset mobility does not necessarily

38. See, e.g., Bates and Lien, "A Note on Taxation"; North and Weingast, "Constitutions and Commitment"; Tilly, *Coercion, Capital, and European States*.

39. Bueno de Mesquita et al., *The Logic of Political Survival*; see also Clark et al., "Why Do Autocrats Overachieve?"

help victims escape the threat to their person presented by a negligent or oppressive ruler.

Since resource-cursed states are typically military dictatorships, single-party states, or rigged-election democracies, ordinary citizens have few opportunities to threaten a ruler's political survival by shifting allegiance to a political rival. The ruler's ability to retain power is rarely affected by electoral outcomes, so curse victims can rarely exit by voting for a rival.⁴⁰ Moreover, leadership struggles are typically confined to a small elite and are largely insulated from ordinary citizens by the traditional hierarchies often found in resource-cursed societies.⁴¹ This means that those who are able to credibly shift political allegiances often threaten to do so to extract greater *private* goods.

Finally, although resource-cursed countries often experience armed civil conflict,⁴² ordinary citizens are not usually in a place to take up armed resistance against nefarious rulers. The latter typically have a massive military advantage, which makes armed resistance too costly for citizens living at subsistence levels. Moreover, as illustrated by civil wars in Angola (1975–1991, 1992–2002), Sierra Leone (1991–2002), and Sudan (1983–present) — in which resource control was often a central issue — ordinary citizens are often casualties rather than participants in civil conflict, typically caught between elite groups vying for political power.

So asset mobility, armed resistance, and shifting one's allegiance are examples of exit options, but making these credible for curse victims poses serious difficulties. Another obvious candidate is migration. If a victim is permitted to migrate from her resource-cursed conditions to a country with stable democratic governance and stable economic opportunities, she can thereby improve her life prospects. Of course, emigration is costly to the migrant. The associated legal and administrative costs can be quite high. More

40. But, as the case of Venezuela shows, being able to do so might not make a difference.

41. Bates, *When Things Fell Apart* provides illustrations.

42. Although, as noted in chapter 5 at footnote 2, whether resource abundance *causes* an increase in civil war onset is controversial.

importantly, leaving one's country often involves leaving behind family and friends, not to mention the sense of dislocation associated with relocating to a new society with (potentially vastly) different social and cultural norms and expectations. There are also the costs of being an "outsider," of being alienated and treated with suspicion by immigration officials, potential employers, and neighbors. Whether emigration presents a credible exit option depends on whether the *net* benefits of migration are greater than the net costs. This means that the credibility of emigration as an exit option declines as the costs increase. There is potential here, as some of the costs associated with migration are under the control of developed countries. But future research on the availability of emigration options for curse victims should be cognizant of these costs. It should also be cognizant of the extent to which mitigating these costs is *feasible*. Immigration imposes costs on the residents of destination countries as well, and residents of developed countries typically resist bearing these costs. This is especially true when the immigrants are desperately poor and largely unskilled, as victims of the resource curse typically are. This fact supplies us with a powerful rhetorical point: for all our talk about helping the global poor — for all our calls for foreign aid and humanitarian intervention — our resistance to admitting the poor into our own countries reveals that we are unwilling to do the things that could really help.⁴³ Although damning, this point does not yield a promising proposal to address the curse.

The central point here is a schematic one: future proposals to address the resource curse should focus on granting citizens credible exit options. This is so for two reasons. First, it is a necessary condition for effectively overcoming the curse; second, exit options effectively compensate curse victims for the deprivation caused by unjust resource transactions. The prospects for providing curse victims with exit options may not look good

43. Thanks to Bill Clark for raising this point.

at a first glance. Yet, because extant theorists have largely neglected the importance of exit options, the possibilities are underexplored. Hence, future prescriptive work on the resource curse should focus on improving citizens' exit options.

6.5 Where To Now? A Tentative Proposal

Fortunately, the preceding challenges do not yet warrant despair. The theory presented in chapter 5 implies that it is not resource revenue per se that causes development failures, but any factor that diminishes rulers' reliance on citizens. Hence, potential solutions to the resource curse need not focus directly on resource revenue. We might look at other levers that potentially affect rulers' reliance on citizens and try to pull on those in an effort to counteract the deleterious effects of resource revenue. The analysis in section 5.4 suggests a place to look: any form of income for which rulers need not put in political and organizational effort to receive, "unearned" revenue as Moore calls it. These certainly include foreign aid disbursements;⁴⁴ the Venezuelan case study suggests sovereign lending might be another. Do countries with vast resource revenues receive significant amounts of foreign aid or take on significant amounts of public debt? Could these factors give us enough leverage to mitigate the resource curse?

For a limited number of cases, they are. Among countries classified as resource dependent by Haber and Menaldo, three were among the top ten recipients of official development aid (ODA) from 2004 to 2006: Nigeria (US\$6.1 billion per year), Democratic Republic of Congo (\$1.9 billion per year), and Sudan (\$1.6 billion per year).⁴⁵ Additionally, Egypt and Indonesia are among the top ten ODA recipients since 1970, receiving

44. See footnote 37 in this chapter.

45. OECD, *Development Aid at a Glance*, Organisation for Economic Co-operation and Development, 2008, 23 (Table 1.2.7). *Official development aid* is defined as official flows to countries aimed at promoting economic development and welfare; see *ibid.*, 9.

Table 6.1: General government gross debt, percent of GDP, 1990–2010

	1990	1995	2000	2005	2010
Bahrain	8.3 (4.5)	16.5 (5.8)	29.3 (8)	28.7 (13.5)	32.8 (21.7)
Guinea	99.6 (2.7)	92.6 (3.7)	118.7 (3.1)	150.2 (2.9)	91.4 (4.3)
Qatar	10.9 (7.4)	42.4 (8.1)	55 (17.8)	18 (43)	25.6 (126.5)

SOURCE: IMF World Economic Outlook Database, 2010

GDP in US\$ billions reported in parentheses.

averages of US\$3 billion and \$1.9 billion per year respectively.⁴⁶ While many resource dependent countries have decreasing public debt, a few have shown upward trends during the last twenty years, including Bahrain, Guinea, and Qatar (see table 6.1). And, of course, Venezuela experienced a debt crisis despite the availability of resource revenue. All this is to say that we do find resource dependent countries receiving aid and loans despite their vast resource revenues.

A strategy targeting foreign aid and sovereign loan disbursements has definite limits. In the first place, we are talking about a relatively small number of countries. Secondly, the money received in aid and loans is often much lower than that received from resource extraction, so these might not be influential levers to pull.

Yet, a strategy targeting aid and debt can have some promise. Several of the listed countries are among the most populous, with Indonesia, Nigeria, Egypt, and the DRC among the twenty most populous countries in the world. The list also comprises some of the worst countries in the world, with Sudan, the DRC, Guinea, and Nigeria among the top fifteen states in *Foreign Policy*'s "Failed States Index."⁴⁷ So targeting aid and loans has the

46. OECD, *Development Aid at a Glance*, 25 (Table 1.2.10).

47. Available at http://www.foreignpolicy.com/articles/2010/06/21/2010_failed_states_index_interactive_map_and_rankings.

potential to lift vast numbers of people out of extreme deprivation. Finally, although small compared to resource revenues, the amounts disbursed in aid and loans is not insignificant. Aid and loans may provide enough of a handle to gain some leverage on the problem.

I end with a tentative outline for a strategy to mitigate the resource curse that targets foreign aid and loan disbursements to resource dependent countries. Given the conditions for an effective proposal, my strategy has two prongs. One aims to increase rulers' fiscal reliance on citizens by restricting their access to aid and loans; the other aims to increase citizens' exit options by opening up opportunities for citizens to withhold their income from domestic taxation.

To increase rulers' fiscal reliance on citizens, I propose that we look for ways to make development aid and government loans conditional on reaching a threshold level of fiscal reliance on non-resource tax revenue. We should require states to draw some threshold percentage of their revenue from taxes on income and profits earned domestically by individuals and businesses outside the resource sector. This rule should be general; it should apply to all states. But the threshold should be set in a way that discriminates between non-resource dependent states and resource dependent countries. The aim of the rule is to induce resource dependent states to increase their fiscal reliance on citizens, not punish those who have already achieved this. The threshold should also be set in a way that accounts for the aid or loans to be disbursed; states should achieve the threshold even when we include the money to be disbursed.

I anticipate that, if enforced, the rule would induce two changes in resource dependent states. First, they would need to develop broad tax extraction capacity in addition to the more narrow capacity to milk revenue from resource companies specifically. The former capacity requires tax authorities that penetrate deeply into the population. This, in turn, requires development of a coherent, disciplined, and meritocratic bureaucracy; given the

difficulty and cost of collecting taxes from multifarious sources, rulers will prefer a civil service that collects taxes honestly and efficiently to maximize the state's take. By closing holes that permit the waste and corruption that so often plagues resource dependent states, creation of such a bureaucracy could reduce revenue misallocation and thereby improve development outcomes.

Second, resource dependent states would need to cultivate a broad domestic tax base that is capable of providing revenue. Simply, tax revenue requires that citizens earn an income and businesses turn a profit. Enabling citizens and businesses to earn money that can be taxed requires developing an economic environment that induces and facilitates productive activity. To increase its tax revenue, the state will need to provide certain public goods: rule of law, physical security, secure property rights, public infrastructure, and education and health care. These investments would advance citizens' interests markedly.

Inducing productive activity further requires setting the tax rate at a place where citizens prefer work to leisure. Thus, cultivating a broad tax base involves an implicit bargain between the state and civil society. This brings us to the issue of citizens' exit threats. Citizens' ability to exit the workforce places constraints on a ruler's ability to collect tax revenue. The ruler cannot simply set a high tax rate to quickly achieve the specified tax reliance threshold; at a sufficiently high rate, people would simply stop working, eroding the ruler's tax base. Instead, the ruler must set tax rates low enough to induce productive activity. The tax yield might be relatively low initially, further inducing the ruler to make growth enhancing investments. To collect enough tax revenue at rates that are low enough to encourage economic activity, the ruler will need to make investments that enable more people to make more money. Again, such investments would advance citizens' interests markedly.

We must provide citizens with additional exit options; the threat to stop working is not

credible so long as a person's net income under predatory tax rates provides more than the life of bare subsistence to which she is likely to be relegated if she stops working. Suppose citizens could earn income outside the ruler's jurisdiction, leaving it out of the ruler's reach (at least, out of his reach for the purpose of satisfying the proposed rule). This would diminish the ruler's eligible tax revenue, which would jeopardize his ability to attract the aid and loans he needs. To maintain his tax base, the ruler would need to cultivate an environment where citizens prefer to earn taxable income domestically rather than abroad. This would give citizens additional leverage in bargaining with the ruler over taxation.

Can we provide poor individuals with opportunities to earn income abroad? Here's a possibility. Peer-to-peer (P2P) lending networks, such as ProsperTM and Lending Club, are a recent innovation that enable individuals to lend and borrow money directly without the mediation of banks. To participate, an individual opens an account with the network. To borrow money, one must fill out a survey used to screen borrowers and determine a credit rating as defined by network rules. This rating is used to set the terms of the loan and permits lenders to assess the risks and rewards associated with lending to any particular borrower. Lenders survey the outstanding loan requests and decide which borrowers to invest in. Lenders are free to contribute as much as they choose above a minimum (typically \$25) to any particular loan. This permits individuals to invest small sums of money and earn a relatively tidy return, even after the fee charged by the network to facilitate the transactions.⁴⁸

If a ruler is making investments that enable people to earn an income, then P2P lending could present attractive investment opportunities even for people with low income.

48. Returns range from roughly 5.5% to more than 25%, depending on the terms of the loan and the borrower's rating. Networks typically charge an annual servicing fee that works out to around 1% of the lender's total return. For further details, see <http://www.prosper.com> or <http://www.lendingclub.com>.

But two changes would be required for P2P lending to provide resource curse victims with credible exit options. First, we would need to establish *international* P2P lending networks; currently, P2P lending networks are limited to domestic contexts. Plausibly, this could be done; Kiva, a quasi-P2P microfinance organization, connects lenders and borrowers across borders. (Although there is an important difference between the Kiva and P2P models: established microfinance institutions actually disburse the loans in the Kiva case, whereas loans are disbursed directly by lenders in the P2P model. Perhaps this makes a difference for implementation.⁴⁹) Second, P2P lenders are often legally required to meet minimum income or net worth thresholds that are out of reach for low income individuals. For example, Idaho, New Hampshire, Oregon, Virginia, and Washington require P2P lenders to have an annual gross income of at least \$70,000 plus a net worth of at least \$70,000 (excluding house, home furnishing, and vehicle) or to have a net worth of at least \$250,000 (with the same exclusions).⁵⁰ Such requirements would obviously exclude the target individuals in the present. At a glance, reducing the minimum requirements to make P2P accessible to as many poor individuals as possible seems doable. Of course, determining the extent to which this is possible would require full consideration of the legal rationale for retaining relatively high minimum requirements. I must reserve this task for another occasion.

In sum, I propose the following general strategy as a means to mitigating the resource curse in a limited number of countries. First, we implement a requirement that official aid and sovereign loan disbursements be conditional on meeting a minimum threshold of fiscal reliance on domestic tax revenue. To be eligible to satisfy the requirement, the revenue should derive from taxes on income earned domestically by individual citizens and profits earned by businesses outside the resource sector. In implementing this rule, we induce

49. See <http://www.kiva.org> for further details of their model.

50. See <http://www.prosper.com/legal/compliance.aspx>.

rulers to make investments that advance citizens' interests in an effort to increase his eligible tax base. Second, we establish international peer-to-peer lending networks that could facilitate opportunities for citizens to invest their limited resources abroad. This could enable citizens to reduce the amount of their income that is eligible for taxation that meets the rule, thereby providing them with leverage in bargaining with the ruler over taxation. To keep citizens' investments in the country, the ruler may have to concede oversight of taxation and public spending to citizens, which would further advance citizens' interests.

I should reiterate the tentative nature of my proposal; its sole purpose is to outline a framework for a potential solution to the resource curse in a limited number of cases. Its strongest claim is to have satisfied the conditions for an effective prescription in broad strokes. Serious questions remain concerning implementation and feasibility. For example, we must settle the question of who is to implement and enforce the rule. Foreign aid and sovereign lending are decentralized phenomena; states receive loans and aid from multifarious sources. Implementing the rule is likely to require a complex (and costly) regulatory apparatus. We would need to identify an agent that is both capable and willing to oversee implementation. Perhaps international organizations such as the World Bank or IMF are candidates, but whether this is so requires further research. On the feasibility side, there remains the question of whether the proposal is incentive compatible. Randall Stone presents evidence that IMF conditionality lacks credibility because powerful donor countries often interfere with rigorous enforcement when the recipient country has influence with powerful patron states.⁵¹ A similar problem could apply here.

My proposal is not without its limits. But I submit that it points in a direction that is worth exploring further. Perhaps most importantly, given the argument in section 6.4, it illustrates how we might integrate a focus on exit options into our prescriptions to avert development failures.

51. Randall W. Stone, "The Political Economy of IMF Lending in Africa," *American Political Science Review* 98, no. 4 (2004): 577–591.

6.6 A Last Word

In the introduction, I announced my commitment to cosmopolitan moral principles. This leads me to affirm that we — the residents of developed countries — have duties of justice to address the resource curse and enable resource curse victims to live minimally decent lives. But I want to make clear that one need not share my cosmopolitan commitments to agree that we have some sort of duty to resource curse victims. Thomas Nagel — who is decidedly not a cosmopolitan — distinguishes duties of justice from duties of humanitarian assistance. On his view, duties of justice arise only among members of the right sort of association. Duties of assistance follow from “some minimal concern we owe to fellow human beings threatened with starvation or severe malnutrition and early death from easily preventable diseases”; they are supposed to be less demanding than duties of justice since they do not depend on one’s relationship to the deprived.⁵² Curse victims’ privation is severe enough to guarantee that, at the very least, we have a minimal humanitarian duty to help them avoid their poverty and oppression.

We need not settle whether we have duties of justice or duties of humanitarian assistance to resource curse victims to affirm that we are, at minimum, obliged to provide them with credible exit options. If we have duties of justice, institutions that ensure just treatment of individuals must provide them with credible exit options vis-à-vis their political leaders; if we only have duties of assistance, credible exit options will be among the most effective ways to help curse victims avoid poverty and oppression.

My institutional proposal also meets the stringent demands of cosmopolitan justice while remaining consistent with minimal humanitarian obligations. If it is true, as I concede in section 6.3, that the resource curse constitutes an injustice, then we are straightforwardly required to avert the injustice. This requires that we either avoid the actions

52. Nagel, “The Problem of Global Justice,” 118.

that cause the injustice or establish institutions that prevent those actions from generating unjust outcomes. I present some reasons to think that we cannot simply stop buying natural resources from dictators. Failing this, the strategy I outline satisfies the necessary conditions for effectively mitigating the unjust consequences of transacting with dictators. Thus, it satisfies the demands of cosmopolitan justice.

In lieu of granting resource curse victims entry into developed countries, the strategy I outline also meets the minimum requirements for effectively assisting curse victims avoid their plight, as specified in section 6.3. We could, as humanitarians frequently prescribe, merely continue to give humanitarian aid directly to curse victims. This might alleviate the symptoms of their problem. But it would not address the underlying condition, namely, citizens' weak bargaining leverage vis-à-vis their ruler. Genuine aid targets the underlying condition. Since my proposal does this, it (or something like it) would be required by a minimal humanitarianism.

CHAPTER 7

Conclusion: Toward a Pragmatic Statism

The preceding chapters lay the groundwork for a distinctively pragmatic approach to questions of global justice. In contrast to the conventional approach, the failure analysis approach integrates architecture and engineering, the normative and the empirical. Rather than defining our obligations to the global poor by reference to an ideal of justice, it seeks practical remedies to development failures. We need not look hard to identify development failures; malnutrition, disease, oppression, and violence abound. We need not know what obligations we would have in an ideal world to know that morality requires we do something to mitigate these development failures. The statistics, images, and stories are too jarring to permit moral complacency. We are hence faced with an urgent question: *What must we do to avert or mitigate development failures?*

The conventional wisdom is that development failures persist because we — developed states and their constituents — do not intervene enough to provide developing states with much-needed assistance. Simply, poverty persists because we do not provide enough development and humanitarian assistance;¹ violence and genocide persist because we do not intervene militarily to prevent it.² Cosmopolitans take this theme one step further. Not

1. See, e.g., Sachs, *End of Poverty*.

2. See, e.g., Collier, *The Bottom Billion*; International Committee on Intervention and State Sovereignty, *The Responsibility to Protect* (Ottawa: International Development Research Centre, 2001).

only should we intervene more often to prevent development failures; we should institutionalize the prescribed interventions by establishing a layer of suprastate institutions with mandatory global jurisdiction and the authority to restrict activity within state borders.³

My argument implies that it is too quick to step from an obligation to mitigate development failures to an obligation to intervene to provide assistance. At a minimum, any case for intervention must examine the effect the prescribed intervention is likely to have on citizens' bargaining position vis-à-vis their leaders. When some intervention can be shown to improve citizens' bargaining leverage, then there is a strong case in favor of the intervention. But there is empirical evidence that shows that official foreign aid and armed intervention in particular are not often employed to empower citizens but rather to strengthen the position of rulers.⁴ In such cases, developed countries' intervention undercuts the bargaining dynamic that underlies successful development, thereby freeing rulers to neglect citizens' interests. This suggests that, in these cases, development failures arise because we intervene *too much* rather than too little. Perhaps our obligation to mitigate development failures requires that we refrain from our *current* interference in sovereign states.

Regarding cosmopolitan institutional proposals, my argument in chapter 5 suggests an even stronger claim, namely, that cosmopolitan institutions are *self-defeating* in an important sense. Recall from the introduction that we can reduce arguments for cosmopolitan institutional schemes to a common schema:

1. The cosmopolitan commitment generates an obligation to realize the cosmopolitan objective.
2. Cosmopolitan institutional arrangements are required to realize the cosmopolitan objective.

3. See my discussion in the introduction on this point.

4. See Alberto Alesina and David Dollar, "Who Gives Foreign Aid to Whom and Why?" *Journal of Economic Growth* 5 (2000): 33–63; Bruce Bueno de Mesquita and George W. Downs, "Intervention and Democracy," *International Organization* 60 (2006): 627–649; Bueno de Mesquita and Smith, "A Political Economy of Aid"; Morrison, "Oil, Nontax Revenue, and Regime Stability."

3. The cosmopolitan commitment generates an obligation to establish cosmopolitan institutional arrangements.

An important question is whether cosmopolitan institutions would, if implemented, realize cosmopolitan objectives. Since cosmopolitan objectives — secure rights, sufficient material well-being, and so on — are realized when successful development occurs, this question is transposed to the following in light of chapter 5: *would cosmopolitan institutional arrangements improve citizens' bargaining leverage vis-à-vis their leaders?* The answer, I think, is no. I conjecture that cosmopolitan institutions would induce state leaders to seek support from third parties (other states, or regional and global agencies) at the expense of their dependence on citizens for support, thereby undermining citizens' bargaining leverage.

Consider two issues that are much discussed in the global justice literature: global distributive justice and limits on state sovereignty (which includes humanitarian intervention and the practice of state recognition).

Proposals for global redistributive institutions — e.g., Brock's global taxation schemes, Pogge's Global Resources Dividend, or the numerous calls to increase foreign aid⁵ — share a common feature. Each of them proposes to transfer “unearned” resources to developing countries; rulers need not expend effort vis-à-vis their citizens to receive it.⁶ Rulers who need not earn their income have little incentive to interact with citizens, nevermind negotiate with citizens over policy. In short, global redistributive institutions diminish rulers' fiscal reliance on citizens. In light of the theory presented in chapter 5, this implies that global redistributive institutions undermine a necessary condition for inducing rulers to advance citizens' interests, hence, for realizing cosmopolitan objectives.

5. Brock, *Global Justice*, ch. 5; Pogge, *World Poverty and Human Rights*, ch. 8; see Shue, *Basic Rights* for an example of an argument to increase official aid transfers.

6. Recall my discussion of earned vs. unearned income in sections 5.4 and 6.5.

Or consider cosmopolitan arguments in favor of military intervention to prevent gross and systematic human rights violations,⁷ or prescriptions to reform international institutions to reduce the need for intervention.⁸ The key shared feature here is that a ruler's retention of power ultimately depends on the decisions and actions of outsiders rather than citizens. If third parties can credibly threaten to forcibly remove a ruler, he will be inclined to seek outside support rather than internal support. The more a ruler depends on outside political support, the less he depends on citizens.⁹ Since policy is shaped in negotiations with political supporters, conditioning a ruler's authority on outside support induces rulers to bargain over policy with noncitizens rather than citizens. Ultimately, when authority is conferred from outside, rulers are accountable to noncitizens.

It is true that cosmopolitan institutional proposals typically condition redistributive transfers or respect for sovereignty on the realization of cosmopolitan objectives. Typically, these proposals require developing countries to make progress on improving people's lives to receive their distributive share or avoid military intervention. Fine. But this ignores an important fact about the world in which the proposed institutions are supposed to operate. Those who make redistributive or intervention decisions (third party states, international organizations) have their own agendas and use their control over resources as bargaining chips in pursuit of those agendas. Sure enough, cosmopolitans have in mind that these outside authorities would use this influence to provoke rulers to fulfill

7. For example, Brock, *Global Justice*, ch. 7; Caney, *Justice Beyond Borders*, ch. 7; Fernando R. Tesón, *Humanitarian Intervention*, 3rd ed. (Ardsley, NY: Transnational, 2005).

8. For example, Buchanan, *Justice, Legitimacy, and Self-Determination*; Thomas W. Pogge, "Pre-empting Humanitarian Interventions," in *Freedom, Power, and Morality: Essays for Felix Oppenheim*, ed. Ian Carter and Mario Ricciardi (London: Palgrave, 2001); Thomas W. Pogge, "Moralizing Humanitarian Intervention: Why Jurying Fails and How Law Can Work," in *Humanitarian Intervention, Humanitarian Intervention*, ed. Terry Nardin and Melissa S. Williams, NOMOS XLVII (New York and London: New York University Press, 2006).

9. Consider how third party states or outside organizations intervene to prop up allied dictators despite their domestic unpopularity; e.g., Saddam Hussein (before the late-80s), Marcos, Mobutu, Pahlavi, Suharto. When dictators fall, it is often because third parties intervene to forcibly remove them — e.g., Hussein — or step aside to let domestic politics take over when intervention would be too costly — e.g., Mubarak, Pahlavi.

cosmopolitan objectives. But this requires that these outside authorities reliably promote the interests of another country's citizens, sometimes to the detriment of their own private interests. This becomes especially worrisome once we recognize that those who fill the offices of these outside agencies (be they other states or global agencies) are constrained by the interests of their own political supporters. I find it implausible to expect that *these latter interests* would be reliably aligned with the interests of citizens in another country.

Suppose we can rely on global agencies to consistently use their authority to prod developing country rulers to fulfill cosmopolitan objectives. Even still, this remains fragile assurance that the latter will be induced to advance citizens' interests. Mill highlights two problems with relying on others to promote an individual's interests: insufficient motivation and lack of relevant knowledge.

[The individual] is the person most interested in his own well-being; the interests which any other person, except in cases of strong personal attachment, can have in it, is trifling, compared with that which he himself has; the interest which society has in him individually. . . is fractional, and altogether indirect; while with respect to his own feelings and circumstances, the most ordinary man or woman has means of knowledge immeasurably surpassing those that can be possessed by any one else.¹⁰

The point is that the citizens themselves will be most motivated to press their interests with their ruler, and they will know best which interests or needs are most urgent. Hence, even if we can to rely on global agencies to promote individuals' interests (though I'm dubious), we obtain a stronger guarantee for realizing cosmopolitan objectives if we pursue courses of action that *enlarge* the number of citizens on whose support rulers depend and *intensify* the degree to which rulers rely on that support.

The preceding examples provide some reasons to think that, in cases where rulers come to rely more on outside agencies for support than on their own citizens, cosmopolitan institutional arrangements are liable to undermine the fulfillment of cosmopolitan objectives.

10. John Stuart Mill, *On Liberty*, in *The Basic Writings of John Stuart Mill*, ed. Dale E. Miller (New York: Random House, 2002), 79.

In these cases, institutional cosmopolitanism is self-defeating.

Stated generally, the central lesson is this:

In a world where politics is predominantly strategic, cosmopolitan institutional arrangements are liable to undermine state leaders' need to rely on their citizens for political support, thereby compromising citizens' bargaining leverage vis-à-vis their leaders.

When citizens' bargaining leverage vis-à-vis their leaders is compromised, cosmopolitan objectives are liable to go unfulfilled.

My arguments throughout the dissertation crucially assume that the relevant political agents act strategically. By this, I don't mean that political agents are ultra-rational, self-interested agents without any concern for others. All I mean is that political agents undertake actions in service of their agendas, which are largely shaped by the agendas of their core supporters. Should we endorse this assumption when asking questions about global institutional design?

It depends on the normative question we seek to answer. If we seek to identify a normative standard by which we can assess current institutions and outcomes, we need not. One way to arrive at an evaluative standard is to specify the institutional arrangements that would be realized in a fully just world. It seems appropriate to assume that political agents in a fully just world are motivated largely by moral considerations.¹¹

But the institutional prescriptions I have considered throughout this essay are clearly interested in advancing cosmopolitan objectives in our decidedly unjust world. That political agents act strategically seems to be a deep and enduring feature of our world. No, not all political agents act strategically; and those who do don't do so all the time. But a sufficient number of political agents are motivated by strategic considerations often enough to ensure that outcomes are reliably shaped by strategic, rather than normative, considerations.

11. See the discussion at the end of section B.2.

This fact is surely lamentable from a moral perspective; but it is likely a fact that we can't escape, at least not any time soon. The upshot is that our reasoning about institutional design should account for the fact that the particular agents who will occupy positions of authority in any institutional scheme will typically be moved by strategic considerations. Thankfully, this does not imply that we should give up devising institutional reforms meant to advance cosmopolitan objectives. Instead, we must seek institutional remedies that shape political agents' incentives in a way that aligns their interests with the fulfillment of cosmopolitan objectives.¹²

This is a difficult task, to be sure. But my arguments in chapters 5 and 6 suggest a direction in which to look. In general, those who seek to mitigate development failures should attend to the likely effects of an institutional prescription on individuals' bargaining relationship with their political leaders. More specifically, they should devise strategies that (1) increase (rather than reduce) rulers' dependence on a broad coalition of citizens for support, and (2) provide citizens with credible exit options. Simply put, successful remedies for development failures are likely to realize their moral objectives when they improve citizens' capacity to press their demands with their rulers.

Contrary to typical cosmopolitan offerings, I conjecture that these conditions are more likely to be satisfied in a system of sovereign states than in a system of vertically dispersed authority with global organizations residing at the top. This is decidedly not a recommendation to maintain the status quo; in the current system, rulers' reliance on citizens is too often undermined, as powerful states use threat of force or withdrawal of cooperation to keep rulers in the developing world in line. Instead of a system in which the bound-

12. For elaboration of this point, see Geoffrey Brennan and James Buchanan, "The Normative Purpose of Economic 'Science': Rediscovery of an Eighteenth-Century Method," *International Review of Law and Economics* 1 (1981): 155–166; Geoffrey Brennan and Philip Pettit, "The Feasibility Issue," in *Oxford Handbook of Contemporary Philosophy*, ed. Frank Jackson and Michael Smith (Oxford: Oxford University Press, 2005).

aries of political communities are created and maintained by a few global powers, what is needed is a system in which sovereign political entities are confined within boundaries that are contested by members of political communities, and emerge from a process of negotiation between prospective rulers and those they intend to rule.¹³ This requires institutions that effectively prevent global powers from interfering in the domestic political processes of political communities so that the outcome will arise from *a process of internal contestation among members of the community*. Such a process of internal contestation requires prospective rulers to secure the support of domestic citizen groups, which, in turn, gives those domestic groups influence over the policy making process. In short, realizing cosmopolitan objectives requires institutions that will *enforce* rather than dilute strong sovereignty norms.

I do not assert that states or nations have moral value *per se*.¹⁴ I tentatively endorse a system of states for engineering reasons; a system of sovereign states is most likely to satisfy the conditions under which development failures would be averted and cosmopolitan objectives would be realized in our world. Since cosmopolitans are centrally concerned with establishing the conditions needed to fulfill the requirements of their moral commitments, my arguments suggest that they would be wise to adopt a pragmatic statism.

13. Cf. Bates, *Prosperity and Violence*; Tilly, *Coercion, Capital, and European States*.

14. Cf. David Miller, *On Nationality* (New York: Oxford University Press, 1995); Michael Walzer, "The Moral Standing of States: A Response to Four Critics," *Philosophy & Public Affairs* 9, no. 3 (1980): 209–229.

APPENDICES

APPENDIX A

Technical Appendix

I now present the reasoning from section 5.3 formally. To do this, I extend the “Exit, Voice, and Loyalty” game presented by Clark, Golder, and Golder.¹ The strategic logic presented here is the same as in that model. My contribution is the formalization of this strategic logic when the ruler faces more than one group of citizens with differential bargaining strength and competing policy objectives.

The players. There are three players who bargain over an indivisible prize.

- A ruler (R), who sets policy, first, to insure his retention of political power and, second, to maximize discretionary income.
- An elite class (E), whose members engage in rent-seeking activity; they demand institutions and policies that maximize their share of the prize.
- A citizen class (C), whose members demand public goods, i.e., institutions and policies that secure their rights and promote their general well-being.

The timeline.

1. The ruler chooses to: neglect the interests of both the elites and citizens (PREY); advance elites’ interests (SUPPORT E); or advance citizens’ interests (SUPPORT C).
2. The elites and citizens each choose one of the following actions simultaneously²: to lobby the ruler for a favorable change (VOICE); to comply with the ruler’s choice and absorb the consequences (LOYALTY); to take action to avoid the consequences of the ruler’s choice (EXIT). If neither the elites nor the citizens choose VOICE, the game ends and the players receive their payoffs.

1. Clark, Golder, and Golder, “The Balance of Power.”

2. Although the timing of elites’ and citizens’ response to the ruler’s initial choice doesn’t matter. In all cases, at least one player has a (weakly) dominant response.

3. If either the elites or the citizens choose VOICE, the ruler must either CONCEDE (i.e., advance the interests of a lobbying player) or IGNORE (i.e., persist in implementing the initial choice). Since elites and citizens prefer opposing policies, the ruler can concede to only one of the two players if both elites and citizens choose VOICE. If the ruler concedes to either player, the game ends and the players receive their payoffs.
4. If the ruler ignores any exercise of voice, then the player(s) who exercised voice choose(s) either EXIT or LOYALTY. If the ruler reverses a choice that is initially favorable for i , then i chooses either EXIT or LOYALTY. The game ends and the players receive their payoffs.

The payoffs.

- If the ruler preys or both the elites and citizens exit, he receives a benefit of 1; otherwise, he receives 0. If the elites remain loyal, the ruler receives an additional payoff L_E ; if the citizens remain loyal, the ruler receives an additional payoff L_C . The ruler pays a cost K_R if either the elites or the citizens use voice.
- If the ruler supports the elites and the latter remain loyal, then the elites receive 1; otherwise, they receive 0. If the elites exercise voice, they pay a cost K_E and receive 1 if the ruler concedes and 0 if the ruler ignores. If the elites exit, they receive X_E .
- If the ruler supports citizens and the latter remain loyal, then the citizens receive 1; otherwise, they receive 0. If the citizens exercise voice, they pay a cost K_C and receive 1 if the ruler concedes and 0 if the ruler ignores. If the citizens exit, they receive X_C .

Assumptions.

1. The players have full information.
2. $K_i, L_i > 0$ for $i \in \{E, C\}$.
3. $1 > X_i$ for $i \in \{E, C\}$.
4. i will exit only if $X_i > 0$ for $i \in \{E, C\}$
5. The ruler will respond only if $L_i > 1$ for $i \in \{E, C\}$
6. If $L_i, L_j > 1$, the ruler responds to i if $L_i > L_j$ for $i, j \in \{E, C\}$; $L_i \neq L_j$
7. i will not use voice if $X_i > 1 - K_i$ for $i \in \{E, C\}$

I solve the model for the subgame perfect Nash equilibrium. Since the point of the model is to determine the conditions under which rulers attend to citizens' interests, the main claims of interest concern the ruler's initial policy choice. These are stated below as propositions

A.7 and A.8. Before proving these propositions, I prove a series of lemmas concerning equilibrium play in the various subgames.

Lemma A.1. *For $i, j \in \{E, C\}$, $j \neq i$, if the ruler ignores i or concedes to j following i 's use of voice, then*

- a. *i remains loyal if $X_i \leq 0$;*
- b. *i exits otherwise.*

Proof. Suppose the ruler ignores i or concedes to j . $U_i(\text{loyalty}) = 0$. Suppose $X_i \leq 0$. Then $U_i(\text{exit}) = X_i \leq U_i(\text{loyalty})$. So i remains loyal.

Now suppose $X_i > 0$. Then $U_i(\text{exit}) \geq U_i(\text{loyalty})$ and i exits. \square

Lemma A.2. *For $i \in \{E, C\}$, if the ruler concedes to i following i 's use of voice, i remains loyal.*

Proof. Suppose the ruler concedes to i . Then $U_i(\text{loyalty}) = 1 > U_i(\text{exit}) = X_i$ (by assumption). So i is always loyal if the ruler concedes. \square

Lemma A.3. *For $i \in \{E, C\}$, if i uses voice, then*

- a. *the ruler ignores i if $X_i \leq 0$;*
- b. *the ruler ignores i if $X_i > 0$ and $L_i \leq 1$, except when $i \neq j \in \{E, C\}$ remains loyal following an initial policy of supporting j and $X_j \leq 0$ or $L_j < L_i \leq 1$; then the ruler concedes to i .*
- c. *the ruler concedes to i if $X_i > 0$ and $L_i > 1$, except when either (1) both i and j use voice, or (2) j is loyal following an initial policy of supporting j , and $X_j > 0$ and $L_j > L_i$; then the ruler ignores i .*

Proof. I prove each subitem separately.

- a. Suppose $X_i \leq 0$. Then i remains loyal if ignored (by lemma A.1). Thus, $U_R(\text{concede}) = L_i \leq U_R(\text{ignore}) = 1 + L_i$. So the ruler ignores i .

- b. Suppose $X_i > 0$ and $L_i \leq 1$. Then $\min U_R(\text{ignore}) = 1$ (by lemma A.1), which is greater than $U_R(\text{concede}) = L_i$. So the ruler ignores i .

The preceding holds *except* in either of the following cases when $X_i > 0$ and $L_i \leq 1$.

Suppose j remains loyal following SUPPORT j .

Case 1. Suppose $X_j \leq 0$. Then j is loyal when the ruler concedes to i (by lemma A.1). So $U_R(\text{concede to } i) = L_i + L_j - K_R > U_R(\text{ignore } i) = L_j - K_R$ (by lemmas A.1 and A.2). Thus, the ruler concedes to i .

Case 2. Suppose $L_j < L_i$. Whether j exits or not, $\min U_R(\text{concede to } i) = L_i - K_R$ when the ruler concedes to i (by lemma A.2), whereas $U_R(\text{ignore } i) = L_j - K_R$ (by lemma A.1). Since $L_j < L_i$, $U_R(\text{concede to } i) > U_R(\text{ignore } i)$. Thus, the ruler concedes to i .

- c. Suppose $X_i > 0$ and $L_i > 1$. Then $U_R(\text{ignore } i) = 1 - K_R < U_R(\text{concede to } i) = L_i - K_R$ (by lemmas A.1 and A.2). So the ruler concedes to i .

The preceding holds *except* in either of the following cases. Suppose (in addition) that $X_j > 0$ and $L_j > L_i$.

Case 1. Suppose both i and j use voice. Then $U_R(\text{concede to } j) = L_j - K_R > U_R(\text{concede to } i) = L_i - K_R > U_R(\text{ignore}) = 1 - K_R$ (by lemmas A.1 and A.2). Thus, the ruler ignores i .

Case 2. Suppose j is loyal following SUPPORT j . Then $U_R(\text{ignore } i) = L_j - K_R > U_R(\text{concede to } i) = L_i - K_R$ (by lemmas A.1 and A.2). Thus, the ruler ignores i .

This concludes the proof. \square

Lemma A.4. For $i \in \{E, C\}$, if the ruler initially supports i , then i remains loyal.

Proof. Let s_j denote j 's strategy for $i \neq j \in \{E, C\}$. $U_i(\text{exit} \mid \text{support } i, s_j) = X_i$. Suppose the ruler supports i in his initial policy choice. If j uses voice following SUPPORT i , then the ruler can ignore j or concede to j (by lemma A.3b).

Case 1. Suppose the ruler ignores j ; then $U_i(\text{loyalty} \mid \text{support } i, s_j) = 1 > U_i(\text{exit} \mid \text{support } i, s_j)$.

So i is loyal whenever the ruler supports i and ignores j .

Case 2. Suppose the ruler concedes to j . If $X_i \leq 0$, i remains loyal following CONCEDE TO j and exits otherwise (by lemma A.1). If $X_i \leq 0$, then

$$U_i(\text{loyalty} \mid \text{support } i) = \begin{cases} 0 & \text{if } j \text{ uses voice} \\ 1 & \text{if } j \text{ exits or remains loyal} \end{cases}.$$

Since $U_i(\text{exit} \mid \text{support } i, s_j) \leq 0$, $U_i(\text{loyalty} \mid \text{support } i, s_j) \geq U_i(\text{exit} \mid \text{support } i, s_j)$; so loyalty weakly dominates exit. Thus, if the ruler concedes to j following j 's use of voice and $X_i \leq 0$, i responds to SUPPORT i with loyalty whatever j does. If $X_i > 0$, then

$$U_i(\text{loyalty} \mid \text{support } i) = \begin{cases} X_i & \text{if } j \text{ uses voice} \\ 1 & \text{if } j \text{ exits or remains loyal} \end{cases}.$$

Again, $U_i(\text{loyalty} \mid \text{support } i, s_j) \geq U_i(\text{exit} \mid \text{support } i, s_j)$; so loyalty weakly dominates exit. Thus, if the ruler concedes to j following j 's use of voice and $X_i > 0$, i responds to SUPPORT i with loyalty whatever j does. Since i is loyal whether $X_i \leq 0$ or $X_i > 0$, i responds to SUPPORT i with loyalty whenever the ruler concedes to j following j 's use of voice.

Since i is loyal whether the ruler concedes to j or ignores j following j 's use of voice, i is loyal whenever the ruler supports i . \square

Lemma A.5. For $i \in \{E, C\}$, if the ruler initially supports j , then

- a. i remains loyal if $X_i \leq 0$;
- b. i uses voice if $0 < X_i \leq 1 - K_i$ and either $X_j \leq 0$ or $L_i > L_j$ for $i \neq j \in \{E, C\}$;
- c. i exits otherwise.

Proof. Suppose the ruler supports j in his initial policy choice. By lemma A.4, j remains loyal.

- a. Then $U_i(\text{loyalty}) = 0$. Suppose $X_i \leq 0$. Then $U_i(\text{exit}) = X_i \leq U_i(\text{loyalty})$. So i remains loyal.
- b. Now suppose $X_i > 1 - K_i$. Since $\max U_i(\text{voice} \mid \text{support } j) = 1 - K_i$, $U_i(\text{voice} \mid \text{support } j) < U_i(\text{exit} \mid \text{support } j)$ and i exits if the ruler supports j .
- c. Finally, suppose $0 < X_i \leq 1 - K_i$.

Case 1. Suppose $X_j \leq 0$. Then j remains loyal if the ruler concedes to i when i uses voice (by lemma A.1). Thus, $U_i(\text{voice} \mid \text{support } j) = 1 - K_i \geq U_i(\text{exit} \mid \text{support } j)$. So i uses voice in response to SUPPORT j if $0 < X_i \leq 1 - K_i$ and $X_j \leq 0$.

Case 2. Suppose $L_i > L_j$. Whether j exits or not, the ruler concedes to i (by lemma A.3b, case 2). So $U_i(\text{voice} \mid \text{support } j) = 1 - K_i \geq U_i(\text{exit} \mid \text{support } j)$. So i uses voice in response to SUPPORT j if $0 < X_i \leq 1 - K_i$ and $L_i > L_j$.

Case 3. Suppose $X_j > 0$ and $L_j > L_i$. Then the ruler ignores i 's voice (by lemma A.3c, case 2). Since $X_i > 0$ (by assumption), i exits if ignored and $U_i(\text{voice} \mid \text{support } j) = X_i - K_i \geq U_i(\text{exit} \mid \text{support } j) = X_i$. Thus, i exits in response to SUPPORT j if $0 < X_i \leq 1 - K_i$, $X_j > 0$, and $L_j > L_i$.

This concludes the proof. \square

Lemma A.6. For $i \in \{E, C\}$, if the ruler initially preys, then

- a. i remains loyal if $X_i \leq 0$;
- b. i uses voice if $0 < X_i \leq 1 - K_i$ and either $L_i > \max\{1, L_j\}$ or $L_i > 1$ and $X_j > 1 - K_j$ for $i \neq j \in \{E, C\}$;
- c. i exits otherwise.

Proof. Suppose the ruler preys in his initial policy choice. Then $U_i(\text{loyalty}) = 0$.

- a. Suppose $X_i \leq 0$. Then $U_i(\text{exit}) = X_i \leq U_i(\text{loyalty})$. So i remains loyal if $X_i \leq 0$.

- b. Now suppose $X_i > 1 - K_i$. Since $\max U_i(\text{voice} \mid \text{prey}) = 1 - K_i$, $U_i(\text{voice} \mid \text{prey}) < U_i(\text{exit} \mid \text{prey})$ and i exits if $X_i > 1 - K_i$.
- c. Now suppose $0 < X_i \leq 1 - K_i$. Let s_j denote j 's strategy.

Case 1. Suppose $L_i > \max\{1, L_j\}$. By lemma A.3c, the ruler concedes to i if i uses voice. Thus, $U_i(\text{voice} \mid \text{prey}, s_j) = 1 - K_i \geq U_i(\text{exit} \mid \text{prey}, s_j) > U_i(\text{loyalty} \mid \text{prey})$. So i uses voice if $0 < X_i \leq 1 - K_i$ and $L_i > \max\{1, L_j\}$.

Case 2. Suppose $L_i > 1$ and $X_j > 1 - K_j$. Consequently, j exits if the ruler preys (by lemma A.6b) and the ruler concedes to i if i uses voice (by lemma A.3c). Thus, $U_i(\text{voice} \mid \text{prey}, s_j = \text{exit}) = 1 - K_i \geq U_i(\text{exit} \mid \text{prey}, s_j = \text{exit}) > U_i(\text{loyalty} \mid \text{prey}, s_j = \text{exit})$. So i uses voice if $0 < X_i \leq 1 - K_i$, $L_i > 1$, and $X_j > 1 - K_j$.

Case 3. Suppose $L_j > L_i > 1$ and $0 < X_j \leq 1 - K_j$. Consequently, j uses voice (by lemma A.6c, case 1) and the ruler concedes to j (by lemma A.3c). Thus, $U_i(\text{voice} \mid \text{prey}, s_j = \text{voice}) = X_i - K_i \leq U_i(\text{exit} \mid \text{prey}, s_j = \text{voice})$ and $U_i(\text{loyalty} \mid \text{prey}, s_j = \text{voice}) = 0 < U_i(\text{exit} \mid \text{prey}, s_j = \text{voice})$. So i exits if $0 < X_i \leq 1 - K_i$, $L_j > L_i > 1$, and $0 < X_j \leq 1 - K_j$.

Case 4. Suppose $L_i \leq 1$. Then the ruler ignores i (by lemma A.3b). Thus, $U_i(\text{voice} \mid \text{prey}, s_j) = X_i - K_i \leq U_i(\text{exit} \mid \text{prey}, s_j = \text{voice})$ and $U_i(\text{loyalty} \mid \text{prey}, s_j = \text{voice}) = 0 < U_i(\text{exit} \mid \text{prey}, s_j = \text{voice})$. So i exits if $0 < X_i \leq 1 - K_i$ and $L_i \leq 1$.

Proposition A.7. *For $i \in \{E, C\}$, the ruler preys if and only if $X_i \leq 0$ or $L_i \leq 1$.*

Proof. If: Case 1. Suppose $X_i \leq 0$ for all $i \in \{E, C\}$. Then i remains loyal whatever the ruler does (by lemmas A.4, A.5a, A.6a). Thus, $U_R(\text{prey}) = 1 + L_E + L_C > U_R(\text{support } i) = L_i + L_j$. So the ruler preys.

Case 2. Suppose $X_i \leq 0 < X_j$ and $L_j \leq 1$ for $i \neq j \in \{E, C\}$. Then i remains loyal whatever the ruler does (by lemmas A.4, A.5a, and A.6a). Further, j exits in response to

PREY (by lemma A.6b and c, case 4) or SUPPORT i (by lemma A.5b and c, case 3) but remains loyal when SUPPORT j (by lemma A.4). Thus, $U_R(\text{prey}) = 1 + L_i \geq U_R(\text{support } j) = L_i + L_j > U_R(\text{support } i) = L_i$. So the ruler preys.

Case 3. Suppose $X_i > 0$ and $L_i \leq 1$ for all $i \in \{E, C\}$. Then i exits in response to PREY (by lemma A.6b and c, case 4) or SUPPORT j (by lemma A.5b and c, case 3) but remains loyal when SUPPORT i (by lemma A.4). Thus, $U_R(\text{prey}) = 1 \geq U_R(\text{support } i) = L_i$. So the ruler preys.

Since the ruler preys in all cases, the ruler preys if $X_i \leq 0$ or $L_i \leq 1$.

Only If: Suppose the ruler preys. Prove the contrapositive to show a contradiction. Suppose it's not the case that $X_i \leq 0$ or $L_i \leq 1$ for $i \in \{E, C\}$. By De Morgan's law, $X_i > 0$ and $L_i > 1$. It follows from lemmas A.1, A.2, A.3, and A.6 that $\max U_R(\text{prey}) = \max\{L_i - C_G, 1\}$. It follows from lemma A.4 that $\min U_R(\text{support } i) = L_i$. Since $L_i > 1$ by assumption, $\min U_R(\text{support } i) > \max U_R(\text{prey})$. So the ruler supports i . Thus, we get a contradiction by assuming $X_i > 0$ and $L_i > 1$. Hence, if the ruler preys, either $X_i \leq 0$ or $L_i \leq 1$. \square

Proposition A.8. *For $i \in \{E, C\}$, the ruler supports i if and only if $X_i > 0$ and $L_i > 1$ and either $X_j \leq 0$ or $L_j \leq L_i$ for $i \neq j \in \{E, C\}$.*

Proof. If: Suppose $X_i > 0$ and $L_i > \max\{1, L_j\}$. Then i exits in response to PREY (by lemma A.6b and c, case 4) or SUPPORT j (by lemma A.5b and c, case 3) but remains loyal when SUPPORT i (by lemma A.4).

Case 1. Suppose $X_j \leq 0$. Then j remains loyal whatever the ruler does (by lemmas A.4, A.5a, and A.6a). Thus, $U_R(\text{support } i) = L_i + L_j > U_R(\text{prey}) = 1 + L_j > U_R(\text{support } j) = L_j$. So the ruler supports i if $X_i > 0$ and $L_i > 1$ and $X_j \leq 0$.

Case 2. Suppose $X_j > 0$ and $L_i > \max\{1, L_j\}$. Then j exits in response to PREY (by lemma A.6b and c, case 4) or SUPPORT i (by lemma A.5b and c, case 3) but remains loyal

when SUPPORT j (by lemma A.4). Thus, $U_R(\text{support } i) = L_i > \max\{U_R(\text{support } j) = L_j, U_R(\text{prey}) = 1\}$. So the ruler supports i .

Since the ruler supports i in all cases, the ruler supports i if $X_i > 0$, $L_i > 1$ and either $X_j \leq 0$ or $L_j \leq L_i$.

Only If: Suppose the ruler supports some $i \in \{E, C\}$. Prove the contrapositive to show a contradiction. Suppose it's not the case that $X_i > 0$ and $L_i > 1$ and either $X_j \leq 0$ or $L_j \leq L_i$ for $i \neq j \in \{E, C\}$. By De Morgan's law, either it's not the case that $X_i > 0$ and $L_i > 1$ or it's not the case that $X_j \leq 0$ or $L_j \leq L_i$ for $i \neq j \in \{E, C\}$.

Case 1. Suppose it's not the case that $X_i > 0$ and $L_i > 1$ for some $i \in \{E, C\}$. By De Morgan's law, $X_i \leq 0$ or $L_i \leq 1$. We can ignore j 's action, so the ruler receives $L_j \in \{0, x\}$ for $x \in \mathfrak{R}$. If $X_i \leq 0$, it follows from lemma A.6 that $\min U_R(\text{prey}) = 1 + L_i + L_j$ and from lemma A.4 that $\max U_R(\text{support } i) = L_i + L_j$ for all $i \in \{E, C\}$. Since $\min U_R(\text{prey}) > \max U_R(\text{support } i)$, the ruler preys if $X_i \leq 0$. If $L_i \leq 1$, it follows from lemma A.6 that $\min U_R(\text{prey}) = 1 + L_j$ and from lemma A.4 that $\max U_R(\text{support } i) = L_i + L_j$ for all $i \in \{E, C\}$. Since $\min U_R(\text{prey}) > \max U_R(\text{support } i)$, the ruler preys if $L_i \leq 1$. Hence, the ruler preys if $X_i \leq 0$ or $L_i \leq 1$.

Case 2. Suppose it's not the case that $X_j \leq 0$ or $L_j \leq L_i$ for $i \neq j \in \{E, C\}$. By De Morgan's law, $X_j > 0$ and $L_j > L_i$. If $L_j > 1$, it follows from the "If" part of this proof that the ruler supports j . Now hold i 's action constant, so the ruler receives $L_i \in \{0, x\}$ for $x \in \mathfrak{R}$. If $L_j \leq 1$, then it follows from lemma A.6 that $\min U_R(\text{prey}) = 1 + L_i$ and from lemmas A.3–A.5 $\max U_R(\text{support } j) = \max U_R(\text{support } i) = L_j + L_i$. Since $\min U_R(\text{prey}) > \max U_R(\text{support } i)$, the ruler preys if $L_j \leq 1$. Thus, the ruler does not support i if $X_j > 0$ and $L_j > L_i$.

In either case, the ruler does not support i . Thus, we get a contradiction by proving the contrapositive. Hence, if the ruler supports i , $X_i > 0$ and $L_i > 1$ and either $X_j \leq 0$ or

$L_j \leq L_i$ for $i \neq j \in \{E, C\}$. \square

In words, proposition A.7 says the following.

1. The ruler preys upon and neglects both groups' interests just in case both elites and citizens either lack credible exit options, or possess loyalty that is worth less than whatever benefits accrue to the ruler from retaining discretion over the good.

In words, proposition A.8 says the following.

2. The ruler advances citizens' general interests just in case citizens have credible exit options and either (a) loyalty that is both more valuable than the elites' loyalty and more valuable than whatever benefits accrue to the ruler from retaining discretion over the good, regardless of the elite's exit options; or (b) loyalty that is more valuable than whatever benefits accrue to the ruler from retaining discretion over the good and the elite have no exit options. In the latter case, the elite's loyalty could be more valuable than the citizens.
3. The ruler advances elites' interests just in case the latter have credible exit options and either (a) loyalty that is more valuable than the citizens' loyalty and more valuable than whatever benefits accrue to the ruler from retaining discretion over the good, regardless of citizens' exit options; or (b) loyalty that is more valuable than whatever benefits accrue to the ruler from retaining discretion over the good and the citizens have no exit options. In the latter case, the citizens' loyalty could be more valuable than the elites'.

APPENDIX B

An Analysis of Feasibility

My aim in this appendix is to articulate a conceptual framework that helps discipline our evaluation of feasibility claims.¹ I claim that feasibility should be analyzed in terms of *constrained possibility*: a political proposal ϕ is feasible just in case ϕ is possible in view of certain relevant facts. My analysis draws on work in the semantics of modals. On the standard view, possibility claims express existential quantification over a contextually restricted set of possible worlds.² Hence, more precisely, ϕ is feasible just in case there exists a world at which ϕ is realized among the set of possible worlds that are similar to the actual world in certain respects. This set is composed by ordering and placing constraints on worlds along those comparative dimensions that are deemed relevant for thinking about feasibility. My objective here is to sketch a formal apparatus that models how we compose this set and, hence, how we should evaluate feasibility claims.

That the concept of feasibility should be given a modal analysis is likely not surprising. But the formal apparatus is helpful for locating the areas where substantive theorizing is

1. For other discussions of feasibility, see Brennan and Pettit, “The Feasibility Issue”; Tyler Cowen, “The Importance of Defining the Feasible Set,” *Economics and Philosophy* 23, no. 1 (2007): 1–14; Mark Jensen, “The Limits of Practical Possibility,” *Journal of Political Philosophy* 17, no. 2 (2009): 168–184; Juha Räikkä, “The Feasibility Condition in Political Theory,” *Journal of Political Philosophy* 6, no. 1 (1998): 27–40; Sanjay G. Reddy, “The Role of Apparent Constraints in Normative Reasoning: A Methodological Statement and Application to Global Justice,” *The Journal of Ethics* 9 (2005): 119–125.

2. Angelika Kratzer, “The Notional Category of Modality,” in *Words, Worlds, and Contexts: New Approaches in Word Semantics*, ed. Hans Jürgen Eikmeyer and Hannes Rieser (Berlin: de Gruyter, 1981), 38–74; Angelika Kratzer, “Modality,” in *Semantics: An International Handbook of Contemporary Research*, ed. Arnim von Stechow and Dieter Wunderlich (Berlin: de Gruyter, 1991), 639–650.

required when evaluating feasibility claims. My analysis suggests that the real theoretical work consists in identifying the considerations we should take as the salient dimensions for comparing possible worlds. To this end, I present a reasonably comprehensive list of such dimensions as an initial offering; *inter alia*, this list includes technological, motivational, and institutional factors. Since this analysis enables us to be flexible with regard to where we locate the constraints along these dimensions, I also offer some broad guidelines for thinking about how restrictive we should be when evaluating the feasibility of particular kinds of political proposals. The result is a conceptual framework to guide our assessment of feasibility claims in political philosophy.

B.1 A general schema

Desiderata for an analysis. Here, I simply stipulate the desiderata that any analysis must meet. (I justify these desiderata at some length in a longer version of this paper.)

Any analysis of feasibility must meet at least the following four conditions.

1. *Continuity.* Feasibility is most plausibly treated as a matter of degree. Any analysis should treat feasibility as a continuous rather than a dichotomous notion.
2. *Multidimensionality.* Numerous types of considerations affect a political proposal's feasibility. To wit, economic, cultural, environmental, and institutional considerations are distinct types of constraints, each of which impinges on a proposal's realization. Any analysis of feasibility should account for this multidimensionality. (I offer an initial list of the relevant dimensions below section B.2.)
3. *Flexibility.* The concept of feasibility is differentially employed depending on the normative question at hand. Any analysis of feasibility should be flexible enough to accommodate these different applications.
4. *Determinacy.* For any context, feasibility should have a determinate meaning. When a theorist employs the concept of feasibility to rule out a particular political prescription, we should have a reasonably precise grasp of, for example, what sorts of prescriptions would count as feasible beyond the particular proposal under consideration. Put differently, a theorist should be able to specify her understanding of feasibility with enough determinacy to enable her audience to apply her notion of feasibility to proposals not under consideration. Thus, any analysis of feasibility should permit one to specify the concept as applied to a particular context as determinately as is possible (acknowledging the unavoidable imprecision of our assessment of feasibility constraints).

The intuitive picture. As a first pass, say that a proposal ϕ is feasible just in case ϕ 's realization is practically possible. Let's treat ϕ as a proposition that describes the state of affairs required by the proposal. On the standard view of the semantics of modals, possibility claims express existential quantification over a contextually restricted set of possible worlds. In this case, the restriction is *circumstantial*: ϕ is feasible just in case ϕ 's realization is possible in view of certain facts.³ More precisely:

Initial pass: ϕ is *feasible* just in case there exists a world at which ϕ is realized among the set of worlds that are similar enough to the actual world at t ,

where t denotes the time at which the transition to ϕ is to start.

Two complications arise here. First, there does not seem to be any obvious way to capture the continuity of feasibility using this framework. Existential quantification is a dichotomous matter; either a world exists at which ϕ is realized or not. We need a way to delimit the set of worlds over which we quantify so as to make apparent that feasibility is a matter of degree. Second, I have spoken throughout of feasibility being a matter of “processes” and “pathways”; but this first pass fails to capture this. Whether there exists a world at which ϕ is realized is distinct from whether there is a pathway by which we can transition from a world to one at which ϕ is realized. So we need to amend this first pass to accommodate talk of processes and pathways.

Some terminology before continuing. A world w' is *circumstantially accessible* from another world w if w' is compatible with certain specified facts at w ; that is, w' is circumstantially accessible from w if the relevant facts at w do not rule out w' . For my purposes, say w' is circumstantially accessible from the actual world if (1) w' satisfies logical, nomological, and biological constraints and (2) w' shares a common history with the actual world up to t , the time at which the transition to ϕ is supposed to start. Elaborating on (1), w' is circumstantially accessible from the actual world if it does not entail any logi-

3. Kratzer, “Modality.”

cal contradictions, if it does not entail any violations of the actual world's physical laws, and if it does not entail that the structures, functions, or developmental processes of living organisms are fundamentally different from those of the organisms in the actual world.

To satisfy the multidimensionality condition, let's treat modal space as a N -dimensional metric space, with each of the N dimensions corresponding to a class of considerations we think represents a relevant constraint on our ability to transition to any political proposal. We can think of each of the N dimensions as an axis of the metric space. Suppose we arrange each world that is circumstantially accessible from the actual world along each of these axes according to the extent to which it deviates from the actual world along that dimension at t . In this way, we can treat a dimension as representing a similarity relation between worlds at t . A world at which human agents have radically different motivations from agents at the actual world is relatively far away from the actual world along the motivation dimension; a world that is pretty similar to the actual world with respect to technology is placed relatively close to the actual world along the technology dimension. Thus, each world has a location in the space that is defined by that world's distance from the actual world at t along each of its dimensions. (Note that we are judging the similarity of worlds *at* t ; that is, we are judging the extent to which *initial* conditions at a world are similar to *initial* conditions at the actual world. This permits worlds to differ greatly at times after t . Importantly, this leaves open the possibility that proposals that require radical deviations from the actual world at some time hence are feasible, so long as there is a causal path from the specified initial conditions to the conditions required by the proposal.)

To define feasibility in any context, we decide how far from the actual world we are willing to diverge — *at* t — along each dimension before we cease to count a world as feasible. Put differently, we determine which barriers at t we will treat as obstructing any

path to a proposal ϕ 's realization. These bounds define the set of feasible worlds, the set of worlds over which we quantify. To determine whether ϕ is feasible, we look at the set of feasible worlds and determine whether ϕ is realized at any world in the feasible set. If so, then we can say that ϕ is a feasible proposal; if not, not.

More precisely. Suppose we arrange all the worlds that are circumstantially accessible from the actual world in a N -dimensional metric space with the actual world, denoted by α , at the origin. For each dimension i , we stipulate a function $f_i : W \rightarrow \mathbb{R}$, which gives a world's (directed) distance from α along dimension i at t . Each world w can now be characterized as a vector $\mathbf{w} = (f_1(w), \dots, f_N(w))$, which represents the conditions that obtain at w at t . (Note that the space need not be dense; where some conjunction of, say, technological, motivational, and institutional requirements is inconsistent, there will be a gap at the location characterized by that vector. Also notice that there might be more than one world at any location, namely, those worlds that are characterized by the same vector but differ in other (irrelevant, for our purposes) respects.)

Transposing our first pass above, we want to say that ϕ is feasible just in case there exists a world at which ϕ is realized among the set of feasible worlds. Let Θ be the set of feasible worlds; that is, the members of Θ are those worlds that are not only circumstantially accessible from α , but also satisfy some limits on how far conditions at a world can deviate from conditions at α at t . Our present task is to delimit Θ . We can do this in several ways. One straightforward way is to specify a vector of constraints $(\gamma_1, \dots, \gamma_N)$, $\gamma_i \in \mathbb{R}$, which stipulate (as precisely as possible) how close to α (or how similar to α) some world w must be along each dimension to count as feasible. (I stress that these distance judgments are made by comparing worlds at t .) So let each Θ_i be the set of worlds that satisfy the constraint for that dimension: $\Theta_i = \{w : f_i(w) \leq \gamma_i\}$. Then Θ contains the intersec-

tion of the Θ_i s: $\Theta = \bigcap_{i=1}^N \Theta_i$. (This defines a N -dimensional cube, or hypercube, which contains all the worlds taken to satisfy the constraints.)

A more general way to delimit Θ — which permits a more nuanced definition of Θ — is to define a metric for the modal space. A clear candidate here is the Euclidean norm:

$$\|\mathbf{w}\| = \sqrt{f_1(w)^2 + \dots + f_N(w)^2}.$$

Then Θ is the set of worlds that are within some specified distance of the actual world: $\Theta = \{w : \|\mathbf{w}\| \leq \Gamma\}$, where $\Gamma \in \mathbb{R}$ is the maximum distance a world w can be from α to count as feasible. (Γ could be a function of the γ_i s if one desired.) If we want to make it more or less difficult for worlds with higher values along some dimension i to count as feasible, we define a weighted distance function, $d : W \times W \rightarrow \mathbb{R}$:

$$d(w_1, w_2) = \sqrt{\beta_1[f_1(w_1) - f_1(w_2)]^2 + \dots + \beta_N[f_N(w_1) - f_N(w_2)]^2},$$

where $\beta_i \geq 0$. (For example, one might want to weight dimensions according to their respective costliness of transitioning to new conditions.) If $\beta_i = 1$, then the distance along dimension i enters unweighted. Now $\Theta = \{w : d(w, \alpha) \leq \Gamma\}$. If we want to add hard constraints along some dimension i , then let $\Theta = \{w : d(w, \alpha) \leq \Gamma \wedge f_i(w) \leq \gamma_i\}$. The point here is simply to show that we have options in delimiting Θ ; we are not forced to define Θ in any particular way.⁴

Having delimited Θ , we are now in a position to define feasibility more carefully: ϕ is *feasible* just in case there exists a world at which ϕ is realized in Θ . But since we want to make clear the point about transitional processes, we should be a little more precise.

4. Note also that using a metric space and distance function is not necessary. More generally, we could say that f_i generates a partial preorder on worlds with respect to dimension i . We could then define a function F to yield a total order as a function of the f_i s. Finally, we could define a function G to select worlds from the total order to determine the members of Θ . I use a metric space and a distance function because they yield a helpful graphical representation; hence, they are useful for expository purposes. (Thanks to Alex Silk for pointing this out to me.)

Recall that t denotes the time at which the transition to ϕ is supposed to start; we want to know whether ϕ is realized at some time after t . Hence,

Final pass: ϕ is *feasible* if and only if there exists a world in Θ at which ϕ is realized at $t' > t$.⁵

Formally,

$$\phi \text{ is feasible iff } \exists w \in \Theta \text{ and } \exists t' > t : \llbracket \phi \rrbracket^{w,t'} = 1.$$

We can capture the continuity of feasibility as follows. Let \mathcal{P} be the set of worlds at which ϕ is realized (formally: $\mathcal{P} = \{w : \llbracket \phi \rrbracket^w = 1\}$) and let \mathcal{Q} be the set of worlds at which ψ is realized, where ψ is an alternative proposal (formally: $\mathcal{Q} = \{w : \llbracket \psi \rrbracket^w = 1\}$).

- A proposal ψ is *at least as feasible as* (\supseteq) ϕ iff $\forall w \in \mathcal{P} \exists w' \in \mathcal{Q} : d(w', \alpha) \leq d(w, \alpha)$.
- ψ is *more feasible than* ϕ iff $\psi \supseteq \phi \wedge \neg(\phi \supseteq \psi)$.
- Finally, ψ and ϕ are *equally feasible* iff $\psi \supseteq \phi \wedge \phi \supseteq \psi$.

Notice that this analysis satisfies each of the desiderata mentioned above. The analysis is obviously *multidimensional*. The analysis provides *flexibility* because the feasible set's bounds are free to vary depending on the question at hand. When doing ideal theory, we can loosen the constraints on each dimension (or ignore some dimensions altogether), allowing a more inclusive definition of feasibility. When doing nonideal theory or, closer to the ground, making specific policy recommendations for specific problems, we can tighten the constraints on each dimension individually, altering these constraints as is fit given the salience of any particular dimension. This yields a more restrictive definition of feasibility, depending on the case at hand. Notice that this also addresses the complication introduced by process-talk. When we set the constraints that delimit the feasible set, we determine which barriers we will take as blocking a proposal's realization within the salient time frame. Put differently, the worlds that satisfy the specified constraints are those worlds to

5. The problem of how to determine whether ϕ is realized at a counterfactual world arises here. I briefly address this issue at the end of this appendix.

which there are open transitional paths. (Importantly, this does not imply that the transition is easy.)

The analysis provides *determinacy* by requiring theorists to be specific (or, more accurately, as specific as possible) about where they set the constraints on each dimension. This prevents theorists from ruling out (or ruling in) proposals for vaguely-defined reasons; by requiring specificity, theorists must expose their assumptions about what counts as feasible in a given context, thereby subjecting those assumptions to scrutiny.

Finally, that the analysis satisfies *continuity* is straightforward — each of the N dimensions that carve up the space is a continuum and circumstantially accessible worlds can be located anywhere along each continuum (provided that some location does not entail an impossible world). Moreover, worlds are free to vary along these dimensions independently, as are the constraints we set on any dimension. In this way, feasibility is *fully continuous*. It is true that any *particular* definition of the feasible set treats feasibility as dichotomous; the definition draws a bright line around those worlds that are set apart for consideration. But this is only for the purposes of deliberation; the line is set to rule out the worlds we deem too distant for practical consideration. In principle, the line is free to vary in a continuous fashion across deliberative contexts. Moreover, across the whole space, we can intelligibly judge some worlds more or less feasible than others to the extent that the former are closer to the actual world than the latter.

B.2 Putting the schema to work

The foregoing provides only the skeleton of an analysis of feasibility. A further advantage of this framework is that it indicates where substantive theorizing must take place when defining feasibility. This work must occur at two places. First, we must figure out which dimensions — which classes of considerations — represent relevant constraints on

our capacity to transition to any given political prescription. Second, we must articulate some broad guidelines for determining how inclusive the feasible set should be depending on the normative question of interest. I take up the first question in the next section and leave the second to the following section.

Which dimensions? Recall that each dimension corresponds to an axis of the modal space, and that each dimension can be treated as a similarity relation among worlds — we organize worlds along a dimension according to their degree of similarity to the actual world at t with respect to some set of considerations that affect our capacity to realize a political proposal. But which sets of considerations should we use to define the dimensions of the modal space? Common sense tells us that a proposal ϕ is practically impossible for us and therefore infeasible if ϕ requires the realization of a logically contradictory state of affairs, or the violation of the physical laws of our universe, or that living organisms have structures, functions, or developmental processes that are fundamentally different from living organisms in the actual world. Our evaluation of feasibility claims can safely ignore all worlds that violate these logical, nomological, and biological constraints. Hence, we do not need any dimensions that correspond to these hard constraints to organize the space of possible worlds; we simply stipulate that only those worlds that meet these constraints are circumstantially accessible from the actual world and populate the modal space with only those worlds. Nor do we need a history dimension, since we are only interested in determining whether ϕ can be reached from the specified initial conditions, which presupposes the actual history. Thus, we ignore all worlds whose histories up to time t diverge from the history of the actual world. Aside from logic, physics, biology, and history, what other types of considerations influence ϕ 's feasibility?

Here, I only canvass what I take to be a reasonably comprehensive list of feasibility

considerations. This is not meant to be the definitive list, so I do not make much of a case for the inclusion of any particular dimension. Whether we should attend to any particular dimension depends on whether that dimension enriches our evaluation of particular feasibility claims. If we attend to some set of considerations and it consistently turns out not to make much difference to our assessment, that is some evidence that we do not need a distinct dimension corresponding to that set of considerations.

The following list characterizes the dimensions—or similarity relations among possible worlds—that I take to be representative of reasonably distinct constraints on our capacity to realize ϕ .

1. *Environmental constraints.* Environmental constraints comprise things like a location's climate, natural resource profile, and topography. These factors affect the realization of, for example, particular development strategies. Agriculturally-based development becomes less feasible as the local environment becomes less hospitable to agriculture; certain industries become more viable as certain natural resources become increasingly available locally and as the location has increased access to natural transportation routes (e.g., navigable waterways).
2. *Technological capabilities.* Technological capabilities encompass the tools, techniques, and organizational schemes by which we implement ϕ . Consider the following examples of technological constraints on our ability to realize some ϕ . The extent to which we can implement a proposal to redistribute global wealth depends on the means available to collect and distribute tax revenue, as well as the means available to monitor and enforce compliance. The feasibility of widespread cooperation depends on the available communication technology, while the feasibility of broad-based economic development depends on the available productive technology.
3. *Cognitive capabilities.* Human agents are subject to multifarious cognitive biases and other cognitive limitations. This means that our ability to implement political proposals is often limited by attending to irrelevant or misleading considerations or making choices that are otherwise less than optimally rational. Moreover, we lack the computational capacity and often the social scientific knowledge required to anticipate the consequences of our reform choices with any accuracy. These limitations surely influence the feasibility of any ϕ , as well as our ability to specify the conditions that must arise to make ϕ feasible.
4. *Intrinsic motivational capacities.* Intrinsic motivational capacities comprise the intrinsic features of individual human agents that influence their ability to be motivated to undertake those actions that are required by ϕ . These include preferences, the capacity to put oneself in another's position, and capacities to overcome various affective biases (e.g., kin preference), prejudices, and fears. To illustrate the point, it is a common refrain among nationalists that people cannot be motivated

to pursue cosmopolitan political objectives, thereby undermining the feasibility of cosmopolitan proposals.⁶ The general point is that some ϕ is feasible only insofar as the relevant agents can be motivated to undertake the actions required by ϕ given certain intrinsic constraints on motivational capacity.

5. *Extrinsic motivational factors.* The set of extrinsic motivational factors consists of features of an agent's environment that interface with her intrinsic motivational capacities and affect her motivation to undertake the actions required by ϕ . These comprise social norms, which dictate appropriate conduct, as well as the incentive structure to which agents respond—i.e., the structure of rewards and punishments meant to provoke or elicit certain types of conduct. The incentive structure subsequently affects the extent to which ϕ is *incentive compatible*, the extent to which the structure of rewards and punishments required to induce the actions required by ϕ accords with the relevant agents' intrinsic motivational limitations.⁷
6. *Resource availability (budget constraints).* Clearly, ϕ is feasible only to the extent to which the resources required to realize ϕ —including money and labor—are available. This is why the economic concept of a *budget constraint* is helpful here. In economic analysis, individuals' consumption options are limited to those that are affordable given a particular income level. A similar notion applies here: our policy or institutional options are limited to those that can be implemented given the available resources.
7. *Institutional factors.* Extant institutional capacities and constraints affect the extent to which the policies required by some ϕ can be enacted. To cite two examples: the number and distribution of veto points in an institutional structure—that is, the number of offices that can prevent deviation from the status quo—affect the extent to which reform legislation can be passed;⁸ the extent to which citizens can hold a political official accountable for the latter's policy choices affects the provision of public goods, such as security, individual liberties, infrastructure, and investment in economic development.⁹ Hence, any evaluation of ϕ 's feasibility should attend to the institutional conditions required to realize ϕ .
8. *Social factors.* Social factors comprise those features of a society or of social relationships that affect individuals' ability (not motivation) to undertake the actions required by ϕ . These include the extent to which individuals are united in their pursuit of a common goal, the extent to which individuals' personal goals are harmonized, and the depth and extent of civil conflict. Since these factors are influenced by

6. Miller, *On Nationality*. For an empirical evaluation of the claim, see Gillian Brock and Quentin D. Atkinson, "What can Examining the Psychology of Nationalism Tell Us About Our Prospects for Aiming at the Cosmopolitan Vision?" *Ethical Theory & Moral Practice* 11, no. 2 (2008): 165–179.

7. For a discussion of feasibility that treats it almost entirely in terms of incentive compatibility, see Brennan and Pettit, "The Feasibility Issue." Frey argues that focusing entirely on designing incentive schemes to prevent opportunistic behavior "crowds out" civic virtue; thus, cultivating appropriate social norms is important for motivating cooperative behavior. See Bruno Frey, "A Constitution for Knaves Crowds out Civic Virtues," *The Economic Journal* 107, no. 443 (1997): 1043–1053; Bruno Frey, "Institutions and Morale: The Crowding-Out Effect," in *Economics, Values, and Organization*, ed. Avner Ben-Ner and Louis G. Putterman (Cambridge: Cambridge University Press, 1999), 437–460.

8. George Tsebelis, *Veto Players: How Political Institutions Work* (Princeton: Princeton University Press, 2002).

9. Bueno de Mesquita et al., *The Logic of Political Survival*.

pervasive cultural and religious practices, I take social factors to subsume cultural factors. The importance of social factors for feasibility is reflected, for example, in Rawls's move to make the "fact of reasonable pluralism" — that is, the fact that citizens within a liberal society hold "incompatible yet reasonable comprehensive doctrines" about what constitutes a worthwhile life—a central constraint on the choice of political principles.¹⁰

I take the foregoing list as nothing more than a start, although a reasonably comprehensive start, on the task of specifying the dimensions we use to frame our evaluation of feasibility claims. I said earlier that I take this task to be where the "real work" needs to be done. By this, I mean that any evaluation of feasibility claims fundamentally rests on the considerations we take to be salient for evaluating the feasibility of particular political proposals. Much of the confusion perpetuated by the political philosophy literature on this score is tied to a failure to say precisely which sets of considerations are taken to affect the feasibility of particular proposals. The aim of the foregoing analysis is to provoke theorists to be precise about the considerations they take to be salient and, further, to specify as precisely as possible the constraints they are placing along each of the salient dimensions when declaring a proposal (in)feasible. But the flexibility afforded by my analysis of feasibility permits theorists to precisify the notion of feasibility in accordance with the particular normative question they engage. In the next section, I offer some guidance on how to tailor the notion of feasibility to suit one's theoretical purposes.

Feasibility in political philosophy. The motivation for requiring that feasibility be given an analysis that permits flexible use is this: theorists use a more or less inclusive notion of feasibility depending on their normative question of interest. I have already given several examples of this: we use a more inclusive notion to evaluate the feasibility of organizing state institutions in accordance with Rawls's two principles of justice than to evaluate the feasibility of Pogge's proposed amendment for averting coups; we evaluate

10. John Rawls, *Political Liberalism* (New York: Columbia University Press, 1996), xvi and *passim*.

the feasibility of a particular proposal to democratize international institutions such as the World Trade Organization differently depending on whether our aim is to identify near-term policy solutions, institutional arrangements that are practicable under conditions of pervasive injustice more generally, or institutional arrangements to which we should aspire as a long-term goal. Which dimensions we take to be salient and where to set the constraints along these dimensions will differ across domains of inquiry. Can we identify some general rules-of-thumb to be applied in distinct domains?

Before proposing some general rules, I first distinguish between four broadly-defined types of domain. The first is what I have called *ground-level policy analysis*. The objective here is to identify practicable near-term policy reforms to address particular social problems (e.g., an embargo on Sudan to address ongoing armed civil conflict and genocide there). It is important for ground-level policy analysis that the proposed reforms be practicable under conditions as we find them in the actual world.

The second and third types I call *clinical theory* and *nonideal aspirational theory*. Note that these are species of what is conventionally called *nonideal theory*. The conventional nonideal theory label is ambiguous between three distinct conceptions of the task of nonideal theory: (1) theorizing that identifies intermediate institutional reforms to help us transition from actual institutions to fully just institutions; (2) theorizing that identifies institutions that we should aspire to implement under nonideal conditions (i.e., under conditions of persistent injustice); and (3) theorizing that prescribes near-term or midrange institutional solutions to general classes of actual injustice (e.g., institutional solutions to address the problem of genocide more generally, not some particular genocide). We might call these (1) transitional theory,¹¹ (2) nonideal aspirational theory, and (3) clinical theory

11. We should take care to distinguish this use of “transitional” from its use in the burgeoning transitional justice literature. The latter addresses the moral, legal, and political issues that arise in the wake of war and civil conflict; e.g., the appropriate role of war crimes tribunals in facilitating the transition to a cooperative society.

respectively. Failing to disambiguate *nonideal theory* has led political theorists to conflate the various objectives; for example, theorists commonly identify clinical theory with transitional theory.¹² But we should be careful to differentiate types of theorizing that are animated by distinct objectives; this is especially true for the present discussion, since each of these might appeal to more or less inclusive notions of feasibility.

(I do not treat transitional theory separately here, for two reasons. First, transitional theory is not a worthwhile enterprise. As I argue in chapter 3, the institutional arrangements we identify as fully just offer, at best, misleading guidance for the task of identifying morally desirable institutions under nonideal conditions. Second, it is not clear to me that the task of transitional theory requires a notion of feasibility that is more or less restrictive than either nonideal aspirational theory or clinical theory.)

The last type of theoretical domain is what I have been calling *ideal theory*. The objective here is to examine what social and political arrangements might be like “under reasonably favorable but still possible historical conditions, conditions allowed by the laws and tendencies of the social world.”¹³ Put differently, ideal theory aims to work out fully just institutions given reasonably optimistic assumptions about what is possible for human agents who are like us in essential respects and who live in a world that resembles ours in certain fundamental respects. What institutions could arise in a world of modest but not acute resource scarcity, where technological and social limitations do not significantly bind our ability to bring about just states of affairs, and where individual agents can be motivated to pursue just social arrangements for their own sake? That is a question for ideal theory.

Distinguishing between domain types as above lends itself fairly straightforwardly to

12. See my discussion above in sec. 3.2.

13. John Rawls, *Justice as Fairness: A Restatement*, ed. Erin Kelly (Cambridge, MA: Harvard University Press, 2001), 4.

the business of articulating general guidelines for evaluating feasibility claims in each of these domains. Ground-level policy analysis resides at the restrictive end of the spectrum. Since we are here interested in entertaining proposals that are implementable in the near-term under actual conditions, we should restrict our attention to those worlds that very closely resemble the actual world along each of the dimensions enumerated above. (Recall that all similarity judgments are made with respect to t , the time at which the implementation of the proposal is to begin.) We appropriately reject as infeasible those proposals that are not realized in any worlds that are very close to our own.¹⁴ We can afford to be slightly more inclusive when doing clinical theory. This is because our interest in doing clinical theory extends to midrange solutions; thus, we should consider worlds at which conditions obtain that are likely to arise given actual conditions. We should continue to employ each of the dimensions discussed above, but we can afford to set the constraints to include worlds that are a little further away from the actual world.

At the other extreme, we can (and perhaps should) employ a fairly inclusive notion of feasibility when doing ideal theory. Here, we can consider worlds that depart quite substantially from our own, yet are still recognizable to us as salient possibilities. This permits us to set the constraints on each of the dimensions to include worlds that are quite distant from the actual world, and perhaps even ignore some dimensions altogether. (For example, for certain questions, we might ignore technological or social limitations altogether.)

Nonideal aspirational theory falls somewhere between clinical theory and ideal theory. This is because the former shares features with each of the latter. On the one hand, we want to entertain political proposals that are to operate under conditions of persistent injustice; thus, we must assume the presence of some agents who are unwilling or unable to meet their obligations due to some notable limitations. On the other hand, we seek proposals

14. Cf. Rääkkä's notion of political feasibility in "The Feasibility Condition," 28f.

that are worthy to be called aspirational; we want to explore the limits of what is possible under conditions of persistent injustice. Thus, we must assume conditions that are less restrictive than those that we find in the actual world—for example, that a greater number of agents are subject to fewer motivational limitations or that institutions possess greater capacity to overcome collective action problems. The flexibility afforded by the analysis I have offered permits us to get the right mix of constraints and capacities. We can be more restrictive along some dimensions while being less restrictive along others, depending on the issue at hand; alternatively, we can uniformly set the constraints at a middle range between the two extremes.

The general point in this section is that the notion of feasibility we should employ in any given domain of inquiry should be as restrictive or inclusive as we need for the purpose at hand. We should set our constraints to include only worlds that are quite close to ours when we appraise proposals that are meant to be implemented under conditions that closely resemble those that obtain in the actual world, and move the constraints outward as our question of interest leads us to consider proposals that are meant to operate under conditions that diverge from our own.

B.3 Concluding remark

The practical utility of this framework rests on our ability to determine whether some proposal is realized in a counterfactual world, and I promised above to briefly address the question of how we might do this. Although it is beyond the scope of this paper to defend a response at any length, I conjecture that formal (mathematical) game theoretic models can be put to good use here. Formal models have two virtues that lend them to this task. First, models are useful *isolating devices*; that is, models serve as rigorous “thought experiments” that control for causal noise through the use of idealizing or exaggerating

assumptions with the aim of isolating the causal relationship of interest and placing it in an environment where the causal connection is stable.¹⁵ Given that our analysis of a proposal's feasibility focuses on the effect of select conditions rather than the totality of potential causal factors (however remote or indirect their effect), a model's capacity to isolate is helpful. Second, models are useful for investigating counterfactual worlds.¹⁶ The modeller has wide latitude in setting the model's parameters; thus, within certain limits (e.g., the model must be mathematically tractable), the modeller can construct (models of) counterfactual worlds by setting the model's parameters to match conditions in the counterfactual world of interest. This enables us to investigate a wide range of parameters and determine necessary or sufficient conditions for a proposal's implementation. We can then compare these conditions to those we find in the actual world to determine how closely the former resemble the latter. Clearly, I have said too little to make a persuasive point; but I hope I have at least indicated a promising line of research.

15. Cf. Uskali Mäki, "Models are Experiments, Experiments are Models," *Journal of Economic Methodology* 12, no. 2 (2005): 303–315; Uskali Mäki, "MISSing the World. Models as Isolations and Credible Surrogate Systems," *Erkenntnis* 70, no. 1 (2009): 29–43.

16. Cf. Robert Sugden, "Credible Worlds: The Status of Theoretical Models in Economics," *Journal of Economic Methodology* 7, no. 1 (2000): 1–31; Robert Sugden, "Credible Worlds, Capacities and Mechanisms," *Erkenntnis* 70, no. 1 (2009): 3–27.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Acemoglu, Daron, Simon Johnson, and James A. Robinson. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91, no. 5 (2001): 1369–1401.
- Alesina, Alberto, and David Dollar. "Who Gives Foreign Aid to Whom and Why?" *Journal of Economic Growth* 5 (2000): 33–63.
- Amnesty International. *Report 08: The State of the World's Human Rights*. London: Amnesty International, 2008.
- Anand, Sudhir, Paul Segal, and Joseph Stiglitz, eds. *Debates on the Measurement of Global Poverty*. Oxford: Oxford University Press, 2009.
- Anderson, Lisa. "The State in the Middle East and North Africa." *Comparative Politics* 20, no. 1 (1987): 1–18.
- Auty, Richard, ed. *Resource Abundance and Economic Development*. New York: Oxford University Press, 2001.
- Axelrod, Robert, and Robert O. Keohane. "Achieving Cooperation under Anarchy: Strategies and Institutions." *World Politics* 38, no. 1 (1985): 226–254.
- Baland, Jean-Marie, and Patrick Francois. "Rent-Seeking and Resource Booms." *Journal of Development Economics* 61 (2000): 527–542.
- Barry, Christian. "Applying the Contribution Principle." *Metaphilosophy* 36, nos. 1/2 (2005): 210–227.
- . "Understanding and Evaluating the Contribution Principle." In *Real World Justice*, edited by Andreas Follesdal and Thomas W. Pogge. Dordrecht: Springer, 2005.
- Bates, Robert H. *When Things Fell Apart: State Failure in Late-Century Africa*. New York: Cambridge University Press, 2008.
- . *Prosperity and Violence: The Political Economy of Development*. 2nd ed. New York and London: W. W. Norton & Co., 2009.
- Bates, Robert H., and Da-Hsiang Donald Lien. "A Note on Taxation, Development, and Representative Government." *Politics & Society* 14, no. 1 (1985): 53–70.

- Beblawi, Hazem. "The Rentier State in the Arab World." In *The Rentier State: Nation, State and the Integration of the Arab World*, edited by Hazem Beblawi and Giacomo Luciani. London: Croom Helm, 1987.
- Beitz, Charles R. *Political Theory and International Relations*. 2nd ed. Princeton: Princeton University Press, 1999.
- Bell, Joseph, Paul Collier, Robert Conrad, Thomas Heller, et al. *Natural Resource Charter*, 2010. URL: <http://www.naturalresourcecharter.org>.
- Blake, Michael. "Distributive Justice, State Coercion, and Autonomy." *Philosophy & Public Affairs* 30, no. 3 (2001): 257–296.
- Brambor, Thomas, William Roberts Clark, and Matt Golder. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14 (2006): 63–82.
- Brennan, Geoffrey, and James Buchanan. "The Normative Purpose of Economic 'Science': Rediscovery of an Eighteenth-Century Method." *International Review of Law and Economics* 1 (1981): 155–166.
- Brennan, Geoffrey, and Philip Pettit. "The Feasibility Issue." In *Oxford Handbook of Contemporary Philosophy*, edited by Frank Jackson and Michael Smith. Oxford: Oxford University Press, 2005.
- Brewer-Carías, Allan R. *Dismantling Democracy in Venezuela: The Chávez Authoritarian Experiment*. New York: Cambridge University Press, 2010.
- Brock, Gillian. *Global Justice: A Cosmopolitan Account*. New York: Oxford University Press, 2008.
- Brock, Gillian, and Quentin D. Atkinson. "What can Examining the Psychology of Nationalism Tell Us About Our Prospects for Aiming at the Cosmopolitan Vision?" *Ethical Theory & Moral Practice* 11, no. 2 (2008): 165–179.
- Brunnschweiler, Christa, and Erwin Bulte. "Natural Resources and Violent Conflict: Resource Abundance, Dependence, and the Onset of Civil Wars." *Oxford Economic Papers* 61, no. 4 (2009): 651–674.
- Buchanan, Allen. *Justice, Legitimacy, and Self-Determination: Moral Foundations for International Law*. New York: Oxford University Press, 2004.
- . "Institutionalizing the Just War." *Philosophy & Public Affairs* 34, no. 1 (2006): 2–38.
- Buchanan, Allen, and Robert O. Keohane. "The Preventive Use of Force: A Cosmopolitan Institutional Proposal." *Ethics & International Affairs* 18, no. 1 (2004): 1–22.
- Bueno de Mesquita, Bruce, and George W. Downs. "Intervention and Democracy." *International Organization* 60 (2006): 627–649.
- Bueno de Mesquita, Bruce, and Alastair Smith. "Foreign Aid and Policy Concessions." *Journal of Conflict Resolution* 51, no. 2 (2007): 251–284.

- Bueno de Mesquita, Bruce, and Alastair Smith. "A Political Economy of Aid." *International Organization* 63, no. 2 (2009): 309–340.
- Bueno de Mesquita, Bruce, James D. Morrow, Randolph M. Siverson, and Alastair Smith. "Political Competition and Economic Growth." *Journal of Democracy* 12, no. 1 (2001): 58–72.
- Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson, and James D. Morrow. *The Logic of Political Survival*. Cambridge, MA: MIT Press, 2003.
- Byers, Michael, and Simon Chesterman. "Changing the Rules About Rules? Unilateral Humanitarian Intervention and The Future of International Law." In *Humanitarian Intervention: Ethical, Legal, and Political Dilemmas*, edited by Jeffrey L. Holzgrefe and Robert O. Keohane. Cambridge: Cambridge University Press, 2003.
- Cabrera, Luis. *Political Theory of Global Justice: A Cosmopolitan Case for the World State*. New York: Routledge, 2004.
- Caney, Simon. "Cosmopolitan Justice and Equalizing Opportunities." *Metaphilosophy* 32, nos. 1/2 (2001): 113–134.
- . *Justice Beyond Borders: A Global Political Theory*. New York: Oxford University Press, 2005.
- . "Cosmopolitan Justice and Institutional Design: An Egalitarian Liberal Conception of Global Governance." *Social Theory and Practice* 32, no. 4 (2006): 725–756.
- Caselli, Francesco, and Tom Cunningham. "Leader Behaviour and the Natural Resource Curse." *Oxford Economic Papers* 61, no. 4 (2009): 628–650.
- Chatterjee, Deen K., and Don E. Scheid, eds. *Ethics and Foreign Intervention*. Cambridge: Cambridge University Press, 2003.
- Chaudry, Kiren Aziz. *The Price of Wealth: Economies and Institutions in the Middle East*. Ithaca and London: Cornell University Press, 1997.
- Chayes, Abram, and Antonia H. Chayes. "On Compliance." *International Organization* 47, no. 2 (1993): 175–205.
- Chesterman, Simon. *Just War or Just Peace? Humanitarian Intervention and International Law*. New York: Oxford University Press, 2001.
- Ching, Francis D.K. *Architecture: Form, Space, and Order*. 3rd ed. Hoboken, NJ: Wiley, 2007.
- CIA. "Rank Order – Oil Production." *The CIA World Factbook*. Central Intelligence Agency, 2008.
- Clark, William Roberts, Matt Golder, and Sona Golder. "The Balance of Power Between Citizens and the State: Democratization and the Resource Curse." University of Michigan and Penn State University, unpublished manuscript, 2008.

- Clark, William Roberts, Paul Poast, Thomas Flores, and Robert R. Kaufman. "Why Do Autocrats Overachieve? Political Competition and Material Well-Being in Comparative Perspective." University of Michigan and Rutgers University, unpublished manuscript, 2010.
- Coady, C.A.J. "War For Humanity: A Critique." In Chatterjee and Scheid, *Ethics and Humanitarian Intervention*.
- Cohen, G.A. "Marxism and Functional Explanation." In *Analytical Marxism*, edited by John Roemer. Cambridge: Cambridge University Press, 1986.
- . "Facts and Principles." *Philosophy & Public Affairs* 31, no. 3 (2003): 211–245.
- Cohen, Joshua. "Philosophy, Social Science, Global Poverty." In Jaggar, *Pogge and His Critics*.
- Collier, Paul. *The Bottom Billion: Why the Poorest Countries are Failing and What Can Be Done About It?* New York: Oxford University Press, 2007.
- Collier, Paul, and Anke Hoeffler. "On Economic Causes of Civil War." *Oxford Economic Papers* 50 (1998): 563–573.
- . "Greed and Grievance in Civil War." *Oxford Economic Papers* 56, no. 4 (2004): 563–595.
- Coram, Bruce Talbot. "Second Best Theories and the Implications for Institutional Design." In *The Theory of Institutional Design*, edited by Robert E. Goodin. New York: Cambridge University Press, 1996.
- Corrales, Javier. "Hugo Boss." *Foreign Policy*, no. 152 (2006): 32–40.
- Corrales, Javier, and Michael Penfold. *Dragon in the Tropics: Hugo Chávez and the Political Economy of Revolution in Venezuela*. Washington, D.C.: Brookings Institution Press, 2011.
- Cowen, Tyler. "The Importance of Defining the Feasible Set." *Economics and Philosophy* 23, no. 1 (2007): 1–14.
- Crisp, Brian F. *Democratic Institutional Design: The Powers and Incentives of Venezuelan Politicians and Interest Groups*. Stanford: Stanford University Press, 2000.
- Dewey, John. *The Public and its Problems*. Athens, OH: Ohio University Press, 1954.
- . *Liberalism and Social Action*. Amherst, NY: Prometheus Books, 2000.
- Dollar, David, and Aart Kraay. "Growth is Good for the Poor." *Journal of Economic Growth* 7 (2002): 195–225.
- Downs, George W., David M. Rocke, and Peter N. Barsoom. "Is the Good News about Compliance Good News about Cooperation?" *International Organization* 50, no. 3 (1996): 379–406.

- Dunning, Thad. *Crude Democracy: Natural Resource Wealth and Political Regimes*. New York: Cambridge University Press, 2008.
- Easterly, William. *The White Man's Burden: Why the West's Efforts to Aid Have Done So Much Ill and So Little Good*. New York: Penguin, 2006.
- Elster, Jon. "Marxism, Functionalism, and Game Theory: The Case for Methodological Individualism." *Theory and Society* 11, no. 4 (1982): 453–482.
- . *Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press, 1989.
- . *Explaining Social Behavior: More Nuts and Bolts for the Social Sciences*. New York: Cambridge University Press, 2007.
- Farrelly, Colin. "Justice in Ideal Theory: A Refutation." *Political Studies* 55, no. 4 (2007): 844–864.
- Fearon, James D. "Bargaining, Enforcement, and International Cooperation." *International Organization* 52, no. 2 (1998): 269–305.
- . "Primary Commodity Exports and Civil War." *Journal of Conflict Resolution* 49, no. 4 (2005): 483–507.
- Fish, M. Steven. *Democracy Derailed in Russia: The Failure of Open Politics*. New York: Cambridge University Press, 2005.
- Freedom House. *Freedom in the World—Venezuela Country Report*. 2010. URL: <http://www.freedomhouse.org/template.cfm?page=363&year=2010>.
- Frey, Bruno. "A Constitution for Knaves Crowds out Civic Virtues." *The Economic Journal* 107, no. 443 (1997): 1043–1053.
- . "Institutions and Morale: The Crowding-Out Effect." In *Economics, Values, and Organization*, edited by Avner Ben-Ner and Louis G. Putterman, 437–460. Cambridge: Cambridge University Press, 1999.
- Friedman, Milton. *Capitalism and Freedom*. Chicago: University of Chicago Press, 1962.
- Geertz, Clifford. "Deep Play: Notes on the Balinese Cockfight." In *The Interpretation of Cultures*. New York: Basic Books, 1973.
- Gelb, Alan, and Associates. *Oil Windfalls: Blessing or Curse?* New York: Oxford University Press, 1988.
- Geuss, Raymond. *Philosophy and Real Politics*. Princeton and Oxford: Princeton University Press, 2008.
- Goodin, Robert E. "Political Ideals and Political Practice." *British Journal of Political Science* 25, no. 1 (1995): 37–56.
- Gordon, Joy. "A Peaceful, Silent, Deadly Remedy: The Ethics of Economic Sanctions." *Ethics & International Affairs* 13, no. 1 (1999): 123–142.

- Haber, Stephen, and Victor Menaldo. "Do Natural Resources Fuel Authoritarianism? A Reappraisal of the Resource Curse." *American Political Science Review* 105, no. 1 (2011): 1–26.
- Hafner-Burton, Emilie, and Kiyoteru Tsutsui. "Justice Lost! The Failure of International Human Rights Law To Matter Where Needed Most." *Journal of Peace Research* 44 (2007): 407–425.
- Hassoun, Nicole. "Making the Case for Foreign Aid." *Public Affairs Quarterly* 24, no. 1 (2010): 1–20.
- Hathaway, Oona A. "Do Human Rights Treaties Make a Difference?" *Yale Law Journal* 111, no. 8 (2002): 1935–2042.
- . "Why Do Countries Commit to Human Rights Treaties?" *Journal of Conflict Resolution* 51 (2007): 588–621.
- Hedström, Peter, and Richard Swedberg, eds. *Social Mechanisms: An Analytical Approach to Social Theory*. New York: Cambridge University Press, 1998.
- Held, David. *Democracy and the Global Order: From the Modern State to Cosmopolitan Governance*. Stanford: Stanford University Press, 1995.
- Herb, Michael. "No Representation without Taxation? Rents, Development, and Democracy." *Comparative Politics* 37, no. 3 (2005): 297–316.
- Holtzapple, Mark T., and W. Dan Reece. *Foundations of Engineering*. 2nd ed. New York: McGraw-Hill, 2003.
- Human Security Report Project. *Human Security Report*. Simon Fraser University, 2009/2010.
- Humphreys, Macartan. "Natural Resources, Conflict, and Conflict Resolution: Uncovering the Mechanisms." *Journal of Conflict Resolution* 49, no. 4 (2005): 508–537.
- Humphreys, Macartan, Jeffrey D. Sachs, and Joseph E. Stiglitz, eds. *Escaping the Resource Curse*. New York: Columbia University Press, 2007.
- International Committee on Intervention and State Sovereignty. *The Responsibility to Protect*. Ottawa: International Development Research Centre, 2001.
- Jaggar, Alison M., ed. *Thomas Pogge and His Critics*. Malden, MA: Polity Press, 2010.
- Jamieson, Dale. "Duties to the Distant: Aid, Assistance, and Intervention in the Developing World." *Journal of Ethics* 9 (2005): 151–170.
- Jensen, Mark. "The Limits of Practical Possibility." *Journal of Political Philosophy* 17, no. 2 (2009): 168–184.
- Jensen, Nathan, and Leonard Wantchekon. "Resource Wealth and Political Regimes in Africa." *Comparative Political Studies* 37, no. 7 (2004): 816–841.
- Kant, Immanuel. *Toward Perpetual Peace*. In *Practical Philosophy*, edited and translated by Mary J. Gregor, 311–352. Cambridge: Cambridge University Press, 1996.

- Karl, Terry Lynn. *The Paradox of Plenty: Oil Booms and Petro-States*. Berkeley and Los Angeles: University of California Press, 1997.
- Keohane, Robert O. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton and Oxford: Princeton University Press, 2005.
- Koremenos, Barbara, Charles Lipson, and Duncan Snidal. "The Rational Design of International Institutions." *International Organization* 55, no. 4 (2001): 761–799.
- Kratzer, Angelika. "The Notional Category of Modality." In *Words, Worlds, and Contexts: New Approaches in Word Semantics*, edited by Hans Jürgen Eikmeyer and Hannes Rieser, 38–74. Berlin: de Gruyter, 1981.
- . "Modality." In *Semantics: An International Handbook of Contemporary Research*, edited by Arnim von Stechow and Dieter Wunderlich, 639–650. Berlin: de Gruyter, 1991.
- Kuran, Timur. *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Cambridge, MA: Harvard University Press, 1995.
- Lam, Ricky, and Leonard Wantchekon. "Political Dutch Disease." Yale University, unpublished manuscript, 2002.
- Le Billon, Philippe. "The Political Ecology of War: Natural Resources and Armed Conflicts." *Political Geography* 20 (2001): 561–584.
- Levi, Margaret. *Of Rule and Revenue*. Berkeley and Los Angeles: University of California Press, 1988.
- Lipsey, R.G., and Kelvin Lancaster. "The General Theory of Second Best." *The Review of Economic Studies* 24, no. 1 (1957): 11–32.
- Mahdavy, Hussein. "The Patterns and Problems of Economic Development in Rentier States: The Case of Iran." In *Studies in Economic History of the Middle East*, edited by M. A. Cook. London: Oxford University Press, 1970.
- Mäki, Uskali. "Models are Experiments, Experiments are Models." *Journal of Economic Methodology* 12, no. 2 (2005): 303–315.
- . "MISSing the World. Models as Isolations and Credible Surrogate Systems." *Erkenntnis* 70, no. 1 (2009): 29–43.
- Marshall, Monty, and Keith Jagers. "Polity IV Project: Political Regime Characteristics and Transitions, 1800–2006." University of Maryland, unpublished manuscript, 2008.
- McGillivray, Fiona, and Alastair Smith. *Punishing the Prince: A Theory of Interstate Relations, Political Institutions, and Leader Change*. Princeton and Oxford: Princeton University Press, 2008.
- Mehlum, Halvor, Karl Moene, and Ragnar Torvik. "Institutions and the Resource Curse." *The Economic Journal* 116, no. 1 (2006): 1–20.

- Mill, John Stuart. *On Liberty*. In *The Basic Writings of John Stuart Mill*, edited by Dale E. Miller. New York: Random House, 2002.
- . *The Basic Writings of John Stuart Mill*. Edited by Dale E. Miller. New York: Random House, 2002.
- . “The Subjection of Women.” In *The Subjection of Women*, edited by Dale E. Miller. New York: Random House, 2002.
- Miller, David. *On Nationality*. New York: Oxford University Press, 1995.
- . *National Responsibility and Global Justice*. New York: Oxford University Press, 2007.
- Miller, Richard W. “Respectable Oppressors, Hypocritical Liberators: Morality, Intervention, and Reality.” In Chatterjee and Scheid, *Ethics and Humanitarian Intervention*.
- Mills, Charles W. “‘Ideal Theory’ as Ideology.” *Hypatia* 20, no. 3 (2005): 165–184.
- Milner, Helen V. *Interests, Institutions, and Information*. Princeton: Princeton University Press, 1997.
- Moellendorf, Darrel. *Cosmopolitan Justice*. Boulder: Westview Press, 2002.
- . “The World Trade Organization and Egalitarian Justice.” *Metaphilosophy* 36, nos. 1/2 (2005): 145–162.
- Moore, Mick. “Political Underdevelopment: What Causes ‘Bad Governance’?” *Public Management Review* 3, no. 3 (2001): 385–418.
- Morrison, Kevin M. “Natural Resources, Aid, and Democratization: A Best-Case Scenario.” *Public Choice* 131 (2007): 365–386.
- . “Oil, Nontax Revenue, and the Redistributive Foundations of Regime Stability.” *International Organization* 63 (2009): 107–138.
- Morrow, James. “Modeling the Forms of International Cooperation: Distribution Versus Information.” *International Organization* 48, no. 3 (1994): 387–423.
- Nagel, Thomas. “The Problem of Global Justice.” *Philosophy & Public Affairs* 33, no. 2 (2005): 113–147.
- North, Douglass C. *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press, 1990.
- North, Douglass C., and Robert Paul Thomas. *The Rise of the Western World: A New Economic History*. Cambridge: Cambridge University Press, 1973.
- North, Douglass C., and Barry Weingast. “Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England.” *Journal of Economic History* 49, no. 4 (1989): 803–832.
- Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.

- Nussbaum, Martha C. *Women and Human Development*. Cambridge: Cambridge University Press, 2000.
- OECD. *Development Aid at a Glance*. Organisation for Economic Co-operation and Development, 2008.
- Olson, Mancur. "Dictatorship, Democracy, and Development." *American Political Science Review* 87, no. 3 (1993): 567–576.
- O'Neill, Onora. *Bounds of Justice*. Cambridge: Cambridge University Press, 2000.
- Pahl, Gerhard, Wolfgang Beitz, Jörg Feldhusen, and Karl-Heinrich Grote. *Engineering Design: A Systematic Approach*. 3rd ed. Edited by Ken Wallace and Lucienne Blessing. London: Springer, 2007.
- Petroski, Henry. *To Engineer is Human: The Role of Failure in Successful Design*. New York: Vintage Books, 1992.
- . *Success Through Failure: The Paradox of Design*. Princeton and Oxford: Princeton University Press, 2006.
- Poast, Paul. "(Mis)Using Dyadic Data to Analyze Multilateral Events." *Political Analysis* 18, no. 4 (2010): 403–425.
- Pogge, Thomas W. *Realizing Rawls*. Ithaca, NY: Cornell University Press, 1989.
- . "An Egalitarian Law of Peoples." *Philosophy & Public Affairs* 23, no. 3 (1994): 195–224.
- . "Pre-empting Humanitarian Interventions." In *Freedom, Power, and Morality: Essays for Felix Oppenheim*, edited by Ian Carter and Mario Ricciardi. London: Palgrave, 2001.
- . *World Poverty and Human Rights*. Malden, MA: Polity Press, 2002.
- . "'Assisting' the Global Poor." In *Ethics of Assistance: Morality and the Distant Needy*, edited by Deen K. Chatterjee. Cambridge: Cambridge University Press, 2004.
- . "Real World Justice." *The Journal of Ethics* 9 (2005): 29–53.
- . "Recognized and Violated by International Law: The Human Rights of the Global Poor." *Leiden Journal of International Law* 18, no. 4 (2005): 717–745.
- . "Severe Poverty as a Violation of Negative Duties." *Ethics & International Affairs* 19, no. 1 (2005): 55–83.
- . "Moralizing Humanitarian Intervention: Why Jurying Fails and How Law Can Work." In *Humanitarian Intervention, Humanitarian Intervention*, edited by Terry Nardin and Melissa S. Williams. NOMOS XLVII. New York and London: New York University Press, 2006.
- Räikkä, Juha. "The Feasibility Condition in Political Theory." *Journal of Political Philosophy* 6, no. 1 (1998): 27–40.

- Ravallion, Martin. "Growth, Inequality, and Poverty: Looking Beyond Averages." *World Development* 29, no. 11 (2001): 1803–1815.
- . "How Not to Count the Poor? A Reply to Reddy and Pogge." In Anand, Segal, and Stiglitz, *Global Poverty*.
- Rawls, John. *Political Liberalism*. New York: Columbia University Press, 1996.
- . *A Theory of Justice*. 2nd ed. Cambridge, MA: Harvard University Press, 1999.
- . *The Law of Peoples*. Cambridge, MA: Harvard University Press, 1999.
- . *Justice as Fairness: A Restatement*. Edited by Erin Kelly. Cambridge, MA: Harvard University Press, 2001.
- Reddy, Sanjay G. "The Role of Apparent Constraints in Normative Reasoning: A Methodological Statement and Application to Global Justice." *The Journal of Ethics* 9 (2005): 119–125.
- Reddy, Sanjay G., and Thomas W. Pogge. "How *Not* to Count the Poor." In Anand, Segal, and Stiglitz, *Global Poverty*.
- Risse, Mathias. "What We Owe to the Global Poor." *The Journal of Ethics* 9 (2005): 81–117.
- Robeyns, Ingrid. "Ideal Theory in Theory and Practice." *Social Theory and Practice* 34, no. 3 (2008): 341–362.
- Robinson, James A., Ragnar Torvik, and Thierry Verdier. "Political Foundations of the Resource Curse." *Journal of Development Economics* 79, no. 2 (2006): 447–468.
- Rodrik, Dani, Arvind Subramanian, and Francesco Trebbi. "Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development." *Journal of Economic Growth* 9 (2004): 131–165.
- Ross, Michael L. "The Political Economy of the Resource Curse." *World Politics* 51 (1999): 297–322.
- . "Does Oil Hinder Democracy?" *World Politics* 53 (2001): 325–361.
- . "Does Taxation Lead to Representation?" *British Journal of Political Science* 34, no. 2 (2004): 229–249.
- Ross, Michael L., Phil Keefer, Steve Knack, and Miriam Lowi. "Does Resource Wealth Cause Authoritarian Rule?" University of California, Los Angeles, unpublished manuscript, 2000.
- Rosser, Andrew. "The Political Economy of the Resource Curse: A Literature Survey." *IDS Working Paper*, no. 268 (2006).
- Roth, Leland M. *Understanding Architecture: Its Elements, History, and Meaning*. Boulder: Westview Press, 1993.

- Rubenstein, Jennifer. "Distribution and Emergency." *The Journal of Political Philosophy* 15, no. 3 (2007): 296–320.
- Sachs, Jeffrey D. *The End of Poverty: Economic Possibilities for Our Time*. New York: Penguin, 2005.
- Sachs, Jeffrey D., and Andrew M. Warner. "Natural Resource Abundance and Economic Growth." *NBER Working Paper Series*, no. 5398 (1995).
- . "The Curse of Natural Resources." *European Economic Review* 45 (2001): 827–838.
- Sangiovanni, Andrea. "Global Justice, Reciprocity, and the State." *Philosophy & Public Affairs* 35, no. 1 (2007): 3–39.
- Sen, Amartya. *Development as Freedom*. New York: Anchor Books, 1999.
- . "What Do We Want From A Theory of Justice?" *Journal of Philosophy* 103, no. 5 (2006): 215–238.
- . *The Idea of Justice*. Cambridge, MA: Harvard University Press, 2009.
- Shifter, Michael. "In Search of Hugo Chávez." *Foreign Affairs* 85, no. 3 (2006): 45–59.
- Shue, Henry. *Basic Rights: Subsistence, Affluence, and U.S. Foreign Policy*. Princeton: Princeton University Press, 1980.
- Simmons, A. John. "Ideal and Nonideal Theory." *Philosophy & Public Affairs* 38, no. 1 (2010): 5–36.
- Singer, Peter. "Famine, Affluence, and Morality." *Philosophy & Public Affairs* 1, no. 3 (1972): 229–243.
- . "Achieving the Best Outcome: Final Rejoinder." *Ethics & International Affairs* 16, no. 2 (2002): 127–128.
- . *One World: The Ethics of Globalization*. New Haven and London: Yale University Press, 2002.
- . "Poverty, Facts, and Political Philosophies: A Reply to 'More Than Charity'." *Ethics and International Affairs* 16, no. 2 (2002): 121–124.
- . *The Life You Can Save: Acting Now To End World Poverty*. New York: Random House, 2009.
- Snyder, Richard, and Ravi Bhavnani. "Diamonds, Blood, and Taxes: A Revenue-Centered Framework for Explaining Political Order." *Journal of Conflict Resolution* 49, no. 4 (2005): 563–597.
- Stemplowska, Zofia. "What's Ideal About Ideal Theory?" *Social Theory and Practice* 34, no. 3 (2008): 319–340.

- Stone, Randall W. "The Political Economy of IMF Lending in Africa." *American Political Science Review* 98, no. 4 (2004): 577–591.
- Sugden, Robert. "Credible Worlds: The Status of Theoretical Models in Economics." *Journal of Economic Methodology* 7, no. 1 (2000): 1–31.
- . "Credible Worlds, Capacities and Mechanisms." *Erkenntnis* 70, no. 1 (2009): 3–27.
- Swift, Adam. "The Value of Philosophy in Nonideal Circumstances." *Social Theory and Practice* 34, no. 3 (2008): 363–387.
- Tesón, Fernando R. *Humanitarian Intervention*. 3rd ed. Ardsley, NY: Transnational, 2005.
- Tilly, Charles. "War Making and State Making as Organized Crime." In *Bringing the State Back In*, edited by Peter Evans, Dietrich Rueschemeyer, and Theda Skocpol. Cambridge: Cambridge University Press, 1985.
- . *Coercion, Capital, and European States, AD 990–1992*. Revised ed. Malden, MA: Blackwell, 1992.
- Tornell, Aaron, and Philip R. Lane. "The Voracity Effect." *American Economic Review* 89, no. 1 (1999): 22–46.
- Torvik, Ragnar. "Natural Resources, Rent Seeking, and Welfare." *Journal of Development Economics* 67 (2002): 455–470.
- Tsebelis, George. *Veto Players: How Political Institutions Work*. Princeton: Princeton University Press, 2002.
- UNDP. *Human Development Indicies*. United Nations Development Programme, 2008.
- UNHCR. *2009 Statistical Yearbook*. United Nations High Commissioner for Refugees, 2010.
- UNICEF. *State of the World's Children – Statistical Tables*. United Nations Children's Fund, 2011.
- Unwin, Simon. *Analysing Architecture*. 2nd ed. New York: Routledge, 2003.
- Valentini, Laura. "On the Apparent Paradox of Ideal Theory." *Journal of Political Philosophy* 17, no. 3 (2009): 332–355.
- Veblen, Thorstein. *The Theory of the Leisure Class*. New York: Penguin Books, 1994.
- Walzer, Michael. "The Moral Standing of States: A Response to Four Critics." *Philosophy & Public Affairs* 9, no. 3 (1980): 209–229.
- Wantchekon, Leonard. "Why do Resource Abundant Countries Have Authoritarian Governments?" *Journal of African Finance and Economic Development* 5, no. 2 (2002): 57–77.

- Watkins, J.W.N. "Methodological Individualism: A Reply." *Philosophy of Science* 22, no. 1 (1955): 58–62.
- Wenar, Leif. "What We Owe to Distant Others." *Politics, Philosophy & Economics* 2, no. 3 (2003): 283–304.
- . "Property Rights and the Resource Curse." *Philosophy & Public Affairs* 36, no. 1 (2008): 2–32.
- . "Feasible Reform of International Trade in Resources." In Jaggar, *Pogge and His Critics*.
- Weyland, Kurt. "Will Chavez Lose His Luster?" *Foreign Affairs* 80, no. 6 (2001): 73–87.
- World Bank. *World Development Report 2000/2001: Attacking Poverty*. New York: World Bank / Oxford University Press, 2001.
- . *World Development Indicators*. International Bank for Reconstruction and Development, 2010.
- World Health Organization. *World Health Statistics*. World Health Organization, 2010.
- Young, Iris Marion. "Violence Against Power: Critical Thoughts on Military Intervention." In Chatterjee and Scheid, *Ethics and Humanitarian Intervention*.