Navigating by Judgment: Organizational Structure, Autonomy, and Country Context in Delivering Foreign Aid

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:17467366">http://nrs.harvard.edu/urn-3:HUL.InstRepos:17467366</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Navigating by Judgment: Organizational Structure, Autonomy, and Country Context
In Delivering Foreign Aid

A dissertation presented
by
Daniel J Honig
to
The Committee on Higher Degrees in Public Policy

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Public Policy

Harvard University
Cambridge, Massachusetts
May 2015
Navigating by Judgment: Organizational Structure, Autonomy, and Country Context
In Delivering Foreign Aid

Abstract

This dissertation examines when initiatives by International Development Organizations (IDO) are more, and less, successful. The core argument is that allowing field-level agents to drive initiatives – what I call organizational Navigation by Judgment – will often be the most effective way to deliver aid. This inverts what a classical application of the principal agent model – the workhorse of studies of public management and bureaucracy – would predict, with better performance resulting from less control. In the delivery of foreign aid the costs of monitoring to the principal are often overshadowed by the deleterious effects of the monitoring itself.

The core of the argument is that development implementation requires soft information, tacit knowledge, and flexibility that are crowded out by tight controls or an organizational navigation strategy focused on short term measurement and targets. As a result there are increasing returns to Navigation by Judgment in environments that are uncertain or difficult to understand from the outside and tasks where outputs are difficult to observe and/or poorly correlated with long term intervention goals. Insecure political authorizing environments which constrain the autonomy of IDOs prevent these organizations from Navigating by Judgment in situations where this is the best strategy.

Empirically, this dissertation examines a cross-IDo dataset of projects (including over 14,000 projects over 50 years over 9 organizations), which I have assembled. It also examines eight cases of development interventions in Liberia and South Africa. These cases are matched pairs comparing the performance and navigation strategies of the US Agency for International Development (a low autonomy IDO) and the UK’s Department for International Development (a higher autonomy IDO) in capacity building and health sector interventions.
# Table of Contents

Acronyms and Abbreviations ................................................................................................. v

Acknowledgements ............................................................................................................... vii

Chapter 1: Introduction ........................................................................................................ 1
  1.1: The Argument ............................................................................................................. 6
  1.2: Structure of the Dissertation ...................................................................................... 9

CHAPTER 2: Theory & Empirical Strategy ........................................................................ 12
  2.2: Establishing Boundary Conditions for Optimal Navigation Strategy ..................... 42
  2.3: International Development as a Context of Study .................................................. 48
  2.4: Empirical Predictions in the Context of International Development ..................... 57
  2.5: Empirical Strategy ................................................................................................... 60
  2.6: Generalizability ...................................................................................................... 66
  2.7: Conclusion .............................................................................................................. 68

CHAPTER 3: Large-N Analysis of Project Outcomes ......................................................... 70
  3.1: Data and Specifications ............................................................................................ 71
  3.2: Quantitative Results ............................................................................................... 85
  3.3: Conclusion .............................................................................................................. 116

CHAPTER 4: Comparing USAID and DFID in South Africa ............................................ 118
  4.1: Introduction ............................................................................................................ 118
  4.2: Capacity Building .................................................................................................. 126
  4.3: Health & PMTCT ................................................................................................... 154
  4.4: Investigating Channels: The Flow of Constraint & Impacts of Same .................... 184
  4.5: South Africa Conclusion – Striking Parallels, Important Divergences ................... 190

CHAPTER 5: Comparing USAID and DFID in Liberia .................................................... 194
  5.1: Introduction ............................................................................................................ 194
  5.2: Capacity Building .................................................................................................. 201
  5.3: Health ..................................................................................................................... 225
  5.4: The Impacts of Constraint & Navigation Strategy .................................................... 258
  5.5: Liberia Conclusion – Similar Objectives, Different Methods & Results .................. 262
  5.6: Comparing Across Countries: Bringing Liberia and South Africa Cases Together .... 267

Chapter 6: Conclusion .......................................................................................................... 279
  6.1: The Argument Restated ........................................................................................... 280
  6.2: The Evidence .......................................................................................................... 281
  6.3: Alternative Explanations of IDO Success ................................................................ 286
  6.4: Navigation by Judgment and the Literature ............................................................. 292
  6.5: The Way Forward ................................................................................................... 295

Appendix I: Interview Sources & Methods ...................................................................... 308

References .............................................................................................................................. 319
# ACRONYMS AND ABBREVIATIONS

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-PLAN</td>
<td>Accelerated Plan for PMTCT (DFID-funded)</td>
</tr>
<tr>
<td>ANC</td>
<td>African National Congress (SA political party)</td>
</tr>
<tr>
<td>AOR</td>
<td>Agreement Officer Representative (USAID)</td>
</tr>
<tr>
<td>ASI</td>
<td>Adam Smith International, implementer of CISCAB for DFID</td>
</tr>
<tr>
<td>AWARE</td>
<td>Action for West Africa Region (USAID Liberia HIV program)</td>
</tr>
<tr>
<td>BASICS</td>
<td>Basic Support for Institutionalizing Child Survival – (USAID health program)</td>
</tr>
<tr>
<td>CDC</td>
<td>Center for Disease Control (US)</td>
</tr>
<tr>
<td>CISCAB</td>
<td>Civil Service Capacity Building Program (DFID Liberia)</td>
</tr>
<tr>
<td>CISREP</td>
<td>Civil Service Reform Program (DFID Liberia – name after revision)</td>
</tr>
<tr>
<td>CMTP</td>
<td>Consolidated Municipal Transformation Programme</td>
</tr>
<tr>
<td>COGTA</td>
<td>Cooperative Governance &amp; Traditional Affairs (DPLG’s successor)</td>
</tr>
<tr>
<td>CoP</td>
<td>Chief of Party (Head of project on behalf of contracted firm)</td>
</tr>
<tr>
<td>COP</td>
<td>Country Operation Plan (PEPFAR annual operational plan)</td>
</tr>
<tr>
<td>COR</td>
<td>Contracting Officer Representative (USAID)</td>
</tr>
<tr>
<td>CSA</td>
<td>Civil Service Agency (Liberia)</td>
</tr>
<tr>
<td>CSRD</td>
<td>Civil Service Reform Directorate</td>
</tr>
<tr>
<td>CTO</td>
<td>Cognizant Technical Officer</td>
</tr>
<tr>
<td>DAI</td>
<td>Development Alternatives International – contractor implementing BRDG</td>
</tr>
<tr>
<td>DFID</td>
<td>Department for International Development (United Kingdom)</td>
</tr>
<tr>
<td>DG</td>
<td>Director General</td>
</tr>
<tr>
<td>DPLG</td>
<td>Department for Provincial and Local Government (COGTA predecessor)</td>
</tr>
<tr>
<td>EGPAF</td>
<td>Elizabeth Glaser Pediatric AIDS Foundation (US PEPFAR implementer)</td>
</tr>
<tr>
<td>FARA</td>
<td>Fixed Asset Reimbursement Agreement (USAID health sector financing mechanism)</td>
</tr>
<tr>
<td>FHI</td>
<td>Family Health International (USAID lead contractor on AWARE)</td>
</tr>
<tr>
<td>GBP</td>
<td>Great Britain Pounds</td>
</tr>
<tr>
<td>GC/GRC</td>
<td>Governance Commission/Governance Reform Commission</td>
</tr>
<tr>
<td>HLSP</td>
<td>A UK-based consulting firm (MSP/RRHF implementer); not an acronym</td>
</tr>
<tr>
<td>ICAP</td>
<td>International Center for Aids Care &amp; Treatment Program, Columbia University</td>
</tr>
<tr>
<td>IDP</td>
<td>Integrated Development Plan (SA municipal planning instrument)</td>
</tr>
<tr>
<td>IDO</td>
<td>International Development Organization (e.g. DFID, USAID, WB)</td>
</tr>
<tr>
<td>INGO</td>
<td>International Non-Governmental Organization</td>
</tr>
<tr>
<td>JSI</td>
<td>John Snow International (Firm contracted by USAID to implement RBHS)</td>
</tr>
<tr>
<td>ISF</td>
<td>Integrated Services Facilitator (DFID CMTP Advisor)</td>
</tr>
<tr>
<td>ISLGS</td>
<td>Increasing Sustainable Local Government Services</td>
</tr>
<tr>
<td>LGSP</td>
<td>Local Governance Support Programme (LGSP)</td>
</tr>
<tr>
<td>LIPA</td>
<td>Liberia Institute of Public Administration</td>
</tr>
<tr>
<td>LISGIS</td>
<td>Liberia Institute of Statistics and Geoinformation Services</td>
</tr>
<tr>
<td>LRDC</td>
<td>Liberia Reconstruction and Development Committee (Lead Donor-Government coordination body, Liberia)</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>MatCH</td>
<td>Maternal, Adolescent, and Child Health; Department of Obstetrics and Gynaecology; University of the Witwatersrand</td>
</tr>
<tr>
<td>M&amp;E</td>
<td>Monitoring and Evaluation</td>
</tr>
<tr>
<td>MoHSW</td>
<td>Ministry of Health and Social Welfare, Liberia</td>
</tr>
<tr>
<td>MRC</td>
<td>Medical Research Council (South Africa)</td>
</tr>
<tr>
<td>MSP</td>
<td>HIV/AIDS Multi-Sectoral Program (DFID)</td>
</tr>
<tr>
<td>NAC</td>
<td>National AIDS Commission (Liberia)</td>
</tr>
<tr>
<td>NACP</td>
<td>National AIDS Control Program (Liberia)</td>
</tr>
<tr>
<td>NDOH</td>
<td>National Department of Health (South Africa)</td>
</tr>
<tr>
<td>NGO</td>
<td>Non-Governmental Organization</td>
</tr>
<tr>
<td>NiC</td>
<td>National Investment Commission, Liberia</td>
</tr>
<tr>
<td>OBE</td>
<td>Order of the British Empire – An honor bestowed by the British monarch</td>
</tr>
<tr>
<td>OECD</td>
<td>Organization for Economic Cooperation and Development</td>
</tr>
<tr>
<td>OFDA</td>
<td>Office of Foreign Disaster Assistance</td>
</tr>
<tr>
<td>OFM</td>
<td>Office of Financial Management of MoHSW, Liberia</td>
</tr>
<tr>
<td>OGAC</td>
<td>Office of the Global AIDS Coordinator (US PEPFAR administration hub)</td>
</tr>
<tr>
<td>OTI</td>
<td>Office of Transitional Initiatives (USAID)</td>
</tr>
<tr>
<td>PAO</td>
<td>Principal Administrative Officer</td>
</tr>
<tr>
<td>PARI</td>
<td>Public Affairs Research Institute (South Africa)</td>
</tr>
<tr>
<td>PBC</td>
<td>Performance-based Contracting</td>
</tr>
<tr>
<td>PBF</td>
<td>Performance-based Financing</td>
</tr>
<tr>
<td>PEPFAR</td>
<td>President’s Emergency Plan For AIDS Relief (US Gov’t program)</td>
</tr>
<tr>
<td>PF</td>
<td>Pooled Fund (Liberia; initiated by DFID)</td>
</tr>
<tr>
<td>PMTCT</td>
<td>Prevention of Mother to Child Transmission of HIV/AIDS</td>
</tr>
<tr>
<td>PSPPD</td>
<td>Programme to Support Pro-Poor Policy Development (South Africa)</td>
</tr>
<tr>
<td>PwC</td>
<td>Price Waterhouse Coopers (DFID contractor for OFM, PF)</td>
</tr>
<tr>
<td>RBHS</td>
<td>Rebuilding Basic Health Services (USAID Liberia)</td>
</tr>
<tr>
<td>RRHF</td>
<td>Rapid Results Health Fund (DFID vehicle that financed A-Plan)</td>
</tr>
<tr>
<td>RTC</td>
<td>Right to Care (SA-based major PEPFAR implementer)</td>
</tr>
<tr>
<td>SA</td>
<td>South Africa</td>
</tr>
<tr>
<td>SALGA</td>
<td>South African Local Government Association</td>
</tr>
<tr>
<td>SANAC</td>
<td>South African National AIDS Commission</td>
</tr>
<tr>
<td>SARRAH</td>
<td>Strengthening South Africa’s Response to HIV and Health (DFID project)</td>
</tr>
<tr>
<td>STI</td>
<td>Sexually Transmitted Infections</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>US</td>
<td>United States</td>
</tr>
<tr>
<td>USD</td>
<td>United States Dollars</td>
</tr>
<tr>
<td>UN</td>
<td>United Nations</td>
</tr>
<tr>
<td>UNDP</td>
<td>United Nations Development Programme</td>
</tr>
<tr>
<td>UNHCR</td>
<td>The UN Refugee Agency</td>
</tr>
<tr>
<td>UNICEF</td>
<td>United Nations Children’s Fund</td>
</tr>
<tr>
<td>USAID</td>
<td>United States Agency for International Development</td>
</tr>
<tr>
<td>WB</td>
<td>World Bank</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

There are so many people to whom I owe thanks. Peter Hall, my dissertation chair, has been unflaggingly generous with his time and wisdom. Beth Simmons, Ryan Sheely, and Matt Andrews have been wonderful committee members, providing guidance and direction as well as shaping my work via their reactions. The guidance of my committee has informed my research substantively, teaching me lessons I hope will make me a better scholar and a clearer thinker in all my endeavors. They are also object lessons in how a scholar can make an impact – one I aspire to in my career to come.

Michael Woolcock, Lant Pritchett, Dustin Tingley, Dan Carpenter, Jenny Mansbridge, Eric Werker, and many others have all helped to make this work what it is today. Parts of this dissertation have also benefited from feedback from seminar participants at the Kennedy School, the Government department, and external conferences and seminars. Any errors that remain in this work are, of course, my own.

This work would not have been possible without interview participants in Liberia and South Africa and the kind souls in aid organizations across the globe that went out of their way to share results data. I am thankful to all for their time and efforts. I am also thankful for documents from (anonymous) primary sources in Liberia and South Africa as well as those acquired via UK Freedom of Information and US Freedom of Information Act requests that inform this work. I also thank the National Science Foundation Graduate Research Fellowship Program (GRFP) for its financial support under grant# DGE-1144152.

Harvard Kennedy School has proven a wonderful home for me these past years. Many thanks to Nicole Tateosian, Michael Weinbeck, the 3rd floor Taubman community, and many others for making HKS feel like home. Special thanks to the Inequality and Social
Policy program and to Pam Metz, Kathy Edin, and Bill Wilson, among others, for this valuable part of my HKS experience. My colleagues in the Public Policy Ph.D. program and Harvard more broadly have been unflagging in their kindness and support. When I think of my Public Policy cohort I realize that there is no member of it who has not at some time or another lent a hand. In a few cases – particularly those of Tara Grillos, Gabe Chan, Bill Skimmyhorn, Ariel Stern, and Trisha Shrum – I doubt I would be in a position to submit this dissertation without their assistance when the going got tough. Jeff Friedman was my primary guide in how to approach the Ph.D.; indeed, these acknowledgements, like so much else, benefit from his example.

My wife Özsel has been unflagging in her support, and I could not – would not – have traveled this road in the same way, and with the same confidence, in her absence. She is my foundation, my strength. Her tolerance of my noting concepts employed in this dissertation (soft information, Navigation by Measurement, etc.) whenever I perceive them has been nothing short of remarkable. I thank our son Dylan (now 8 months) for his many energy-giving smiles, and for reminding me that life’s most rewarding moments rarely happen in front of computer screens.

I dedicate this book to someone whose name appears nowhere on its pages, but in whose debt its seed lies. Many years ago, while running an NGO in East Timor, I had a volunteer staff member who is still my object lesson in the uncodifiability, and critical importance, of soft information and tacit knowledge. This work is dedicated Vicente Brito and all those who make the world better via their Navigation by Judgment - be it in the mountains of Liquica, the halls of aid organizations, or anywhere else in their professional and personal lives.
CHAPTER 1: INTRODUCTION

Although the uniqueness of the foreign aid agency’s task has been recognized and understood, the organizational environment that such a task requires has never been specified.... I ascribe problem results to an organizational, rather than a historical, determinism. – Judith Tendler (Tendler 1975, p. 9, 110)

Friedrich Hayek and James Scott are not commonly seen as kindred spirits, as intellectual fellow travelers.¹ Both Scott and Hayek agree, however, that control from above is unlikely to produce the best outcomes, particularly when the individuals at the top do not have all the relevant information.² Hayek famously decried the tyranny of central planning; Scott critiques “schemes to improve the human condition” imposed from above. (Scott 1998) They both believe that institutions need to have “respect for the existing state of affairs.” (Scott 2013)³

These two eminent scholars diagnose similar problems but prescribe quite different remedies. Hayek champions the market; prices will supply the data needed for decentralized decision making. Hayek notes “the most significant fact about this [price]

¹ Hayek was a Chicago school economist; he is a hero of the Libertarian movement. Scott is an anthropologist and political scientist currently at Yale; a recent paper describes him as a “Marxist (of the American academic type)”. (Pritchett and Sandefur 2013, footnote 1)

² Many thanks to Nils Hagerdal for this insight. In an effort to coin a name for this consensus I looked up the geographic midpoint between University of Chicago (for Hayek) and University of Wisconsin (a broadly liberal institution where Scott spent many years). The midpoint between the two is, somewhat oddly, Harvard, Illinois. Both Hayek and Scott might appreciate Harvard, IL’s motto: “Success Comes Naturally Here.”

³ This is a quote from an interview in which Scott describes what he perceives as the overlap between his ideas and those of Hayek.
system is the economy of knowledge with which it operates, or how little the individual participants need to know in order to be able to take the right action.” (Hayek 1972, p. 86-7) Scott’s suggestion is not to simplify things at all, but rather to incorporate the complexity of the world as it is via institutions that incorporate mêtis. Mêtis is “a wide array of practical skills and acquired intelligence in responding to a constantly changing natural and human environment” and “denotes the knowledge that can come only from practical experience”. (Scott 1998; p. 313, p. 6)

Scott and Hayek’s central insight that control from the top has substantial costs which are often unseen by those who champion them undergirds this work, which will examine the navigation strategies of International Development Organizations (IDOs) – organizations such as the World Bank, US Agency for International Development (USAID), and UK’s Department for International Development (DFID). The two organizational navigation strategies that will serve as archetypes in this work have parallels in Scott and Hayek’s contrasting remedies for centralized decision-making. Scott’s focus on non-codifiable knowledge and skills parallels what I will describe as tacit knowledge and soft information-laden Navigation by Judgment. Hayek’s focus on the power of numbers as signals and organizers has elements of Navigation by Measurement. When these numerical targets are not an end in themselves but rather a means to an end – e.g. the counting of process or intermediate outputs to a broader goal - this Navigation by Measurement is, I will argue, not in service of Hayekian autonomy but rather in conflict with it. I will frame

---

4 As quoted in Machan 1988.
this as a form of Navigation from the Top – of ensuring agents do what principals intend via controls from above.⁵

International development is big business; OECD estimates over 210 billion USD in concessional development financing was disbursed by members of the OECD Development Assistance Committee in 2013. (OECD 2015, p. 101)⁶ These flows go to activities from education, health, and infrastructure to public financial management, anti-corruption efforts, and judicial reform. The basic unit of these activities – and unit of analysis in this dissertation – is the project. An aid project has a discrete time horizon (rarely less than a year or more than five) and a discrete objective. Implementation of projects is conducted by IDOs themselves, by personnel of the recipient country, and by contracted nonprofit and for-profit firms.

The broad spread of IDOs and their projects across much of the developing world and a wide range of sectors provides an ideal research environment for a study of comparative organizational performance: one can compare the functioning of the same IDO across countries and tasks as well as compare the functioning of different organizations within environments and task domains. This work explores the role of IDO organizational

⁵ The parallel in Hayek’s case is the focus on quantitative indicators and lack of reliance on the independent knowledge and wisdom of actors on the ground. I do not mean to imply that Hayek would likely support Navigation by Measurement when it is used as a means of central control.

⁶ This excludes multilateral and regional banks (e.g. the World Bank and African Development Bank), who (in combination with the UN) OECD estimates spent an additional 41.5 billion in 2013. Both these figures exclude non-OECD DAC donors – most notably China.
strategy – what I call navigation strategy – and the relationship of these strategies to project performance. This study focuses in particular on how environmental legibility and uncertainty on the one hand and task observability and contractibility on the other affect the optimal level of organizational autonomy and via autonomy the menu of available organizational navigation strategies.

Less autonomy and more control have the advantage of clear performance standards and more consistent agent action. This is particularly useful if we think of the problem in a classical principal-agent framework, with principals needing to invest in constraining agents. If the agents are Niskanen-style (1968) bureaucrats, interested primarily in maximizing budgets and not in accomplishing the organization’s goals, then this monitoring is likely necessary, if costly. On the other hand, voices from Economics Nobel Laureate Elinor Ostrom (Gibson et al. 2005) to former USAID Administrator Andrew Natsios (Natsios 2010) have suggested that development requires local knowledge, agent discretion, and adaptability. These views suggest agents utilize granted autonomy for good, not ill; that at the margin less monitoring will result in better performance in addition to being less costly for the principal. This work seeks to explore which side of this debate is right, when, and why. Put another way, there is dysfunction to be had both from too much control and (cf. Barnett and Finnemore 2003) from too little. This dissertation explores which effect dominates in any given case and what the systematic components of the variance between these cases may be.
In the world of foreign aid optimal control, autonomy, and planning (as opposed to adaptation) is very much a subject of active debate. Conventional wisdom is that measurement and management regimes that set targets and manage towards them are signs of a high-performing organization. Organizations are increasingly moving towards measurement and management through targets and indicators which are linked to externally verifiable data so as to improve accountability and work towards what many in development term the “results agenda”. (Gulrajani 2011) A number of leading IDOs have come together, in fact, to form a “Global Partnership on Output-Based Aid”, which looks to contract directly on output targets. 7 I challenge this conventional wisdom, arguing that under certain conditions measurement regimes are not just an indicator of conditions for low performance but also directly causal of reduced performance towards organizational ends.

This argument fits with those of a small number of scholars have been arguing for a more iterative, agent judgment-driven approach to foreign aid delivery. (Andrews, Pritchett, and Woolcock 2012; Barder 2009; Booth 2013; Easterly 2014; Ramalingam 2013) These scholars argue for less control and focus on the adoption and achievement of pre-specified targets, instead allowing organizations to adapt to the soft, contextual information of recipient country environments.

7 The GPOBA is at www.gpoba.org.
1.1: The Argument

The core argument of this work will be that Navigation by Judgment — empowering field level agents to make critical decisions about aid interventions — is likely to be a superior organizational strategy in many international development tasks. This is because Navigation by Judgment allows the incorporation of soft information — observable non-verifiable (and thus non-contractible) features of the situation. It also allows agents to make use of tacit knowledge, things agents know how to do but cannot codify for transmission up a hierarchy. The returns to Navigation by Judgment will be increasing the less externally observable and monitorable the task — the more difficult it is to Navigate from the Top, the alternative strategy.

The returns to Navigation by Judgment will also be larger in more uncertain and in what Scott (1998) might call less “legible” environments. The value of soft information and tacit knowledge will be greater as the need to adapt to changing circumstances increases and the environment is less inherently tractable to understanding by actors outside the context. There will also be an organizational effect of failing to Navigate by Judgment, wherein disempowered agents become demotivated or exit the organization.

Political authorizing environments influence and constrain the navigation strategies of IDOs, preventing IDOs from Navigating by Judgment when beneficial to do so. Organizations facing unstable political authorizing environments will have less autonomy than those with more stable and secure political authorizing relationships. These constraints will flow down to field level agents, who will have smaller scope for
independent action than field level agents of more autonomous IDOs. The area for independent action will be constrained both by top-down controls and by measurement. As such, more autonomous organizations will engage in Navigation by Judgment more readily where it is appropriate to do so; they will consequently perform better in more unstable and opaque environments and in tasks that are harder to monitor externally.

**Measurement and Targets**

Navigation by Measurement - the setting of targets and orientation of projects to them – will be a focus of this dissertation. Measurement and ex-ante target setting does not necessarily crowd out autonomy over means, though it always fixes the ends towards which interventions are focused. Where the outcomes desired are stable and measurable, or good measures can be found with these stable outcomes – that is, measures that are not subject to “gaming” by actors and are well correlated with the ultimate goals of the intervention – Navigation by Measurement is likely to be successful. In these cases Navigation by Measurement will facilitate autonomy (and thus the use of soft information and tacit knowledge) regarding means (but not ends), serving to improve intervention performance.

I do not expect target setting and measurement to be limited to such contexts, however. In practice I expect to find measurement even where it serves to distort agent behavior, crowding out autonomy as it orients agents towards the externally observable and quantifiable at the expense of the unobservable. I expect the measures themselves will
in some cases be quite poorly correlated to the actual goals of interventions, employed instead to allow for the IDO to demonstrate performance in the face of an insecure political authorizing environment. In these cases Navigation by Measurement will be a form of Navigation from the Top – a management tool from above which forms a constraint on agent behavior, crowding out tacit knowledge and soft information. Navigation by Measurement, then, will be more successful the more externally observable and contractible the task and the more stable, predictable, and legible the environment — that is, roughly the inverse of when Navigation by Judgment is expected to be superior.

**Similar Arguments & Existing Empirics**

This argument has some echoes in that of other scholars who have argued, contrary to conventional wisdom, that under particular conditions we may get more results with fewer controls. Trailblazers from whom I take inspiration on this counterintuitive path include Jenny Mansbridge’s selection model regarding political representatives (Mansbridge 2009), Jim Scott’s call for states to reduce their use of, and perhaps collection of, data (Scott 1998), and Bohnet, Frey, and Huck’s suggestion that it may be possible to get “More Order with Less Law”. (2001) In the management literature it has echoes in recent work suggesting organizational goals may proliferate inappropriately (Ordóñez et al. 2009) and that there may be a “transparency paradox”, with less supervision and monitoring of agents sometimes auguring for better results. (Bernstein 2012)
Few organizations will fully forsake either control or judgment, of course. Avoiding these extremes is almost certainly in the agency’s best interest; some combination of the two may often be the best solution. (e.g. Canales 2013) There is nonetheless a tradeoff between organizational control and agent autonomy, and thus organizational Navigation by Judgment. There is a continuum between Navigation from the Top and Navigation by Judgment, and organizations are arrayed along that continuum.

While there are some empirics regarding the constituents of autonomy (e.g. Dobbin and Boychuk 1999), there has been little work that looks at the effect of autonomy and organizational structure in a broad, systematic way across different levels of environmental uncertainty and different task environments. While such an investigation may contribute to a variety of literatures, theoretical and empirical, it’s also of great practical importance. Managers need a strategy for determining when output measurement and/or autonomy are more likely to be optimal. Boards of directors and political authorizers need to know if their attempts at control lead to less, not more, of what they want.

1.2: Structure of the Dissertation

Chapter 2 will provide the theoretic foundations for the core argument. It will explore autonomy, Navigation by Judgment, soft information, tacit knowledge, and measurement. The chapter will then describe the empirical strategy to be employed in
testing the proposition developed and the case selection of interventions examined in Chapters 4 and 5.

Chapter 3 will investigate the relationship between autonomy and political authorizing environment stability - Navigation by Judgment’s necessary preconditions - and project success. In doing so, it will employ regression analysis of a novel dataset composed of over 14,000 project evaluations from nine IDOs. This dataset, which I have assembled for this work, is the first of its kind to include performance information from multiple development organizations. As such this is the first quantitative examination of IDO performance able to make inter-IDO comparisons. Chapter 3 will examine whether differences in IDO autonomy have differential effects across country and task in the manner hypothesized in chapter 2.

Chapters 4 and 5 shift focus from quantifiable autonomy to observable but non-quantifiable navigation strategy. These chapters examine the performance of a low autonomy IDO with an insecure political authorizing environment (USAID) and a relatively more secure and autonomous IDO (DFID) in Liberia and South Africa, respectively. These chapters examine the navigation strategy employed by each IDO in the design, implementation, and revision (if any) of each project as well as the success of each intervention. By comparing across roughly comparable pairs of projects that vary in task domain and recipient country environment these chapters provide a test of chapter 2’s theory linking political authorizing environments, autonomy, navigation strategy, and comparative performance.
These chapters also provide nuance regarding mechanisms and the environment IDOs face. This includes providing illustrations of Navigation by Judgment and Navigation from the Top via Navigation by Measurement and tracing constraints flowing from political authorizing environments via organizational and field level autonomy to intervention performance. Chapter 5 concludes with a section comparing performance and navigation strategy across the South African and Liberian cases.

Chapter 6 brings the dissertation to a close, synthesizing the evidence described in chapters 3, 4, and 5 in light of the theory articulated in chapter 2. The chapter discusses where this work leaves us – what it illuminates and what is left in shadow. Implications for future research and policy are also addressed.
CHAPTER 2: THEORY & EMPIRICAL STRATEGY

This chapter articulates the expected advantages of Navigation by Judgment as compared to Navigation from the Top. This chapter also establishes boundary conditions – not merely to instantiate that Navigation by Judgment is good in some cases and bad in others but explore where those limits lay. In practice, that means an examination of the costs and benefits of greater agent autonomy and how the balance of these costs and benefits are affected by elements of the task and environment.8

This chapter proceeds as follows: Section 1 explores the relationship between authorizing environments, autonomy, and navigation strategy. It also explores how measurement can constrain autonomy in service of Navigation from the Top or enable autonomy in service of Navigation by Judgment. Section 2 establishes boundary conditions, suggesting when Navigation by Judgment is likely to be a superior organizational strategy.

Section 3 describes the world of international development assistance (foreign aid), motivating why it is an excellent context for studying these dynamics and surveying relevant literature on international development assistance and the organizations that deliver it. Section 4 brings 2 and 3 together, instantiating the more general propositions in

---

8 The reader may wonder what special magic exists in Navigation by Judgment and Navigation from the Top rather than simply discussing agent autonomy. I would argue that agent autonomy is representative of a deeper difference in organizational strategy – the degree to which management on the one hand and field agents on the other dictate the actions of organizations. To make this merely a discussion of agent autonomy is to diminish the extent to which agent autonomy speaks to who controls the organization from a functional standpoint.
the context of international development. Section 5 outlines the empirical strategy and methods.

The final short sections discuss the implications of the theory beyond the empirical setting developed. Section 6 discusses applicability beyond public management to the private sector and to other domains of our lives. Section 7 concludes.

2.1: Authorizing Environments, Autonomy, Navigation Strategy, & Measurement

The public sector is no stranger to the debate regarding optimal levels of autonomy, discretion, and delegation. Indeed, this has become a core topic of the public management (both US and international) and bureaucratic politics literatures as well as the literature on international organizations. (e.g. Calvert, McCubbins, and Weingast 1989; Carpenter 2001, 2010; Hawkins and Jacoby 2006; Hawkins et al. 2006; Hawkins 2006; Huber and McCarty 2004; Huber and Shipan 2002, 2006; Johnson and Urpelainen 2014; Martin 2006; Nielson and Tierney 2003; Rasul and Rogger 2013; Keiser 1999) The conventional way of modeling bureaucracies, particularly by economists and econometrically minded political scientists, has been via principal-agent models. As Francis Fukuyama (2013) recently put it, “approaches favored by economists sought to conceptualize governance in a principal-agent framework, and have sought to control corruption and bad administration through manipulation of incentives.” (p. 349)
The principal-agent model incorporates both direct controls and schemes such as pay-for-performance, shifting agent incentives towards measured and rewarded targets. Less easily conceptualized in this framework – which looks ‘down’ from the perspective of principals – is the notion that agents may perform better in the absence of controls or manipulation of incentives. Autonomy may emerge from such a model, but it is as often as not a bug (e.g. Barnett & Finnemore 2003) creating problems rather than a potential solution to difficult environments. In this regard the principal-agent literature stands at odds with its private sector cousin; the organizational behavior and management literatures have long suggested that autonomy may be beneficial in many situations. (e.g. Lawrence and Lorsch 1967; Thompson 1967)

While there are clearly ways to foster agent autonomy and for organizations to promote it (cf. Carpenter 2001, 2010; Huber and Shipan 2002), at core operating autonomy is typified by absence not presence; by the absence of constraint. This means that autonomy is a bit like Tolstoy’s (1899) unhappy families in Anna Karenina – there is great variety in the ways agents and organizations can be made un-autonomous. That autonomy is in some sense a residual state – the absence of constraint – does not mean it is undifferentiated.

Verhoest et al. (2004) provide a very useful framework for thinking about organizational autonomy in the public sector, while also noting that ambiguity around definitions of autonomy plays some role in the ambiguous and conflicting findings examining the relationship between public agency autonomy and performance. They break
down autonomy into both having decision-making competencies over day-to-day managerial and broader policy decisions as well as the exemption of constraints on the actual use of those competencies (or de facto ability to make use of de jure autonomy), with structural, financial, legal, and interventional constraints potentially diminishing the exercise of decision-making competencies.

That autonomy is differentiable does not mean these categories are unrelated, of course. If a lack of managerial autonomy changes what an organization or its agents do, that then might affect the organization’s ability to take advantage of its policy autonomy. At the level of constraints, they may be additive. For example, financial and structural constraints might have a cumulative effect in limiting the exercise of autonomy. Constraints might also induce one another. For example, structural constraints (e.g. an organization having to report to a higher body or board) might lead to a greater presence of interventional constraints (e.g. that board intervening in the internal workings of the organization). Measurement can be thought of as another form of constraint in this sense. Other constraints may lead an organization to measure more and do so even where such measurement is destructive rather than productive in an attempt to seek legitimacy and security in an unstable authorizing environment.

**Linking Organizations and Agents**

Both organizations and employees can have a greater or less degree of decision-making competencies (or, as I think of it, de jure autonomy) over both means (day-to-day
managerial) and ends (policy). Both can have constraints or not on the exercise of these decision-making competencies. The content, however, of these constraints differs. Only one of these classes of actors is likely to face direct conversation with political authorizers over financial appropriations; only one of these classes of actors can telework.

While organizational and agent autonomy are certainly conceptually distinct, they are invariably linked. While an autonomous organization can choose to make its agents more, or less, autonomous an organization whose autonomy is constrained has no choice but to impose those constraints on its agents. If the organization needs to follow orders it cannot allow its agents substantial discretion.

Put another way, constraints on agents need not flow from constraints on organizations; a fully autonomous organization could impose significant de facto constraints on the exercise of an agent’s autonomy. However a fully autonomous, secure agency is less likely to constrain agents when those constraints are sub-optimal; organizational autonomy conditions employee autonomy, affecting not just the presence or absence of employee autonomy but the contexts in which it will be found. Table 2.1 describes this articulation graphically:
Table 2.1: Relationship Between Organizational Autonomy and Navigation Strategy

<table>
<thead>
<tr>
<th></th>
<th>Navigation by Judgment</th>
<th>Navigation from the Top</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High Organizational Autonomy</strong></td>
<td>Likely, if autonomy optimal for task; org will behave ‘rationally’</td>
<td>Will likely only occur if optimal for task; org will behave ‘rationally’</td>
</tr>
<tr>
<td><strong>Low Organizational Autonomy</strong></td>
<td>Very unlikely; Organization needs to “manage up”</td>
<td>Likely; HQ will need to ensure agents deliver what is needed by constrained organization</td>
</tr>
</tbody>
</table>

This suggests, then, that one way to identify when and where autonomy is likely to be sub-optimally constrained is to identify contexts in which that constraint flows from authorizing environments – contexts in which organizations or agencies have diminished autonomy because of the need to justify themselves. In these organizations the need to engage in legitimacy seeking will mean at least some of what the organization does (and thus the context the employee faces) is in fact “myth & ceremony” rather than motivated by functional considerations. (Meyer and Rowan 1977)

What value does agent autonomy and ability to Navigate by Judgment hold that would require us to think carefully about whether it is being fostered or constrained and how those constraints interact? There are four broad advantages to autonomy which will be examined in this work: 1) The collection and incorporation of soft information & tacit knowledge; 2) Flexibility; 3) The attraction and retention of better quality agents; and 4)
Improved organizational learning. Set against these benefits are potential costs to autonomous agents and Navigation by Judgment, including 1) Behavior by agents undesired by their principals; 2) The inability of agents to make effective use of autonomy; 3) A lack of standardization; and 4) The absence of high-powered performance incentives.9

**Soft Information & Tacit Knowledge**

Aghion & Tirole (1997) provide a framework for thinking about what may be lost as autonomy is circumscribed, framing it as “The Basic Trade-off between Loss of Control and Initiative.” (p. 10) They argue that agents who do not have autonomy have no incentive to gather contextual information about the environment they face; agents will take no initiative in either decision-making or in acquiring information that would inform organizational decisions in a future planning process.

For this concern to be relevant, it must be the case that the principal cannot just incentivize agents to collect the relevant information or do the right thing — there must be important features of the environment that cannot be directly contracted on. This does not mean that those features cannot be observed, however. As Dixit (2002) puts it, “We must distinguish between two senses of observability. In the stronger sense, known as verifiability or contractibility, the outcome can be proved to third parties, so contracts contingent on it can be written and enforced. In the weaker sense, the parties to the

9 The nature of the empirical strategy will make it easier to identify effects that vary within organization across environment; that is, the benefits accruing from attraction and retention of better quality agents and improved organizational learning and the costs of a lack of standardization will be harder to identify in this work. Nonetheless this work will speak to these organization-level dynamics where possible.
relationship observe the outcome, and can use it to draw inferences and take future actions that they are free to choose, but the outcome cannot be proved to outsiders and so cannot be the basis of a contingent contract” (p. 697-698)

For example, while we may not be able to build an output contract that specifies the quality of a doctor's visit (thus making the quality of the doctor's visit difficult to contract on), the patient can observe the quality of the visit. This information about quality, then, is observable though not verifiable. Observable non-verifiable information is foundational to theory about incomplete contracting and is a well-theorized, if still contentious, concept in contract theory. (e.g. Aghion and Bolton 1992; Baiman 1990; Hart and Moore 1988; Laffont and Tirole 1988)

Soft information is an extension of this observable non-verifiable notion from contract terms to any information that might be critical to an organization's work and/or an agent's task. Stein (2002) defines this as

Information that cannot be directly verified by anyone other than the agent who produces it. For example, a loan officer who has worked with a small-company president may come to believe that the president is honest and hardworking – in other words, the classic candidate for an unsecured ‘character loan.’ Unfortunately, these attributes cannot be unambiguously documented in a report that the loan officer can pass on to his superiors. (p.1892)

Soft information’s benefits are not only for the organization’s planning process, they also directly influence the organization’s performance in the short term; the bank where the loan officer works will profit from unsecured character loans properly awarded.
Tacit knowledge is the related idea that “we can know more than we can tell.” (Polanyi 1966) If soft information is data, then tacit knowledge is method. How to interrogate a data set or blink one’s eyes are examples of tacit knowledge – knowledge that is beneficial but is difficult to codify (e.g. write down) for transmission to someone else. In international development there are many examples of soft information and tacit knowledge about, and they are often intertwined. An IDO field agent, in visiting a project field site, recognizes early warning signs of project failure before any reports have been filed or data reported regarding implementation. The data that indicates project failure is soft information; the ability to recognize it is likely a skill honed via experience that would be difficult to codify, or tacit knowledge.

In the implementation of international development projects soft information and tacit knowledge come together in assessments of Ministry personnel and their motivations; identifying ways to structure or revise a project to maximize its likelihood of being in the interests of important political actors and thus being fully implemented; or assessing whether a project under implementation is headed in the right direction. Many things that are hard to codify and communicate up a hierarchy are plausibly critical to the success of development projects. More autonomy effectively puts the organization’s locus of control closer to the ‘ground’, to front line agents and (for the public sector) street level bureaucrats. (Lipsky 1980) This allows the organization to take advantage of non-codifiable soft information and tacit knowledge in the implementation and revision of projects. It also affects the design of projects – a more autonomous agent augurs for an
organization that acquires, and is more able to incorporate, soft information. This soft information allows the organization to make decisions that rely on this information, to Navigate by Judgment, in planning future projects.

Soft information and tacit knowledge are critical to making good decisions in a wide variety of tasks. Soft information is critical to understanding illegible environments; planning new projects; and making optimal day-to-day management decisions.

**Flexibility**

Autonomous agencies are more rapidly able to adapt when adaptation is needed. (Chandler 1977; Lawrence and Lorsch 1967; Thompson 1967; Williamson 1983) Rather than need to appeal to higher levels of management and provide them with information that might facilitate a decision, agents in the field can in a more autonomous organization make a greater number of decisions based on soft information and tacit knowledge. This makes decision making both faster and more responsive to environmental information. Navigation by Judgment will thus be associated with a greater ability to adapt to changing circumstances in implementation.

This flexibility channel interacts with that of soft information and tacit knowledge described above. They are, nonetheless, distinct – flexibility is in greater demand when environments change more rapidly, whereas the direct returns to soft information and tacit knowledge may persist irrespective of the velocity of change in the environment.
Agent Quality

Autonomy also has effects on workplaces and the employees who enter and exit a firm. Autonomy is positively associated with employee motivation (Hackman and Oldham 1980; Lopes, Lagoa, and Calapez 2014; Spector 1986) and job satisfaction. (Blegen 1993; Mortimer and Lorence 1989; Nguyen, Taylor, and Bradley 2003; Pink 2011) Indeed, autonomy is one of the key elements in job characteristics theory (Hackman and Oldham 1975), arguably the dominant framework for understanding work design and employee motivation in the organizational behavior literature.

If better quality employees prefer control over their work and a sense of having control over relevant parts of an organization's activities, they will be more likely to stay in organizations that grant them more autonomy. There is evidence supporting this notion, in that perceived autonomy is associated with employee retention. (Hannay and Northam 2000; Mitchell et al. 2001; Samuel and Chipunza 2009) If better quality employees were able to produce better organizational outcomes, then, this would provide another channel between autonomy and organizational performance.

The preceding paragraph suggests that there may be differential attrition out of organizations. There may also be selection in to certain occupations in ways that augur for differential optimal levels of autonomy, as greater autonomy attracts different kinds of agents. In the specific organizational context investigated here – IDOs – there is good reason to think that agents may be motivated in part by intrinsic, pro-social motivations.
Many employees of IDOs are not maximizing financial rewards in their current positions given their qualifications.

The literature on public service motivation (e.g. Perry and Wise 1990) suggests that employees may well be attracted by a desire to deliver on organizations’ missions, thus lessening the gains to be had from tighter monitoring. In other words, there may be less of a need to ensure that agents do not act in ways counter to the principal’s desires via Navigation from the Top if the agents share the principal’s objectives. Research on pro-social motivation, a related concept, highlights that these positive motivations are not constant; motivation may be kindled by meaningful work (Grant 2007, 2008) or extinguished by too much monitoring and control (Belle and Ongaro 2014; Hoey 2015).

This suggests that a common objection put forward to granting agents more autonomy – that the agents are not of adequate quality to be able to make use of autonomy – may reverse cause and effect. Agents may be of lower quality in part because of the lack of autonomy; giving agents autonomy may help to retain better agents among the existing pool, as well as attract additional high-quality agents.

In the long run, then, one could imagine two equilibria; one with low-quality agents with little autonomy, the other with higher-quality agents with a greater level of autonomy. As such, a shift towards a greater level of autonomy may be beneficial to an organization’s performance in the long term even where, due to low quality agents, it is not beneficial in the short term.
Organizational Learning

There are also effects of agent autonomy on the ability of an organization to learn. The Aghion & Tirole (1997) dynamic suggests that more constrained agents will take less initiative, reducing the amount of soft information the organization has to learn from; both organizational theory and empirics support this notion, with autonomy and greater freedom for agents associated with greater degrees of tacit knowledge acquisition and organizational learning.\(^\text{10}\) (Lam 2000; Marks and Louis 1999)

Benefits in Sum

Agent autonomy, then, can allow an organization to take more initiative in gathering and incorporating soft information and tacit knowledge, improve flexibility, increase motivation and retention, and potentially increase employee quality over time. An organization that Navigates by Judgment – that empowers field agents to make consequential decisions based on their instincts and locally gathered information – will see more initiative taken by agents. The cost of this initiative is as Aghion & Tirole describe – less control from the top. The next section articulates some of the costs of failing to have these controls.

\(^{10}\) McGrath (2001) finds a contingent effect, with autonomy beneficial to organizational learning under certain conditions.
The Potential Downsides of Agent Autonomy & Navigation by Judgment

Agents Behaving Badly

An agent who has more autonomy will find it easier to engage in a range of actions, including those that may be illegal or undesired. This includes both outright corruption as well as capture. (Tirole 1994) Capture is when an agent ceases to become an effective representative of their organization as they become unduly influenced by (in standard models of regulatory capture) the actors of a particular firm or industry they are attempting to regulate. (Dal Bo 2006; Laffont and Tirole 1991; Levine and Forrence 1990)

Capture is very much a live issue in development circles. Development organizations are concerned that as they send individual agents to distant locales where they spend much of their time interacting with domestic agents in the recipient country, their agents' interests will become aligned with something other than organizational best interest. Capture is sometimes articulated as a reason for staff rotation policies. More autonomous agents would, presumably, both be more susceptible to capture as well as be able to do more damage were they to in fact be captured.

Less constrained organizations and agents may also simply act in a manner other than that desired by their principals. This is why Aghion & Tirole (1997) frame the other side of the autonomy tradeoff as a loss of control. Niskanen’s classic treatment of bureaucrats in the American setting suggests bureaucrats wish to maximize their budgets irrespective of what organizational superiors or political authorizers might wish to see
transpire. (Niskanen 1968) While Niskanen himself has fallen out of favor,\(^\text{11}\) Niskanen points to the more general problem fundamental to principal-agent models. Agents may have objective functions that do not map those of principals and may wish to do different things. Depending on the size of the mis-alignment between principal and agent interests, tightly controlling agents may be the optimal course of action for principals even if in so doing there are costs as outlined in the previous section.

*Poor Quality Agents*

It is also possible that agents lack the capacity to make use of autonomy – that if Navigating by Judgment they will make bad judgments. I think of this as the “action movie runaway plane” scenario. We do not normally think it a good idea to have the distant control tower give pilots step-by-step directions and leave them without discretion. But when a passenger has to land a runaway plane we may think it is a very good idea to give that passenger as much guidance as possible, leaving as little as possible to Navigation by Judgment. Agent quality concerns presumably could be offset somewhat by training and progressive grants of autonomy – a relaxing, rather than simply removing, of constraints. Nonetheless, there is the potential for poor quality agents to move in the wrong direction simply because their judgment is flawed.

\(^{11}\) In the words of Barnett & Finnemore (2004), “Adopting the rather battered Niskanen hypothesis seems less than promising given the glaring anomalies” (p. 706).


Lack of Standardization

In many domains there is a benefit to consistency – of treating similar situations in similar manners. Doing so so can avoid accusations of favoritism and maintain brand consistency. In addition, if there is an optimal way to perform a task, any deviations from the established standard will lead to a loss of efficiency. If there is a “right” way of doing things, there is no reason to do things in any other way. In a parallel to the classic tradeoff between unitary (U-form) and multi-divisional (M-form) organizations, where flexibility is critical variation may be beneficial; where standardization is critical variation may be detrimental to organizational performance. (Chandler 1977; Williamson 1983)

Lack of Appropriate Performance Incentives

The section that follows will argue that measurement, like autonomy, is neither universally beneficial nor universally deleterious. That implies, of course, that measurement and performance incentives are sometimes a very good thing. An organization that provides its agents more autonomy will need to measure something other than day-to-day outputs in order to provide agents with the slack to act in the manner they see as best. Depending on the nature of the task and environment, giving agents autonomy may necessitate providing few if any measurable targets at all, which precludes providing the high-powered performance incentives that contract theory tells us is the first-best solution, particularly for repetitive, highly observable, routinizable tasks.
One of the key ways public sector organizations – including IDOs - Navigate from the Top is via measurement. Measuring what people do and motivating them accordingly has a clear and compelling intuition in its favor, one that resonates with anyone who has ever seen a child’s eyes get brighter at the promise of something they desire after their homework is completed. Aligning tangible rewards with production aligns employees' incentives with firms, ensuring that employees are contributing to a firm's bottom line with greater rewards for a greater contribution. In the last few decades new public management (NPM) has increasingly carried over private sector incentive schemes into the public sector. (Christensen and Laegreid 2011; Hood 1991, 2004; Lorenz 2012; Pollitt and Bouckaert 2011) At the employee-employer level this is often framed as Pay-for-Performance or Performance-Related Pay, incorporating (measured) performance into employee compensation.

The logic for importing these private sector practices to the public sector is often articulated as one of catalyzing better performance. As the US law instituting performance measurement for Federal programs puts it, “Federal managers are seriously disadvantaged in their efforts to improve program efficiency and effectiveness, because of insufficient articulation of program goals and inadequate information on program performance; and congressional policymaking, spending decisions and program oversight are seriously handicapped by insufficient attention to program performance and results.” (Government Performance and Results Act of 1993, 2.a.2-3)
One reason organizations may engage in Navigation from the Top – particularly via Navigation by Measurement - is for legitimacy. (Meyer and Rowan 1977) By measuring things and meeting targets, even when those targets are not well linked to ultimate organizational goals, organizations can appear to be performing to political authorizers. This dynamic is hinted at by the text of the US Government’s performance measurement act, which lists as its first purpose to “improve the confidence of the American people in the capability of the federal government, by systematically holding Federal agencies accountable for achieving program results,” (US Government 1993, 2.b.1) The Federal measurement regime, then, is explicit that one of its goals is legitimacy-seeking.

Measurement has increasingly become critical to justifying continued funding and build legitimacy in the public sector as part of a broader discourse on accountability and control. The spread of New Public Management can be read as a form of mimetic isomorphism (cf. DiMaggio and Powell 1983), with measurement spreading for these legitimacy-seeking reasons. (Dunleavy and Hood 1994; Hood 2004; Lynn Jr 1998; Modell 2004; Oliver 1991) When agent autonomy is constrained and measures contracted on in an attempt to speak to outside stakeholders (be they stock market analysts or political authorizers) additional potential problems are introduced. As Ittner and Larcker (1998) put it, “Institutional theories argue that in organizations such as government agencies, whose survival depends primarily on the support of external constituents and only secondarily on actual performance, managers will implement the mandated [measurement] systems in order to appear modern, rational and but will not actually use the systems to improve performance.” (p. 233) One could certainly read the text of the
Government Performance and Results Act quoted above as, in part, suggesting reasons other than direct performance for measurement; namely, promoting public confidence and giving the Congress better ability to control projects. A defender of the act could argue that the goal is to improve performance via Congressional oversight, with better performance that will improve public confidence in government as a result. But what if in at least some instances oversight, measurement, and diminished autonomy hurt performance?

Organizations measuring for isomorphic or legitimacy-seeking reasons may take a sub-optimally low level of risk. If an organization needs to meet measures in the short term in order to e.g. receive continued funding, the organization may not take the risks necessary to achieve long term ends – an organization’s ‘risk appetite’ may be sub-optimally constrained. While there is little empirical work on this question directly, there is some suggestion that organizations that are too tightly supervised may be sub-optimally risk averse. (Bozeman and Kingsley 1998; Singh 1986)

The organizational learning dynamic may also be complicated by organizational incentives for activities to be seen by external stakeholders as successful. If authorizing environments are such that organizations will be penalized for failure, then organizations will not be able to codify their failures and learn from them. By not documenting where things may have gone wrong, actors in the authorizing environment will not be able to hold the organization to account. One can imagine two individuals in a US Government agency or Wall Street bank conversing about a difficult subject matter rather than writing an email, lest that email become a document that Congress or regulators could ask to view. In this
way fear regarding an organization’s relationship with its authorizing environment can interfere with an organization’s ability to communicate with itself, to understand its own actions and failures.

**Measurement & Autonomy**

Measurement and autonomy are not *necessarily* opposed to one another. At one extreme, measurement in packing warehouses (e.g. those of Amazon) are places where staff are instructed by computer and measured on per-second productivity. The measurement regime is in this case intimately connected to one reporter’s recent description of this environment as that of “warehouse wage slavery” and highly constrained autonomy. (McClelland 2012) On the other hand, proponents of “Cash on Delivery” in foreign aid – paying developing countries in exchange for their performance – hail the *increase* in autonomy that would result from a shift to Cash on Delivery as one of its key strengths. (Birdsall et al. 2010) In the former case, measurement is in service of Navigation from the Top; in the latter, Navigation by Judgment.

The key distinction is when measurement occurs – at what level, with what frequency. The contract theory and ‘pay for performance’ literature commonly have in mind a frequently measured output, something like a garment worker’s piece rate or hospital worker’s daily attendance and productivity. By contrast the Results-Based Financing literature has in mind measuring results; final outcomes like firm profits, patient’s health, or student’s achievement. With clear, well aligned, and clearly observable
outcome measures, autonomy can be complementary to high-powered performance incentives. (Holmstrom and Milgrom 1994)

Contracting on measures in either the short-term (pay-for-performance, output) or long-term (results based financing, outcome) sense can have distortionary effects and be subject to Goodhart’s Law. Both short- and long-term measurement can be beneficial or deleterious, and in both cases theory suggests that the alignment of contracted measures to what is actually desired is critical, as are issues of asymmetric information and observability. The effect of long-term measurement on autonomy is to loosen short-term task controls. If a factory is measured on annual rather than daily production, the factory management will be better able to adjust the time, place, and manner of work, and may well innovate with regards to the production process. Short-term measures, however, tighten short-term controls. This is what Austin (2013) in the business literature has in mind when he describes measurement as the opposite of delegation; as shifting control from agent to principal.

Examples of inappropriate measures or timing abound in the private and public sector. While by no means the only cause, that bankers were contracted on short-term performance is related to their engaging in riskier behavior than would have otherwise been the case; as Crotty (2009) put it, “Since [financial firm operatives] do not have to return their bubble-year bonuses when the inevitable crisis occurs... they have a powerful incentive to pursue high-risk, high-leverage strategies.” (565) Measures, then, need not just
to be well aligned to desired organizational performance but to by synchronous in terms of periodicity to avoid trading long-term success for short-term output.

   Measurement, then, may very well *enhance* autonomy and agent ability to Navigate by Judgment by giving organizations and agents control over how to meet goals. If, however, the attempt to measure is done at an inappropriate level or time, measurement will likely form a constraint on agent autonomy. Put simply, an organization will get what it measures; when this measure is an appropriate summary statistic for an activity, this will be empowering and performance enhancing. When it is not, this will be constraining, crowding out soft information as agents and organizations meet measurement targets rather focus on the task at hand.

   While contracting on measures – giving the measures ‘teeth’ – makes more extreme both potential good and bad effects of measurement regimes, there is also power to measurement, to numbers, that goes beyond their use in management by others, and are harder to explain in a strictly rational choice framework. The thrust of the commensuration literature is in part that metrics can have powerful individual and social consequences beyond their direct ‘teeth’ in performance monitoring or evaluation. (Espeland and Sauder 2007; Espeland and Stevens 1998) Anyone who has ever “cheated” on a Weight Watchers diet by recording fewer points than actually consumed or pushed through pain at the gym to meet a target knows that numbers can have weight in decision making and reasoning even in entirely independent tasks, without any monitoring by others.
When output measures are contracted on – when they are tied to organizational rules and/or performance measurement – measures become all the more powerful. As individuals respond to these measures they necessarily behave in a manner different than they might have in the absence of the measure; the question at bar is when this differential behavior is net positive for organizational goals, and when it is net negative. In either case, the counterfactual – the path that measurement moves us away from – is what agents, and agencies, acting autonomously, would have done both in the management and policy sense in the absence of output-based numerical supervision schemes. This includes decisions that would have incorporated observable non-verifiable information (soft information) or abilities (tacit knowledge) that are unavailable to agents & organizations but cannot be made due to measurement constraints. As Aghion & Tirole (1997) argue, this will also mean agents will not collect soft information (having no ability to use it and thus no incentive to invest effort), leaving the organization without the knowledge necessary to make policy decisions. A need to ‘manage up’ to political authorizing environments is one of the key drivers of inappropriate measurement dynamics.

**Technical Challenges of Measurement**

Measurement is also not a simple thing to get right, leading to many possible distortions; these distortions are more likely when the measurement is done for legitimacy seeking, rather than performance-linked, reasons. Much depends on what is measured, when, and for what purpose. There are five basic technical issues with measurement as a
management tool: 1) The cost of finding tractable (where measurement is both timely and accurate) indicators; 2) Applying measurement in multitask environments; 3) Applying measurement where the goal cannot be articulated ex-ante; 4) The stability (in the sense of Goodhart’s Law) of the relationship between measures and ultimate organizational goal; and 5) The effect of measurement on intrinsic motivation. Each will be discussed briefly below. These technical challenges then combine with authorizing environment constraints in determining the right level of measurement (and the constraints on same).

Finding Timely and Accurate Measures

The public sector differs most fundamentally from the private in that there is no overarching outcome (e.g. profit) which can be observed, and which can thus be contracted on (in some cases) as an accurate metric of organizational performance. (Dixit 2002) In the public sector this is rarely the case, and thus what can be measured (e.g. test scores) is often merely a proxy for what’s desired (e.g. student achievement or teacher performance); these measures can often be mis-aligned with outcomes, sometimes described as the folly of rewarding A while hoping for B. (Kerr 1975)

For a measure to be useful, it needs to not just be accurate but also be collectable in a timely and reliable manner. A medical test may provide an excellent measure for a given ailment, and thus whether an intervention aimed at reducing that ailment is succeeding. If, however, bad roads make it impossible to administer the test with frequency, or a lack of cold storage diminishes the reliability of the test it will not be possible to successfully Navigate by Measurement using such a metric.
Organizations with insecure authorizing environments are constrained with regards to the potential periodicity of measurement. It will not be possible to wait for long-term successes to develop if metrics are needed to justify a project’s continued funding or an executive’s proper stewardship of the firm. Instead, short-term output measures will need to be sought in order to generate data in a timely manner, increasing the likelihood of measurements being poorly correlated with outcomes and/or distortionary.

*Measurement & Multitask: You Get What You Contract For*

Measurement leads to greater production of whatever is measured; the question is in some ways one of when that is likely to be a good thing, and when bad. There is substantial work that suggests that if a job is "multi-task", having both measurable and unmeasurable components, measurement will be distortionary. (Dewatripont, Jewitt, and Tirole 2000; Holmstrom and Milgrom 1991) Agents will invest in what is being evaluated, but will consequently under-invest in what is not being measured. The implication, then, is that when a task can be fully summarized by a measure that measure will not have a multitask problem; if one is constructing a school, measuring the construction quality and presence will avoid multitask problems. If, however, one is training teachers, it will be hard to avoid multitask difficulties wherein observable outputs of the training (e.g. hours trained or content retained) are measured and thus delivered while unobservable components (e.g. communicating the ‘soft skills’ of reading a classroom) are under-invested in.
**Applying Measurement Where the Goal Cannot be Articulated Ex-ante**

Determining a target against which to assess progress necessitates having a fixed goal – being able to determine ex-ante what will constitute success ex-post. In some sectors – e.g. health, education, and infrastructure – this may be relatively unproblematic. In others – e.g. reforming the functioning of a Ministry of Planning – it is less obvious what ought be specified ex-ante as constituting success. Advocates of the “Doing Development Differently” have, in their manifesto, argued that successful international development interventions “focus on solving local problems that are debated, defined and refined by local people in an ongoing process” and “blend design and implementation through rapid cycles of planning, action, reflection and revision”.¹² (Doing Development Differently 2014)

This brings to mind Lindblom’s (1959) “muddling through” – the view that successive limited comparisons may be an efficient and desirable strategy when needs are complex and goals difficult to specify. It is difficult to muddle through if targets have been contracted on ex-ante.

---

**Goodhart’s Law and the Breakdown of the Correlation with Outcomes**

In the public sector the breakdown of association between measures and their correlates once they are contracted on is so fundamental that it has led to a series of oft-invoked and related axioms (Goodhart’s Law, Campbell’s Law, and the Lucas Critique)¹²

¹² I was a participant at the 2014 conference from which this statement comes, and am a signatory to the statement. I did not, however, influence the writing of the statement or contribute any language to same.
describing the incentive-distorting element of contracting on measures. As Charles Goodhart, former Chief Advisor of the Bank of England, put it “any observed statistical regularity will tend to collapse once pressure is placed upon it for control purposes.”\textsuperscript{13} In the absence of pressure, output measures may be well correlated with outcomes – the success of a clinic in improving community health, for example, may be well correlated with the number of patients the clinic sees. Imagine, then, if pressure is put upon this measure for control purposes – e.g. if clinics move to a measurement regime wherein number of patients seen is a measure of performance or is directly incentivized via a pay-for-performance scheme. The number of patients seen may cease to be correlated with community health as the clinic sees more patients at the expense of quality care or need for patients to be seen (that is, seeing patients who are not ill to boost numbers).

\textit{Measurement and Intrinsic Motivation}

There is substantial research that suggests pay for performance and a focus on extrinsic targets has the effect of crowding out intrinsic motivation, shifting focus to instrumental (reward-linked) goals. (Frey and Jegen 1999; Frey 1994; Gneezy, Meier, and Rey-Biel 2011; Gneezy and Rustichini 2000; Wrzesniewski et al. 2014) This is another way in which measurement can shift individuals and organizations towards contracted aims and instrumental rewards at the potential expense of the ultimate ends to which these contracted measures are intended to be aligned. This also echoes back to McGregor’s foundational work on “Theory X vs. Theory Y” half a century ago; to the notion that

\textsuperscript{13} Goodhart – now an emeritus professor at LSE – has endorsed the application of his law to the theory articulated in this work.
measurement will make employer-employee relationships transactional, reducing performance and ‘crowding out’ intrinsically motivated agents when applied to tasks that cannot be so easily measured. (McGregor 1960)

These technical challenges do not doom measurement and Navigation by Measurement to failure. They, do, however, potentially impose costs which need to be balanced against the benefits of measurement in aligning agents towards accomplishing tasks, particularly where direct process monitoring is difficult (e.g. because agents are distant from principals, working in relatively isolated environments). Whether measurement is viable will depend on the nature of the task and the environment – as will the benefits of autonomy.

**Special Public Sector Challenges for Navigation by Measurement**

Prominent economists have argued that the multitask and difficult to monitor nature of public sector jobs augurs for lower powered incentives and separating tasks so they are as narrowly constructed as possible. (Dewatripont, Jewitt, and Tirole 1999) Public sector firms also differ in terms of governance. They are often answerable to a complex web of actors – effectively, multiple principals – that change the nature of what incentives are possible, and the effect of incentives once they are put in place. (Dixit 1997) James Q. Wilson argues that in addition to living in this web of politics, government tasks are differentially observable, and differentially monitorable; the implication is that contracting on measures is likely to be differentially effective based on the monitorability of the task.
(Wilson 1989) Dixit notes, however, that there is little systematic empirical research on the
effect of incentives and pay-for-performance in the public sector, suggesting that such work
“should not seek sweeping universal findings of success or failure of performance-based
incentives or privatization, but should try to related success or failure to specific
characteristics like multiple dimensions and principals, observability of outputs and inputs,
and so on.” (Dixit 2002)

While there is some evidence that US Government agencies employing performance
measurement improved performance prior to performance measurement being mandated
by law in 1993, there is little evidence that agencies responding to the mandate improved
their performance via measurement. (Cavalluzzo and Ittner 2004) While control and
contracting on measures sometimes leads to improvements in the public sector, it does so
with a large degree of Goodhart’s Law problems, as measures cease being as well-
correlated with outcomes as they were prior to contracting on measures. (Propper and
Wilson 2003) Multitask problems do lead to difficulties in setting up an optimal
measurement regime; attempts to rectify this lead to some outcomes (largely financial
outcomes) being achieved but more mixed performance in other areas. (Kloot and Martin
2000; Rantanen et al. 2007) A recent review of pay for performance in the public sector
finds mixed but broadly positive results. (Hasnain, Manning, and Pierskalla 2014) This
work also suggests that studies of pay for performance are overwhelmingly concentrated
in jobs where outcomes are relatively observable, suggesting that observability may be a
necessary condition for the success of these kinds of contracts. (Hasnain, Manning, and
Pierskalla 2014)
While private sector organizations are often accountable to executive boards and shareholders, public sector organizations are accountable to a more diverse set of actors with a broader range of preferences about a public agency’s optimal role and function. In this debate measurement can play a legitimating role, helping an agency to justify itself in the manner suggested by Meyer and Rowan (1977). This resonates with some observers of public sector measurement, with these measures sometimes based in ‘myth’ or fad, not data (Modell 2004). Firms then might engage in isomorphic mimicry in the style of Dimaggio and Powell (1983), copying measurement structures to further engage in legitimacy seeking. Thiel and Leeuw (2002) suggest that even though (on their view) the motivation for the push for measurement in the public sector is often about accountability and not performance, it can still be beneficial if done right.

**Other Techniques of Navigation from the Top**

The past pages have focused on measurement as it is one of the key under-explored levers via which Navigation by Judgment is precluded; it is far from the only potential means of control, however. As alluded to above, financial dependency and procedural rules can also constrain and organization’s, and agent’s, ability to act autonomously. In addition, career concerns models (e.g. Dewatripont, Jewitt, and Tirole 1999; Holmstrom 1999) underline that Navigation from the Top need not have ‘teeth’ only via control during implementation; another way to drive agents towards attempting to ‘manage up’ to the top is via promotion and salary rewards which can be determined after a project’s close. All of
these control schemes have the effect of causing agents and organizational managers to point their compasses ‘up’ to the top – to satisfying their principals – rather than ‘down’, towards the soft information to be found only at ground level. Other scholars have also found that M&E systems more broadly can have deleterious consequences (Belle and Ongaro 2014; Hoey 2015); this work will attempt to disentangle to some extent what elements of Navigation from the Top seem to be driving differential performance.

2.2: Establishing Boundary Conditions for Optimal Navigation Strategy

The section above has served to introduce the complex relationship between autonomy and control, particularly measurement, with organizational success. That there are both advantages and disadvantages to organizational Navigation by Judgment relative to organizational Navigation from the Top does little to tell us precisely what a manager or organizational designer ought do to improve performance. When should we expect Navigation from the Top to be a superior strategy? When will a scheme that constrains agent autonomy, including by contracting on the measurable – contracting on outputs that are both observable and verifiable – be the best way to achieve results? And when should we expect an organizational scheme that relies on agents’ judgment, that gives agents autonomy and incorporates soft information, to be relatively more successful? The paragraphs below seek to develop the conditions under which, precisely, a move towards greater Navigation by Judgment and autonomy is likely to be most useful.
The notion that context matters for optimal strategy, with autonomy as one of the key potential levers, is an old one in organizational behavior. (Lawrence and Lorsch 1967; Thompson 1967) It has also been extended to public sector agency, with the notion that different policy areas require different delegation strategies. (e.g. Huber and Shipan 2002)

Table 2.2: Summary of Benefits and Costs of Agent Autonomy

<table>
<thead>
<tr>
<th>Benefits</th>
<th>Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>The collection and incorporation of soft information &amp; tacit knowledge</td>
<td>Behavior by agents undesired by their principals</td>
</tr>
<tr>
<td>Flexibility</td>
<td>The inability of agents to make effective use of autonomy</td>
</tr>
<tr>
<td>The attraction and retention of better quality agents</td>
<td>A lack of standardization</td>
</tr>
<tr>
<td>Improved organizational learning</td>
<td>The absence of high-powered performance incentives</td>
</tr>
</tbody>
</table>

Table 2.2 summarizes the advantages and disadvantages of agent autonomy outlined above. Where Navigation by Judgment is the superior strategy is a function of which set of effects dominates. This will be a function in part of the context in which the organization is intervening. How certain, or predictable, is the environment? To what degree is soft information required to make optimal decisions? How great a mis-alignment is there between the interests of the principal and that of the agents?

*Environmental Uncertainty & Legibility*

One of the key drivers of optimal strategy is the degree of environmental uncertainty. The literature argues that greater environmental volatility (and thus less
potential for task routinization) is associated with a higher optimal level of agent discretion and autonomy (Burns and Stalker 1961; March and Simon 1958; Perrow 1967; Thompson 1967). Much like autonomy, environmental uncertainty is a somewhat vague notion. Milliken (1987) helpfully differentiates between 1) state uncertainty – a lack of knowledge regarding the state of the external environment; 2) effect uncertainty – the impact of the environment on the organization’s activities; 3) response uncertainty – what the reaction of the environment will be to the organization’s activities.

This fits nicely with the theory on autonomy and measurement articulated above; in an environment with greater effect and response uncertainty flexibility will be more highly prized, and in an environment with greater state, effect, and response uncertainty it will be much harder to specify ex-ante appropriate output targets, ensuring that they remain well correlated with desired outcomes and are not distorted by Goodhart’s Law problems. More unpredictable contexts – those with greater effect and response uncertainty – will be contexts in which soft information (including soft information to address state uncertainty) is more likely to be a necessary part of the production process, tacit knowledge more likely to be relevant. As such an organization whose locus of control is further from these field-level agents will not fare as well, with it difficult to make appropriate decisions from the center for these contexts.

There is also the more direct matter of the legibility of an environment – the extent to which environments can be understood from a distance. In the context of international development this might be understood as the correlation of de jure structures with de facto
reality. In those developing countries characterized by higher levels of legibility the name on the door of a government unit is well correlated with the activities that take place within. In others, this is not the case. The greater the gap between structures and reality, the greater returns to soft information.

Navigation by Judgment, then, will have increasing returns when environments are more uncertain and illegible. In the most uncertain & illegible environments, Navigation from the Top is likely to be the superior strategy, ceteris paribus; in the most unpredictable environments, Navigation by Judgment is expected to be the superior strategy.

Task Observability

As suggested by Wilson (1989), not all tasks are created equal. If an organization is constructing a building, there are clear standards of quality that can be observed or contracted on. If an organization is training teachers it is much harder to develop appropriate measures against which results can be measured; as such an attempt to measure will prove more distortionary in the latter context than the former. The notion that tasks are inherently different and pose different measurement challenges is well-articulated in the Management Control Systems literature in private sector contexts and has also proved a critical part of some of the most prominent theorizing in the public administration literature on bureaucratic functioning and contracting. (Brown and Potoski 2003, 2005; Wilson 1989)
As explored above, tasks can be multi-task to a differing degree, with good, stable, non-distortionary measures more or less available. Tasks also demand differing levels of soft information, as some tasks are more routinizable than others. A good proxy for considering a task’s incorporation of soft information is to ask – ignoring the bounds of technology and cost – whether it would be possible to program a robot to perform the task. The greater the likelihood such a program would be able to account for all possible contingencies, incorporating all necessary environmental feedback and adapting accordingly, the less the task incorporates soft information. The more “programmable” the task, the greater the extent to which high-powered measures can be contracted upon as performance outcomes.

One of the key margins here is how observable the relevant parts of the task are; if the task is more like constructing a building or training teachers, and thus how closely a measurement regime can approximate the first best solution of identifying and contracting on outcomes. The greater the wedge between this first best solution and the actual realized environment, the more likely it is that autonomy and Navigation by Judgment are optimal organizational choices.

Deviations from Optimal Navigation in Equilibrium

If all firms fully endogenized the relevant concerns, choosing navigation according to the nature of the task, environment, and other relevant concerns, then the theory developed here would be merely descriptive rather than potentially prescriptive. There are many reasons firms may not engage in optimal management choices, including path-
dependent historical legacies, mis-evaluation of the nature of the environment, or a simple lack of awareness to path-dependent retention. As argued above, one substantial and systematic cause is legitimacy-seeking behavior, including (but not limited to) isomorphic mimicry of dominant management trends. If autonomy were a way of signaling that one’s organization was high performing, I would then predict that we would see inappropriately high levels of autonomy as organizations; at the present moment, particularly as regards public management, the pendulum is closer to the other extreme. As such, we should expect to see autonomy sub-optimally constrained.

To the extent, then, that we can identify contexts in which the motivation for constraint (including constraint via measurement) is legitimacy seeking, the constraint of the organization is more likely to redound to constraints on agents that are not optimal for achievement of the organization’s goals. I am arguing, then, that constraints travel ‘downhill’, with constraint at the top translating into the (lack of) autonomy of agents.

*Implications of the Theory in Sum*

In sum, then, I am arguing that an organization that Navigates by Judgment will fare better when tasks are difficult to routinize and/or environments are relatively unpredictable and it is hard to define appropriate targets ex-ante or find good measures.
2.3: **International Development as a Context of Study**

The relative paucity of empirics regarding the precise contours of autonomy and control's contingent relationship with performance is eminently understandable; it is difficult to identify appropriate research contexts for such a question, particularly ones with consistent outcome variables and variation on the relevant independent variables. While the propositions articulated above are falsifiable, this does not mean they are readily testable; these are not easily tractable questions, one of the reasons why despite drawing on well-established theory some of which is a half-century old these propositions have yet to be examined comprehensively.

*Necessary Features of An Optimal Context for Empirics*

An appropriate context to test the theory articulated here would need, first, some kind of relatively consistent and reliable measure of performance that could be compared across environments and tasks. It also requires some way of measuring organizational autonomy, including possibly some measure of the authorizing environment constraint that is particularly likely to constrain organizations, and consequently agents, in a sub-optimal manner. If this context could provide a venue for examining the link between organizational and agent constraint that would also be helpful in providing suggestive evidence for the mechanisms posited here.

A proper research domain would also need to avoid the selection problem – the tendency of organizations to engage only in contexts where they are relatively likely to be
successful. It would be optimal for organizations to respond to their own organizational constraints by only engaging in environments or sectors where they are more likely to perform well; however, this acts to effectively censor the data, in that one cannot then see results where organizations might perform least well.

This last issue augurs towards investigation of a situation where the engagement or investment decision is made independent of the organization; where the organization’s engagement is dictated by rule or higher power. This suggests a public sector context may be most appropriate, but one in which individual organizations engage in a variety of different tasks in a variety of different environments. Foreign aid – efforts by public agencies to catalyze development by implementing activities in a wide range of recipient country environments and task domains – is just such a context.

*International Development Assistance as an Appropriate Context*

The administration of Official Development Assistance is largely via large public bureaucracies, both national and international – agencies like the US Agency for International Development, the UK’s Department for International Development, and the multilateral World Bank. Output measurement has increasingly become a core part of international development, with organizations increasingly moving towards measurement and management through targets and indicators which are linked to externally verifiable data so as to improve accountability and work towards what many in international development assistance term the “results agenda”. (Gulrajani 2011) There is also a stream
of work from both the academy and practitioners that much of what development, and particularly development measurement, produces can be thought of as legitimacy seeking rather than results oriented. (Easterly 2002; Eyben 2013) This has been connected to political authorizing environments and the fight for continued funding and resources. (M. Barnett 2009; Bush 2011; Martens 2002; McMahon 2001)

International development also provides a cogent way of considering the periodicity of measurement and the interrelationship of short- and long-term results via what is sometimes called a “logical framework” (logframe) or “results chain”. While there is no universally accepted format, one version (in this case as articulated by DFID) is depicted in Figure 2.1.

Figure 2.1: Results Chain from DFID (UK Department for International Development (DFID) 2011, p. 11; figure from the original work)

The example provided in DFID’s articulation shows the various measures (amount of funding, finalization of strategy, number of schools built, number of children completing
school, literacy rate) that could be employed in determining the effectiveness of a hypothetical education project. These measures occur in different time scales – while strategy finalization can be verified immediately following the strategy’s completion, literacy rates will be much slower to change. These measures are also differentially subject to Goodhart’s law. While there is a small possibility for Goodhart’s law problems when measuring literacy rates (e.g. if evaluators were biased), there is clearly much greater Goodhart’s law potential with building schools – it would be fairly easy for a project to build schools but for those constructed schools not to translate into improved literacy; e.g. the buildings may be unused or staffed with untrained teachers. If number of schools constructed is used as a performance measure, actors will have an incentive to build schools irrespective of these concerns; thus a measure (number of new schools) that in the absence of monitoring/contracting on measures we believe is likely to be correlated with improving area literacy may cease to be accurate.

*Political Authorizing Environments and the Choice of Measures in International Development*

Is, then, international development a context in which external political authorizing environments augur for legitimacy-seeking Navigation from the Top, including Navigation by Measurement?

The complex political authorizing environments of aid givers (de Mesquita and Smith 2009) and the distortions they sometimes give rise to (Barnett and Finnemore 2003) naturally provide variation which can be empirically exploited regarding the
characteristics of aid agencies. Numerous scholars have framed international development agencies as organizations in which the interplay between political principals and IDOs, and IDOs and their agents, play critical roles in organizational functioning and outputs. (Hawkins et al. 2006; Nielson and Tierney 2003) As for bureaucrats in such agencies, there is good reason to think that they substantially influence what occurs, and matter critically to organizational rules and success. (Johns 2007; Johnson and Urpelainen 2014)

It is surprising, then, that there has been so little empirical work on international organizations that animates the agents, rather than the principal, despite calls to do so. (Hawkins and Jacoby 2006) This work responds to Wilson’s call to begin to both focus on organizational systems and begin with front line workers in understanding organizations. (Wilson 1989; 23, 33-34) It aims to respond to what some have called “Wilson’s Challenge”, namely “Ignorance of variation and complexity, and the consequent failure to recognize the importance of internal organization.” (Chang, Figueiredo, and Weingast 2001) This work is also among the first to take up Dixit’s (2002) call for empirics that do “not seek sweeping universal findings of success or failure of performance-based incentives or privatization, but should try to related success or failure to specific characteristics like multiple dimensions and principals, observability of outputs and inputs, and so on.” (p.724)

Criticism also constrains what international development organizations do, and imagine themselves as capable of doing; as Tendler (1975) puts it regarding the US Agency for International Development (USAID):
It has been generally recognized that criticism of the foreign aid program weakened [USAID] and kept it from doing what it wanted to do. Less understood is the fact that the process of living with criticism profoundly affected what the agency wanted to do and what it was capable of doing. (p. 40)

Political authorizing environment constraints change the incentive structure of aid organizations and their agents, shifting toward Navigation by Measurement and away from Navigation by Judgment. It is no surprise, then, that aid organizations often focus on changing what can be easily measured (policy, structures) in recipient countries, focusing on areas where outputs can be measured at the expense of areas where measurement is more difficult. (Andrews 2011; Eyben 2013; Pritchett and Woolcock 2004)

A recent report surveying British non-governmental organizations (NGOs) that implement projects in the humanitarian space discusses evaluation in practice. (Guerrero, Woodhead, and Hounjet 2013) The report links the NGO reporting to DFID’s increasing demands for accountability. Perhaps most instructively, the report classifies 1,680 indicators used by 11 NGOs in measuring their performance. The results of this classification using the same definitions at the DFID results chain pictured in Figure 2.1 are presented in Figure 2.2.
What is most striking about this data is how the great majority (greater than 80%) of measured indicators are at the input, process, and output levels, which also are the levels where we might expect these measures to be least well-correlated with ultimate goals and most susceptible to Goodhart’s Law problems.

While it is surely the case that this distortion towards the more immediate and short-term is driven in part by the availability and timeliness of data for an agency’s own ends, it also may be that the need to report ‘up’ to funders and political authorizers also plays a role. There is some preliminary evidence that as reliance on numbers intensifies in international development the quality of those metrics is degrading in some cases; that numbers are becoming less reliable as implementers have increasing incentive to provide inaccurate data in reporting to funders and authorizers. (Sandefur and Glassman 2013)
There is perhaps no more eloquent description of these dynamics than that of Andrew Natsios, the former head of USAID. Natsios (2010) describes what he calls Obsessive Measurement Disorder (OMD), an intellectual dysfunction rooted in the notion that counting everything in government programs (or private industry and increasingly some foundations) will produce better policy choices and improve management... demands of the oversight committees of Congress for ever more information, more control systems, and more reports have diverted professional USAID (and now MCC14) staff from program work to data collection and reporting requirements. (p. 8)

Natsios relates the inappropriateness of measurement directly to the nature of the difficulty of measuring outcomes, invoking Wilson’s (1989) framework on bureaucracy referenced above to do so.

*Soft Information in International Development*

Natsios’ prescription to resolve these issues is one that dovetails nicely with this work’s focus on autonomy and soft information. He argues that “USAID should decentralize aid programming and decision making to the lowest possible organizational level, where officers have the greatest knowledge of what is happening on the ground” (p. 72) He also notes that staff are often frustrated by the lack of autonomy and “risk aversion” that results from this environment. (p. 57-58)

These insights echo those of Elinor Ostrom and her team, who argue in an analysis of the Swedish International Development Agency (SIDA) that “the broader institutional context of the donor agency has profound effect on the relationships between recipient and

14 Another US aid agency, the Millennium Challenge Corporation.
beneficiary organizations, contractors, and the individuals working with the aid agency”.
(Gibson et al. 2005, p. 156) They also argue for decentralization to the field, in part to give
staff the autonomy and incentive to overcome what they see as “significant asymmetries” of
local knowledge (tacit knowledge or soft information). (ibid, p. 42)

The theoretic frame described in earlier sections, then, fits the world of foreign aid
quite well. So, too, do the data generating dynamics. Aid organizations face differential
political environments and have different levels of organizational and staff autonomy.
Agencies work in a variety of different task domains, from road building to judicial reform,
which are differentially verifiable and contractible. The countries in which these
organizations work range from the relatively stable and predictable to the fragile and
unpredictable, including those where conflict is ongoing and its outcome uncertain. The
biggest organizations work in all of the most fragile countries in the world, suggesting an
absence of ability to select out of unpredictable environments (empirics in later chapters
will further support this assertion). Finally, many aid organizations not only measure
project targets but also have holistic ex-post reviews of projects, providing a vast amount of
outcome data that, while by no means uncomplicated, can provide leverage on results. In
addition, one can compare the functioning of the same organization across environments
and tasks, different organization within environments and tasks, etc., providing analytical
leverage for both quantitative and qualitative investigations.

Contracting in International Development
Another notable feature of IDO behavior is the use of contractors to implement projects. The targets and control mechanisms discussed as part of development’s Management from the Top are often formalized in formal contracts with implementing organizations. These organizations (which are often – but by no means exclusively – not for profit entities) bid for work and then implement projects on behalf of IDOs. On the one hand this complicates the role of soft information, which can flow through contractors as well as directly from IDO field staff; it also introduces variation in how IDOs supervise their field staff. On the other hand, to the extent that contracts (by design) introduce inflexibility, it allows for an examination of how differential Navigation from the Top affects contractors as well as IDO agents and thus development outcomes. All contracts may, or may not be, created equal. Whether contracting diminishes or eliminates the benefits of IDO autonomy, or whether there is variation in how IDOs manage contractors and the success of projects that are delivered via these contracts, will be investigated via this work.

**2.4: Empirical Predictions in the Context of International Development**

Large IDOs face heterogeneous political authorizing environments, with heterogeneous levels of insecurity and consequently autonomy. I theorize that IDOs with less room to maneuver in their political authorizing environments will respond by focusing on Navigation from the Top and on managing ‘up’, on responding to politics and the concerns of those who authorize the organization’s funding and thus carefully justifying the organization’s actions and programs to a greater extent than is the case for a more secure, more autonomous IDO. This will in turn put constraints on the actions of field-level agents,
constraining their autonomy and ability to Navigate by Judgment. As a result of these dynamics the decisions of a less autonomous IDO will incorporate lower levels of soft information. I hypothesize that there will be a positive relationship between organizational autonomy – an organization’s de facto ability to exercise managerial and policy autonomy – and the autonomy of field agents of those organizations. More autonomous IDOs will have a greater propensity to Navigate by Judgment.

IDOs with lower levels of autonomy, then, will engage in greater levels of measurement where it is not performance-enhancing, and contract to a greater extent on those measures – making them binding for agents. The extent to which lower levels of autonomy and a greater focus on measurement is deleterious to performance will differ by task; the nature of the task itself will make measurement more appropriate in some contexts than in others. In sectors where outputs that can be measured easily, frequently, and quickly (e.g. the distribution of a vaccine) are tightly linked to desired outcomes (e.g. the acquisition of immunity), output measurement can be of great benefit in cutting through the complexity of process and ensure that aid achieves desired outcomes.

When the gap between the observable and thus contractible output and the desired outcome is greater, however (e.g. when focusing on governance reforms, or when seeking to improve a health system rather than build health clinics), a control regime that circumscribes agencies and agents’ zone of independent action (via either tighter explicit supervision or intense application of measurement) will fare less well. Agents will be able to use a variety of means to achieve their goals – that is, Navigation by Measurement will
allow Navigation by Judgment in service of the measures. However, as these measures will often be inappropriate (due to multitask, Goodhart’s Law, and flexibility considerations, among others) organizations that navigate in this manner will perform less well, with performance degrading the more inappropriate Navigation by Measurement. In environments with more state, effect, and/or response uncertainty a more autonomous IDO will fare better due to their ability to Navigate by Judgment. This argument is depicted graphically in Figure 2.3.

Figure 2.3: Hypothesized relationship between environmental uncertainty/unpredictability and project success for IDOs of differing levels of autonomy

The Potential for the Marginal Effect of Autonomy to Vary by Level of Autonomy

The marginal effect of autonomy need not be constant at all levels of autonomy; autonomy’s relationship to performance need not be linear, even as task and environment remain constant. The benefits of increased autonomy may outweigh the costs at low levels of current autonomy, even as the benefits of decreased autonomy outweigh the costs at
high levels of current autonomy. This means that any empirical examination of these phenomena will have its external validity constrained by the range of autonomy present in the universe of cases. The further the projected level of autonomy from the universe of cases, the less any given empirical study can serve as a basis for predicting what will occur if levels of autonomy are modified. This also means that a finding that more autonomy is beneficial is unlikely to augur in favor of entirely unconstrained organizations and agencies, just as a finding that autonomy is detrimental is unlikely to augur in favor of giving organizations and their agents no autonomy whatsoever.

My prior is that there are few agencies at, much less beyond, the optimal level of Navigation by Judgment; that even the least constrained IDOs still exercise too much Navigation from the Top. If this is correct, and the relationship between autonomy and performance is that of an inverted parabola, I expect all the existing data to come from the ‘left’ of the parabola’s inflection point. As such, I expect the relationship to be found by investigating the existing universe of cases to approximate a linear relationship – with more autonomy unambiguously positive – than is the case in the abstract.

2.5: **Empirical Strategy**

I will engage in two broad strategies to test the hypotheses articulated above: Examination of project outcome data and carefully chosen case studies of select
development projects. These strategies are chosen to be mutually reinforcing – the shortcomings of the quantitative strategy are, to the extent possible, addressed by the qualitative strategy (and vice versa).

Quantitative Strategy

In addition to the (largely output) measures against which project performance is evaluated in the short term, IDOs also engage in ex-post evaluation of projects, often substantially after the fact. These evaluations allow in many cases for direct observation of project goals. To this end I have compiled a database of over 14,000 unique projects in 178 countries over the past 50 years from nine donor agencies. The nine agencies are the EC, DFID, AsDB, GFATM, KfW, WB, JICA, GiZ, and IFAD. This dataset is unique in systematically including project performance data; issues relating to data collection and quality are discussed in detail in the chapter that utilizes these data.

15 Details of the interview protocol and methods employed in the qualitative investigation are described in Appendix I.

16 I thank the European Commission (EC), the UK’s Department for International Development (DFID), the Asian Development Bank (AsDB), the Global Fund for AIDS, Tuberculosis, and Malaria (GFATM), and KfW (the German Development Bank) for their provision of data. World Bank (WB) data used in this analysis are publicly available; Japanese International Cooperation Agency (JICA), GiZ (the German Society for International Cooperation), and International Fund for Agricultural Development (IFAD) data were assembled from individual project completion reports by Odesk-contracted research assistants under my supervision, with the data compiled in this manner then sent back to the originating agency for their comment and/or correction. GiZ was kind enough to respond with corrections, which were incorporated; JICA wished it to be made clear these data were generated by me rather than JICA, and JICA is not responsible for them. I am currently in discussions with potential archives regarding how best to institutionalize the maintenance and updating of these data in order to serve as a resource to other researchers and practitioners.
This econometric examination allows us to see broad trends; whether organizations with lower levels of autonomy fare better or worse in more uncertain environments or more unpredictable task domains. By measuring autonomy/constraint at the political authorizing, organizational, and agent levels we can also investigate the inter-correlation between these measures, suggestive of the linking of these levels.

**Qualitative Strategy**

The case studies provide an opportunity to go to a finer granular level in investigating the relationship between political authorizing environment, organizational autonomy, organizational navigation strategy (as seen in agent autonomy and Navigation by Judgment), and project success, including the ways in which autonomy seem to translate to success, hypothesized to be better planning, better day-to-day management, and greater ability to revise appropriately when needed. They also provide a much more direct test of the notion that organizational and agent constraints are interrelated; that agents will be constrained by organizational rules and measurement schema that are linked to organizational needs for legitimacy-seeking from political authorizing environments. It also allows for a much more direct examination of the role of mechanisms via which control, autonomy, and measurement interrelate. Do agents and organizations miss things as a result of control structures? Do staff feel compelled to meet metrics? Does Navigation from the Top alter agent behavior and agents’ ability to utilize (and thus incentive to collect) soft information? Does Navigation by Measurement allow autonomy over how to
meet measures? Does Navigation from the Top, including Navigation by Measurement, create inflexibilities that make adaptation more difficult?

In order to maximize variation on the critical independent variables I investigated a total of 8 case studies, with 2 IDOs in each of 2 countries and 2 case studies in domains of varying contractibility. To maximize IDO variation I chose one IDO with a relatively stable political authorizing environment, DFID (and thus, I would predict, lower levels of inappropriate measurement as well as other constraints, and thus higher levels of autonomy) and one IDO with a relatively unstable political authorizing environment, USAID (and thus, I would predict, higher levels of inappropriate measurement as well as other constraints, and thus lower levels of autonomy). I chose two countries, one relatively predictable and legible (South Africa) and one less so (Liberia). In each of these country-IDO pairs I chose two projects, attempting to ‘match’ them across IDOs to the extent possible; one with outputs that were relatively verifiable (and thus contractible) in addition to being observable (health), the other in a relatively difficult to verify sector (capacity building) where outputs were nonetheless observable – that is, where we would expect soft information to play a more important role. A schematic of these case studies is presented in Figure 2.4.
The cases selected also allow for some additional intra-IDO variation in structure; e.g. USAID’s Office of Transitional Initiatives (OTI), by far USAID’s most autonomous unit, managed the USAID capacity building project in Liberia. This allows an examination of whether unit autonomy can mitigate larger organizational effects. In addition, in South Africa USAID’s implementation of the President’s Emergency Plan for AIDS Relief (PEPFAR) is complemented by the US Center for Disease Control’s implementation of PEPFAR projects, allowing an examination of how two very different organizations (USAID and CDC) implement the same funding stream.
Table 2.3 highlights differences between USAID and DFID, providing contrast regarding the political authorizing environment (including financial and structural constraints) each IDO faces. The last column gives the rank of each agency on the organization-level autonomy measure employed in the econometric chapter; this ranking does not directly incorporate the items in the other columns, and thus is suggestive that these differences in political authorizing environment are associated with predicted differences in organizational autonomy. These differences will be explored in more detail in Chapters 4 and 5.

Table 2.3: Comparison of USAID and DFID’s Political Authorizing Environment

<table>
<thead>
<tr>
<th></th>
<th>Political Status of Aid Agency Head</th>
<th>Budget Security</th>
<th>Response to 2008 Financial Crisis</th>
<th>Workplace Satisfaction Surveys</th>
<th>Rank (out of 33) on Autonomy Measure Used in Chapter 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DFID</strong></td>
<td>Full ministerial rank, limited coordination with Foreign Affairs</td>
<td>Three-year budget allocations; few earmarks</td>
<td>Only ministry spared from across-the-board cuts; budget has continued to increase</td>
<td>Top 2%</td>
<td>3</td>
</tr>
<tr>
<td><strong>USAID</strong></td>
<td>Head of USAID (Administrator) reports to State Department</td>
<td>Yearly, often delayed; USAID budget heavily earmarked</td>
<td>Cutting aid funding promises literally the first thing mentioned by Obama ticket (as candidate)</td>
<td>Bottom third</td>
<td>29</td>
</tr>
</tbody>
</table>

Sources: 2012 US Federal Employee Viewpoint Survey Global Satisfaction Index (USAID 25th of 36); 2013 UK Civil Service People Survey Employee Engagement Index (DFID tied for 2nd of 98); Biden-Palin Debate, October 2 2008; Author
Empirical Strategy in Sum

The qualitative and quantitative strategies are designed to be mutually supportive. The quantitative data does not allow direct inter-organizational comparisons of interventional success; the qualitative strategy does just this. The qualitative cases, while chosen to maximize variation, could be outliers in unobservable ways; the broad coverage of the quantitative data does not suffer from the same concerns. If the two strategies lead to mutually supportive results in line with this chapter’s predictions, this will provide strong evidence for the underlying theory.

2.6: Generalizability

It is not just in the public sector or international development where autonomy and soft information matter, or where measurement may be used for legitimacy seeking reasons and thus constitute an inefficient means for autonomy. A senior official at a major financial organization recently estimated to me that no more than 10% of the units in his organization generated a measure (such as profit) that allowed management to directly contract on performance. (Interviews) Issues of legitimacy-seeking and isomorphic mimicry come from the private sector organizational behavior and management literature, which is also no stranger to the notion that executive boards my sub-optimally constrain organizations. While environmental uncertainty and task complexity are foundational independent variables in organizational theory, what empirical tests there have been are
usually partial and single-industry. International development is relatively unique both in the broad range of contexts and tasks undertaken by IDOs and the (relative) lack of selection out of those tasks and countries when performance may not meet desired standards.

These data and this investigation provide in many ways the cleanest test to date of some of the foundational theories in organizational behavior, including those of Perrow, Thompson, and Lawrence & Lorsch. (Lawrence and Lorsch 1967; Perrow 1967; Thompson 1967) In addition, in this age of multinational corporations and global expansion organizations increasingly need to adapt to unfamiliar and uncertain (in state, effect, and response uncertainty senses) contexts. The role of soft information/tacit knowledge in such adaptation and how best to avoid suffocating agents’ and thus organizations’ ability to acquire and thus incorporate same in managerial and policy decisions is critical to these questions of adaptation.

I intend for this work to problematize measurement, which has increased in our lives at work, at home, and at play. I hope for this work to contribute to a broader view of measurement as something like a screwdriver; the perfect tool for some problems (screwing in a screw), adequate for others when no better tools are available (hammering a nail), and deeply inappropriate – and in fact destructive if employed – for others (inflating a tire).
2.7: Conclusion

Control and autonomy have been substantially theorized in a variety of literatures, including those of organizational behavior, business strategy, public administration, and economics. These conceptualizations have differences, however, particularly as regards agent autonomy. There have also not been adequate empirics investigating the heterogeneous consequences, rather than determinants, of organizations’ propensity to Navigate by Judgment or Navigate from the Top (including Navigation by Measurement) in any of these literatures. IDOs provide a unique opportunity for study, with identical organizations to engage in a broad variety of tasks and environments. Perhaps the classic principal-agent model is correct, and agents need to be constrained and incentivized; perhaps the organizational behavior literature is correct, and whether this is true depends on features of the case, which this work can help elucidate.

A better, empirically grounded understanding of the topography that augurs for the relative success of an emphasis on one of these navigation techniques rather than the other is sorely lacking. Such an investigation has the potential to speak to scholars in providing a framework within which to unify the now somewhat isolated bits of theory on delegation, autonomy, and contracting in both the business and public sector literatures, framing the debate as one about determining the relevant features of the organization and context (among other domains) along which measurement or autonomy are likely to be most successful. It also has clear practical implications for organizations, managers, and political authorizers involved in foreign aid, the broader public sector, and beyond.
The chapters that follow attempt to move in this direction via systematic exploration of foreign aid organizations’ performance. The next chapter utilizes a novel database of foreign aid organization project performance covering over 14,000 unique projects across a variety of task domains and recipient country environments. It allows us to examine of the observable implication of this theory — whether autonomy and project performance have a contingent relationship, one mediated by task domain and recipient country environment.
CHAPTER 3: LARGE-N ANALYSIS OF PROJECT OUTCOMES

This chapter examines a key observable implication of the theory: that IDOs which are more autonomous are systematically more successful in more fragile states and harder to observe/contract on than their less autonomous peers. To this end it brings to bear an original cross-IDO dataset, incorporating over 14,000 unique development projects into what is now the world’s largest cross-organizational database to incorporate development outcomes. Because of the nature of the data, it is not possible to directly examine the mechanism, the channels, via which organizational autonomy affects project success; subsequent chapters will attempt to address this.

This chapter, rather, focuses on establishing the broad relationships hypothesized in chapter 2, exploring the returns to autonomy in differential environmental opacity and unpredictability and in differentially measurable task domains. Navigation by Judgment will be most useful when tasks are difficult to routinize and/or environments are relatively unpredictable and opaque. As less autonomous IDOs are unable to Navigate by Judgment, they should fare less well in situations where Navigation by Judgment would be more appropriate.

While the movement for aid information transparency has made impressive strides in the past few years, most of the progress to data has been on inputs – on spending data and financial flows. No other source (including the International Aid Transparency Initiative, the OECD Development Assistance Committee’s Creditor Reporting System, and the AidData archive) includes systematic information on the results of projects in a way tractable to quantitative analysis for any donor other than the World Bank, which also makes these data public and easily accessible from the Bank’s website (the only such donor to do so).
3.1: Data and Specifications

It would be ideal to have time-varying data on organizational autonomy for every IDO, including variation at the country (or even project) level. The data available only varies at the IDO level and is time-invariant. This work therefore cannot test directly for the effect of autonomy on success directly, as different IDOs have different measurement standards; a rating of 4 given by the German Development Bank (KfW) may or may not mean a project is more successful than one that received a rating of 3 from the International Fund for Agricultural Development. This work can, however, examine the differential performance of IDOs with varying levels of autonomy in interaction with other explanatory variables, thus leveraging the idea that a rating of 4 given by KfW means a project succeeded better than a project assigned a 3 by KfW.

I examine differential returns to autonomy in a dataset comprising over 14,000 unique projects in 178 countries carried out by nine IDOs over the past 50 years. The nine agencies are the European Commission (EC), the UK’s Department for International Development (DFID), the Asian Development Bank (AsDB), the Global Fund for AIDS, Tuberculosis, and Malaria (GFATM), the German Development Bank (KfW), the World Bank

---

18 This study’s focus on measurement at the organizational level is not intended to suggest there is not recipient and recipient-year variation in autonomy, only that this is the level at which measurement is most clean and broad. Controls below ensure that my results are not biased by these other levels of variation in autonomy.

(WB), the Japanese International Cooperation Agency (JICA), the German Society for International Cooperation (GiZ), and the International Fund for Agricultural Development (IFAD). This dataset is unique in systematically including project performance data, discussed in greater detail in footnote 17 above. In cases where results ratings were extracted from individual project documents I have either extracted this data myself or audited the extractions of research assistants. In cases where IDOs provided data in summary form, evaluation documents have been located where possible for a subset of projects to confirm the accuracy of the transmitted data.

It is also possible, of course, for the data to be accurate in the sense of correctly reflecting an organization's assessment, but for that assessment to bear little connection to the actual performance of the project. The reliability of these data and the econometric means of systematically testing it will be discussed below; however, to the extent possible, I have also attempted to validate these evaluations by returning to primary documentation. The World Bank archives uniquely allows access, following an extended vetting and declassification process, to primary project documents, including correspondence between

---

20 I thank the European Commission, the UK's Department for International Development, the Asian Development Bank, the Global Fund for AIDS, Tuberculosis, and Malaria, and the German Development Bank for providing data. World Bank data used in this analysis are publicly available. Data for the Japanese International Cooperation Agency (JICA), the German Society for International Cooperation (GiZ), and the International Fund for Agricultural Development were assembled from individual project completion reports by Odesk-contracted research assistants under my supervision, with the compiled data then sent back to the originating agency for comment and/or correction. GiZ was kind enough to respond with corrections, which were incorporated; JICA wished it to be made clear that these data were generated by me rather than by JICA and that it is not responsible for them. I am currently in discussions with potential archives regarding how best to institutionalize the maintenance and updating of these data as a resource for researchers and practitioners.
project staff and between World Bank staff and national governments, back-to-office reports and (often handwritten) notes by those monitoring projects, detailed financial and performance indicators, and the detailed evaluation reports that draw in part on these documents and which generate the outcome data for inclusion in this data set.

For a small handful of World Bank projects (approximately a dozen), I have reviewed archival documents at length, focusing on cases in which similar projects (such as the first and second phases of a particular project in a particular country) received quite different ratings and one might therefore be particularly doubtful about the reliability of those ratings. In reviewing the archival documents (which in every case occurred many months after identifying the projects to be reviewed), I intentionally proceeded without knowledge of which projects were more or less successful and attempted to generate my own rating from the primary documentation. I cannot say that my rating on a six-point scale always matched the World Bank Independent Evaluation Group’s score precisely; indeed, this would be troubling if true, since the Independent Evaluation Group also engages in conversations with project personnel, recipient government officials, and project beneficiaries, transcripts of which are not included in the archives. However, there were no cases in which my archivally generated rating differed by more than one point from the World Bank’s official six-point rating. In short, success and failure do seem to be different and do map onto real features of the projects, at least in this sample. Figure 3.1 shows the distribution of projects across countries.
Data Collection

There is no existing cross-IDO database of project outcome data. These data therefore had to be collected from each IDO in the sample individually. I approached every OECD bilateral aid agency in the top 10 in terms of the volume of official development assistance aid delivered directly (not via a multilateral agency) in 2010 (the last available data when this research commenced). This includes agencies in the US, Germany, the UK, France, Japan, Canada, Norway, Australia, Sweden, and Denmark. All of the biggest multilateral aid agencies (the European Commission, UN Development Programme, World Bank, African and Asian Development Banks, and Global Fund) were approached, as were other agencies with which I had links (for example, Irish Aid, International Fund for Agricultural Development, Food and Agriculture Organization, and International Monetary Fund).
There were two basic reasons to exclude an agency: either it did not collect project-level outcome data with a holistic project outcome rating (e.g., Canada, United States, Sweden, UNDP) or I could not get access to that data despite repeated attempts (e.g., African Development Bank).

*Project Success*

The key dependent variable in the analysis below is overall project success, a holistic rating undertaken by independent evaluators (either external evaluation contractors or independent evaluation units) or by project staff in project completion reports. For most IDOs, project success is an ordinal variable ranging from 1 to 6, with 6 being “Highly Satisfactory” and 1 being “Highly Unsatisfactory.” Some organizations evaluate projects on alternative scales (such as a four-point scale, with 4 being best); I transform all scales to be on a consistent six-point scale and employ IDO fixed effects in all models that use this six-point scale. I also employ a z-transformed version of this variable in the analysis when IDO fixed effects are absent. This process effectively de-means project success, just as employing IDO fixed effects would do.

The generation of z-scores and the use of IDO fixed effects helps to avoid spurious interpretations by putting each IDO’s project results on an identical parallel scale. Interpreting directly between IDOs (for example, determining which IDO is most

---

21 These are the World Bank’s designations. No IDO has significantly different names/standards in this regard, which would in any case be removed by IDO fixed effects.
successful) is not possible with these data, given that they are based on separate measurement frameworks used by different IDOs. This chapter limits itself to claims about comparative relative performance; e.g. the performance of more autonomous IDOs is less affected by environmental unpredictability than is that of their less autonomous peers.

The underlying construct employed by different IDOs for measuring the success of projects is relatively consistent, with an OECD-wide standard for bilateral IDOs. A given project’s rating is intended to incorporate a project’s relevance, effectiveness, efficiency, sustainability, and impact. (OECD 2000) Multilateral IDOs in the sample either use this standard explicitly or something closely related.\footnote{E.g. The World Bank focuses on impact, sustainability, and quality of preparation and implementation.}

**Autonomy**

Organizational autonomy is measured at the IDO level and is proxied in two ways: by a scale drawn from the Paris Declaration monitoring indicators and by a direct field survey of aid experts. These measures focus on organizational and field staff autonomy, as described above.

To build the autonomy scale I take five measures from Paris Declaration monitoring surveys, a mechanism designed to monitor the commitments made by parties (including IDOs) to this international agreement to improve aid quality and impact. The measures used are indicative of either an IDO’s propensity to devolve control over project
implementation to recipient countries or the degree of autonomy the agency itself has relative to its political authorizing environment. The first group includes indicators of the extent to which an organization values control (and is thus a proxy for the field-level autonomy of the staff): the use of recipient-country public financial management (PFM) systems; the use of recipient-country procurement systems; and the avoidance of parallel implementation units.\footnote{Parallel implementation units are separate units inside recipient countries that use donor standards and thus give donors more control/separation of funds or procurement.}

The second group includes indicators of the autonomy of the agency itself relative to its political authorizing environment, which, in turn, constrains the autonomy of the field staff. These indicators are, first, the degree to which aid is untied; that is, the extent to which it is not required that funds be spent on goods and services produced by the donor country. A high level of tying is a sign of an IDO’s need to build political consensus for aid by serving domestic political constituencies and thus reflects more insecure footing in the IDO’s political authorizing environment. The second is the predictability of the aid; that is, the extent to which ex-ante estimates of aid volume are proved accurate ex-post. Research suggests that variations are very donor-dependent and linked to IDO funding insecurity and political meddling in same (Celasun and Walliser 2008; Desai and Kharas 2010).

The two subscales are reasonably well correlated (.42) and principal components analysis yields a single component with relatively equal primary principal component
loading from each measure. The overall scale has a Cronbach’s alpha of .798.\textsuperscript{24} This provides some confidence that these measures and the two subscales map the same essential facts regarding IDOs and thus provides suggestive evidence for my conjecture that the two levels of autonomy measured here are linked, in that field-level autonomy is largely endogenous to an organizations’ relationship with its political authorizing environment. The results presented below are robust to dropping either subscale as well as to dropping any single measure. Figure 3.2 is a dendrogram with the scale’s component mapping, graphically demonstrating that scale components interrelate as predicted given the underlying theory. The scale used here is a time-invariant measure formed from the average of the three waves (2005, 2007, and 2010) of the Paris Declaration survey.\textsuperscript{25}

\textsuperscript{24} This is for the full autonomy scale with all IDOs; restricting the sample to IDOs with project outcome data, the Cronbach’s alpha is .742.

\textsuperscript{25} The autonomy scale is a simple average of the five measures except in the case of multilaterals (AsDB, WB, IFAD, EC), for which tied aid is not calculated; in these cases, the scale is an average of the remaining four measures. The three waves of Paris Declaration surveys (2005, 2007, 2010) are averaged here, in keeping with expert advice that these were effectively multiple mappings of the same facts, with insufficient time for organizations to change significantly between the first wave in 2005 and the last wave in 2010. Results are robust to using any wave and dropping any wave of the survey.
Given the critical role of measurement of autonomy to the empirical strategy, I attempted to validate the Paris Declaration scale with more direct measurement. I conducted a small-scale direct field survey of aid experts—individuals who have substantial development experience or whose jobs bring them into contact with a wide variety of donors. A typical role for one of these respondents would be a senior position in the aid management unit of a recipient government’s ministry of finance. Respondents rated a number of development agencies (including but not limited to those in the sample) on a scale of 1 to 7 in response to the following question:

---

26 The survey has a concentration of nationals and internationals with expertise in Liberia and South Africa (as these are case study countries for my related qualitative work). The survey N is limited by the small number of individuals in any given country who can make expert inter-donor comparisons (this generally excludes employees of development agencies, who can only speak intelligently regarding their own organization).
To what degree do you believe the in-country field office/bureau of the agencies listed below (presented in random order) are enabled to make decisions with a significant impact on the direction, nature, or quality of development projects? Please only respond for those agencies you have had exposure to either via working with the agencies or discussions with colleagues.

The survey N is 28, with varying coverage for different donors. This is a small but well-informed sample; methodological studies suggest small numbers of high-quality respondents will prove more accurate than significantly larger samples that lack expertise (Leuffen, Shikano, and Walter 2012). Moreover, this survey is well correlated with the Paris Declaration-based scale (.71), providing an additional level of confidence in the accuracy of the Paris Declaration-based measure.

Environment Unpredictability & Opacity

Environmental unpredictability and opacity is measured via the State Fragility Index (SFI) of the Polity IV/Integrated Network for Societal Conflict Research (Center for Systemic Peace 2012). This index incorporates security, governance, economic development, and social development measures and has two subscales: effectiveness and legitimacy. The two subscales are highly correlated (.66) and Cronbach’s alpha (.78)

---

27 This is the remaining N after removing surveys which were not substantively responsive or gave indications of nonsense answers; the two largest reasons for exclusion were (a) rating the Asian Development Bank despite stating that all relevant development-related work experience was in an African country (where the Asian Development Bank does not function) or (b) rating the survey’s anchoring vignettes such that the most autonomous text was evaluated as being just as autonomous or less autonomous than the least autonomous text.
suggests that they map the same underlying construct. While the analysis below looks at the aggregate SFI measure, results are robust to dropping either subscale.

More fragile contexts are inherently less predictable; predictability and fragility are often linked explicitly in development practice, with practitioners speaking about the difficult and unpredictable nature of fragile states (Ghani, Lockhart, and Carnahan 2005; Institute of Development Studies 2014; Weijer 2012). Fragility is in some sense the likelihood that the current equilibrium will break down or change rapidly, but makes no claim as to what positive state of the world will replace it. The 2011 World Bank World Development Report implicitly makes this link between state fragility and environmental unpredictability while also suggesting the link to Navigation by Measurement hypothesized in chapter 2, saying “Standard development measures... are excellent long-term goals and indicators, but they are not always helpful in fragile situations in the short term. These indicators move too slowly to give feedback on the speed and direction of progress.” (World Bank 2011, pgs. 209-210) In addition, more fragile states are much more likely to have changed rapidly in the past, be emerging from conflict, and to generally be environments less legible to external actors.

**Sector**

In order to determine project sectors for observability and contractibility, I use OECD Development Assistance Committee (DAC) sector and purpose codes, standard classifications that are usually assigned by the IDOs themselves in their databases/project

---

28In the sample data.
reports or their reports on aid flows to DAC. Even the more specific of these (the five-digit purpose codes) leave much to be desired. One can’t look, for example, at the delivery of antiretroviral drugs to HIV/AIDS patients specifically, as the relevant sector (“sexually transmitted disease control including HIV/AIDS”) includes such things as public awareness and social marketing campaigns, strengthening of countries’ HIV/AIDS response programs, and projects that focus on prevention in addition to treatment, as well as entirely unrelated STDs such as syphilis. One might wish to zero in on vaccine delivery, but this is under a code (“basic health care”) that also includes such things as nutrition services, support for nursing care, and strengthening of rural health systems. Thus, this chapter cannot systematically code sectors as observable or unobservable and will instead examine sectors (largely infrastructure) in which observability/contractibility is relatively clear and compare the results to those of related sectors that are less observable.

Addressing Potential Organizational Selection Out of Difficult Contexts/Sectors

In the original dataset employed below, two organizations—the Global Fund for Aids, Tuberculosis, and Malaria and the International Fund for Agricultural Development—work in particular sectors. Of the rest, all IDOs have projects in 10 of 16 of the broad sectors (Education, Health, and so on) coded in the data. Four broad sectors have participation from all but one IDO. Only the two smallest sectors, “Communications” and

---

29 In a small number (fewer than 5%) of cases, codes are assigned by me or by research assistants whom I supervised, based on the detailed contents of project reports.

30 By “broad sectors,” I mean the two-digit sectors of the DAC’s sectoral classification scheme, excluding here debt relief and humanitarian assistance.
“Business and Other Services”—accounting for only 3% (342 of 10,857) of the total projects for which sector codes are available—fail to have projects from two IDOs. We see, then, that donors are doing similar things across sectors. They are also doing them in the same countries. Excluding regional donors (in this sample, the Asian Development Bank and the Japanese International Cooperation Agency, which do most of their work in Asia), all the IDOs in this sample work in almost every developing country; the vast majority of projects in this sample occur in countries in which all the IDOs work.

Table 3.1 tackles this question in a more systematic way. After constructing a dataset with the number of observations from each IDO in each country in each sector, I replicate the main analysis (to be presented below) while using this observation data as the dependent variable. This allows us to see if there is any potential for selection bias in the results; if, for example, more autonomous IDOs systematically select into more (or less) fragile states or different sectors than do their less autonomous peers. The analysis finds no selection along the main dimension of inquiry (the interaction between autonomy and state fragility), which further suggests that IDO selection of sectors and/or countries is not a systematic problem for this analysis.
Table 3.1: Examining IDO Selection of Projects Along Dimension of Inquiry

<table>
<thead>
<tr>
<th></th>
<th>DV: # of observations by donor-country-sector</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>State Fragility Index (SFI)</td>
<td>0.0000166 (0.000807)</td>
<td>-0.000728 (0.00139)</td>
<td>0.0000166 (0.000807)</td>
<td>-0.000728 (0.00139)</td>
<td></td>
</tr>
<tr>
<td>Autonomy*SFI</td>
<td>0.000411 (0.00120)</td>
<td>0.000411 (0.00120)</td>
<td>0.000411 (0.00120)</td>
<td>0.000411 (0.00120)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.00461** (0.00101)</td>
<td>0.0117 (. )</td>
<td>-0.000685 (0.00257)</td>
<td>0.00393 (. )</td>
<td></td>
</tr>
<tr>
<td>IDO Fixed Effects</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Recipient Fixed Effects</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Sector Fixed Effects</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>$R^2$-Within</td>
<td>0.000 0.007 0.008</td>
<td>0.015</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2$-Between</td>
<td>0.090 0.089 0.089</td>
<td>0.089</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>957096 957096 957096</td>
<td>957096</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.1 demonstrates a remarkable lack of selection into and out of countries and sectors in response to realized organizational performance. This lack of selection provides a unique context for empirical examination of the links between organizational autonomy and performance across task and environment, as alluded to in Chapter 1. The empirical models employed below will still include controls for both sector and recipient-country fixed effects so as to test the robustness of the findings to considering only within-sector and within-recipient variation.

Summary Statistics of Key Variables

Table 3.2 presents summary statistics for the variables that form the core of the analysis.
Table 3.2: Summary Statistics for Key Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Project Success (6 pt scale)</td>
<td>14610</td>
<td>4.235</td>
<td>1.203</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Overall Project Success (z scores)</td>
<td>14610</td>
<td>0</td>
<td>1</td>
<td>-3.53</td>
<td>2.011</td>
</tr>
<tr>
<td>State Fragility Index</td>
<td>9546</td>
<td>12.486</td>
<td>4.996</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>Project Size (USD Millions)</td>
<td>9957</td>
<td>29.194</td>
<td>74.299</td>
<td>.004</td>
<td>4015</td>
</tr>
<tr>
<td>Autonomy (from Paris Declaration scale)</td>
<td>14961</td>
<td>.654</td>
<td>.058</td>
<td>.564</td>
<td>.79</td>
</tr>
<tr>
<td>Autonomy (from expert survey)</td>
<td>13389</td>
<td>3.96</td>
<td>.516</td>
<td>3</td>
<td>6</td>
</tr>
</tbody>
</table>

The coverage of the State Fragility Index, one of the key covariates, only begins in 1994, thus limiting the analysis to the nearly 10,000 projects of that time period. This also limits the mismatch between the periodicity of this data and the Paris Declaration monitoring surveys from which the autonomy scale is drawn, which were conducted from 2005-2011.

Any (constant) systematic differences amongst IDO evaluation criteria or measurement standards are addressed in two ways: by including IDO fixed effects in econometric models (generating results which leverage intra-IDO comparisons across projects) and by normalizing project ratings using IDO-specific z-scores where fixed effects are not employed.

3.2: Quantitative Results

This section lays out the primary findings then addresses potential econometric concerns. Findings below are from fitting OLS models onto six-point scales of project success. In some cases, IDOs do not use a six-point scale, instead using, for example, a four-point scale; for this analysis, all scales are standardized to a six-point measure. The model for project $i$ in recipient country $j$ implemented by IDO $k$ generalizes to

---

31 Style inspired by Faye and Niehaus (2012).
Project Success_{i,j,k} = \beta_1 \text{Environmental Unpredictability}_j + \beta_2 \text{Environmental Unpredictability}_j \times \text{Autonomy}_k + \beta_3 \text{Controls}_i + \text{Fixed Effects}_j + \text{Fixed Effects}_k + \varepsilon_t.

**Autonomy and Recipient Fragility**

Table 3.3 reports the core findings. As expected, there is a robust and statistically significant negative relationship between level of state fragility and project success; environmental unpredictability is associated with less successful projects. This relationship is mitigated by IDO autonomy. More autonomous organizations have less pronounced negative relationships between state fragility and project success. These relationships are robust to the inclusion of project size as a control variable (under the logic that agencies might place differential attention—or give systematically different success ratings—to projects of different sizes).

**Table 3.3: Main Results on Unpredictability with Recipient, Sector FE s, and Project Size**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DV: Project Success (6-pt scale)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State Fragility Index (SFI)</td>
<td>-0.186***</td>
<td>-0.185**</td>
<td>-0.159**</td>
<td>-0.156**</td>
<td>-0.116**</td>
<td>-0.117**</td>
</tr>
<tr>
<td>Autonomy*SFI</td>
<td>0.228**</td>
<td>0.227**</td>
<td>0.201**</td>
<td>0.197**</td>
<td>0.117*</td>
<td>0.118*</td>
</tr>
<tr>
<td>Project Size (USD Millions)</td>
<td>0.000690**</td>
<td>0.000625**</td>
<td>0.000829***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>4.729***</td>
<td>4.742***</td>
<td>5.050***</td>
<td>4.786***</td>
<td>6.065***</td>
<td>6.174***</td>
</tr>
</tbody>
</table>

IDO Fixed Effects
- Y
- N
- N

Recipient Fixed Effects
- Y
- N
- N

Sector Fixed Effects
- Y
- N
- N

$R^2$-Within
- 0.029
- 0.024
- 0.080
- 0.081
- 0.087
- 0.093

$R^2$-Between
- 0.048
- 0.086
- 0.062
- 0.101
- 0.277
- 0.513

Observations
- 9313
- 7248
- 9313
- 7248
- 7371
- 5447

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Models 3 and 4 in Table 3.3 incorporate recipient-country fixed effects, indicating that results are not being driven by the unique features of the heterogeneous distribution of each IDO’s projects across countries. Models 5 and 6 do the same for sector fixed effects, controlling for sectors at the most fine-grained level available, the 223 unique five-digit OECD Development Assistance Committee Creditor Reporting System (CRS) purpose sectors. Findings are robust when focusing on differences in state fragility within countries over time or within sectors. These results should provide confidence that selection into and out of countries and sectors is not driving either the results or the consistency of the results.

Figure 3.3: Returns to Autonomy in Countries of Differential Environmental Fragility
Figure 3.3 draws from Model 1 of Table 3 to graphically represent differential performance by autonomy, using the extremes on the autonomy scale in the sample. Given the lack of common evaluation standards across IDOs, one cannot interpret the results above as making any absolute claims regarding the superior or inferior performance of any IDO relative to any other.

Both high- and low-autonomy IDOs perform better in less fragile (and thus more predictable and legible) contexts than they do in more fragile environments. More autonomous IDOs perform better than their less autonomous peers in more fragile contexts relative to their own performance in other contexts. While an IDO with autonomy comparable to that of the EC sees a bit over six-tenths of a point (or approximately 10% of the six-point outcome scale) difference between its performance in a state like Turkey (SFI=7, or one standard deviation more stable than the mean) and its performance in a state like Rwanda (SFI=17, or one standard deviation below the mean), an IDO with autonomy comparable to that of DFID sees only about .03 of a point (or approximately .05% of the six-point outcome scale) in performance differential.

The models in Table 3.3 do not incorporate a base term for IDO autonomy; as a time-invariant measure at the IDO level, it is collinear to IDO fixed effects. As such this work makes no claims about the direct effect of autonomy on project success. The inclusion of IDO fixed effects precludes any bias in the interaction term that might otherwise result from a failure to include the base term in the interaction. Table 3.4 replicates Table 3.3, save that it incorporates the IDO autonomy base term and drops IDO fixed effects. While
we cannot learn much from the coefficient on the base term (as the IDO-specific z-score outcome measure precludes any direct comparison between IDOs), it is worth noting that without IDO fixed effects, it becomes much easier to interpret the $R^2$ terms; Model 1 of table 3.4 suggests that autonomy and state fragility (and their interaction) are jointly explaining a remarkably large share ($R^2$-between=.54) of the variance in differential normalized project success between IDOs.

Table 3.4: Main Results Including Base Autonomy Term with Recipient-Country, Sector Fixed Effects

<table>
<thead>
<tr>
<th>DV: Project Success (Z-score)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autonomy (PD Scale)</td>
<td>-1.859**</td>
<td>-2.295***</td>
<td>-1.892***</td>
<td>-2.184***</td>
<td>-0.331</td>
<td>-0.584</td>
</tr>
<tr>
<td></td>
<td>(0.664)</td>
<td>(0.493)</td>
<td>(0.403)</td>
<td>(0.294)</td>
<td>(0.675)</td>
<td>(0.714)</td>
</tr>
<tr>
<td>State Fragility Index (SFI)</td>
<td>-0.141***</td>
<td>-0.159***</td>
<td>-0.133***</td>
<td>-0.142***</td>
<td>-0.0924**</td>
<td>-0.103***</td>
</tr>
<tr>
<td></td>
<td>(0.0260)</td>
<td>(0.0260)</td>
<td>(0.0230)</td>
<td>(0.0183)</td>
<td>(0.0290)</td>
<td>(0.0310)</td>
</tr>
<tr>
<td>Autonomy*SFI</td>
<td>0.170***</td>
<td>0.194***</td>
<td>0.165***</td>
<td>0.177***</td>
<td>0.0934*</td>
<td>0.107*</td>
</tr>
<tr>
<td></td>
<td>(0.0392)</td>
<td>(0.0353)</td>
<td>(0.0343)</td>
<td>(0.0279)</td>
<td>(0.0452)</td>
<td>(0.0475)</td>
</tr>
<tr>
<td>Project Size (USD Millions)</td>
<td>0.000617***</td>
<td>0.000436*</td>
<td>0.000518*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000129)</td>
<td>(0.000209)</td>
<td>(0.000208)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>1.587***</td>
<td>1.892***</td>
<td>1.207**</td>
<td>1.401**</td>
<td>1.578</td>
<td>1.760</td>
</tr>
<tr>
<td></td>
<td>(0.433)</td>
<td>(0.361)</td>
<td>(0.454)</td>
<td>(0.436)</td>
<td>(1.018)</td>
<td>(1.026)</td>
</tr>
<tr>
<td>IDO Fixed Effects</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Recipient Fixed Effects</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Sector Fixed Effects</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>$R^2$-Within</td>
<td>0.026</td>
<td>0.022</td>
<td>0.078</td>
<td>0.070</td>
<td>0.088</td>
<td>0.095</td>
</tr>
<tr>
<td>$R^2$-Between</td>
<td>0.536</td>
<td>0.054</td>
<td>0.003</td>
<td>0.057</td>
<td>0.105</td>
<td>0.337</td>
</tr>
<tr>
<td>Observations</td>
<td>9313</td>
<td>7248</td>
<td>9313</td>
<td>7248</td>
<td>7371</td>
<td>5447</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.5 adds a series of fixed effects to the main findings. Inclusion of time fixed effects (either yearly or in five-year periods) does nothing to diminish the association between autonomy and recipient fragility. The result remains robust to including time*IDO fixed effects and time*recipient fixed effects. These results should allay any concerns that the primary results are driven by heterogeneous IDO project performance over time or by heterogeneous entry of IDOs into and out of recipient countries over time.
Table 3.5: Expanding Fixed Effects for Robustness\textsuperscript{32}

<table>
<thead>
<tr>
<th>DV: Project Success (6-pt scale)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>State Fragility Index (SFI)</td>
<td>-0.185***</td>
<td>-0.183***</td>
<td>-0.100*</td>
<td>-0.0934*</td>
</tr>
<tr>
<td></td>
<td>(0.0333)</td>
<td>(0.0347)</td>
<td>(0.0342)</td>
<td>(0.0359)</td>
</tr>
<tr>
<td>Autonomy*SFI</td>
<td>0.227**</td>
<td>0.222**</td>
<td>0.161**</td>
<td>0.151**</td>
</tr>
<tr>
<td></td>
<td>(0.0475)</td>
<td>(0.0499)</td>
<td>(0.0347)</td>
<td>(0.0369)</td>
</tr>
<tr>
<td>Constant</td>
<td>4.720***</td>
<td>4.294***</td>
<td>3.317***</td>
<td>2.947***</td>
</tr>
<tr>
<td></td>
<td>(0.0847)</td>
<td>(0.0379)</td>
<td>(0.437)</td>
<td>(0.459)</td>
</tr>
<tr>
<td>IDO Fixed Effects</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Year Fixed Effects</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Year*IDO Fixed Effects</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>5-yr ‘bin’ Fixed Effects</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Recipient Fixed Effects</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Recipient* 5-yr bin FEs</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>IDO*5-yr bin FEs</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>$R^2$-Within</td>
<td>0.031</td>
<td>0.048</td>
<td>0.145</td>
<td>0.149</td>
</tr>
<tr>
<td>$R^2$-Between</td>
<td>0.063</td>
<td>0.014</td>
<td>0.073</td>
<td>0.128</td>
</tr>
<tr>
<td>Observations</td>
<td>9313</td>
<td>9313</td>
<td>9313</td>
<td>9313</td>
</tr>
</tbody>
</table>

\textsuperscript{32}The inclusion of time*recipient effects necessitates using five-year periods rather than individual years; at approximately 180 recipients*30 years, this generates nearly 5000 dummy variables and thus would severely restrict degrees of freedom/analytic leverage, not to mention requiring advanced computing capacity to generate output. These models do not include project size (though all findings are robust to its inclusion), as missing data on project size leads to significantly smaller samples when it is included and project size is of little substantive significance to the relationship between the key independent variables and project success.

Disentangling Environmental Features

This analysis has employed the State Fragility Index to incorporate both the opacity and unpredictability of the environment. This combines two related arguments from chapter 2 – that more autonomous IDOs and Navigation by Judgment both allow interventions to be more flexible and to make better sense of opaque environments. While
these are related channels, and both speak to soft information and tacit knowledge, they are nonetheless distinct.

Tables 3.6 and 3.7 attempt to distinguish between these two mechanisms by examining environmental features which speak – albeit noisily – to different causal pathways between environment and success. Table 3.6 looks at political stability as evaluated by the World Bank’s World Governance Indicators (which includes propensity for violence and regime change) as well as the EM-DAT International Disasters database. These variables speak to the likely (political stability) and actual (disaster database) occurrence of rapid change that would require adaptation – that is, the flexibility channel. Table 3.7 examines features of the environment that would make it less legible but are less connected to unpredictability. These are the ethno-linguistic fractionalization (ELF) index and the World Bank’s IDA Resource Allocation Index measure on property rights and rule-based governance (PRRG), which assesses “the extent to which private economic activity is facilitated by an effective legal system and rule-based governance structure” (Teorell et al. 2013, p. 111). The PRRG measure speaks to rule-based governance in the private sector, a plausible proxy for the gap between de jure institutions and de facto reality and thus environmental legibility to external actors. More ethno-linguistic fragmentation implies a greater number of groups in the society who may have conflicting interests, and thus an environment harder to understand from afar.

All the measures in tables 3.6 and 3.7 move in the hypothesized directions, though not all have point estimates significant at the 95% level. Indeed, in Table 3.6 none of the
variables are significant at this threshold; in Table 3.7 both the PRRG and ELF interactions with autonomy are significant at this threshold when recipient fixed effects are included.

Table 3.6: Environmental Unpredictability & Autonomy

<table>
<thead>
<tr>
<th>DV: Project Success (6-pt scale)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WGI Political Stability</td>
<td>0.529</td>
<td>0.306</td>
<td>(0.294)</td>
<td>(0.269)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Autonomy*Political Stability</td>
<td>-0.641</td>
<td>-0.209</td>
<td>(0.403)</td>
<td>(0.375)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Disaster Deaths</td>
<td></td>
<td></td>
<td>-0.00000287</td>
<td>0.00000912</td>
<td>(0.00000771)</td>
<td>(0.00000563)</td>
</tr>
<tr>
<td>Autonomy*Disaster Deaths</td>
<td></td>
<td></td>
<td>0.00000716</td>
<td>-0.0000140</td>
<td>(0.0000112)</td>
<td>(0.00000827)</td>
</tr>
<tr>
<td>Disaster Damage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.61e-08</td>
<td>1.92e-09</td>
</tr>
<tr>
<td>Autonomy*Disaster Damage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(8.20e-09)</td>
<td>(1.06e-08)</td>
</tr>
<tr>
<td>Constant</td>
<td>4.402***</td>
<td>4.942***</td>
<td>4.233***</td>
<td>5.414***</td>
<td>4.226***</td>
<td>5.414***</td>
</tr>
<tr>
<td>IDO Fixed Effects</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Recipient Fixed Effects</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>$R^2$-Within</td>
<td>0.007</td>
<td>0.079</td>
<td>0.000</td>
<td>0.066</td>
<td>0.002</td>
<td>0.066</td>
</tr>
<tr>
<td>$R^2$-Between</td>
<td>0.048</td>
<td>0.019</td>
<td>0.006</td>
<td>0.027</td>
<td>0.035</td>
<td>0.028</td>
</tr>
<tr>
<td>Observations</td>
<td>7502</td>
<td>7502</td>
<td>14610</td>
<td>14610</td>
<td>14610</td>
<td>14610</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
These results provide stronger support for the opacity channel than that of unpredictability and adaptation to changing environments. That said, all these measures are rough proxies; suffice to say that all these measures moving in a direction consistent with the hypothesis supports the general thrust of these results.

Extensions

The theory articulated in chapter 2 argues that the gathering and incorporation of soft information and the use of tacit knowledge is one of the channels through which autonomy impacts project performance. We might expect, then, the returns to having an in-country office to be higher for more autonomous IDOs, who are thus better able to

34 The PRRG measure is reverse coded, so higher numbers are associated with lower levels of rule-based governance. This is for ease of interpretation so that higher numbers indicate more difficult environments, consistent with all other scales used in this chapter.
incorporate soft information into decisions. Table 3.8 provides suggestive support for this hypothesis using a subset of six IDOs for which country office data is available.\footnote{The six IDOs are the AsDB, DFID, IFAD, JICA, KfW, and GiZ. This data is quite messy, with it frequently difficult to determine when precisely in-country offices opened or closed. The analysis presented in Table 3.6 assumes that where opening or closing dates are unknown offices presently open always existed. This is surely inaccurate in many cases, and thus adds additional noise.}

<table>
<thead>
<tr>
<th>DV: Project Success (Z-score)</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>State Fragility Index (SFI)</td>
<td>-0.155***</td>
<td>-0.152***</td>
</tr>
<tr>
<td></td>
<td>(0.0151)</td>
<td>(0.0335)</td>
</tr>
<tr>
<td>autonomy*office</td>
<td>0.824**</td>
<td>1.114**</td>
</tr>
<tr>
<td></td>
<td>(0.261)</td>
<td>(0.350)</td>
</tr>
<tr>
<td>Autonomy*SFI</td>
<td>0.190***</td>
<td>0.194***</td>
</tr>
<tr>
<td></td>
<td>(0.0224)</td>
<td>(0.0468)</td>
</tr>
<tr>
<td>Autonomy (PD Scale)</td>
<td>-2.791***</td>
<td>-3.054***</td>
</tr>
<tr>
<td></td>
<td>(0.545)</td>
<td>(0.711)</td>
</tr>
<tr>
<td>office</td>
<td>-0.568**</td>
<td>-0.752**</td>
</tr>
<tr>
<td></td>
<td>(0.210)</td>
<td>(0.243)</td>
</tr>
<tr>
<td>Constant</td>
<td>2.236***</td>
<td>2.078***</td>
</tr>
<tr>
<td></td>
<td>(0.396)</td>
<td>(0.621)</td>
</tr>
</tbody>
</table>

IDO Fixed Effects          | N       | N       |
Recipient Fixed Effects     | N       | Y       |
$R^2$-Within                | 0.026   | 0.080   |
$R^2$-Between               | 0.657   | 0.008   |
Observations                | 7992    | 7992    |

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

While the net effect of office is somewhat ambiguous over the sample as a whole (a post-hoc F test finds office and the office-autonomy interaction are not jointly significant), the interaction between autonomy and the presence of an office suggests that there are indeed increasing returns to having an office present for more autonomous IDOs. Put
another way, having an office has no observed relationship with project success overall, but there is indication that when an agency is more autonomous having an in-country office does indeed lead to projects in that country performing better relative to that IDO’s projects in other countries where it does not have an office.

While far from ironclad, the increased returns to opening an office for more autonomous IDOs provides some suggestive evidence that soft information and the ability of an IDO to incorporate same may indeed be operative in generating the observed relationship between autonomy and project success.

A thread of recent scholarship has argued that IDO support stimulates isomorphic mimicry in recipient-country governments, with the result of de jure reform but little de facto progress and a divorcing of formal organizational form from function (Andrews 2011, 2013; Buntaine, Buch, and Parks 2013). One could interpret this finding as suggestive evidence that the same is true of the IDOs themselves; that while many IDOs open offices, it is only for the more autonomous IDOs that offices actually lead to improved project performance, presumably via better incorporation of soft information by properly placed field agents. If field agents are less autonomous, it is more difficult to translate the de jure organizational form of having an in-country office into something that contributes to de facto improvement in project performance.
Environmental unpredictability is not the only relevant factor in estimating the anticipated returns to soft information, and hence to autonomy. An anti-corruption program is very difficult to evaluate and measure and is therefore a context in which we should expect to see quite large returns to incorporating soft information; this is less true of power plant construction, where each part of the process can be easily defined and measured. An IDO attempting to build a power plant can simply contract on observable quantifiable metrics, incentivizing staff to deliver; this would mitigate the need for soft information and thus for autonomy. For such tasks, Navigation from the Top might indeed be the more effective strategy. Delivering dams and promoting democracy are very different tasks that may well call for different delivery mechanisms and levels of measurement relative to staff autonomy; that is, for a different optimal point on the Navigation from the Top — Navigation by Judgment continuum.

Being able to contract on outcomes does not necessarily mean an IDO will do so, which adds noise to any attempt to observe the relationship between task-domain observability and the role of soft information. Indeed, significant forces in the aid community—including the World Bank’s focus on Performance-Based Financing, the Center for Global Development-initiated push for Cash on Delivery, and, one might argue, much of the thrust of both the Gates Foundation and the US President’s Emergency Plan for AIDS Relief—have argued that IDOs insufficiently contract on outcomes when they can and ought do so. Bill Gates, for example, has highlighted the importance of measuring vaccine transmission and coverage rates rather than simply sending out health personnel to
conduct vaccine drives (Gates 2013). If IDOs do not, in fact, manage based on observable outcomes when they can—perhaps focusing instead on input-based metrics—it is more ambiguous how we might expect the relationship between autonomy and project success to vary across the observability of task domain.

The messiness of foreign aid sector classifications further complicates this picture, as discussed above; sectors commonly include both the observable (such as antiretroviral drug delivery) and the less observable (such as public awareness and social outreach campaigns around HIV) in the same sector. The sectors are most straightforward with regard to tangible infrastructure, which is relatively externally observable and contractible. Road and power line construction are clearly task domains for which audits and performance incentives can work and for which we can use the first best solution of contracting on outcomes.

Tables 3.9 and 3.10 therefore focus (in Table 3.9) on purpose codes related to infrastructure construction or observable service delivery for which we might not expect to see as strong a relationship between autonomy and outcome and (in Table 3.10) on purpose codes which focus on related policy or administration tasks but are more difficult to observe. Focusing on related but difficult-to-observe domains helps to ensure that the

36 It is worth noting that in the same document Gates seems to implicitly endorse this work’s view that measurement is only contingently beneficial, depending on the nature of the measure and task. He says, for example, “You can achieve amazing progress if you set a clear goal and find a measure that will drive progress toward that goal” (p.1), which seems to imply that a well-aligned measure is a necessary condition for measurement to be beneficial.
results are not driven by differential ease of performance in certain broad sectors than others.

Table 3.9: Relationship between Autonomy and State Fragility by Sector (Outcomes Easily Observed; Sector by CRS Code)

<table>
<thead>
<tr>
<th>DV: Project Success (6-pt scale)</th>
<th>(1) Road Infrastructure &amp; Transport</th>
<th>(2) Building Power Transmission Lines</th>
<th>(3) Agricultural &amp; Irrigation &amp; Water</th>
<th>(4) Basic Drinking Water Supply &amp; Sanitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>State Fragility Index (SFI)</td>
<td>-0.262</td>
<td>0.586*</td>
<td>-0.516</td>
<td>-0.298</td>
</tr>
<tr>
<td></td>
<td>(0.352)</td>
<td>(0.128)</td>
<td>(0.343)</td>
<td>(0.152)</td>
</tr>
<tr>
<td>Autonomy*SFI</td>
<td>0.356</td>
<td>-0.958*</td>
<td>0.735</td>
<td>0.386</td>
</tr>
<tr>
<td></td>
<td>(0.561)</td>
<td>(0.201)</td>
<td>(0.536)</td>
<td>(0.233)</td>
</tr>
<tr>
<td>Constant</td>
<td>5.010***</td>
<td>5.120***</td>
<td>4.588***</td>
<td>4.621***</td>
</tr>
<tr>
<td></td>
<td>(0.161)</td>
<td>(0.0796)</td>
<td>(0.152)</td>
<td>(0.0486)</td>
</tr>
</tbody>
</table>

IDO Fixed Effects

<table>
<thead>
<tr>
<th>R²-Within</th>
<th>Y</th>
<th>Y</th>
<th>Y</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>R²-Between</td>
<td>0.030</td>
<td>0.031</td>
<td>0.024</td>
<td>0.054</td>
</tr>
<tr>
<td>Observations</td>
<td>469</td>
<td>167</td>
<td>165</td>
<td>271</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

* p < 0.05, ** p < 0.01, *** p < 0.001

Table 3.10: Relationship between Autonomy and State Fragility by Sector (Outcomes Difficult to Observe; Sector by CRS Code)

<table>
<thead>
<tr>
<th>DV: Project Success (6-pt scale)</th>
<th>(1) Transportation Management</th>
<th>(2) Agricultural Policy &amp; Administration</th>
<th>(3) Social/Welfare Services (Administration, Capacity Building)</th>
<th>(4) All Administration/Policy Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>State Fragility Index (SFI)</td>
<td>-1.030***</td>
<td>-0.670***</td>
<td>-0.371***</td>
<td>-0.151***</td>
</tr>
<tr>
<td></td>
<td>(0.0271)</td>
<td>(0.123)</td>
<td>(0.0178)</td>
<td>(0.0125)</td>
</tr>
<tr>
<td>Autonomy*SFI</td>
<td>1.716***</td>
<td>0.928**</td>
<td>0.561***</td>
<td>0.192***</td>
</tr>
<tr>
<td></td>
<td>(0.0407)</td>
<td>(0.182)</td>
<td>(0.0305)</td>
<td>(0.0195)</td>
</tr>
<tr>
<td>Constant</td>
<td>2.978***</td>
<td>4.587***</td>
<td>4.508***</td>
<td>4.554***</td>
</tr>
<tr>
<td></td>
<td>(0.0266)</td>
<td>(0.246)</td>
<td>(0.0288)</td>
<td>(0.0210)</td>
</tr>
</tbody>
</table>

IDO Fixed Effects

<table>
<thead>
<tr>
<th>R²-Within</th>
<th>Y</th>
<th>Y</th>
<th>Y</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>R²-Between</td>
<td>0.234</td>
<td>0.077</td>
<td>0.025</td>
<td>0.019</td>
</tr>
<tr>
<td>Observations</td>
<td>39</td>
<td>55</td>
<td>160</td>
<td>1530</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

* p < 0.05, ** p < 0.01, *** p < 0.001

There is no relationship between autonomy and project success in the first set of task domains, where the focus is on constructing something or delivering a tangible and relatively easily monitorable service, but the relationship is relatively strong in related administrative sectors. These results are consistent with my contention that task domain mediates the relationship between project success and environmental unpredictability.
One might worry that these results are driven by idiosyncratic features of the distribution of donors across project domains or by some nonsystematic mechanism other than sector observability. To alleviate this concern, Table 3.11 creates dummy variables for those sectors described as observable/unobservable in the tables above and considers them in the context of the data as a whole. The results confirm those in Tables 3.7 and 3.8. Looking at the triple-interactions of the observable/unobservable dummies with autonomy and state fragility, there is no consistent relationship between the interaction of state fragility and autonomy in the relatively observable sectors. That relationship is, however, present in the related sectors where outcomes are more difficult to observe.
While there is no evidence that Navigation from the Top is the better organizational strategy in more observable task domains, neither is there evidence that Navigation by Judgment is better. This provides further evidence that considering the effects of measurement is critical in determining where measurement is likely to have a negative  

Table 3.11: Sector Observability in the Full Model\textsuperscript{37}

<table>
<thead>
<tr>
<th>DV: Project Success (6pt scale)</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observable<em>Autonomy</em>sfi</td>
<td>0.0721</td>
<td>0.0938</td>
</tr>
<tr>
<td></td>
<td>(0.116)</td>
<td>(0.119)</td>
</tr>
<tr>
<td>Unobservable<em>Autonomy</em>sfi</td>
<td>0.0357***</td>
<td>0.0325**</td>
</tr>
<tr>
<td></td>
<td>(0.00654)</td>
<td>(0.00820)</td>
</tr>
<tr>
<td>State Fragility Index (SFI)</td>
<td>-0.189***</td>
<td>-0.162**</td>
</tr>
<tr>
<td></td>
<td>(0.0341)</td>
<td>(0.0373)</td>
</tr>
<tr>
<td>Autonomy*SFI</td>
<td>0.227**</td>
<td>0.202**</td>
</tr>
<tr>
<td></td>
<td>(0.0482)</td>
<td>(0.0545)</td>
</tr>
<tr>
<td>Observable*sfi</td>
<td>-0.0430</td>
<td>-0.0498</td>
</tr>
<tr>
<td></td>
<td>(0.0727)</td>
<td>(0.0763)</td>
</tr>
<tr>
<td>Unobservable*sfi</td>
<td>-0.00817</td>
<td>-0.00669</td>
</tr>
<tr>
<td></td>
<td>(0.00858)</td>
<td>(0.00696)</td>
</tr>
<tr>
<td>Observable*Autonomy</td>
<td>-0.0117</td>
<td>-0.159*</td>
</tr>
<tr>
<td></td>
<td>(0.104)</td>
<td>(0.0642)</td>
</tr>
<tr>
<td>Unobservable*Autonomy</td>
<td>-0.198</td>
<td>-0.167*</td>
</tr>
<tr>
<td></td>
<td>(0.0946)</td>
<td>(0.0722)</td>
</tr>
<tr>
<td>Constant</td>
<td>4.753***</td>
<td>5.056***</td>
</tr>
<tr>
<td></td>
<td>(0.0485)</td>
<td>(0.0303)</td>
</tr>
</tbody>
</table>

IDO Fixed Effects | Y | Y  
Recipient Fixed Effects | N | Y 
$R^2$-Within | 0.030 | 0.082 
$R^2$-Between | 0.040 | 0.050 
Observations | 9313 | 9313 

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

\textsuperscript{37} Only the sectors used in Tables 4 and 5 are coded as observable and unobservable in this analysis.
effect on project success—that is, in harder-to-observe task domains—and where its
effects are likely to be more ambiguous and potentially beneficial. Soft information and
tacit knowledge seem to matter to development success, with more autonomous agencies
thus better able to manage more unpredictable contexts and task domains less tractable to
Navigation from the Top. This suggests that autonomy can have positive effects inasmuch
as it provides support for the acquisition and use of soft information and tacit knowledge.

**Robustness**

One might be concerned that the autonomy measure is not actually mapping autonomy. As noted in the data description, I conducted a small survey of aid experts in the
field who come into contact with a wide range of IDOs (largely as consultants or as employees of developing country governments) and thus can make expert inter-IDO assessments. The correlation between this survey measure and the autonomy scale drawn from the Paris Declaration surveys is .71. Table 3.12 substitutes the survey measure of autonomy for that of the Paris Declaration-based measure; the results are similar, which should increase confidence in the Paris Declaration-based autonomy scale.
Table 3.12: Robustness of Results to Use of Survey Autonomy Measure

<table>
<thead>
<tr>
<th>DV:</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>State Fragility Index (SFI)</td>
<td>-0.101***</td>
<td>-0.0845***</td>
<td>-0.0750***</td>
<td>-0.0717***</td>
</tr>
<tr>
<td></td>
<td>(0.0170)</td>
<td>(0.0144)</td>
<td>(0.0205)</td>
<td>(0.0173)</td>
</tr>
<tr>
<td>SFI*Autonomy (Survey)</td>
<td>0.0167***</td>
<td>0.0144***</td>
<td>0.0121**</td>
<td>0.0117**</td>
</tr>
<tr>
<td></td>
<td>(0.00417)</td>
<td>(0.00352)</td>
<td>(0.00467)</td>
<td>(0.00401)</td>
</tr>
<tr>
<td>Autonomy (Survey)</td>
<td>-0.140**</td>
<td>-0.130*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0478)</td>
<td>(0.0536)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>4.743***</td>
<td>0.882***</td>
<td>5.109***</td>
<td>0.501</td>
</tr>
<tr>
<td></td>
<td>(0.0326)</td>
<td>(0.193)</td>
<td>(1.076)</td>
<td>(0.263)</td>
</tr>
</tbody>
</table>

IDO Fixed Effects           | Y | N | Y | N |
Recipient Fixed Effects      | N | N | Y | Y |
$R^2$-Within                 | 0.025 | 0.022 | 0.076 | 0.073 |
$R^2$-Between                | 0.129 | 0.449 | 0.228 | 0.213 |
Observations                 | 8314 | 8314 | 8314 | 8314 |

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

One might also worry, particularly given the small number of IDOs in this multilevel model, whether results are driven by features of the modeling. To address this concern, Table 3.13 examines the relationship between autonomy and project success nonparametrically, summarizing the relationship between state fragility and project success for each donor in isolation; that is, using only data from one donor at a time and implementing nine different regressions.\(^{38}\) In each case, the model is of the form

\[
\text{Project Success}_{i,j} = \beta_1 \times \text{State Fragility Index}_j + \varepsilon_i.
\]

IDOs are listed in order of ascending autonomy for ease of interpretation.

\(^{38}\) This is intuitively similar to a rank-based regression.
As expected, greater state fragility has a more negative and statistically significant relationship with project success for less autonomous donors. This confirms—using an approach that does not rely on the parameterization of the interaction term—that higher levels of autonomy mitigate the inverse relationship between the State Fragility Index and project success. These results are robust to dropping any individual IDO, and thus are not driven by a single outlier, a reasonable concern given the modest 2nd-level N (with 9 IDOs

39 The Paris Declaration monitoring survey does not differentiate between institutions from a single country; thus GiZ and KfW (both arms of the German government) have the same autonomy score.
in the sample). A graphical version of the data underlying table 3.11 is included as Figure 3.4.

![Figure 3.4: Graph of IDO-by-IDO slopes](image)

One might be worried that results are driven by quirks in the variance of outcomes. Table 3.14 examines this concern in a simple nonparametric manner, by dividing state fragility and autonomy scores at their respective means and then examining the variance in project success z-score by autonomy and state fragility quadrant, and finds no substantively large differences. Table 3.14 also shows another nonparametric way of testing the intuition underlying the core findings. Both low- and high-autonomy IDOs do better in contexts of lower state fragility; the gap between low- and high-SFI contexts is larger for low-autonomy IDOs (approximately .26 SD) than for high-autonomy IDOs (.1 SD).

---

40 This graph has only 8 points as GiZ and KfW have the same autonomy score, as noted in the preceding footnote. Credit to Chris Kilby, who as a discussant at NEUDC 2014 first generated this graph.
This should give further confidence that the main results are not driven by idiosyncratic features of the modeling.

<table>
<thead>
<tr>
<th></th>
<th>Low Autonomy</th>
<th>High Autonomy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low SFI</td>
<td>.181 (.882)</td>
<td>.076 (.966)</td>
</tr>
<tr>
<td>High SFI</td>
<td>-.080 (1.038)</td>
<td>-.039 (1.014)</td>
</tr>
</tbody>
</table>

Figure 3.5, combining the insights of Table 3.13 with that of 3.14, demonstrates that the variance also behaves in a consistent manner; all organizations see higher variance in more unpredictable (high SFI as compared to low SFI) environments. However, high autonomy IDOs see a slightly smaller gap in variances than do low autonomy IDOs, suggestive of their better ability to deal with more unpredictable environments.

---

41 By calculation, the Z-score outcome measure has mean 0 and standard deviation 1 for each donor. The table is designed primarily to examine if the variance in this measure differs systematically along SFI and Autonomy axes, thus potentially distorting the interpretation of OLS results. The question, then, is whether any of the quadrants deviate substantially enough from 1 to cause concern. Given the large N, the analysis can of course confirm that that these variances are not equal (even the difference between 1.038 and 1.014 is significant at the 95% level); the question is whether they are substantively different enough to potentially bias results. I would argue the answer to this is in the negative.
Evaluation Data Reliability

Another concern might be that these data rely on evaluations of project success made by the agencies themselves. One might worry that an agency with a fragile relationship with its political authorizing environment would, in addition to being less autonomous, have a greater incentive to self-evaluate projects to have been successes. Anecdotally, interviews suggest that such behavior occurs in at least some cases. Alternately, one might think that career-concerned agents have every incentive to evaluate their own projects as successes.

Either of these dynamics would reduce the variation in the outcome measure—the differential performance between agencies and within agencies across context, domain,
and time. Either of these dynamics would therefore reduce the likelihood of Type I error (false positives) while increasing the likelihood of Type II error (false negatives) and thus ought not to diminish our confidence in the principal findings.

The involvement of independent evaluation units also mitigates against this type of dynamic. In cases where projects are evaluated by implementation staff, the frequent rotation of IDO staff also means that it is by no means certain that the staff involved in project evaluation would see their careers best served by positive evaluations. In any case, Table 3.15 controls for the type of evaluation; that is, whether the data source is an internal review by project staff, a review conducted by an IDO’s own independent evaluation unit, or a review conducted by an externally contracted evaluator.
Interestingly, Table 3.15 suggests that none of these particular types of evaluator evaluates projects systematically differently than any other. The relationship between autonomy and state fragility remains unchanged, giving some comfort that evaluation bias is not driving the results. It is also not the case that different IDOs evaluate projects at different times. As Table 3.16 indicates, while projects in more fragile states see more time between completion and evaluation, this is not differential by IDO autonomy or the interaction of autonomy and state fragility.
Table 3.16: Examining Lag Time to Evaluation

<table>
<thead>
<tr>
<th>DV: Lag time to eval (days)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>State Fragility Index (SFI)</td>
<td>4.381</td>
<td>4.381</td>
<td>-47.84</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.539)</td>
<td>(2.540)</td>
<td>(45.07)</td>
<td></td>
</tr>
<tr>
<td>Autonomy (PD Scale)</td>
<td></td>
<td>-1561.4</td>
<td>-808.2</td>
<td>-1793.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3187.6)</td>
<td>(2748.3)</td>
<td>(3239.6)</td>
</tr>
<tr>
<td>Autonomy*SFI</td>
<td></td>
<td>83.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(70.22)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>702.3***</td>
<td>1771.8</td>
<td>1224.2</td>
<td>1843.5</td>
</tr>
<tr>
<td></td>
<td>(161.5)</td>
<td>(2050.9)</td>
<td>(1805.9)</td>
<td>(2123.3)</td>
</tr>
</tbody>
</table>

IDO Fixed Effects
- N N N N
- $R^2$-Within: 0.000 0.000 0.001 0.015
- $R^2$-Between: 0.025 0.024 0.013 0.010
- Observations: 7745 12657 7745 7745

Standard errors in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

Placebo Tests

One might be concerned that, despite the survey of aid experts, what this paper calls autonomy is in fact mapping a more general construct of good donor practice. If this were the case, the results might provide reassurance that the consensus wisdom on what constitutes good development—articulated, in part, by the very Paris Declaration from whose monitoring surveys the autonomy measure employed above is constructed—is on point. These results would not, however, suggest that organizational autonomy is an important factor, nor necessarily that soft information and tacit knowledge are critical in aid delivery.

To address this I run a series of placebo tests, examining whether other measures of good donor conduct yield the same relationship with the data observed for the autonomy measure. Table 3.17 gives summary statistics on two alternate scales that aim to measure...
and compare IDOs’ practices: the Commitment to Development Index (CDI) and the Quality of Official Development Assistance (QuODA) (Birdsall and Kharas 2010).\footnote{The CDI is an annual product of the Center for Global Development; the QuODA is an occasional product of the Brookings Institution in collaboration with the Center for Global Development (the last wave was in 2009). The CDI has a number of components (Aid, Investment, Migration, Environment, Security, and Technology) that assess the commitment of nations (multilateral organizations such as the World Bank are not included) to assisting the developing world. The QuODA has four components: Maximizing Efficiency, Transparency and Learning, Reducing Burden, and Fostering Institutions. All components of both the CDI and the QuODA involve a variety of submeasures. CDI is available here; QuODA is available here.} In both cases, I also look at the subscales that seem most relevant—CDI’s Aid component and QuODA’s Maximizing Efficiency and Fostering Institutions subscales. There is some overlap between these measures and my autonomy scale (which is repeated below for ease of reference). The CDI aid index penalizes tied aid (a component of the autonomy scale); untied aid is also a component of QuODA’s Maximizing Efficiency measure. QuODA’s Fostering Institutions component draws from the Paris Declaration monitoring surveys as well, incorporating avoidance of project implementation units and use of recipient-country systems.\footnote{This last measure combines the procurement and public financial management measures I use in the autonomy scale.}

Table 3.17: Summary Statistics for Alternate Scales

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autonomy scale (from Paris Declaration Surveys)</td>
<td>14961</td>
<td>.654</td>
<td>.058</td>
<td>.564</td>
<td>.79</td>
</tr>
<tr>
<td>Commitment to Development Index (CDI) 2012 Overall</td>
<td>4999</td>
<td>5.226</td>
<td>.763</td>
<td>3.4</td>
<td>5.7</td>
</tr>
<tr>
<td>Commitment to Development Index (CDI) 2012 Aid</td>
<td>4999</td>
<td>4.679</td>
<td>1.839</td>
<td>1.6</td>
<td>6.8</td>
</tr>
<tr>
<td>Quality of Development Assistance (QuODA) 2009 Overall</td>
<td>14831</td>
<td>.528</td>
<td>.138</td>
<td>.043</td>
<td>.655</td>
</tr>
<tr>
<td>Quality of Development Assistance (QuODA) 2009 Maximizing Efficiency</td>
<td>14831</td>
<td>.154</td>
<td>.268</td>
<td>-.89</td>
<td>.51</td>
</tr>
<tr>
<td>Quality of Development Assistance (QuODA) 2009 Fostering Institutions</td>
<td>14831</td>
<td>.39</td>
<td>.279</td>
<td>-.71</td>
<td>.93</td>
</tr>
</tbody>
</table>

Table 3.18 re-runs the primary model employed above (Table 3.3, Model 1), substituting each of these measures in turn for the autonomy scale; scales are standardized to allow for direct comparison across scales.
None of the other measures have anywhere near the strength of association of the autonomy scale. In interaction with state fragility, a better QuODA overall score and a better score on QuODA’s Maximizing Efficiency subscale moves in the opposite direction as that of autonomy, with higher scores associated with a stronger relationship between state fragility and evaluated project success. The Maximizing Efficiency subscale contains measures such as the ratio of project administrative costs to total project costs, which one could think of as a sign of Navigation from the Top rather than Navigation by Judgment, as Navigation from the Top requires a greater (and possibly more costly) investment in monitoring; it is therefore not entirely surprising that this scale moves in the opposite direction, with higher scores on Maximizing Efficiency associated with greater declines in performance in more fragile states. QuODA’s Fostering Institutions measure and CDI’s Aid measure—the two measures below whose indicators most overlap with those of the Paris Declaration-based autonomy scale—move in the same direction as autonomy but with very small point estimates which are not statistically significant.

### Table 3.18: Relationship between Project Success and (Normalized) Alternative Scales in Interaction with State Fragility

<table>
<thead>
<tr>
<th>Scale in Column Title</th>
<th>Autonomy</th>
<th>CDI Overall</th>
<th>CDI Aid</th>
<th>Quoda Overall</th>
<th>Quoda Max Eff</th>
<th>Quoda Foster Inst</th>
</tr>
</thead>
<tbody>
<tr>
<td>State Fragility Index (SFI)</td>
<td>-0.186***</td>
<td>-0.0167</td>
<td>-0.208*</td>
<td>-0.0357***</td>
<td>-0.0379***</td>
<td>-0.0343***</td>
</tr>
<tr>
<td>Scale in Column Title</td>
<td>0.228**</td>
<td>0.00788</td>
<td>0.0143</td>
<td>-0.0110</td>
<td>-0.0111</td>
<td>0.0121</td>
</tr>
</tbody>
</table>

| IDO Fixed Effects | Y | Y | Y | Y | Y | Y |
| Recipient Fixed Effects | N | N | N | N | N | N |
| $R^2$-within | 0.03 | 0.01 | 0.00 | 0.02 | 0.02 | 0.02 |
| $R^2$-between | 0.06 | 0.03 | 0.03 | 0.03 | 0.14 | 0.17 |

| Observations | 9313 | 3627 | 3627 | 9205 | 9205 | 9205 |

* Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
This should give reassurance regarding the uniqueness of the autonomy measure’s relationship with project success in conditions of differential environmental predictability, and thus the importance of soft information in the development production process.

In addition to the robustness checks discussed here, the results above are robust to:

- Using ordered logit models on six point project outcome scales (rather than OLS)
- Using z-scores as outcomes (rather than the six-point scale where employed)
- Compressing success and failure to a binary outcome and employing logit models
- Employing quantile regression (were results to be driven by only part of the distribution)
- Restricting SFI to common support; that is, only the range of SFI realized in all donors’ data (2-22, rather than 0-25 in the main analysis)
- Dropping the latter two waves of the Paris Declaration survey in generating the autonomy measure (to allay concerns that donors responded to measurement by changing their practices)
- Dropping any individual IDO from the sample
- Double-clustering standard errors at the IDO-recipient level (rather than clustering on IDO alone)\(^44\)
- Dropping either subscale of state fragility (legitimacy or effectiveness)

\(^{44}\) Double-clustering is achieved via Cameron, Gelbach, and Miller’s \textit{cgmreg} (2006).
• Using any of the four domains of state fragility (security, political, economic, or social)

Sample Selection and Generalizability

As described in the data section, inclusion in the sample requires a willingness to make results public or to disclose them to me. It also requires that an agency actually collect a holistic project-level indicator of success, which not all of them do. (USAID and UNDP could not be included because they do not generate such an indicator.) This makes the IDOs included in this analysis a convenience sample and thus raises concerns regarding broader generalizability.

To the extent that both organizational measurement decisions and, particularly, the willingness to make data public are plausibly correlated to an agency's autonomy—and Table 3.19, which lists agencies by the autonomy measure employed in this chapter, suggests they may well be, with the IDOs included all towards the top of the autonomy scale. This certainly is a threat to generalizability that must be considered in examining these quantitative results in isolation. That said, the most straightforward effect of this type of sample selection would be to bias the sample away from the least autonomous agencies, who would have the least stable relationships with their political authorizing environments and as a result might be least likely to make information public. If the theory articulated
here is accurate, this would most likely make findings in favor of the mooted hypotheses less likely by underpowering the analysis.\textsuperscript{45}

\textsuperscript{45} if and when additional IDO data becomes available rerunning this analysis incorporating additional IDO data would provide insight on whether this supposition is accurate.
Table 3.19: Full List of Countries’ Autonomy Score (in-sample donors’ scores/ranks in **BOLD**)

<table>
<thead>
<tr>
<th>Donor</th>
<th>Autonomy Score</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ireland</td>
<td>0.8499132</td>
<td>1</td>
</tr>
<tr>
<td>Norway</td>
<td>0.7996858</td>
<td>2</td>
</tr>
<tr>
<td><strong>DFID</strong></td>
<td><strong>0.7901573</strong></td>
<td><strong>3</strong></td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.7806054</td>
<td>4</td>
</tr>
<tr>
<td><strong>IFAD</strong></td>
<td><strong>0.7206322</strong></td>
<td><strong>5</strong></td>
</tr>
<tr>
<td>Sweden</td>
<td>0.7131852</td>
<td>6</td>
</tr>
<tr>
<td>IMF</td>
<td>0.7058333</td>
<td>7</td>
</tr>
<tr>
<td>Finland</td>
<td>0.7003072</td>
<td>8</td>
</tr>
<tr>
<td>Denmark</td>
<td>0.6982759</td>
<td>9</td>
</tr>
<tr>
<td><strong>KIW/GIZ</strong></td>
<td><strong>0.6742818</strong></td>
<td><strong>10</strong></td>
</tr>
<tr>
<td>Canada</td>
<td>0.6672894</td>
<td>11</td>
</tr>
<tr>
<td><strong>JICA</strong></td>
<td><strong>0.6614253</strong></td>
<td><strong>12</strong></td>
</tr>
<tr>
<td><strong>AsianDB</strong></td>
<td><strong>0.6515805</strong></td>
<td><strong>13</strong></td>
</tr>
<tr>
<td>France</td>
<td>0.6469732</td>
<td>14</td>
</tr>
<tr>
<td><strong>WB</strong></td>
<td><strong>0.621796</strong></td>
<td><strong>15</strong></td>
</tr>
<tr>
<td>New Zealand</td>
<td>0.6033334</td>
<td>16</td>
</tr>
<tr>
<td>Switzerland</td>
<td>0.6032289</td>
<td>17</td>
</tr>
<tr>
<td><strong>GFATM</strong></td>
<td><strong>0.6030172</strong></td>
<td><strong>18</strong></td>
</tr>
<tr>
<td>Austria</td>
<td>0.5852491</td>
<td>19</td>
</tr>
<tr>
<td><strong>EU</strong></td>
<td><strong>0.5644109</strong></td>
<td><strong>20</strong></td>
</tr>
<tr>
<td>Spain</td>
<td>0.5397114</td>
<td>21</td>
</tr>
<tr>
<td>Belgium</td>
<td>0.528046</td>
<td>22</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>0.5074713</td>
<td>23</td>
</tr>
<tr>
<td>AfricanDB</td>
<td>0.5063793</td>
<td>24</td>
</tr>
<tr>
<td>Italy</td>
<td>0.5037701</td>
<td>25</td>
</tr>
<tr>
<td>Portugal</td>
<td>0.4723678</td>
<td>26</td>
</tr>
<tr>
<td>Australia</td>
<td>0.4676092</td>
<td>27</td>
</tr>
<tr>
<td>Korea</td>
<td>0.3994828</td>
<td>28</td>
</tr>
<tr>
<td>United States</td>
<td>0.3564023</td>
<td>29</td>
</tr>
<tr>
<td>Inter-American DB</td>
<td>0.3320402</td>
<td>30</td>
</tr>
<tr>
<td>GAVI Alliance</td>
<td>0.3291667</td>
<td>31</td>
</tr>
<tr>
<td>Turkey</td>
<td>0.2852682</td>
<td>32</td>
</tr>
<tr>
<td>United Nations</td>
<td>0.2649928</td>
<td>33</td>
</tr>
</tbody>
</table>
3.3: **Conclusion**

This chapter finds that more autonomous IDOs see their performance decline less in more fragile contexts than does the performance of their less autonomous peers. These findings provide support for the relationship between autonomy and performance hypothesized in chapter 2. Variation in political authorizing environments and in the lack of “slack” between organization and authorizers seem to account for much of these differences in realized organizational autonomy, with quite substantial potential impacts on development outcomes and consequently on developmental trajectories and conflicts. These results suggest that in some instances less autonomous and more autonomous organizations may fare equally well; when working in relatively predictable environments and relatively observable task domains, these data provide no evidence that Navigation from the Top is *always* the inferior strategy, with autonomy having no clear relationship with performance in these contexts.

There are, of course, countless sources of variance in project outcomes. Even on the most charitable reading of the results, autonomy and state fragility jointly explain no more than 55% of the variance in (normalized) inter-IDO project success in the sample. Poor institutional environments, lack of political will, and corruption are commonly mooted as causal of foreign aid delivery failure and this chapter does nothing to suggest they are not important and impactful elements of aid delivery.

While the results in this chapter should give us confidence that there is something to exploring autonomy in foreign aid performance, the limitations of the data mean these
analyses limit themselves to relative performance within IDOs rather than absolute inter-IDO performance. This leaves open at least one central question: do more autonomous IDOs in fact perform better than their less autonomous peers? These data also do little to elucidate the channels via which autonomy affects performance.

The chapters that follow use case study data to address these questions. Chapters 4 and 5 further explore the timing and nature of this autonomy effect while also directly testing the hypotheses and theory of chapter 2, shifting from what could be quantified and examined in this framework (autonomy) to broader navigation strategies. What are the mechanisms via which political authorizing environments and/or IDO autonomy influence project behavior? What is the experience of the organization at the ‘coal face’ of IDOs with clients? How do different IDOs fare in a direct comparison of their project performance and their aggregate organizational impact on development?
CHAPTER 4: COMPARING USAID AND DFID IN SOUTH AFRICA

HIV/AIDS and Municipal Governance Capacity Building Efforts

4.1: Introduction

This chapter examines the performance of US Agency for International Development (USAID) and UK Department for International Development (DFID) projects in South Africa, with a particular focus on the links between that performance and navigation strategy. USAID has a much more insecure political authorizing environment than does DFID; as discussed in chapter 2, this is associated with much lower levels of both organizational and staff autonomy and lower workplace satisfaction for USAID than DFID. Table 2.3 – reproduced below – highlights some of the key observable differences between the two IDOs, including data on workplace satisfaction and autonomy.

Table 2.3: Comparison of USAID and DFID's Political Authorizing Environment

<table>
<thead>
<tr>
<th></th>
<th>Political Status of Aid Agency Head</th>
<th>Budget Security</th>
<th>Response to 2008 Financial Crisis</th>
<th>Workplace Satisfaction Surveys</th>
<th>Rank (out of 33) on Autonomy Measure Used in Chapter 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DFID</strong></td>
<td>Full ministerial rank, limited coordination with Foreign Affairs</td>
<td>Three-year budget allocations; few earmarks</td>
<td>Only ministry spared from across-the-board cuts; budget has continued to increase</td>
<td>Top 2%</td>
<td>3</td>
</tr>
<tr>
<td><strong>USAID</strong></td>
<td>Head of USAID (Administrator) reports to State Department</td>
<td>Yearly, often delayed; USAID budget heavily earmarked</td>
<td>Cutting aid funding promises literally the first thing mentioned by Obama ticket (as candidate)</td>
<td>Bottom third</td>
<td>29</td>
</tr>
</tbody>
</table>

Sources: 2012 US Federal Employee Viewpoint Survey Global Satisfaction Index (USAID 25th of 36); 2013 UK Civil Service People Survey Employee Engagement Index (DFID tied for 2nd of 98); Biden-Palin Debate, October 2 2008; Author
Chapter 3 focused on autonomy – an observable and (chapter 2 has argued) necessary precondition for soft information-utilizing Navigation by Judgment. This chapter and the next focus directly on navigation strategy, the next conceptual link in the chain between political authorizing environment stability and organizational performance. These chapters allow us to look at the ‘coal face’ of implementation, the interaction between IDO field staff, contractors, and recipients in actually delivering interventions. In what ways do USAID and DFID navigate differently, particularly in their propensity to make use of soft information and judgment on the one hand or measurement and top-down controls on the other? Does there seem to be a relationship between autonomy and navigation strategy, or between autonomy/navigation strategy and job satisfaction?

As outlined in Figure 4.1, this chapter covers both municipal government capacity building and interventions focused on the prevention of mother-to-child transmission (PMTCT) of HIV/AIDS in South Africa. These cases allow for a comparison of the performance of DFID and the USAID in a relatively stable development environment: that of substantially post-apartheid South Africa in the 2000’s.
Table 4.1 illustrates the activities on which this chapter primarily focuses, though other engagements will be noted where relevant. The health cases – which focus on the Prevention of Mother to Child Transmission (PMTCT) of HIV/AIDS – focus on the arc of support from these organizations rather than single one-off projects. In part necessitated by the complex, non-projectised form of USAID’s health support via the President’s

---

46 Or, in the case of USAID’s Municipal Governance project LGSP, a single phase of a project. This ‘phase’ distinction is misleading, however; in this case each phase was designed, managed, and tendered (contracted) independently, and so phase 2 is better seen as an independent successor project to phase 1 than a mere continuation or extension.
Emergency Plan For AIDS Relief (PEPFAR), this focus on a broader support arc rather than a single project allows for a clearer view of organizational, rather than project, performance over time.

Table 4.1: Specific Projects Examined in Chapter 4

<table>
<thead>
<tr>
<th>HIV/AIDS Prevention of Mother To Child Transmission</th>
<th>USAID</th>
<th>DFID</th>
</tr>
</thead>
<tbody>
<tr>
<td>President's Emergency Plan For AIDS Relief (PEPFAR) as administered by USAID &amp; CDC, 2004-present; various implementers</td>
<td></td>
<td>HIV/AIDS Multisectoral Programme (MSP), 2003-2008;(^{47}) implemented initially by Futures Group, succeeded by HLSP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rapid Response Health Fund (RRHF), 2008-2009; implemented by HLSP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Strengthening South Africa’s Response to HIV and Health (SARRAH), 2010-2015; implemented by HLSP</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Municipal Governance Capacity Building</th>
<th>USAID</th>
<th>DFID</th>
</tr>
</thead>
</table>

This first section provides a brief introduction to the country context and the projects to be examined. Sections 2 and 3 focus on the cases themselves, attempting to ascertain whether USAID and DFID engage in observably distinct navigation strategies, and whether USAID and DFID’s interventions are differentially successful. These sections also explore the links between navigation strategy and success. Section 2 examines the capacity building cases – first describing the cases and their relative success, then discussing how different navigation strategies contributed to any observed differences in performance.

\(^{47}\) The MSP technically included the RRHF, but this was for administrative, not substantive, reasons; as such the MSP’s official end date is in 2009. Unless otherwise noted references to the MSP in this chapter denote the project that ended at the close of 2008, thus excluding RRHF.
Section 3 follows the same pattern for DFID and USAID’s health cases, also examining how the US Centers for Disease Control (CDC) implemented parallel interventions to those of USAID. This section thus includes an examination of inter-agency variation between US agencies. Section 4 investigates intra-organizational channels, bringing these cases into conversation with chapter 2’s hypotheses regarding the relationship between political authorizing environments, autonomy, soft information, and field agent job satisfaction. Section 5 concludes.

**South Africa in the Mid-2000’s**

In 2004 South Africa celebrated 10 years since the fall of apartheid. A sluggish economy in the first years following apartheid had transitioned to steady growth, at least until the 2008 financial crisis. (World Bank 2013) South Africa remained a very unequal society; in 2009 the bottom half of the population accounted for just 8% of the country’s income. (Statistics South Africa 2012) While whites comprised less than 10% of the population, (Statistics South Africa 2012) white farmers owned 85% of the agricultural land. (World Bank 2013)

South Africa was not without its challenges – political, social, economic – and the legacy of apartheid continued to play a substantial role in South African society and government policy and institutions. That said, South Africa was a stable, growing, legible country; the biggest economy in Africa, on the cusp of hosting the 2010 World Cup, and a member of the Group of 20 (G-20) major economies. Indeed, by 2013 – with South African
GDP above 7,000 USD per capita – DFID announced that it would be gradually phasing out its aid efforts in the country, by no means the first such donor to do so.48

**DFID and USAID in South Africa: History and Trajectory**

Historically a British colony, DFID had been in some sense the ‘lead’ donor in South Africa, though by 2010 aid from the US was more than triple that from the UK; together, the US and UK accounted for over 45% of South Africa’s 1.1 billion USD of total aid disbursements, according to the OECD.49 The large US program was attributed by more than one actor to the US desire to overcome a historic perception of being allied to the apartheid government, and to Nelson Mandela’s request to then-US President Clinton that South Africa for substantial support as apartheid ended.50

Among the ways South Africa is relatively unique among aid recipients is the relatively small share aid plays relative to the national budget; even in 2008 – during the global financial crisis – South Africa’s 1.1 billion USD in aid receipts was a fraction of a percent of its 250+ billion USD economy. Aid’s relative lack of importance has substantial impact on the dynamics of the IDO – Government of South Africa relationship; as one

---

48 Interview 97, 8/15/13. The numerical citation scheme is described in Appendix I.

49 These data are taken from the International Development Statistics (IDS) maintained by the OECD’s Development Co-operation Directorate-Development Assistance Committee (DCD-DAC). In constant 2012 USD, total aid in 2008 to South Africa was 1117.62 million USD; of this the UK contributed 108.98 million USD, the US 400.66 million USD. (OECD)

50 Interview 103, 7/30/13; Interview 87, 9/20/13
individual with long knowledge of donor behavior in South Africa put it, donors are treated by the government with a mixture of “respect and disdain”\textsuperscript{51} A former government official described throwing donor officials whom he felt lacked adequate respect out of his office\textsuperscript{52} It would be very difficult to imagine such a scene in Liberia, where remaining in donors' good graces is critical to success.

As one old aid hand – noting South Africa’s exceptional status as regards donors - put it, “[When a country] is so heavily reliant on donor support and the machinery of Government spends more than half its time thinking about the donors…. the donors aren’t gentle, they demand attention be paid.”\textsuperscript{53} South Africa, then, allows us to explore what happens when the Government can, and sometimes does, push back. As Pam Yako, a former municipal manager and Director General of multiple national ministries put it, “The donor people have got an agenda and even the people on the receiving end have got an agenda, there’s always an agenda, the question is how strong the recipient is in pushing their own agenda and pushing back, and that varies, and so you could have a donor programme that’s completely dominated by what the donor wants, or you could have a donor programme that is the other way round.”\textsuperscript{54}

\textsuperscript{51} Interview 90, 8/2/13; sentiment echoed in interview 94, 7/25/13; interview 103, 7/30/13; interview 100, 11/12/13
\textsuperscript{52} Interview 92, 7/24/13
\textsuperscript{53} Interview 100, 7/31/13
\textsuperscript{54} As this is the first non-anonymized quotation from an interviewee in this work, I will take this opportunity to describe the way I have approached quotations for attribution. In cases where individuals have cleared language for attribution, I will not link the attributed language to the anonymized interview number; to do otherwise would be to allow easy
Does a relatively visible, legible environment – particularly one in which local actors, by pushing back, are likely to contribute the soft information and local knowledge that donors might otherwise lack - allow Navigation from the Top to perform as well, or even better, than Navigation by Judgment? Does this extend to all sectors, or is optimal IDO navigation strategy nonetheless determined by the dynamics of the task – observability & verifiability, soft information returns, etc.?

The health cases look at a realm of health particularly tractable to Navigation by Measurement – the Prevention of Mother to Child Transmission (PMTCT) of HIV/AIDS, which is largely done through drug delivery and education (observable inputs), with success measured based on actual transmission rates to newborns of HIV+ mothers (observable outcomes). How does Navigation by Judgment fare in this context?

The capacity building cases compare USAID and DFID attempts to support municipal governments – in some cases, in the very same municipalities. While capacity building interventions generally may be more difficult for Navigation from the Top because it is harder to find accurate, objective measures and levers of control/monitoring, South Africa is an environment in which this strategy would is relatively likely to be successful. Is it?
4.2: Capacity Building

Municipal government in South Africa was, in some sense, a creation of the post-apartheid Constitution of 1996. While there had been some organized municipalities in the apartheid era, there was no history of (as one actor put it) “wall to wall” municipal government.\textsuperscript{55} The Municipal Government sector was one in which funding wasn’t the constraint;\textsuperscript{56} that is not to say the sector was without problems, however. While a few leading municipalities were thriving, this was the exception rather than the norm.\textsuperscript{57} This unevenness led Ivor Chipkin, the Executive Director of the Public Affairs Research Institute, to remark “to look at the formal organization structures as your end point, as the be all and end all of how the organization works is very, very misleading... [External] consultancies in South Africa are failing more often than succeeding partly because their starting point is the formal structure and where power ostensibly lies within that formal structure... usually that [approach by external actors] doesn’t work because that is not how these organizations work.”\textsuperscript{58}

Strong apartheid municipal governments had also left a complex institutional legacy for the legal framework and enabling environment of municipal government. Just one of

\begin{flushleft}
\textsuperscript{55} Interview 96, 8/5/13; interview 92, 7/24/13. The “wall to wall” phrase is from interview 92.

\textsuperscript{56} Interview 97, 8/15/13; interview 101, 8/2/13

\textsuperscript{57} Interview 96, 8/5/13; interview 92, 7/24/13; interview 103, 7/30/13; interview 100,7/31/13; interview

\textsuperscript{58} PARI, a Johannesburg-based independent research institute – what in the US might be called a think tank - have research looking at municipal governance and differential historical legacies.
\end{flushleft}
many such examples: In apartheid’s waning days white municipal councils had often transferred municipal property to non-public structures (e.g. private trusts) in order to maintain control. In reversing these actions and annulling these transfers, it became very difficult for municipalities to borrow, raise bonds, etc.; after all, if any municipal council could annul the actions of its predecessors, how could creditors expect repayment? By 2004 the framework for municipal borrowing had been addressed by an earlier, largely successful, reform to which donors had contributed.

By 2003 some municipal governments – particularly those in the major metropolitan areas, where there was substantial local revenue collection – were doing very well. These municipalities were unfortunately the exception rather than the rule. Some interviewees went so far as to call South Africa’s attempts at municipal governance largely failed. In many municipalities officials are more accountable to political parties (particularly the ruling African National Congress) than their constituents; municipal government is often seen as a stepping stone to higher, better paying office.

59 This example comes from interview 100, 7/31/13
60 Interview 73, 7/19/13; interview 101, 8/2/13; interview 100, 7/31/13; interview 95, 8/1/13
61 Interview 100, 7/31/13; interview 78, 7/22/13
62 Interview 96, 8/5/13; interview 92, 7/24/13
63 Interview 103, 7/30/13; interview 94, 7/25/13; interview 96, 8/5/13; interview 92, 7/24/13; interview 91, 7/23/13; interview 72, 7/31/13
Project Descriptions

DFID and USAID’s efforts at municipal governance support over this period are remarkable in the similarity both of their diagnosis of the problems and their prescribed remedies.

DFID’s Consolidation of Municipal Transformation Programme (CMTP)\textsuperscript{64} stated its goal as to “consolidate accountable local democracy and pro-poor service delivery. The purpose is to promote the development of effective and efficient municipalities”. (DFID 2003, p. 1) It aimed to do so via four components:

• Support to capacity building for municipal committees, the Auditor General, and journalists
• Placing advisors, known as Integrated Service Facilitators (ISFs), in up to 12 municipalities to address delivery challenges
• Placing experts in up to 12 municipalities to improve budgeting and financial management
• Assisting national government in policy making and capacity building, provincial government in providing support to municipalities, and the South African Local Government Association (SALGA) on awareness activities regarding the role of local government. (DFID 2003)

\textsuperscript{64} DFID Project Reference #104886
The project was initially scheduled to run from 2003-2008 (it ultimately completed in 2009). More than 80% of the programmatic activities budget\textsuperscript{65} was focused on the advisors – both the ISFs and treasury experts. (DFID 2003, p.11) As a result this analysis will focus much of its attention on these components of the project.

USAID’s Local Governance Support Programme (LGSP)\textsuperscript{66} phase 2 – that examined here – built on a phase 1. Phase 1 had had as its goal “Effective and Democratic Local Governance”, and had provided support for developing national policies and programs, direct assistance to municipalities, and codifying lessons learned – that is, learning from project results. (Creative Associates 2004) Phase 2 focused on efforts to “improve municipal planning skills and operating systems to increase effectiveness, transparency and accountability; and increase the revenue stream of selected municipalities ensuring that citizens meet their obligations.” (LGSP/RTI 2006) This would be achieved via support to municipal strategic planning, finance, and revenue systems delivered through “hands-on mentoring and training” (Research Triangle International 2007, p. 24). In practice this meant largely the visiting of municipalities by trainers, what was described by one

\textsuperscript{65} That is, excluding general management and contingency budgets. The figures for this calculation are from the Program Memorandum (the plan for the project before it commenced), which gives a budget of 18.6 million GBP. In the end, CMTP spent 11 million GBP according to the project completion report; (DFID 2009) it’s not clear what proportion of the final expenditure went to which objectives.

\textsuperscript{66} USAID Contract #674-C-00-05-00001-00
interviewee as “targeted training”. In 2007/8 additional funding allowed the placement of advisors in 9 municipalities.

Both LGSP and CMTP were implemented via contractors – in the case of CMTP the contracted party was the consulting firm Deloitte; in the case of LGSP it was the Research Triangle Institute International (RTI), a holder of many USAID contracts. Both also had project leaders with rich experience in South African governance – for Deloitte a former South African municipal government official, for RTI an expat who had long experience in earlier municipal government reform efforts.

LGSP and CMTP are notable in their similarities. Both focused on making local government more effective via capacity building. This capacity building was of both a management and a financial nature. LGSP initially delivered these capacity interventions via trainings delivered by remotely financed experts. In the latter part of the project LGSP

---

67 Interview 85, 11/12/13

68 Initially 7 municipalities (LGSP/RTI 2006), this grew to 9 resident municipal advisors. (Interview 81, 8/1/13) While there are no available budgetary figures for LGSP, it seems safe to assume that the bulk of the programmatic resources went towards activities quite parallel in aim to those of CMTP’s central thrust. A FOIA request is, as of April 2015, still only partially filled 22 months after being filed. The only figure I was able to find for LGSP overall is a 20 million USD figure from a 2005 press release; (USAID 2005) this is almost certainly an understatement, given the additional funding streams which were to be incorporated into LGSP during its implementation.

69 In FY 2006, USAID accounted for 30.5% of RTI’s revenue, or 165.9 million USD. (Research Triangle International 2007, p. 46)

70 Subethri Naidoo and Peter Vaz, respectively. Subethri later worked for DFID on municipal governance, and had earlier worked for USAID in their municipal reform efforts.
placed advisors in local municipalities, just as CMTP did in its initial stage. Both projects also saw additional program areas added during implementation – namely anti-corruption and HIV/AIDS.

**Capacity Building Project Results**

The similar goals and partial overlap of methods for achieving those goals makes comparison relatively straightforward. These projects had the goal of improving the functioning of municipalities. It seems natural to ask, then, did the functioning of the municipalities targeted change in the desired (or any) direction in either the short or long term?

In addition, to what extent did either of these projects have broader systemic impacts? In targeting a handful of municipalities both project aimed to provide examples of municipal transformation that might be replicated. To what extent was this the case? Were there follow on efforts from these donors or others that built on these projects or adopted models or innovations they introduced?

*Direct Impact on Municipalities Served*

Neither CMTP nor LGSP can fairly be framed as an unmitigated success. There was, however, differential success between the two projects.
Until 2007 LGSP’s primary delivery mechanism was via targeted trainings; these trainings were usually every few months and lasted part of a day, or sometimes a whole day, in each municipality.\textsuperscript{71} The only positive statement made about the training component of LGSP was by a single individual who had also been involved in its management; while noting that in many areas the training was a “mixed success”, this individual thought the anti-corruption trainings which became part of the project in its later years had been particularly effective.\textsuperscript{72}

A surprising number of LGSP and USAID personnel made contrasting claims regarding the effectiveness of the training, in doing so echoing the comments of many actors who were not participants in the project.\textsuperscript{73} As one LGSP staff member put it, “I don’t think [training under LGSP] contributed much… because you go there, you don’t have any authority over the people that you training, so if they don’t cooperate you cannot say anything, you go there sometimes, they tell you that we have other priorities, we don’t have time now, those kinds of things.”\textsuperscript{74} Former LGSP Deputy Chief of Party Bongani Matomela said “[LGSP] might have not made the most dent or impact.” A USAID official described LGSP as “classroom oriented”, as it being hard to see much as a result of LGSP in those municipalities he had visited.\textsuperscript{75} This official specifically ascribed LGSP’s failing in part

\textsuperscript{71} Interview 88, 8/1/13

\textsuperscript{72} Interview 85, 11/12/13

\textsuperscript{73} e.g. interview 93, 7/18/13; interview 73, 7/19/13

\textsuperscript{74} Interview 88, 8/1/13

\textsuperscript{75} Interview 86, 7/30/13.
tackling of things like anti-corruption via remote training, thus threatening entrenched local interests.\textsuperscript{76} In so doing he specifically criticized the very component that had been singled out for praise by the sole individual to claim the trainings had substantial impact.

The municipal advisors that USAID placed towards the end of the project were, on the view of at least one LGSP actor, the most successful part of the project; these advisors worked on average 3 days a week, flying to and from the municipalities as needed.\textsuperscript{77} If these advisors were successful it was a quiet success, in that there were no signs of improvement in the fiscal indicators of the LGSP municipalities according to SA municipal audit reports.\textsuperscript{78} In multiple cases national government officials (DPLG/COGTA) didn’t recall the advisory component of LGSP, and in one case a municipal manager whose municipality had received both LGSP training and an LGSP advisor had no memory of the project.\textsuperscript{79} Two

\textsuperscript{76} Interview 86, 7/30/13. The only part of LGSP trainings to be remembered positively by multiple actors was support for municipal HIV/AIDS strategies, which was added later. (Interview 80, 7/22/13; interview 84, 7/22/13; interview 79, 8/5/13; interview 103, 7/30/13) This fairly minor element of the program had overlaps with USAID (and the broader US Government’s) PEPFAR funds, and thus is less about municipal governance and capacity building along the lines of the program’s original intent and more about response to a public health crisis. In other words, the HIV/AIDS strategies saw continued support and funding from PEPFAR, and thus are difficult to evaluate in the medium term as an LGSP accomplishment.

\textsuperscript{77} Interview 81, 8/1/13

\textsuperscript{78} This stylized conclusion comes from an analysis of the panel data to emerge from the series of “General Report of the Auditor-General on Local Government”, issued by the office of the SA Auditor General annually.

\textsuperscript{79} Interview 102, 8/13/13; interview 97, 8/15/13; interview 98, 8/14/13. The municipal manager in question had been in his current role when the advisor was present in the municipality.
of the nine advisors were ‘absorbed’ by the municipality; the municipality continued to pay them, or secured funding for their continuation, after their 6-12 month LGSP-supported stint. This revealed preference for a few of the advisors is perhaps the clearest sign of impact with which LGSP can be credited, though identifying one useful staff member for two municipalities is less than an overwhelming impact.

DFID’s CMTP was not wholly successful, either. That said, the project was far from a total failure. The project received an overall rating of 3 on a 5 point scale on the official project completion report (DFID 2009); as the separately conducted independent evaluation DFID commissioned puts it, “CMTP was a timely, ambitious and comprehensively designed programme with core interventions that delivered impact on the local level.” (Linkd 2009)

CMTP began with advisors – ISFs – based full time in municipalities. Similar to LGSP’s experience, there were some reported cases of resistance from municipalities. Being full time resident there for the long term (2-3 years), ISFs were often – though not always - able to find a way to positively influence municipal systems. Both beneficiaries and project staff reported that CMTP advisors achieved some shifts in municipal practices. Multiple actors noted the permanent status of ISFs in the municipality prevented the

80 Interview 81, 8/1/13
81 Interview 74, 7/30/13
82 Interview 103, 7/30/13; interview 72, 7/31/13; interview 98, 8/14/13; interview 75, 7/25/13; interview 76, 7/29/13
program from being ‘sidelined’ in the way LGSP seems to have been.\textsuperscript{83} At least one ISF later was absorbed into municipal governance going forward.\textsuperscript{84}

In the latter phase of the program it is more difficult to point to specific successes. As advisors’ tours finished, CMTP delivered a number of not fully integrated support elements; the independent evaluation quotes an anonymous senior government official’s description of this latter period as a “fruit salad of low-hanging fruit of no programmatic value” (Linkd 2009), a view supported by interviewees.\textsuperscript{85} In this latter phase it is also the case that some of what CMTP had accomplished in the first phase seemed to be reversed, with some slippage by target municipalities from the better performance experienced while ISFs were present.\textsuperscript{86}

The dedicated municipal advisory component of CMTP seems to have been the most directly impactful. That component was more fully realized for CMTP than LGSP, with advisors resident rather than traveling to and from municipalities, and based in municipalities for a longer time. It would be an overstatement, however, to see CMTP as

\textsuperscript{83} Interview 74, 7/30/13; interview 72, 7/31/13. These actors didn’t make the comparison to LGSP, only the statement about the importance of being resident.

\textsuperscript{84} Interview 76, 7/29/13; this individual now plays an important role in municipal support for the provincial government of Kwazulu Natal. The lack of absorption of advisors is less revealing for CMTP than LGSP; while LGSP focused on local (South African) advisors, the majority of CMTP advisors (who were more highly paid than LGSP advisors) were foreign nationals (interview 76, 7/29/13), with long term absorption into the municipality not part of the strategy.

\textsuperscript{85} Interview 73, 7/19/13; interview 75, 7/25/13; interview 92, 7/24/13

\textsuperscript{86} Interview 75, 7/25/13; interview 103, 7/30/13; interview 76, 7/29/13
having accomplished all, or even most, of what it set out to in terms of direct municipal impact; in some municipalities it accomplished very little.\footnote{Interview 92, 7/24/13; interview 102, 8/13/13} DFID’s project completion report, in giving CMTP a 3 on a 5 point scale, argues “The achievement of CMTP purpose, although uneven in parts, included some highly positive examples in selected municipalities.” (DFID 2009)

*Barriers to Success*

What kept both CMTP and LGSP from being the successes hoped for at inception? Many of the factors pointed to as impediments to intended impacts were common across the two projects – unsurprising given the overlapping aims and mechanisms of the projects. In some cases, these factors seem to have differentially impacted the projects – where this is the case I highlight same, as this data is helpful both in highlighting differential navigation strategies and the consequences of same.

There has been substantial turnover in municipal personnel – for example, of the CMTP municipalities only one had, as of 2013, the same municipal manager since the project commenced.\footnote{Interview 72, 7/31/13; interview 98, 8/14/13.} This turnover was mentioned by almost every individual interviewed.\footnote{A partial list: interview 79, 8/5/13; interview 95, 8/1/13; interview 78, 7/22/13; interview 103, 7/30/13; interview 88, 8/1/13; interview 73, 7/19/13} This meant that LGSP trainings didn’t ‘stick’ – today’s training oven needed
to be repeated tomorrow, as personnel had changed in the interim.\textsuperscript{90} The long term presence of CMTP ISFs allowed them to respond more dynamically to changes at the local level; that said, a number of years after project close it was still the case that a substantial proportion of the individuals with whom ISFs worked had left their posts.

CMTP’s close association with national and (in some cases) provincial government seems to have mitigated, though by no means eliminated, the effect of turnover. Indeed, the national ministry responsible for local government itself transformed radically over this period, with successive ministers and SA Presidents reshaping the Ministry substantially.\textsuperscript{91}

Both DFID and USAID also experienced transitions of personnel over this period; this churn was particularly salient at USAID, where LGSP was managed under three different departments as USAID South Africa’s structure changed.\textsuperscript{92} Multiple actors suggested the personality, temperament, and orientation of supervisory personnel at USAID may have constrained project success as well, with the project not well understood by some managers.\textsuperscript{93}

\textsuperscript{90} Interview 88, 8/1/13

\textsuperscript{91} Interview 81, 8/1/13; interview 101, 8/2/13; interview 93, 7/18/13; interview 73, 7/19/13; interview 84, 7/22/13

\textsuperscript{92} Interview 85, 11/12/13; interview 78, 7/22/13; interview 80, 7/22/13; departments are “programs” in USAID language, e.g. the Democracy and Governance program or the Municipal Governance program.

\textsuperscript{93} Interview 80, 7/22/13; interview 91, 7/23/13; interview 87, 9/20/13; interview 85, 11/12/13.
It was also mooted that there was inadequate attention paid to the project from the donor side; for LGSP this was said regarding USAID itself (rather than RTI, the contractor), for CMTP regarding Deloitte (rather than DFID staff). One actor connected USAID’s failings in this regard to political will in DC, arguing there’s only political will for health and disasters in the US, and thus that USAID can’t do other things consistently.

Broader Systemic Impacts

There is no sign of LGSP having impact on national policy or being cited as an example by others. As one LGSP staff person put it, LGSP was

A real disappointment for us at the end of the programme, that neither the municipalities could sustain it nor the department, at the time even the national department could not sustain that level of intervention that was required, and I think we needed a little bit more, obviously in a lot more municipalities, but the financial model was not sustainable.

By contrast, CMTP acted as a model. DPLG launched a national campaign called Project Consolidate that built in part on CMTP’s delivery mechanisms, placing advisors and direct support in municipalities. This link was noted by DPLG/COGTA, provincial

---

94 USAID: Interview 91, 7/23/13; interview 78, 7/22/13; interview 85, 11/12/13. While not a direct quote, also on point one interviewee who played a direct supervisory role over LGSP didn’t remember the program’s existence; on prompting the interviewee remembered the CoP but stated his biggest strength stemmed from his being a South African national, which the individual in question is not (Interview 83, 9/2/13). Deloitte’s lack of attention: Interview 92, 7/24/13; interview 95, 8/1/13

95 Interview 91, 7/23/13

96 Interview 81, 8/1/13
government, and CMTP actors.\textsuperscript{97} While LGSP did have connections to Project Consolidate, it appears that Project Consolidate influenced LGSP rather than vice versa; on two interviewees’ views it was Project Consolidate that influenced LGSP’s move to advisors.\textsuperscript{98}

CMTP also spent a substantial amount of time focusing on knowledge sharing and on working with its steering committee, which was composed primarily of government actors.\textsuperscript{99} On the view of some, including Government officials, this assisted in CMTP’s impact, allowing CMTP lessons learned to have broader impact and influence.\textsuperscript{100} One LGSP staff member suggested LGSP’s links with the Government were weak, saying LGSP’s connection with DPLG and national government “wasn’t as solid as it should have been.”\textsuperscript{101} Bongani Matomela, former LGSP deputy manager, suggested explicitly that CMTP was the stronger project largely due to LGSP’s lack of government link:

My sense was that it [CMTP] was probably more successful [than LGSP]….. one of the success factors is probably the fact that they [CMTP] were closer to the department [DPLG], they probably did things according to what the department wanted, you know, and they were directly augmenting Project Consolidate whilst I get it that LGSP seems to have been just the

\textsuperscript{97} Interview 103, 7/30/13; interview 75, 7/25/13; interview 76, 7/29/13; 97, 8/15/13; interview 103, 7/30/13

\textsuperscript{98} Interview 81, 8/1/13; interview 102, 8/13/13

\textsuperscript{99} Interview 75, 7/25/13; interview 76, 7/29/13; interview 103, 7/30/13; interview 97, 8/15/13.

\textsuperscript{100} Interview 76, 7/29/13; interview 103, 7/30/13. It is fair to note the independent evaluation (Linkd 2009) notes the steering committee did not meet as frequently as intended, with the ‘technical committee’ becoming the de facto steering committee. LGSP had a somewhat similar steering committee, but only one actor even made mention of the structure. (Interview 85, 11/12/13)

\textsuperscript{101} Interview 81, 8/1/13
programme... Because it [CMTP] was actually anchored within the department and so the chances of certain tools and models that were being developed and tested were then immediately transplanted to the municipality, so I think that probably it's a more sustainable way of doing projects than maybe the LGSP.

This view – that CMTP was the more successful project – was made explicitly or suggested by others who could speak to the success of both projects and would have no obvious reason to speak better of CMTP than LGSP.102 There were no actors who felt LGSP had been the more successful project.

Differences in Navigation Strategies: CMTP and LGSP

That CMTP was more successful that LGSP, and DFID in general more likely to Navigate by Judgment than DFID does not in itself prove that differential navigation strategies played a role in CMTP’s superior performance. CMTP’s relative success could be due to a host of factors that have nothing to do with differential IDO or project-level Navigation by Judgment as opposed to Navigation from the Top.

As described in chapter 2, I hypothesize that navigation matters in three distinct phases of projects: their design, their day-to-day implementation, and their revision (if any). Interviews focused on these three phases, and this section will proceed in similar fashion.

102 Interview 93, 7/18/13; interview 103, 7/30/13; interview 90, 8/2/13
In each phase is there evidence of differential navigation strategies and/or field agent degrees of autonomy? Does differential Navigation by Judgment in design affect the quality of projects, particularly in better incorporating soft information regarding local conditions or what is likely to succeed? In implementation does differential Navigation by Judgment yield projects that flow more smoothly as relationships are less formalized and decisions made on observable but non-contractible/externally verifiable features of the situation? Does differential Navigation by Judgment leads to revisions in projects in productive directions, with activities adapting as necessary to a rapidly changing environment?

The sub-sections that follow explore whether there were, in fact, differential navigation strategies in the design, implementation, or revision of these projects. The connection between these differential strategies and the project’s success will be explored as well.

Design

DFID’s design for CMTP was guided by judgment. Chris Heymans – the South African lead designer of CMTP – had a background working with municipal government, which substantially informed the project in beneficial ways. The notion of working via resident advisers built on the success of advisors in SA Treasury in a municipal capacity building program, where RTI – the party that was to implement LGSP – had been

---

103 Interview 103, 7/30/13; interview 73, 7/19/13; interview 92, 7/24/13
implementing a successful project.\textsuperscript{104} The notion of having CMTP personnel based inside the national Ministry responsible for local government (DPLG) was unproblematic; DFID accepted this as a good idea, and – as noted above – this decision appears to have been related to the project's relative success as compared to LGSP.\textsuperscript{105}

In choosing municipalities with which to work the DFID design team actually visited municipalities which they were considering; in addition to the needs of the municipality the likelihood of the municipality working well with an advisor was a key consideration in support decisions.\textsuperscript{106} One actor described this as designed to avoid the "looks good on paper but won't work in practice problem"; these municipal visits affected the both the selection of municipalicities and the overall design of the project.\textsuperscript{107} As Richard Thomas, DFID's governance advisor at the time of CMTP's design, put it there was a need to figure out "What's really, rather than officially, going on." In Thomas' view, then, South African municipality selection required soft information.

The design of CMTP was led entirely by South Africa-based DFID personnel – a British governance advisor and a South African lead designer.\textsuperscript{108} DFID in London allowed

\textsuperscript{104} Interview 95, 8/1/13

\textsuperscript{105} Individuals who suggested this was a positive CMTP design feature included interview 102, 8/13/13; interview 94, 7/25/13; interview 92, 7/24/13; interview 82, 8/1/13.

\textsuperscript{106} Interview 77, 9/17/13

\textsuperscript{107} Interview 77, 9/17/13

\textsuperscript{108} Interview 95, 8/1/13; interview 77, 9/17/13
most of the key approvals and the primary authority over the design to reside with South
Africa-based staff; one actor suggested this was critical in ensuring quality in design.\textsuperscript{109} The quality of the project as designed was praised by a variety of actors, including a
number not directly involved with the project.\textsuperscript{110}

This is not to say that CMTP had no elements of Navigation from the Top. DFID
headquarters (in conjunction with DPLG) pushed the project to work with some weaker
municipalities.\textsuperscript{111} The inclusion of anti-corruption efforts in the project was driven from
headquarters.\textsuperscript{112} DFID’s concerns about cost constrained the per-consultant cost, and may
have been related to what some perceived as a decline in advisor quality between that
present in design and implementation.\textsuperscript{113} That said, Navigation by Judgment played a
substantial role in CMTP’s design.

The design of LGSP demonstrates less use of contextual knowledge \textit{despite} the fact
that USAID had, if anything, the superior knowledge base on which to draw. RTI had been

\textsuperscript{109} Interview 77, 9/17/13

\textsuperscript{110} Interview 73, 7/19/13; interview 92, 7/24/13; interview 102, 8/13/13

\textsuperscript{111} Interview 95, 8/1/13

\textsuperscript{112} Interview 95, 8/1/13

\textsuperscript{113} Interview 92, 7/24/13; interview 95, 8/1/15. Other actors suggested this gap in quality
was driven by Deloitte effectively enacting a bait and switch, supplying cheaper consultants
while billing the project for more expensive ones (Interview 103, 7/30/13). The relative
positions of the actors leads me to believe the Deloitte manipulation story more likely; it is
also possible both are accurate, with DFID cutting budgets and Deloitte then providing
consultants that were less expensive than the reduced amounts yet billing for the entire sum.
implementing USAID programming while resident in DPLG and the National Treasury;\textsuperscript{114} nonetheless the design of LGSP didn’t draw directly on RTI’s experience.\textsuperscript{115} This is, perhaps, an unfortunate consequence of procurement regulations – as a potential bidder (and, indeed, the eventual recipient of the LGSP contract) it may not have been possible for USAID to leverage RTI’s experience in this way. USAID did, however, have a phase 1 of LGSP to learn from in designing LGSP phase 2 – a phase 1 that had an explicit focus and financing for “lessons learned”, for codifying learning so as to inform future phases of the project.\textsuperscript{116} In reflecting on LGSP phase 1 Creative Associates – the firm who had been managing the project – wrote in a report to USAID

\begin{quote}
The design for the first round of the LGSP used few service providers who clustered assistance to widely dispersed municipalities with the same local government developmental needs. **It did not work.**\textsuperscript{117} **The key reason was that municipal conditions varied and service providers did not necessarily have the local knowledge to adequately deal with implementation issues that arose.** (Creative Associates 2004, p. 28)
\end{quote}

It is striking, then, that USAID pursued the strategy it did in LGSP phase 2. One USAID actor described the strategy this way: “LGSP took the approach of choosing municipalities and then trying to work through a program of capacity building with all of them, your standard sort of package.”\textsuperscript{118} This is nearly identical to the strategy that had so clearly been repudiated in the initial phase.

\begin{footnotesize}
\textsuperscript{114} Interview 93, 7/18/13; interview 100, 7/31/13; interview 105, 8/16/13

\textsuperscript{115} Interview 85, 11/12/13

\textsuperscript{116} That is, the component examined here; as noted above, what I call “LGSP” here is in fact LGSP phase 2, the part of LGSP run as an RTI project.

\textsuperscript{117} Boldfacing mine.

\textsuperscript{118} Interview 86, 7/30/13
\end{footnotesize}
The LGSP phase 1 report also foreshadows some of the other key difficulties faced by LGSP – the rotation of municipal personnel and the difficulty of “political situational management” are also mentioned explicitly. (Creative Associates 2004) It is not clear why USAID largely did not heed these warnings. It may have been related to what one USAID official described as the “bottom up meets top down” project development process.¹¹⁹ The one element of the phase 1 critique not repeated was that the municipalities in phase 2 were all clustered within a few hours’ drive of Pretoria, the national capital. These municipalities were chosen “based primarily on logistics and ease of access and transport”;¹²⁰ one individual involved in LGSP suggested this was to make the municipalities easy to reach so as to allow USAID to engage in monitoring more easily.¹²¹ The selection of winning bids suggests both USAID and DFID valued previously gathered soft information (or its observable proxy, local experience) in the selection process; RTI’s experience in the South African municipal space was a strength of the bid for LGSP.¹²² Deloitte’s use of lead actors with substantial South African experience was similarly seen as critical for success, and DFID in fact insisted that Deloitte manage the project out of South Africa rather than having the project report to a foreign office.¹²³

¹¹⁹ Interview 87, 9/20/13
¹²⁰ Interview 87, 9/20/13
¹²¹ Interview 88, 8/1/13
¹²² Interview 85, 11/12/13
¹²³ Interview 77, 9/17/13
When RTI won the bid it meant that the RTI Chief of Party on LGSP – Peter Vaz, who had been serving as an advisor in DPLG – moved his office out of DPLG at around the same time that CMTP was moving in. USAID specifically insisted that LGSP not be based inside DPLG due to what one actor described as “a warranted worry that whoever was managing the project [LGSP] would be pulled up in too many different directions and not be able to focus on the project.” Whatever the virtues of such a choice, it is a striking one given that a close connection to DPLG, and SA Government more broadly, was directly cited by multiple actors as a factor in CMTP’s success as compared to LGSP.

Implementation

USAID kept relatively tight control on LGSP; one LGSP implementer described it as a process of meeting USAID South Africa “before every decision was made”, of effectively being intermediaries between DPLG and USAID. USAID South Africa was in turn in close contact with Washington on many of these decisions; one USAID actor described the need to get “approvals from Washington on very low threshold of funding and it delayed things.”

---

124 Interview 85, 11/12/13

125 LGSP/RTI did, in fact, consult frequently with DPLG, particularly as Project Consolidate - the SA Government’s municipal capacity effort - became more prominent and RTI sought to keep LGSP aligned to, and supportive of, national government policy. (Interview 97, 8/15/13; Interview 85, 11/12/13) It was nonetheless further from DPLG than was CMTP.

126 Interview 81, 8/1/13

127 Interview 86, 7/30/13; Interview 103, 7/30/13; the quote is from interview 103.
The indicators against which RTI needed to deliver were quite salient as LGSP moved forward; these indicators tended to focus on number of individuals trained and number of trainings delivered. One USAID actor described this as counting “bums on seats”. In the later phases of LGSP, resident advisors would report on meetings held, guidelines drafted, and other such externally observable and verifiable indicators. One LGSP actor described implementing LGSP as “a numbers game... [USAID would say] we want the numbers, we want information.” Michelle Layte, the RTI LGSP Chief of Party at the end of the project, said indicators were chosen “because it was easier to count... but the numbers didn't tell about the impact”. Layte went on to say that, while USAID had been better earlier in the project and understood both South Africa and the programme itself “It was more a number chasing towards the end especially because we needed to reach our target.... USAID’s focus was around meeting numbers as opposed to the impact.”

USAID was also seen as more rule-bound, with process controls and other elements of Navigation from the Top that made it a more difficult system to navigate than DFID's. One LGSP service provider stated that following their experience with LGSP they elected not to work with USAID again because of how heavy the process reporting burden had

128 Interview 88, 8/1/13; Interview 87, 9/20/13; Interview 81, 8/1/13; Interview 79, 8/5/13
129 Interview 87, 9/20/13
130 Interview 82, 8/1/13
131 Interview 82, 8/1/13
132 Interview 103, 7/30/13
been. Subethri Naidoo, who played management roles with both LGSP and CMTP, said “It was easier to work in DFID and I think it was also far more intellectually stimulating for me. You know USAID was very rules driven, compliance driven but there wasn’t kind of a recognition of the need for personal growth and development of your staff... [on the other hand with DFID] there would be a regular kind of intellectual thinking and strategizing about where we are.”

CMTP did not share this focus on Navigation from the Top and Navigation by Measurement. DFID was far more decentralized, with little communication with London in contrast to USAID’s need for contact with Washington. This free reign allowed Deloitte to focus in some cases on profit to the detriment of the project; in one concrete example Deloitte supplied advisors that were less experienced and thus less costly to municipalities while billing DFID for more costly advisor, pocketing the difference as corporate profit. CMTP advisors – in contrast to LGSP trainers or municipal advisers – had a relatively free rein, and while they were reporting on activities did not feel the need to meet certain

133 Interview 79, 8/5/13
134 Boldfacing mine.
135 Interview 103, 7/20/13
136 Interview 92, 7/24/13; interview 103, 7/30/13.
137 Interview 103, 7/30/13
targets, certainly not at the expense of impact. One implementer said that with DFID reporting was “more content-rich; it was not a numbers game.”

One might say that the judgment of these advisors in many ways guided the project as experienced by the individual municipalities, a process supported by DFID. This may have played a role in shaping the view by some DPLG and SA Government actors that the USAID’s LGSP was more “donor driven” than DFID’s CMTP, despite RTI’s substantial consultations with DPLG regarding LGSP.

Revision

As noted above, both these projects changed substantially over the period of implementation. These alterations responded to a changing South African political environment (including the emergence of the Government/DPLG’s Project Consolidate) as well as funding priorities from DFID and USAID headquarters. The different structures, and navigation strategies, of CMTP and LGSP meant they responded to these stimuli in different ways.

---

138 Interview 72, 7/31/13; Interview 74, 7/30/13
139 Interview 82, 8/1/13
140 Interview 102, 8/13/13; interview 97, 8/15/13; Interview 90, 8/2/13
CMTP had been designed in such a way as to give the steering committee that managed the project the authority to revise the project; while DFID sat on this steering committee, the majority of its members were Government actors. Deloitte, by far the actor most invested in the project, seems to have actively managed the steering committee. Flexible funds had been built into the project at design, and Deloitte more than any other actor guided the direction of these funds. This included directing procurements to Deloitte itself, a practice that deservedly raised eyebrows in the independent evaluation of the project. (Linkd 2009) CMTP also responded to Government pressure by altering the program in particular areas; the municipalities where Government pushed back are also some of the locations that are pointed to as particular successes by CMTP personnel and independent reviewers. CMTP advisors in some sense adapted the support on the ground in ways that – given LGSP’s more centralized structure – would have required formal revision for LGSP.

LGSP was less able to adjust in response to information from the field, to soft information. Despite feedback that the initial trainings weren’t working, LGSP didn’t revise the frequency or targeting of the trainings. It seems reasonable to connect the lack of revision in part to LGSP’s Navigation by Measurement and pursuit of quantitative targets.

141 Interview 75, 7/25/13; interview 103, 7/30/13
142 Interview 95, 8/1/13; interview 92, 7/24/13; interview 76, 7/29/13
143 Interview 74, 7/30/13
144 Interview 88, 8/1/13
LGSP did, however, change substantially. One significant alteration was LGSP’s response to Project Consolidate and the move towards advisors; this was largely due to communication between SA Government and high-level USAID actors\textsuperscript{145} rather than project-level implementers, and as such is indicative of modification, and Navigation, from the Top. LGSP accomplished this shift via adding additional funding, approved from above; to have altered LGSP’s scope without increasing its funding would likely have required re-tendering the original contract.\textsuperscript{146} Indeed, LGSP’s substantial alterations – which included, by the end, the operation of Thuthuzela Care Centres for sexual assault survivors, a project hard to reconcile with LGSP’s original mandate - seem to largely have been driven by additional funding streams and changing priorities from USAID senior leadership and HQ-based actors.\textsuperscript{147}

Frequent, unprompted reference was made by actors involved in LGSP to the difficulty of getting approvals from above. From USAID annual budget requests (transmitted via the US ambassador) to shifting funds that were underspent in one area to

\begin{flushright}
\textsuperscript{145} Interview 102, 8/13/13; interview 87, 9/20/13
\textsuperscript{146} Interview 78, 7/22/13
\textsuperscript{147} Interview 87, 9/20/13; interview 78, 7/22/13; interview 81, 8/1/13; interview 85, 11/12/13
\end{flushright}
another to revising activities, process loomed large in the minds of those recalling LGSP.\footnote{Interview 85, 11/12/13; interview 78, 7/22/13; interview 87, 9/20/13; interview 103, 7/30/13} No actor ever mentioned any such complexity as regarded CMTP.

While in these projects we see differential navigation strategies driving revisions – LGSP based on Navigation from the Top, CMTP incorporating Navigation by Judgment - it is not clear that the substantial alterations of either project were necessarily for the best. CMTP was described by one actor as a "projectised ATM" by the end, just dispensing cash with little to show for it.\footnote{Interview 92, 7/24/13} LGSP was similarly described as a series of only loosely connected project elements by the end.\footnote{Interview 81, 8/1/13}

**Differential Municipal Governance Capacity Building Navigation in Sum**

DFID clearly allowed the judgment of the actors it had empowered to guide CMTP’s design process to a greater degree than did USAID regarding LGSP. USAID settled on an initial model that delivered monitorable, measurable training from the outside. This difference continued in implementation and revision, where Navigation from the Top – particularly, in implementation, Navigation by Measurement – played a substantial role in USAID’s intervention but little in DFID’s, where Navigation by Judgment continued to prevail.
While neither of these projects was an unmitigated success, it does seem that the project which Navigated by Judgment to the greater extent proved the more successful, in part because of the ways DFID’s Navigation by Judgment crowded in soft information and USAID’s Navigation from the Top forced attention to constraints from above and away from the environment in which the project occurred. This is not to say there were no costs to CMTP’s greater flexibility and autonomy; one potential such cost (Deloitte’s substitution of lower quality consultants, and thus less expensive, consultants for those originally envisaged) was, in the view of one respondent, one of the key reasons CMTP did not accomplish more than it did. While it is tempting to conclude that CMTP’s failings are a necessary cost in pursuing Navigation by Judgment, it is not at all clear that this is the case; DFID could have been a more active player in LGSP’s management without preventing Deloitte and CMTP from utilizing soft information and judgment in the functioning of the project.

While there may be costs to both navigation strategies, it is clear that in these projects DFID’s relatively greater Navigation by Judgment was integral to CMTP’s

---

151 Interview 103, 7/30/13. Speaking about working with DFID more broadly (not on CMTP), an implementer who had experience with both echoed the ways in which more flexible, judgment-laden reporting could be exploited by implementers, saying “sometimes with [DFID’s] flexibility people then take advantage... you can actually put certain things to a programme ... attribute [to the program] the things that have happened [in the country but not due to the program directly]... [with USAID] that’s much more harder because you need to produce evidence, in terms of numbers because it was in quantitative terms not in qualitative terms.” Interview 82, 8/1/13

152 Interview 95, 8/1/13
comparative success; this strategy from design forward allowed the project to be guided by CMTP’s advisors, those on the ground who could access and incorporate soft information. Indeed, CMTP’s performance declined just as it shifted to a more centralized approach. On the other hand, LGSP’s relative shortcomings seemingly quite related to Navigation from the Top – to the need for countable trainings and items on which to report.\textsuperscript{153}

While municipal governance underwent some structural changes over this period, this seems not to speak to rapid changes and the need for flexible adaptation to same as much as environmental legibility to outside actors – that is, those other than advisors actually resident in the community and thus able to understand the dynamics of those communities. Even in relatively stable, monitorable South Africa the inherent difficulties in pursuing something as messy, as inherently difficult to quantify and manage to by rules, as municipal capacity building augured for Navigation by Judgment as the superior strategy.

\textbf{4.3: Health \& PMTCT}

South Africa has more people living with HIV than any other country; indeed, it has approximately twice as many HIV+ adults as the 2\textsuperscript{nd} place country, Nigeria. Best estimates are that in 2001 around 15\% of South Africa’s population 15-49 was HIV+, or some 4.3 million people; by 2012 this was approximately 18\% of the population, or 6.1 million

\textsuperscript{153} USAID’s next municipal governance program – Increasing Sustainable Local Governance Services (ISLGS) focused on measurable municipal outcomes like water and sanitation systems. This was markedly more successful, perhaps a sign that for USAID success is contingent on projects that match the range of the organization’s possible navigation strategies. This thread is touched upon in Chapter 6 (the concluding chapter).
people. (UNAIDS 2013a) Put another way, there are substantially more South Africans with HIV than Liberia’s entire population; HIV+ South Africans would constitute a country with a population above the median-sized country in the world.

This section focuses much, though not all, of its attention on efforts to prevent mother to child transmission, or PMTCT. This focus on an element of HIV/AIDS response is motivated by the scope of the HIV/AIDS epidemic in South Africa and the donor financing that has been mobilized in response; the US Government’s funding for HIV/AIDS alone was well over 500 million USD annually from 2008-2012 (PEPFAR 2015). As noted above, this funding flowed through both USAID and the US Centers for Disease Control (CDC), among other agencies; as such this section also compares the navigation strategies employed by USAID and CDC.

Navigation from the Top, particularly Navigation by Measurement, are particularly suited to PMTCT tasks. Under the treatment regimen advocated during the period examined here, mothers received treatment while pregnant and breastfeeding; newborns also received treatment for 4-6 weeks.\textsuperscript{154} PMTCT treatment is limited in time and has clear, verifiable, and externally observable outcomes: HIV negative children. As one interviewee with a long history in public health put it,

\begin{quote}
PMTCT is a no brainer because you have got a captive audience, you have got a clear intervention ... it is fairly quantifiable... you can get the
\end{quote}

\textsuperscript{154} That is, under what’s commonly known as “Plan A” and “Plan B” for PMTCT. In 2013 South Africa shifted, contentiously, to “Plan B+”, which has pregnant women continue on treatment indefinitely. One particularly clear explanation of the differences between these treatment regimens can be found in UNICEF 2012, page 3.
results for a particular clinic and demonstrate to people how they are doing, you know you have got a hundred mothers coming, you tested 98 of them, thirty are positive, how many of those did you put on treatment, how many of those did you take a CD 4 [blood test for severity of infection] count for... it is quite easily understandable in a short period of time, you are not following people for years and years and so it is a once off cohort.\textsuperscript{155}

This period provides a quite rare opportunity to investigate the interaction between navigation strategies and recipient government orientations. When the interventions examined here began – 2003/4 – South Africa was engaging in what later came to be known as “denialism” under South African President was Thabo Mbeki and Minister of Health Manto Tshabalala-Msimang.\textsuperscript{156} South African leadership disputed the link between HIV and AIDS, and advocated such remedies as garlic and traditional medicines for those who were HIV+. Denialism came to an end in late 2008, when Thabo Mbeki was replaced by Jacob Zuma and the Minister of Health became Barbara Hogan. Under both Hogan and her successor (Aaron Motosaledi, minister since mid-2009) the National Department of Health (NDOH) abandoned its previous obstructionism, pursuing an aggressive strategy of medical response to HIV/AIDS. These distinct periods allow for an examination of whether, and how, modes of navigation were affected by hostile recipient country environments.

\textsuperscript{155} Interview 117, 7/25/13

\textsuperscript{156} A nice overview of the denialist era is found in Geffen 2006
Intervention Descriptions

Beginning in 2004, the US President’s Emergency Plan for AIDS Relief (PEPFAR)\(^{157}\) had an initial allocation of 15 billion USD over 5 years, known as PEPFAR 1; in 2008 PEPFAR was funded for an additional 5 years and 48 billion USD, known as PEPFAR 2. (Kaiser Family Foundation 2014) It is unarguable that PEPFAR’s scope and ambition was unprecedented in the world of HIV/AIDS response – PEPFAR 1’s goals included treatment of 2 million people and provision of care to 10 million. (Kaiser Family Foundation 2014) PMTCT was a substantial focus of PEPFAR from the beginning; PEPFAR has consistently highlighted PMTCT achievements, including numbers of mothers treated and babies born to HIV+ mothers who are HIV-, in its annual reports. (e.g. PEPFAR 2013)

In South Africa – as in most countries – PEPFAR worked through multiple US agencies, including USAID and CDC, to deliver care; this formed the vast majority of the South African activities for these organizations, with one USAID official estimating PEPFAR as 95% of USAID’s work in South Africa.\(^{158}\) These agencies then make competitive awards to primary recipients - initially mostly international firms - who would in turn sub-contract to a variety of local partners via whom services would be delivered. PMTCT was integrated into the work of some PEPFAR partners, with others dedicated specifically to PMTCT

\(^{157}\) The comments below regarding PEPFAR focus largely PEPFAR implemented via USAID; however, in some cases it was not clear whether interviewees’ references to PEPFAR were as implemented by CDC, USAID, or both. As such I’ve used the generic PEPFAR moniker in the preceding sections.

\(^{158}\) Interview 134, 7/17/13
activities. PEPFAR’s annual Country Operational Plans – documents developed by US embassies & agencies summarizing annual in-country activities from all agencies – consistently show PMTCT spends above 50 million USD annually, all pointed towards a very clear overarching goal: reducing the transmission rate of the virus from HIV+ mothers to their children. (PEPFAR COPs various)

Even in the context of this massive expenditure, many interviewees made the point that South Africa did not, in fact, need donors’ money for health implementation; that said, 500 million USD a year is still a fairly substantial sum. One interviewee described being told by US officials that, had all PEPFAR-supported personnel worked for a single firm, that firm would be the fifth largest employer in South Africa, with over 60,000 employees; another described Right to Care – a South African PEPFAR implementer – as the highest funded US government aid recipient in the world.

PEPFAR was described as essentially a parallel health system both by those inside and outside USAID. This was particularly the case in PEPFAR 1; in PEPFAR 2, a greater attempt was made to implement in conjunction with national systems. By 2013 PEPFAR

---

159 Interview 110, 7/22/13; interview 112, 7/23/13
160 Interview 108, 7/25/13; interview 109, 9/12/13; interview 141, 7/16/13
161 Interview 117, 7/25/13
162 Interview 118, 7/18/13
163 Interview 109, 9/12/13; interview 111, 8/12/13; interview 127, 9/24/13
164 Interview 142, 6/14/13; also PEPFAR 2008d
had announced its intention to scale back in South Africa, and PEPFAR partners were working to actively build capacity in government systems and plan a full scale turnover to NDOH by 2017.\textsuperscript{165}

Over the time period PEPFAR has been operational DFID has, as one interviewee put it, been focusing on working through national government and influencing policymakers.\textsuperscript{166} DFID’s HIV/AIDS Multisectoral Program (MSP)\textsuperscript{167} began in 2003, with an initial budget of almost 20 million GBP, eventually growing to 30 million GBP. (DFID 2006a) The project was designed with the goal of “strengthening South Africa’s national HIV and AIDS response and reducing the impacts of the epidemic” (DFID 2008a, p. 6); it aimed to do this by providing technical assistance to the departments of Health, Social Development, and Defence. (DFID 2008a) The project additionally funded a range of NGOs implementing services and advocating for a change in national policy,\textsuperscript{168} as well as research around PMTCT\textsuperscript{169} among other matters. The project was begun in 2003 with Constella Futures (a DC-based aid contractor) managing the project and Charles Kendall (a UK-based logistics and aid consulting firm) as the procurement agent. These actors were then replaced by HLSP in late 2007.\textsuperscript{170}

\textsuperscript{165} Interview 120, 8/14/13
\textsuperscript{166} Interview 141, 7/16/13
\textsuperscript{167} DFID Project Reference #104938
\textsuperscript{168} Interview 133, 10/25/13
\textsuperscript{169} Interview 132, 7/31/13
\textsuperscript{170} Interview 133, 10/25/13; interview 123, 8/15/13; interview 125, 7/31/13
Shortly after Hogan became Minister of Health in 2008, DFID added 15 million GBP in the form of the Rapid Results Health Fund (RRHF).\footnote{Interview 127, 9/24/13; interview 125, 7/31/13} While viewed by most as a separate project, it was formally added to the existing MSP; by project close, then, the MSP was up to 45 million GBP. (DFID 2009b) The RRHF was a flexible fund designed to assist Hogan in changing the department’s direction. One of the key elements of the RRHF was the Accelerated Plan for PMTCT (A-Plan), an NDOH plan to focus on PMTCT initially in eighteen of South Africa’s 52 districts; DFID funded the pilot program for the initial 5 districts.\footnote{Interview 126, 9/4/13; interview 123, 8/15/13; interview 133, 10/25/13} The A-plan sought to reduce mother to child transmission to less than 5% by 2011, with implementation coordinated by NDOH. (Republic of South Africa National Department of Health 2011)

After RRHF ended in 2009, implementation continued to be funded through its successor, Strengthening South Africa’s Response to HIV and Health (SARRAH).\footnote{DFID Project Reference #200325} SARRAH, a 20 million GBP project as of December 2012 (Griffiths et al. 2014), was in many ways an extension of RRHF; working through HLSP and an embedded advisor in NDOH (who was simultaneously DFID’s health advisor), SARRAH was flexible funding devoted to the priorities of NDOH and the Minister of Health as they evolved.\footnote{Interview 123, 8/15/13;
activities focused around strengthening strategic planning, management, and monitoring and evaluation.\textsuperscript{175}

**PMTCT Intervention Results**

The USAID and DFID interventions over the past fifteen years are of substantially different scales; they aim, however, towards broadly similar objectives, particularly as regards PMTCT: the reduction of mother to child transmission rates, and the establishment of a health system capable of sustaining those reductions.\textsuperscript{176}

On the direct impact side there has been much progress regarding HIV/AIDS treatment in South African in the past decade; progress on PMTCT has been impressive as well. Before 2007 the best available data suggested mother to child transmission rates of 8-9%, with a few studies suggesting substantially higher rates.\textsuperscript{(Goga, Dinh, and Jackson 2010)} By 2010 that had fallen to 3.5%, with the worst performing province at 5.9%. \textsuperscript{(Goga, Dinh, and Jackson 2010)}; by 2011 the rate had fallen to 2.7%. \textsuperscript{(Barron et al. 2013)} According to UNAIDS, between 2009 and 2013 the likelihood of pregnant and breastfeeding women receiving drugs to prevent transmission rose from 63 to 90%; the number of new child infections dropped by a little more than half in the same period, while

\textsuperscript{175} Interview 124, 8/5/13; interview 123, 8/15/13; interview 125, 7/31/13; also supported by the independent mid-term review of SARRAH (Griffiths et al. 2014)

\textsuperscript{176} PEPFAR arguably was concerned with national health systems only starting with PEPFAR 2; DFID’s interest in direct impacts was via improved national health systems.
the number of HIV+ women delivering babies remained virtually unchanged.\footnote{New child infections fell from 33,000 to 16,000 while births to HIV+ women fell from 270,000 to 260,000.} (UNAIDS 2013b)

While these clear outcomes make it easy to conclude that progress has been made, that USAID, CDC, DFID, the South African government, and many others have all pursued this goal still makes complicated how to apportion credit. Yes, PMTCT treatment has been successful – but what of these aid interventions specifically? As DFID pulls out of South Africa and PEPFAR winds down, have these interventions contributed to creating a health system likely to maintain, or even improve on, these gains in the absence of continued donor support?

\textit{Direct Impacts}

DFID’s initial effort from 2003-2008 – the Multisectoral Program (MSP) – was described by multiple actors as a ‘bet hedging’ strategy in response to the hostile South African political environment towards HIV/AIDS treatment; MSP was explicitly designed with the expectation that not all interventions would be successful.\footnote{Interview 133, 10/25/13; interview 127, 9/24/13} While very little progress was made with the Ministry of Health, the work with the Ministry of Social Development on home based care seems to have been relatively more successful.\footnote{Interview 133, 10/25/13; this view is also supported by DFID 2008a.} DFID’s MSP had a substantial PMTCT component; DFID argued that “the MSP can claim its
PMTCT support as one of its most important impacts.” (DFID 2008a, p 16) Specifically, the MSP piloted support for PMTCT via government clinics in three districts of KwaZulu-Natal province through the national government; this combined research and implementation project laid the seeds for the A-Plan to follow.180

While there were seeds of progress, the general view of MSP is overwhelmingly negative. Beset by management challenges, and requiring a change of implementer to avoid a projected 8 million GBP overspend (DFID 2009b) on a project with little to show for it even in the eyes of DFID personnel,181 it is hard to ascribe much by way of direct impacts to MSP. This view was echoed by those who led the project; Robin Gorna, the DFID health lead for DFID over the MSP’s latter period, remarked “I never felt we really got anything wonderful out of MSP”, though she felt there were some pieces of merit within the project. The RRHF, on the other hand, was a marked success – flexible funding under the guidance of a reform-minded minister supporting a range of activities, the RRHF is on the view of many actors one of the most helpful projects in recent years.182 Key among its successes was the A-Plan on PMTCT. Working within the Government – sitting inside NDOH – to develop a national strategy for PMTCT and building on the work done by the MSP (SARRAH 2010),183 the A-Plan is frequently mentioned as key in catalyzing NDOH’s response as well

180 Through the 20000+ component of MSP, implemented by the University of KwaZulu-Natal and the Provincial Department of Health. (MSP 2008)

181 Interview 133, 10/25/13; interview 127, 9/24/13; interview 123, 8/15/13

182 Interview 120, 8/14/13; interview 123, 8/15/13; interview 133, 10/25/13; interview 125, 7/31/13

183 also interview 125, 7/31/13
as providing direct care to PMTCT patients in target districts.\textsuperscript{184} DFID’s project completion report assesses the MSP and RRHF combined as the next to highest score on a 5 point scale.\textsuperscript{185} DFID’s project completion report supports the view of interviewees that a highly successful RRHF compensates for a less than successful MSP. The report argues that “While the outcomes of Outputs 1-4 [those focused on the MSP proper] were mixed in their levels of success, Output 5 in the last year of the programme – which almost exclusively focused on the RRHF – shifted the programme’s total impact.” (DFID 2009b, p.5)

SARRAH, while still ongoing as of mid-2013, has been less clearly impactful thus far. In some sense extending the model of RRHF, SARRAH has been criticized as being diffuse and unfocused,\textsuperscript{186} and the midterm review clearly signals that many components of SARRAH may not meet their targets. (Griffiths et al. 2014) The (now completed) continued support for PMTCT under A-plan – while a very small part of SARRAH – met its goals, however. (Griffiths et al. 2014)\textsuperscript{187}

\textsuperscript{184} Interview 125, 7/31/13; interview 120, 8/14/13; interview 117, 7/25/13; interview 124, 8/5/13. There is a dissenting view here from some NDOH actors, who feel the A-Plan was silo’d and ineffective. (interview 137, 8/12/13; interview 131, 8/12/13) It has been suggested that this bypassing of certain NDOH personnel was an intentional strategy (interview 126, 9/4/13).

\textsuperscript{185} As noted above, the RRHF was administratively technically a part of the MSP, though all existing MSP activities ceased by December 2008, with separate RRHF activities commencing. (DFID 2008a) As such the project score covers the RRHF activities as well as those of the MSP. The final score is a 2 on a 5 point scale, with 1 being highest; this score is taken from the database used for the analysis in chapter 2, released by DFID in response to a public enquiry March 4, 2013.

\textsuperscript{186} Interview 117, 7/25/13

\textsuperscript{187} Annex 13 focuses on PMTCT and the A-Plan; the report also notes that as of 2013 this intervention was projected to be 2% of SARRAH’s total budget.
If DFID's projects throughout this period focused on strategy and policy, USAID and CDC focused substantially on delivery, albeit with a significant research component as well. (PEPFAR COPs various) PMTCT was a significant highlight of the PEPFAR 1 period (2004-2008) with multiple USAID actors highlighting PMTCT work\textsuperscript{188} and PEPFAR’s South African reports stressing PMTCT results. (PEPFAR 2008a)

In the period of denialism PEPFAR’s existence outside government and ability to directly fund implementers meant substantial progress in terms of lives saved and services delivered even as national NDOH dallied. By 2008, when South Africa’s denialism came to an end, 85\% of pregnant mothers were receiving HIV counseling and testing as compared with 46\% in 2004;\textsuperscript{189}(PEPFAR 2008b) this increase is surely attributable in substantial part to PEPFAR’s structure and delivery channels, with PEPFAR reporting 2,211,000 pregnant women receiving counseling and 333,100 HIV-positive pregnant women receiving drugs for PMTCT. (PEPFAR 2008a) As one senior member of South Africa’s medical community put it, “had it not been for PEPFAR and government just had the drugs, we would have treated a few patients and we would have carried on, but we would have been, I would say, ten years behind.”\textsuperscript{190}

\textsuperscript{188} Interview 134, 7/17/13; interview 109, 9/12/13

\textsuperscript{189} These numbers are for counseling and testing and thus are measuring something distinct from the UNAIDS statistics quoted above, which focus on drug delivery.

\textsuperscript{190} Interview 128, 7/24/13
Indeed, even DFID's A-Plan required PEPFAR's support for implementation, for actual delivery to pregnant mothers. As one individual involved in A-plan put it, "At the end of the day we all depended on PEPFAR actually to do the implementation support because they are the ones that are on the ground and they have all the resources and all the people in each of the districts." ¹⁹¹ NDOH officials and non-PEPFAR beneficiaries echoed this, talking about PEPFAR's deep and broad involvement in PMTCT, including in A-plan implementation.²⁰²

Systemic Effects

The relative straightforwardness of PMTCT noted above – its limited time period (pregnancy) and straightforward intervention – makes it not only more tractable to measurement but also easier for a public health system where capacity is wanting to implement. As one South African public health leader put it,

There was a recognition that this was a relatively solvable problem and there was a clear scaling up, a recognition that resources were directed at it, there was a keen championing from the national department [NDOH] and there was an on-going attention paid to data and results and it is easily measurable, so if we couldn't do this it would be a complete indictment of South Africa's health system, you know, all the elements were in place to do it, you have got a very high attendance at antenatal clinics, you have got an excellent laboratory network which can give you results, you have got the infrastructure around the country, it was just a matter of

¹⁹¹ Interview 124, 8/5/13
²⁰² e.g. Interview 137, 8/12/13; interview 117, 7/25/13; interview 125, 7/31/13
getting the drugs out, testing them and then getting them [pregnant mothers] to take the drugs.  

If PMTCT is – from a national health perspective – the ‘low hanging fruit’ of HIV/AIDS response, it is nonetheless fruit that was unplucked, and withering on the vine, when PEPFAR and MSP began. PEPFAR was instrumental in demonstrating that resources and effort could lead to progress on PMTCT, particularly during the denialist era when DFID had little to show for their government-focused support.

DFID’s financially much smaller support is that pointed to by many actors as having been fundamental to as helping government build its response, with the A-plan in particular critical support without which NDOH would not be where it is today. At first glance it is somewhat surprising that DFID’s input was a necessary one; that PEPFAR did not ‘go it alone’ in providing this broader systemic impact, given its substantial implementation budget and research activities. A number of interviewees suggested USAID could have done more, particularly as regards sharing data with NDOH. One interviewee noted that USAID had all the data available to help government actors develop management systems but had failed to do so.

---

193 Interview 117, 7/25/13
194 Interview 119, 8/12/13; interview 113, 8/14/13
195 Interview 137, 8/12/13; interview 124, 8/5/13; interview 120, 8/14/13
196 Interview 111, 8/12/13; interview 126, 9/4/13; interview 124, 8/5/13
197 Interview 117, 7/25/13
Now national policy, what began as the A-plan will, as PEPFAR winds down, end up being the backbone of PEPFAR’s ‘exit strategy’ as regards PMTCT.\textsuperscript{198} This transition for PEPFAR is already underway, with both USAID and CDC shifting towards technical assistance and support of government implementers per the general thrust of PEPFAR 2.\textsuperscript{199} At the same time, the A-plan relied on PEPFAR to initially implement, as noted above; PEPFAR 2’s focus on capacity building and transition\textsuperscript{200} is part of what may strengthen the ability of national health actors to actually maintain the system.

Though it is difficult to speak definitively about the system’s sustainability while PEPFAR funds continue to flow in the quantities they did at the time of these interviews,\textsuperscript{201} it seems fair to conclude that USAID & CDC PEPFAR support as well as DFID’s A-plan (less so other DFID interventions) have had substantial long run impacts. Whether either the US or UK funding stream would have prospered without the other is a difficult question to answer; suffice to say they were mutually reinforcing.

\textit{Barriers to Success & Comparative Performance}

The greatest barrier to progress from 2003-2008 was unquestionably the hostility of national government; this hostility had much more of an impact on DFID’s support than that of USAID and CDC. More than one respondent linked PEPFAR’s success during this

\textsuperscript{198} Interview 120, 8/14/13; interview 118, 7/18/13; interview 117, 7/25/13; interview 125, 7/31/13

\textsuperscript{199} Interview 121, 8/5/13; interview 130, 7/26/13; interview 114, 8/19/13

\textsuperscript{200} Interview 113, 8/14/13; interview 111, 8/12/13

\textsuperscript{201} mid-2013, broadly
period to its focus on targets, on numbers.\textsuperscript{202} This makes a good deal of sense; Navigation by Measurement might help resist political winds, keeping the system oriented towards meeting the numbers – towards delivering the targets.

DFID fared worse in the first (MSP) period – which, given the spirit of denialism at the time, may have been inevitable in any strategy supporting government. Yet in recent years DFID’s projects again seem to have once again been less than fully successful, with some criticism of SARRAH (including independent evaluations, i.e. Griffiths et al. 2014) suggesting the project may be when being too responsive – too flexible. DFID’s initial project – the MSP – was also beset by management difficulties to a much greater degree than USAID & CDC’s.\textsuperscript{203}

With the exception of the A-plan begun under the 1-year RRHF – a period in which DFID reached a level of success potentially comparable to that of USAID/CDC, with DFID’s focus on systems complementing PEPFAR’s direct delivery – DFID’s projects have performed less well throughout.

\textbf{Differences in Navigation Strategies}

Much as was the case for municipal capacity building, it remains an open question whether DFID’s relative lack of success is related to the organization’s navigation strategy.

\textsuperscript{202} Interview 109, 9/12/13; interview 111, 8/12/13; interview 128, 7/24/13
\textsuperscript{203} Interview 133, 10/25/13; interview 127, 9/24/13; interview 125, 7/31/13
The discussion that follows seeks to investigate differences in navigation strategies and how these strategies related to the success, or lack thereof, of each project.

Design

The design process for the RRHF was less a process than an act – a phone call from Robin Gorna, then DFID’s health adviser based in South Africa.\textsuperscript{204} Gorna’s account is that, having received incoming Minister Hogan’s number from a mutual friend she

Picked up my phone to Barbara [Hogan] one night from my kitchen, I remember it very distinctly and said “Hello, I’m a friend of Zackie’s,\textsuperscript{205} could I help you with fifteen million pounds, could you use it, would it help you strengthen things?” Strangely she said yes, and this again is a delightful part of how DFID has been able to work in the past that’s just a little anecdote I cherish but it’s also part of how we as health advisors traditionally had been able to work, this was a rather extreme experience, but I think that a quality of DFID health programmes traditionally was that individuals could spot something and then work it up into a bureaucratically acceptable proposal.\textsuperscript{206}

This is a clear example of design borne of Navigation by Judgment. While approvals were then needed by DFID London for expedited and flexible procurement and the awarding of the contract to HLSP (via MSP) to implement rather than re-tendering,\textsuperscript{207} both the

\begin{flushleft}
\textsuperscript{204} This is slightly exaggerated, of course; following the call there was a bureaucratic process of DFID approving this project and providing waivers to normal tendering processes.

\textsuperscript{205} The mutual friend was Zackie Achmat, head of the Treatment Action Campaign (TAC); an advocacy nonprofit, the TAC famously sued the Government of South Africa to force government initiation of PMTCT treatment, winning their case in 2002.

\textsuperscript{206} Boldfacing mine.

\textsuperscript{207} Interview 125, 7/31/13; interview 127, 9/24/13; interview 133, 10/25/13
\end{flushleft}

170
substantive design and management decisions were made entirely by actors on the ground in South Africa. The design of SARRAH was much the same; design decisions taken entirely in South Africa were then approved by London, with flexible and adaptable projects the result.208

Decisions about PEPFAR design and activities, on the other hand, seem to “all happen in DC”, as one Government actor put it.209 Global targets are set by the Office of the Global AIDS Coordinator in DC;210 those targets then flow through national annual Country Operational Plans (developed by the US Government agencies in country and submitted via the US Ambassador to the country for approval by the Office of the Global AIDS Coordinator) to a series of tenders. In-country personnel have some role in setting direction,211 but ultimately country strategies need to address these global targets. As one senior DC-based PEPFAR official put it,

There’s a lot to be explored here in terms of ownership around goal setting. In some instances, like the Global Fund, the countries which develop the proposals have more ownership in setting goals. In others, like PEPFAR, “targets” are set from above, ie the Global AIDS Coordinators Office in DC. In PEPFAR 1, the original target was 2 -7 – 10. Put 2 million people on treatment, prevent 7 million new infections, and put 10 million people in care. The targets for PEPFAR 2 are 6 – 10 – 10. When targets are set from above, everything then flows down to the agency level, country level, and partner level, so all the funds that we are given, at the different levels, have to add up to meet these arbitrary goals set from above and

208 Interview 127, 9/24/13; interview 123, 8/15/13

209 Interview 107, 8/2/13

210 Interview 142, 6/14/13; interview 130, 7/26/13

211 Interview 122, 8/5/13; interview 130, 7/26/13
there is a heavy quarterly, semi-annual, and annual reporting burden and a complex and arduous annual target setting and planning process to try to meet the goal.\textsuperscript{212}

This Navigation from the Top was intentional; the original PEPFAR strategy makes clear that “the United States Global AIDS Coordinator will oversee and direct all US Government international HIV/AIDS activities in all departments and all agencies of the Federal Government... this will result in more effective and efficient programs.” (PEPFAR 2004, I.1) A few lines later, PEPFAR commits to “establish measurable goals for we will hold ourselves and our partners accountable.” (PEPFAR 2004, I.1)

The navigation strategies that inform the design processes of these UK and US interventions are fundamentally different, consistent with the differences in political authorizing environments and consequently navigation strategies hypothesized in Chapter 2.

\textit{Implementation}

What is true in design is also true in implementation. Interviewee quotes on PEPFAR include “PEPFAR is about those targets, and we will fund you if you can demonstrate you can help us with those targets but that’s all that it’s about, nothing else matters”;\textsuperscript{213} “[PEPFAR] found much a bigger emphasis on the quantitative outputs”;\textsuperscript{214} and

\footnotesize

\textsuperscript{212} Email communication further to interview 142, 6/14/13; boldface mine.

\textsuperscript{213} Interview 125, 7/31/13; under PEPFAR 1 these targets were in fact different than those of the National Department of Health, further underscoring the parallel-ness of the health system. (Interview 124, 8/5/13)
“[with PEPFAR] you had to measure, quantify; and it is quite rigorous reporting guidelines, requirements, that the budget had to matched against”.\textsuperscript{215} The prevailing view was of this reporting as burdensome, costly, and unnecessary; one implementer described it as “donkey work” that served no productive purpose, suggesting PEPFAR would do better with “less attention to minutae”.\textsuperscript{216}

That this Navigation by Measurement was unwelcome does not mean it was unproductive; indeed, it is hard to imagine how else an operation of PEPFAR’s scope could have been managed coherently. One implementer, by way of praising PEPFAR, said “PEPFAR’s been great in the sense that it’s brought the American obsession with counting things to kind of a fuzzy field full of cuddling.”\textsuperscript{217}

Consistent with the discussion in chapter 2, this target-setting in the context of reliable outcome measures not susceptible to Goodhart’s Law (e.g. patients treated) did promote some process autonomy by contractors when implementing. As one implementer put it

   Whatever your number was -- if you said you would do fifty thousand circumcisions, if you contracted with the private sector, whether you went through a sub-grantee, or whether you did it yourself, or whether you went through public health they didn't care they just wanted the fifty thousand

---
\textsuperscript{214} Interview 128, 7/24/13
\textsuperscript{215} Interview 112, 7/23/13
\textsuperscript{216} Interview 108, 7/25/13
\textsuperscript{217} Interview 111, 8/12/13
circumcisions done. And I think that was useful because as I said different people brought different strengths.218

Consistent with this work’s theory, this same implementer noted PEPFAR’s target setting was less effective when dealing with outcomes more susceptible to distortion, saying “One of the things I am critical about is the focus on training because there’s lots of, and it’s something like put fifty backsides on the seats for three hours and you can count them…. You would count the people training but what you didn’t count was what the impact of that training.”219

By contrast, the design for MSP, RRHF, and SARRAH have a common thread that seems incompatible with this Navigation by Measurement: allowing for actors to find, and exploit, opportunities. The view of DFID’s implementation as guided by the judgment of its field-level agents was echoed by a number of actors.220 Communication with London was only around “the political stuff... everything [else] was left to the country office to work it out.”221 Reporting under DFID’s interventions involved some measurement of activities, but there didn’t seem to be pressure on these figures; narrative accounts of progress and the judgment of DFID supervisors and Government counterparts seemed to implementers to

218 Interview 111, 8/12/13
219 Interview 111, 8/12/13
220 e.g. Interview 126, 9/4/13; interview 133, 10/25/13; interview 123, 8/15/13; interview 107, 8/2/13; interview 121, 8/5/13
221 Interview 133, 10/25/13; this view was echoes by interview 123, 8/15/13
matter more than reaching specific targets. As Gugu Ngubane, the lead implementer of the A-Plan under RRHF and SARRAH, put it, DFID “allow you to be the expert in your field and you do what needs to be done… they were supportive and they just let me inform them, I think they gave me the platform to advise them on what is the reality, what can happen, what cannot happen.”

DFID’s Gorna – quoted at length above – argued that MSP’s failure was borne in part from an inappropriate contracting out of what this work calls Navigation by Judgment, in so doing underscoring her belief in and DFID’s enablement of Navigation by Judgment. As she put it

DFID advisors tend to have a particular training which makes them really good at spotting opportunity and creativity and getting in behind exciting new opportunities, when you get caught up in this kind of contracting process of working through consultants, it doesn’t always follow through quite as beautifully… this talent-spotting type of design was great if you had an advisor who was going to be able then pick it up and make it move, but when you contracted out that responsibility, you sort of, you couldn’t contract out the intelligence in the same way.

It is hard to imagine how a project like MSP, RRHF, or SARRAH – with their flexible funding and lack of pre-specification – could possibly persist under the Navigation from the Top/Navigation by Measurement regime of PEPFAR. As one actor involved in implementing under both PEPFAR and DFID’s SARRAH put it, “the DFID way of doing

---

222 Interview 139, 7/31/13; interview 125, 7/31/13interview 123, 8/15/13; interview 124, 8/5/13; interview 126, 9/4/13
things has facilitated that, this [SARRAH] type of work wouldn't, couldn't happen under a PEPFAR type structure.”

Revision

Speaking generally about DFID’s penchant for flexible and responsive health projects, one respondent said that as a result project success “is based on charisma, it depends who is in the [DFID and NDOH] positions at the time. If it is somebody who is well meaning and whose heart is in the right place and who is clear thinking, then it [the project’s success] is probably going to be good, but if it is somebody who is muddled, you know you can get money being dispersed and not strategically used.” As actors inside DFID and NDOH changed so did project focus. When Gorna was replaced as DFID’s health advisor SARRAH was changed substantially by her successor. This has led, on one interviewee’s view, to a mis-match between SARRAH’s HIV-oriented official goal (changing the goal requires re-tendering in DFID’s system) and the health systems-focused outputs (outputs can be altered in a continuing project). Another individual with knowledge of SARRAH described, positively, this transition as “a classic example I think of a development programme where individuals and development agencies can influence it.”

---

223 Interview 125, 7/31/13

224 Interview 117, 7/25/13

225 Interview 132, 7/31/13; interview 117, 7/25/13; interview 127, 9/24/13; interview 123, 8/15/13

226 Interview 132, 7/31/13

227 Interview 125, 7/31/13
of contractors for MSP – a highly unusual move for DFID - was similarly driven by DFID personnel shifts; namely Gorna’s arrival as health advisor.228

While these changes may well have been positive, a design which provided more rules and less flexibility in revising projects – in other words, more Navigation from the Top – might have been of benefit to DFID’s MSP and SARRAH. It would have avoided what one respondent viewed as SARRAH’s effective capture by a particular NDOH official and consequently loss of coherence.229 This view was echoed by someone involved in DFID project administration over the period, who said by way of criticizing DFID’s penchant for alterations at the request of NDOH “Sometimes you do want a funder that will just stick to the contract and say no, boldly to the department [NDOH] irrespective of the political relationships because if it is not in the contract, it is not in the contract.”230

PEPFAR’s process autonomy should not be misconstrued as flexibility or adaptability. Top-down target setting had implications for workplan revision; as one implementer put it, “the way PEPFAR is done where you have to plan your work a year ahead of time.”231 In general revisions to targets or the five year strategy proceeded very

228 Interview 127, 9/24/13; interview 125, 7/31/13
229 Interview 117, 7/25/13
230 Interview 124, 8/5/13
231 Interview 125, 7/31/13
slowly, if at all. That said, some alterations were possible, and additional funds could be found. PEPFAR, then, parallels LGSP in that substantial revision could only happen from the top.

Comparing CDC and USAID

As noted in the opening section of this chapter, PEPFAR allows a unique opportunity to look at how USAID and CDC compare in their PEPFAR processes in South Africa. Reporting ultimately to the same US Congress and supervised by the same Office of the Global AIDS Coordinator some might expect little from this comparison, thinking the constraints of PEPFAR should lead to very similar implementations. My view is, as articulated in chapter 2, that effective political authorizing environment slack, and thus agency autonomy, differs by agency and within agency over time.

As such CDC & USAID may face the same formal DC hierarchy (Congress & OGAC) but nonetheless have different effective authorizing environments, levels of autonomy, and freedom to engage in Navigation by Judgment. This is, arguably, precisely the case here. CDC and USAID are structurally similar; CDC is a unit of the Department of Health and Human Services just as USAID is a unit of the Department of State. They however have very

---

232 Interview 137, 8/12/13; interview 122, 8/5/13; interview 130, 7/26/13; interview 121, 8/5/13

233 Interview 120, 8/14/13; interview 137, 8/12/13

234 e.g. were USAID to become an independent agency again with independent control over its budget, I would predict this greater autonomy would translate to a greater use of Navigation by Judgment where appropriate.
different reputations, which Carpenter (2010) among others has argued is consequential for the power to act independently. The CDC sees itself as under substantially less scrutiny than USAID; staffed by doctors and scientists, CDC exudes an aura of authority that USAID lacks. The CDC is also in distant Georgia rather than DC, which may also assist in providing additional operating slack for the agency. I would expect, then, that the CDC would have more autonomy – and thus a greater propensity to engage in Navigation by Judgment – than USAID.

CDC PEPFAR implementation was described as more flexible by a number of respondents.235 One implementer of both CDC and USAID PEPFAR funds described USAID as “2/3 monitoring, 1/3 implementation”; CDC as “1/3 monitoring, 2/3 implementation”.236 This implementer added that for USAID “everything is by the book... there’s a lot of legality around USAID – there are more regulations, things are more stringent than CDC”.237 CDC was described as less formalized, for better238 or for worse.239 CDC awards were also more hands on; CDC tended to be actively involved with implementers in a way USAID was not.240

---

235 Interview 112, 7/23/13; interview 117, 7/25/13
236 Interview 116, 8/16/13
237 Interview 116, 8/16/13
238 Interview 122, 8/5/13
239 Interview 111, 8/12/13
240 Interview 122, 8/5/13; interview 121, 8/5/13
This more fluid working relationship – closer to DFID’s relationship with contractors in some ways than USAID’s – allowed, in the view of one respondent, for more responsiveness for contractors when things arose that were not foreseen in the original contract; that is, for something like minor informal revisions of agreements. CDC also proved in the eyes of a number of interviewees more responsive to South African government requests and to working through government; this is another way in which CDC’s working relationships, and navigation during implementation, leaned towards DFID’s flexible informality relative to USAID’s rule-based Navigation from the Top. This is despite the fact that field staff in CDC are actually – contrary to what I might have predicted – the less empowered. CDC staff need to return to Atlanta for far more approvals than do USAID staff to DC.

CDC and USAID also differed in a number of other dimensions. CDC’s staff tended to be doctors and nurses and the organization had a priority on research that USAID did not. USAID was praised for its understanding of community work, an area in which CDC was seen as less adept; the opposite was true of working with hospitals. CDC resisted credit claiming, not allowing its logo on anything unless it meets standards of scientific

---

241 Interview 122, 8/5/13
242 Interview 121, 8/5/13; interview 137, 8/12/13; interview 123, 8/15/13
243 Interview 116, 8/16/13; interview 117, 7/25/13; interview 122, 8/5/13
244 Interview 130, 7/26/13; interview 116, 8/16/13; interview 122, 8/5/13
245 Interview 116, 8/16/13
rigor; USAID on the other hand encourages the use of its logo on reports, vehicles, etc. These differences are suggestive of the differential institutional histories of the two organizations; of the CDC as a disease research institution and USAID as one relating to development management. The differential organizational foundations seem to have path-dependent consequences for implementation.

The modal response when interviewees were asked to compare USAID and CDC was to point out the conflict between the organizations, strongest in the early days of PEPFAR. Examples included implementers of one organization being instructed not to attend seminars or workshops hosted by the other as well as direct criticism of one organization by the other. The arrangement in South Africa as of mid-2013 was of geographic separation; USAID led in some geographic areas, CDC others.

---

246 Interview 129, 8/1/13; interview 121, 8/5/13; interview 122, 8/5/13

247 A similar point was occasionally made about contracting firms, with those from academic origins described as weaker on systems and financial management but better on understanding data. (Interview 143, 7/10/13; interview 142, 6/14/13)

248 e.g. interview 114, 8/19/13; interview 115, 8/14/13; interview 111, 8/12/13; interview 116, 8/16/13

249 Interview 115, 8/14/13; interview 114, 8/19/13

250 Interview 111, 8/12/13

251 Interview 130, 7/26/13 It seems almost trivial to point out that this is almost certainly inefficient from the perspective of comparative advantage; if USAID and CDC have relative strengths, they should leverage them. If those comparative strengths are more than offset by the conflict between the organizations that might result, that would seem to argue for the logic of having one or the other leading in South Africa (or on PEPFAR more broadly), but not both.
In any case, CDC and USAID differ. There is also suggestive evidence that they differ in their propensity to Navigate by Judgment – to have flexible arrangements and be less bound by formal rules and regulations. This greater informality by CDC, the less insecure organization, is what the theory in chapter 2 would imply. CDC’s greater centralization and concentration of approvals in distant HQ is the opposite of what might be predicted, however. That CDC is nonetheless the more flexible and more Navigation by Judgment-prone organization suggests that constraints at the top – from political authorizers vis-à-vis HQ - may be more predictive of realized organizational navigation strategy than the formal structures within the agency itself. While exposure to soft information regarding which judgments can be made may be critical, actual formal decentralization of authority may not be determinative of relative flexibility and organizational navigation strategy.\footnote{252}

Differential Health Navigation in Sum

The differences in navigation strategy between PEPFAR (focusing largely, though not exclusively, on USAID) and DFID’s arc of support are quite clear – USAID and PEPFAR more broadly engaged in Navigation from the Top, with Navigation by Measurement the primary mechanism for driving the focus of the system towards the targets which flowed down from the Office of the Global AIDS Coordinator. DFID, on the other hand, engaged in flexible funding driven by the judgments of resident personnel.

\footnote{252 Sections below and in Chapter 5 further explore whether, and how, formal structures on the one hand and political authorizing environments on the other affect field-level behavior and IDO navigation.}
It is possible – following Gorna – that had DFID engaged in Navigation by Judgment even more fully and directly (rather than via contractors) the MSP, by far DFID’s most underwhelming project in this support arc, would have been more successful. The mixed performance of SARRAH – implemented with the DFID health advisor embedded within the National Department of Health and effectively leading implementation – suggests otherwise, however.

In this case Navigation from the Top proved the superior navigation strategy despite its costs in terms of adaptability to changing circumstances and the burden of reporting it placed on implementers. That this is the case is consistent with the theory articulated in Chapter 2; in a relatively predictable environment with externally observable and stable outcomes Navigation from the Top was successfully able to drive performance from above via targets and measurement.

At their best, PEPFAR’s Navigation from the Top and DFID’s Navigation by Judgment interacted via the A-Plan, with both DFID and PEPFAR proving critical to the expansion of PMTCT efforts and transition of same to national government leadership. Whether PEPFAR’s Navigation by Measurement is compatible with the capacity building and government systems-building efforts to which by mid-2013 it was committed to transitioning is an open question. In the context of South African PMTCT – where clear, relatively easily collected metrics of success are resistant to Goodhart’s Law problems in a

253 That is, not susceptible to Goodhart’s Law
context of good data systems and relatively predictable environments - this work’s theory would predict that this transition may well be successful.

4.4: Investigating Channels: The Flow of Constraint & Impacts of Same

The data these cases constitute is useful for investigating another element of the theory hypothesized in chapter 2. Beyond the differences in navigation strategies and project success, these cases help to elucidate the ‘flow’ of constraint from political authorizing environments to the field and the impact of same on agent behavior and quality. That USAID is the organization less likely to Navigate by Judgment doesn’t necessarily mean that it is constrained by the need to manage up to political authorizing environments; it might be the case that USAID simply believes Navigation from the Top the superior strategy. It was, after all, the superior strategy in USAID’s most consequential intervention discussed in this chapter – PEPFAR.

The South Africa interview data suggests otherwise. Political authorizing environments loomed larger for USAID than DFID, constraining autonomy and substantially influencing USAID towards top-down & measurement driven navigation.

---

254 That is, both having the greatest impact and of greatest perceived importance for USAID.
In conducting South Africa interviews I never once mentioned the word “Congress” prior to the respondent’s mention of same; nonetheless “Congress” appears at least thirteen times in respondents’ answers. By contrast, the word “Parliament” appears only once. This case is itself telling - an interviewee with knowledge of both organizations had raised USAID’s reporting to Congress. I asked whether DFID staff ever referred to Parliament. He responded

No I never, no people didn’t talk about, I mean [name redacted] would time and again, maybe talk about the Department of International Development [DFID], you know meaning in London, but it wasn’t a dominant sort of talk about, “we are all doing this so that we can report to the House of Commons” or something like that in the UK Parliament. But the US[AID] is always about “the Congress wants these numbers because they are providing so much for South Africa and they want to know what is it that we have done with their money.”

This interviewee comes from the municipal governance, rather than the health, world in South Africa; this makes his comment about USAID Navigation by Measurement all the more striking, inasmuch as this comment is likely not driven PEPFAR’s top-down orientation or high congressional priority. His view also does not stand alone; numerous USAID implementers and staff variously linked reporting and measurement regimes, project rigidities, and general constraint to US politics generally or Congress specifically. Michelle Layte – quoted above regarding USAID’s focus on numbers to the detriment of impact towards the end of LGSP, USAID’s municipal governance intervention – attempted

---

255 Interview 82, 8/1/13

256 Interview 78, 7/22/13; interview 97, 9/20/13; interview 142, 6/14/13; interview 113, 8/14/13; interview interview 108, 7/25/13
to explain the constraints USAID staff in South Africa faced by saying “Overall I must say they [USAID local staff] were very supportive of what the challenges were and what the changes were but at the end of the day they had to report back to Congress about how the money was spent.” The implication here is clearly that in the absence of Congressional constraint, field level personnel would have had more flexibility to focus on impact rather than numbers, to Navigate by Judgment rather than engage in top down Navigation by Measurement.

Congress also came up regarding the relationship of USAID as an organization to Congress. As Steve Snook, former USAID municipal governance staff, put it for USAID “it is all about the budget, right, the budget battle and the dream of federal agencies is that they could get, put together projects that year in and year out, programmes that year in and year out would have predictable funding, but every year the budget fight is a new adventure”. This “adventure” speaks directly to the insecurity of USAID’s funding and the need to justify itself to Congress. This insecurity also shows up in staffing decisions – in asking why a temporary contractor might be used within USAID (that is, working inside USAID supervising projects)\(^{257}\) rather than permanent USAID staff, one respondent replied

Well, that [all permanent staff] is in the ideal world but **when you have a hostile Congress** and an ineffectual president which was the case pretty much since, I think the last time USAID had any true swagger was under Reagan, **it has been an agency under siege for, I guess it would be going on for on thirty years now** and so what is called OE, operating expenses is consistently cut by Congress forcing USAID to basically outsource its programmes.\(^{258}\)

\(^{257}\) This was the case – temporary personal service contractors in USAID supervisory roles – from time to time during LGSP.

\(^{258}\) Interview 87, 9/20/13; boldfacing mine.
That an agency “under siege” feels the need to justify itself should come as no surprise.

The Consequences of Constraint for Field Staff Behavior & Soft Information

One USAID official described the effect of the restrictions and constraint from above as to “make you cautious.”\textsuperscript{259} This makes sense; insecurity breeds conservatism, the need to ensure that any action taken can be defended. One interviewee very explicitly linked this conservatism and tight supervision by USAID to an inability to make use of what I would call soft information and tacit knowledge:

When a person provides oversight they must provide oversight and not be part of the day to day management and just trust you with the decision that you do the analysis and the calculations of what is needed and what is not needed....[This hurts project performance]\textsuperscript{260} because then you know, some decisions you will take, I mean some will be good, some will be a risk but some decisions you will not make because now you always have this voice in your head that you know is going to say, no not this because maybe they don’t believe in that way and maybe they are not seeing what you are seeing because you are actually managing the programme and your view has maybe deeper areas of focus than the helicopter view and your decision is purely based on that insight which sometimes you cannot translate in words.\textsuperscript{261}

\textsuperscript{259} Interview 109, 9/12/13

\textsuperscript{260} In this ellipsis I asked the follow up question “Do you think that, in addition to being aggravating and frustrating to have that relationship, do you think it actually constrains the performance of projects?”

\textsuperscript{261} Interview 124, 8/5/13
While for USAID the general impression was of Congress as the elephant in the room, for DFID the general impression was of DFID London as only vaguely involved, and Parliament as not at all present. DFID’s Navigation by Judgment was led by relatively autonomous field staff doing the judging; USAID’s Navigation from the Top began all the way up, with Congress’ relationship with USAID HQ influencing USAID staff and contractors. Indeed, it may not be too extreme to say that USAID’s hostile political authorizing environment made Navigation by Measurement an almost inevitable outcome, with numbers allowing USAID to justify itself and bargain for continued funding. This Navigation from the Top pushed USAID staff towards conservative decisions that did not make use of local knowledge, or soft information.

**Work Environment, Job Satisfaction, and Agent Quality Links**

This environment did not make USAID a particularly welcoming place for professional growth or a particularly enjoyable place to work,\(^{262}\) echoing the poor USAID workforce job satisfaction data cited in Chapter 2. By contrast, no DFID employee or implementer had similarly extreme sentiments; the only dissatisfaction raised regarding DFID was that it has become increasingly constrained over time, lessening the judgment field staff can make and thus making the work less rewarding.\(^{263}\) More typical was a former

---

\(^{262}\) Interview 103, 7/30/13; interview 87, 9/20/13

\(^{263}\) Interview 127, 9/24/13; interview 77, 9/17/13. This criticism of DFID is, of course, consistent with the broader theory of this work – that job satisfaction and Navigation by Judgment positively covary, with IDO staff desirous of jobs where their judgment impacts direction. DFID’s perceived change over time will be revisited in the concluding Chapter 6.
USAID implementer describing the transition from USAID to DFID as “A breath of fresh air.”

It is also striking that the USAID staff member most universally praised by others was one of those dissatisfied with USAID’s work environment, and had exited the organization. This is all the more notable as one common response by government counterparts and aid implementers to one of my standard questions – “How would you compare USAID and DFID?” was to state that the answer was driven by the abilities, orientations, and talents of field staff. The implication in some cases was to suggest that differences between people, and thus some of the performance of organizations, was idiosyncratic; but that need not be the case. There is data here suggesting support for four related propositions, which in sum create a causal link between authorizing environment and organizational performance via staff quality:

• Political authorizing environment insecurity contributes to USAID being a less enjoyable and rewarding place to work than is DFID

• An ability to make consequential judgments regarding projects contributes to job satisfaction

• Staff are more likely to exit organizations when they are unsatisfied

264 Interview 125, 7/31/13

265 Interview 103, 7/30/13

266 I only asked this question to those in a position to have been exposed to both organizations. It’s possible the response here is the modal South African response, in fact.

267 Interview 137, 8/12/13; interview 120, 8/14/13; interview 114, 8/19/13; interview 92, 7/24/13; interview 79, 8/5/13
The quality of staff is consequential for project success

Additionally, if staff quality is observable and thus high quality staff are more likely to have outside options and/or (observably or unobservably) high quality staff are more likely to be frustrated by constraints and exit organizational performance should, on average, suffer as a result.

Channels in Sum

There is substantial support in the South African data for the view that USAID’s navigation strategy is intimately related to its political authorizing environment – that the choice to Navigate from the Top in these projects may not have been a choice at all, but rather a necessity borne of a politically insecure organization with a very small degree of autonomy. This navigation decision helped to make USAID a less satisfying place to work, in part by making it harder to incorporate soft information.

4.5: SOUTH AFRICA CONCLUSION – STRIKING PARALLELS, IMPORTANT DIVERGENCES

This chapter has demonstrated that differential navigation strategies were used by USAID and DFID. This chapter’s findings regarding those navigation strategies are consistent with those of Chapter 3, with optimal navigation contingent on the soft information demands and external observability and verifiability of the task and legibility
of the environment. In the context of South Africa’s relatively predictable environment Navigation by Judgment still proved the superior strategy in municipal capacity building, the task where external monitoring was more difficult and soft information – tacit knowledge of the particulars of the case and environment – potentially more critical to success. Navigation by Measurement proved the superior strategy in PMTCT delivery, the more externally observable and verifiable task.

There were costs to these strategies even in the realms in which they were most successful. USAID’s navigation imposed significant transaction costs in reporting and inflexibilities that precluded taking full advantage of opportunities such as Hogan’s arrival at the end of the denalist period. Similarly, DFID’s Navigation by Judgment introduced opportunities for mis-reporting and capture by contractors even in the Municipal Governance arena where it proved the better strategy.

This chapter supports, then, the view that optimal navigation depends on the nature of the task. As one interviewee put it, “the skills that you need to run a technically very sophisticated regional water supply system are very different from the skills you need to facilitate essentially a community and so the capabilities that you need to build are very different as well.” 268 Those organization-level skills seem to have differed across organizations.

268 Interview 93, 7/18/13
These cases suggest that the relationship between South Africa’s ability to ‘push back’ and results is much more complicated and nuanced that a simple positive correlation between country ability to negotiate and success. This description roughly fits DFID’s municipal governance work, with stronger governments associated with better performance; on the other hand, NDOH’s push was forceful but at times pointed in an unproductive direction. One senior South African government official complimented DFID for being “demand driven”; this is a popular buzzword in the world of foreign aid, one with substantial positive connotations. This chapter serves to suggest there can be such a thing as an IDO being unduly driven by recipient government demands to the detriment of intervention effectiveness and impact.

In the municipal sphere the same South African government has been much less able to be an effective advocate for its agenda; that agenda has also been through a number of revisions as Presidents and Ministers have waxed and waned. There is little evidence here that Government advocacy can effectively substitute for IDO navigation; there is, however, evidence that in this extreme of advocacy less Navigation by Judgment may have served DFID well. One respondent in comparing USAID and DFID seemed to echo this view, remarking that DFID

Changes too much... [DFID’s success] is very reliant on there being a country team that is able to not just bend in the wind with every gust but to actually hold its course, whereas with [US]AID that is kind of irrelevant, the country team does some initial design but there is, I think, strong Washington control and then it is out of people’s sight because it is locked into implementation.

\[269\] Interview 107, 8/2/13

\[270\] Interview 93, 7/18/13
That the strength of the Government’s ‘push’ has varied so much across department suggests that “Government”, at least in South Africa, is not monolithic. What is true of South Africa’s government is true of the US government as well; CDC and USAID sit parallel to one another inside PEPFAR, yet see differences in navigation despite PEPFAR’s tight top down management constraints. Even in the extreme it may not be the case that maximal constraint and political authorizing insecurity augurs for no discretion for IDOs and their field staff.

Focusing on South Africa cases can help us determine that optimal navigation is contingent on task and identify DFID & USAID’s different navigation strategies. South African cases alone cannot, however, give us a more nuanced view of the relationship environmental unpredictability plays in this picture. Can Navigation from the Top successfully cut through the increased difficulty of monitoring in a less legible environment, one less tractable to monitoring and complex information systems? How does an aid dependent environment alter IDO-recipient government interactions? The next chapter attempts to answer these queries in the context of Liberia.
CHAPTER 5: COMPARING USAID AND DFID IN LIBERIA

Health Sector and Capacity Building Efforts

5.1: Introduction

This chapter allows for a comparison of the navigation strategies and performance of DFID and USAID in a particularly uncertain, and difficult to observe from afar, environment: that of post-conflict Liberia from 2006-2013. This chapter examines the performance of USAID and DFID health and capacity building interventions in Liberia as outlined in Figure 5.1. Chapter 4 found DFID to Navigate by Judgment to a much greater extent than USAID in South Africa. Is the same true in Liberia? What relationship do any differences in navigation strategies have with differences in performance and success? Chapter 4 found Navigation from the Top – via Navigation by Measurement – to be the superior strategy in the South African health sector, but Navigation by Judgment superior in capacity building. Is the same true in the much more uncertain environment of Liberia?

---

271 I should note my own personal involvement with the Government of Liberia during the period under examination. From 2007-2008 I served as Special Assistant to Liberia’s then-Minister of Finance (now Director of the Africa Region for the International Monetary Fund), Antoinette Sayeh, playing a chief-of-staff-like role; from 2008-2009 I served as Aid Management Advisor to her successor as Minister of Finance (and current Liberian Minister of Foreign Affairs), Augustine Ngafuan. I remained occasionally engaged with Government business for a few years following the close of my full-time engagement in August 2009, and returned to Liberia a few times as a result. None of the Liberia cases in this chapter involve projects regarding which I had been substantially involved (or in some cases projects I had even been aware of) prior to the selection of these cases as part of my dissertation. Having known many of the actors in the Liberian cases (particularly on the Government side) for many years might be seen by some as compromising, as auguring towards something less than complete objectivity. I would argue experience has its benefits in access, honesty in interviewee response, and understanding of the perspectives and motivations of particular actors.
Table 5.1 illustrates the interventions on which this chapter primarily focuses, though other engagements will be noted where relevant. While these cases have been chosen to be as similar in scope and focus to the South African cases in chapter 4 as possible, there are notable differences in focus between the Liberian and South African projects. Liberia has no autonomous municipal governments\textsuperscript{272} the capacity building cases in this chapter focus instead on cross cutting capacity interventions for national ministries. On the health side HIV is much less of a concern in Liberia, with 1.5% of the population HIV+. (LISGIS et al. 2007, p.193) USAID had an HIV-focused project while DFID did not, precluding an HIV-focused comparison between the two IDOs’ health efforts.

\begin{footnotesize}\textsuperscript{272} Liberia is a centralized, unitary state. Decentralization has been proceeding, contentiously, since the return of democratic governance.\end{footnotesize}
This chapter examines USAID and DFID’s support to Liberian health more broadly, focusing on each IDOs’ core interventions which were both directed towards health systems strengthening in parallel, though distinct, ways. USAID’s HIV/AIDS program (AWARE) will also be examined to allow for better cross-country (that is, USAID in Liberia vis-à-vis South Africa) comparisons.

Table 5.1: Specific Projects Examined in Chapter 5

<table>
<thead>
<tr>
<th>Capacity Building</th>
<th>USAID</th>
<th>DFID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rebuilding Basic Health Services (RBHS), 2008-present: Implemented by JSI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed Asset Reimbursement Agreement (FARA), 2011-present</td>
<td>Health Sector Pooled Fund (PF), 2008-present</td>
<td></td>
</tr>
</tbody>
</table>

This first section provides an introduction. Sections 2 and 3 focus on the capacity building and health interventions respectively, examining relative intervention success and differences between USAID and DFID’s navigation strategies regarding these interventions. Section 4 investigates the channels from political authorizing environment-linked IDO navigation and success, bringing these cases into conversation with chapter 2’s hypotheses regarding soft information and field agent job satisfaction and motivation. Section 5
concludes the Liberia section in the main. As this is the end of the case study section of this work, an additional Section 6 compares across Liberia and South Africa cases.

**Liberia in the Mid-2000’s**

In January 2006 Ellen Johnson Sirleaf became Sub-Saharan Africa’s first democratically elected female head of state following a 25 year period of Liberian history dominated by military coups, civil wars, and rule by warlords (most notably Charles Taylor).\(^{273}\) The Liberia of 2006 was one just emerging from war, one in which the only thing more scarce than infrastructure (e.g. no non-generator power – no power grid anywhere in the country) was skills (e.g. approximately 50 physicians for a population of 3.5 million), with those who could seek opportunities outside Liberia’s borders having overwhelmingly done so over the previous quarter-century. (Kay 2014) The last published census was that of 1974 (Government of Liberia 2008a, p. iii), and that census was of dubious quality; while data was unreliable (including on how many Liberians there were and where they were living), the official unemployment statistic was 85%, which seemed if anything an underestimate. The best estimates suggest GDP at the time of the election of 2005 was 160 US per capita, or less than 20% the 1979 level (Radelet 2007).

---

\(^{273}\) Charles Taylor was convicted in 2012 of war crimes and crimes against humanity by the Special Court for Sierra Leone in The Hague. He won the 1997 general election using one of most attention-grabbing, honest and content-rich campaign slogans I’m aware of: “He Killed my Ma, He Killed my Pa, but I Will Vote for Him” – making clear that were he not to be elected, active conflict would resume. Peace held through April 1999, or for little more than two years. By contrast, Ellen Johnson Sirleaf was awarded the 2011 Nobel Peace Prize.
While international aid had been flowing to Liberia prior to Ellen Johnson Sirleaf’s election, these funds were largely focused on emergency response and flowed via NGOs. President Johnson Sirleaf’s election prompted many donors, including USAID and DFID, to reengage with the government. As USAID itself put it, “the inauguration of Africa’s first female head of state in 2006 ushered in a period of hope and high expectations for Liberia’s recovery after decades of instability” (USAID 2014). Reengagement with government led donors to engage in both systems building and more ‘everyday’ issues (e.g. provision of regular health care rather than clinics for those left injured, malnourished, or internally displaced by conflict) in a more substantial way than had historically been the case.

By summer 2013 Liberia had seen stability for the better part of a decade and was one of the fastest-growing countries in the world, with growth of 8.3% in 2012. (IMF 2013, p. 4) The country had clearly transitioned from emergency recovery mode to the broad “development” space occupied by most other sub-Saharan African countries. This did not mean it was stable in the sense one might think of a Western European nation being stable, however. While Ellen Johnson Sirleaf had won reelection in December 2011 to a 2nd term in office, the election was not without violence and disputes. The 2012 State Fragility Index gave Liberia a score of 16 (with higher numbers associated with greater fragility), much closer to Iraq (20) and Afghanistan (22) than Turkey (7) or Thailand (7) or, closer to home in Anglophone sub-Saharan Africa, South Africa (8), Ghana (11), and Kenya (10)

274 Interview 37, 5/27/13. This is often termed “humanitarian assistance” in contrast to development aid.

275 When the bulk of the interviews that inform this chapter were conducted.
(Marshall & Cole 2012). While stability had sprouted, there was much that remained to be done to ensure its roots were of adequate depth to stand firm in the storms the future might bring.

**DFID and USAID in Liberia: History and Trajectory**

Liberia has a history quite closely linked with that of the United States. Settled by freed American slaves under the sponsorship of powerful American whites (including President James Monroe, after whom Liberia’s capital, Monrovia, is named) in 1821, Liberia declared independence in 1847. For the past 150 years the US Government has remained sporadically involved in the country when geostrategic interests have warranted; in a pattern similar to that of more formal colonial relationships in Sub-Saharan Africa (e.g. the UK in Ghana, the French in Mali, etc.), the US is far and away the largest donor in Liberia, with aid flows of approximately 60 million USD in FY 2009/10 (the first year for which cross-donor comparable numbers are available), as compared to under 18 million USD for the UK government. (Liberia Ministry of Finance 2009).

---

276 This settlement was not in a territory which was previously unoccupied. The settler class treated the “natives” in a manner that in many ways reproduced the antebellum South from which they came, albeit with the settlers holding the positions of power. Non-settlers were not citizens until the early 1900’s; a 1930 League of Nations report (the Christy Report) found the Government of Liberia to be sanctioning forced labor and slavery.

277 These figures are early 2009 projections for the Liberian FY 09/10; the Liberian FY runs from July 1-June 30.

278 Due to the US Government’s wider array of projects and funding modalities to deliver assistance in Liberia than the UK, this chapter goes into detail regarding a greater share of UK/DFID funding than is the case for the US/USAID; to do otherwise would make it harder to draw comparisons between the two, and would not allow for a detailed examination of
By contrast, DFID has historically focused more on the former British colony (and Liberia’s neighbor to the west) of Sierra Leone, with Liberia of secondary importance. As one long time DFID official put it, there was “this sort of basic idea that Britain would look after Sierra Leone in so far as you can and the US would look after Liberia.” These differences in Liberia’s relative importance were to have substantial implications for DFID and USAID engagements over this period.

What follows is a description of USAID and DFID’s initial efforts in both the health and capacity building domains. In the case of health, the arc of support will be described for both donors; on the capacity building side focus will be on the initial efforts from both USAID and DFID.

279 Interview 43, 6/25/13
5.2: Capacity Building

There you stand, trying to rebuild a nation in an environment of raised expectations and short patience, because everyone wants to see change take place right away. After all, they voted for you because they had confidence in your ability to deliver – immediately. Only you cannot. Not because of the lack of financial resources, but simply because the capacity to implement whatever change you have in mind does not exist. – H.E. Liberian President Ellen Johnson Sirleaf (Tony Blair Africa Governance Initiative 2011; p. 1)

A focus on Government capacity building was a natural fit for the historical moment in which Liberia found itself. Capacity building could take many forms, from the provision of goods to skills training to systemic redesign; as will be seen in greater detail below, DFID and USAID took quite different strategies, with very different results, in their initial efforts.

Intervention Descriptions

DFID focused on a comprehensive reform of the civil service through their Civil Service Capacity Building (CISCAB) project. As Liberian President Johnson Sirleaf put it, “Economic recovery and reconstruction are constrained by a Civil Service that is highly incapacitated.” (Government of Liberia 2008b; foreword) CISCAB aimed “To help to rebuild and improve the Liberian Civil Service’s contributions to effective governance and service provision” by “helping the Governance Reform Commission, Civil Service Agency, and Liberian Institute of Public Administration to carry out their mandates.” (DFID 2006b, p. 5-
6) While the Civil Service Agency (CSA) was to be the key counterpart agency, the Liberian Institute of Public Administration (LIPA) would lead on training and the Governance Reform Commission and the Governance Reform Commission (GRC)\(^{280}\) on overall structural reforms of the civil service. From the beginning there were conflicts between Dr. William Allen, Director General of the CSA, and Dr. Amos Sawyer, the Chairman of the GRC (and a former Liberian President, having led a Government of National Unity in fighting then-rebel leader Charles Taylor and others from 1990-94) about the direction of the reform, conflicts that were to spill over into implementation.\(^{281}\)

CISCAB initially focused on developing a Civil Service Reform Strategy, a notion that had first been proposed by Allen of the CSA.\(^{282}\) In order to develop the strategy CISCAB brought in a series of short-term consultants to focus on problem analysis and strategy development. Following the strategy’s completion CISCAB shifted formally from a capacity building program (CISCAB) to a program focused on the implementation of the project (CISREP) until its close in May 2010.\(^{283}\) As DFID put it, “Following the approval of the Civil Service Reform Strategy in June 2008, the government requested the DFID to shift the project’s focus from building the capacity of CSA, GRC and LIPA to design and

---

\(^{280}\) The Governance Reform Commission (GRC) was renamed the Governance Commission (GC) while CISCAB was being implemented. This organization is referred to as the GRC throughout for consistency and ease of understanding.

\(^{281}\) Interview 56, 6/4/13; Interview 43, 6/24/13

\(^{282}\) Interview 54, 6/5/13; interview 55, 5/29/13

\(^{283}\) While the project officially changed its name from CISCAB to CISREP at this juncture this work will refer to it as CISCAB throughout for ease of understanding.
implementation of key civil service reforms in line with the CSRP [civil service reform program] and putting in place the building blocks to enable the government to successfully implement civil service reforms coherently and consistently across government over the medium-term.” (UK Department for International Development 2010b, p. 5)

USAID's first project to focus on government capacity after the election of President Ellen Johnson Sirleaf – and thus the one most directly comparable with CISCAB in timing and sectoral focus – was Building Reform and Recovery through Democratic Governance (BRDG, pronounced “Bridge”).284 The main, though not the sole, focus of BRDG was to “Assist the Government of Liberia and other key actors to improve capacity in such areas as effective planning, budgeting, communication, and coordination with relevant counterparts” accomplishing this through “(1) short-term technical assistance (STTA) to key GOL [Government of Liberia] ministries and civil society organizations, and (2) small grants to ministries and local partners.” (Nicolls & Kupperstein 2007, p. 2)285

This project was managed by USAID's Office of Transitional Initiatives (OTI). OTI is often the first USAID unit to deploy in a post-conflict setting, and had been in Liberia a number of years by the time that BRDG began operations in Liberia in late 2006. OTI’s role allows for the examination of intra-organizational variation within USAID; OTI operates quite differently to other USAID units, working with a far leaner and more flexible

284 Contract #DFD-I-00-05-00220, Task Order #3, awarded by USAID OTI to DAI; signed September 12, 2005
285 This was one of 3 objectives of BRDG, but the one to receive the bulk of the funding.
operating structure than do other USAID programs, and traditionally having a relationship with the contractors responsible for implementation that is more a partnership and less a supervisory relationship than that present in the vast majority of USAID projects.\textsuperscript{286} As a Congressional Research Service report on OTI puts it, “OTI staff are actively involved in every stage of project planning and implementation, in contrast with typical USAID staff who, some claim, have become more like contract officers than development experts in recent years.” (Lawson 2009, p. 9)

BRDG Liberia was not competitively bid, instead contracted with Development Alternatives Incorporated (DAI) using a form of expedited contracting known as an Indefinite Quantity Contract (IQC). Under an IQC contractors who wish to pre-qualify bid periodically (normally every 5 years) and, if selected, agree to supply services if contacted by OTI (the services provided are thus of indefinite quantity at the time of selection). While OTI normally uses a different IQC, in this case – given the large component of consulting services envisaged – OTI decided to use BRDG, which in the view of some added complexity and overheads to the project.\textsuperscript{287}

DAI staff deployed to Liberia to lead BRDG in late 2007 for what was initially envisaged as a 12-month (later extended to 16 months) project. The DAI team staffed up

\textsuperscript{286} Interview 70, 5/31/13

\textsuperscript{287} Interview 70, 5/31/13; interview 60, 6/27/13
an office of about 15 people (including 3 internationals) quite quickly; this process was helped by referrals from another (non-OTI) DAI project already running in Liberia.288

BRDG worked through specific activities that were not pre-specified, but rather developed as the project progressed; these activities might involve bringing in short term technical assistance, supporting a training or event, or providing office equipment, among other activities. While the activities were myriad, some oft-recalled activities by BRDG staff included establishing the President’s radio studio, office supplies for the Judiciary, supporting youth day activities held by the Ministry of Youth and Sports, and short term consulting to the Civil Service Agency on pay and grading reform.289

**Capacity Building Project Results**

Seen one way, CISCAB and BRDG were very different interventions. BRDG focused on a wide number of ‘quick wins’; CISCAB was designed to be the beginning of an (inevitably lengthy and complicated) process of reforming the civil service. BRDG was intentionally and explicitly opportunistic, investing where environmental conditions made success likely; CISCAB aimed to tackle a difficult problem and alter environmental conditions.

---

288 Interview 44, 6/6/13; interview 45, 6/10/13; interview 41, 6/10/13

289 Interview 52, 6/4/13; interview 55, 5/29/13; interview 64, 6/19/13;
That said, there are also striking similarities. Both USAID and DFID wanted to support Government functioning following the election of Ellen Johnson Sirleaf. Both projects were of roughly similar size; CISCAB’s total budget was the equivalent of approximately 4.3 million USD, whereas BRDG spent, by my calculations, just over 4 million USD in support of government. Both delivered their support in substantial part via short term consultants brought in to deliver discrete pieces of work.

As both BRDG and CISCAB’s goals make clear, these projects aimed at improving the ability of the Liberian government to function, to deliver. As such this section will apply the same basic standard as used in the previous chapter regarding municipal capacity building. What direct impacts, if any, do we see of the support? And did these improved capacities change trajectories in the long term, or only affect the ‘level’ of progress in the short term? Did BRDG’s ‘quick wins’ open the door for longer term success? DID CISCAB seed reform momentum that carried beyond its time?

290 2.73 million GBP

291 A little under 65% of BRDG’s funded activities went to the government sector, based on my calculations drawn from the almost-final project database (the last version I could locate). Thus assigning overheads proportional to spending, a bit over 4 million USD of BRDG’s 6.4 million USD total budget is attributable to supporting the Government of Liberia in its activities and capacity building. This proportionality is needed because I could identify only 3.2 million of actual activity-by-activity spending; according to one interviewee (Interview 64, 6/19/13) it’s entirely plausible that the remaining monies were consumed by overheads – that is, costs over and above the fees and expenses of consultants and activities (e.g. OTI and DAI salaries, indirect overheads/fees, etc.). While I lack the detailed information to compare this to CISCAB overheads – and thus cannot say with certainty whether BRDG and CISCAB were grossly different in this regard – this strikes me as a surprisingly large percentage of ‘foreign aid’ to not substantially actually benefit the country in any direct fiscal sense.
Direct Impact on Ministries and Agencies

CISCAB’s clearest impact is in the form of Liberia’s Civil Service Reform Strategy, which has served as the roadmap for civil service reform since.\textsuperscript{292} The quality of the strategy received substantial praise, and is universally seen as a significant accomplishment owing much to CISCAB.\textsuperscript{293} DFID increased CISCAB’s project budget by 50\% around this time as well, a seeming sign of confidence in the project’s progress to data and future prospects.\textsuperscript{294} The strategy painted an ambitious agenda, including restructuring the civil service (including changes to political appointment powers), implementing biometric identification for civil servants, rationalizing and standardizing pay and grading, and skills training. It was seen as a very ambitious plan, one intended to be aspirational.\textsuperscript{295} Following the development of the strategy some elements of the project languished. While LIPA benefited in terms of a new building and training, the GRC began mandates and functions review, and the CSA initiated analytic work in line with the strategy,\textsuperscript{296} a number of factors helped contribute to a general slowdown. Among these were tensions between CSA and the GRC\textsuperscript{297} as well as tensions within the ASI team.\textsuperscript{298}

\textsuperscript{292} Interview 65, 6/10/13; interview 55, 5/29/13; interview 50, 6/6/13; interview 40, 6/5/13

\textsuperscript{293} Interview 53, 10/3/13; interview 55, 5/29/13; interview 42, 5/31/13; interview 49, 5/30/13; interview 67, 6/24/13

\textsuperscript{294} Interview 67, 6/24/13

\textsuperscript{295} Interview 38, 6/7/13; interview 54, 6/5/13; interview 55, 5/29/13

\textsuperscript{296} Interview 65, 5/31/13; interview 56, 6/4/13; interview 54, 6/5/13

\textsuperscript{297} Interview 39, 6/12/13; interview 61, 6/7/13; Interview 43, 6/25/13; interview 59, 6/12/13; interview 51, 6/11/13

207
Unlike CISCAB, BRDG didn’t have a specific agency or sectoral focus, and the benefits to each individual beneficiary were modest; as such it’s not surprising that beneficiaries sometimes had to be reminded of specific activities and support before commenting on BRDG.\(^{299}\) In general beneficiary organizations remembered appreciating the support, though not remembering it as particularly substantial or consequential.\(^{300}\) BRDG is remembered for being quick, flexible, and broad, “just a phone call away”,\(^ {301}\) a project that was quick and practical – “a ‘get started’ project”\(^ {302}\) As Sam Hare, former Deputy Minister of Youth and Sports, put it in response to a query regarding other beneficiaries of BRDG, “the question is who was not a beneficiary.” None of the government officials seemed to ‘own’ the project, nor did BRDG coordinate with Government regarding activities selected.\(^ {303}\)

\(^{298}\) Interview 38, 6/7/13; interview 67, 6/24/13; interview 39, 6/12/13

\(^{299}\) Indeed, it’s somewhat surprising that – five years after the project ended – a number of beneficiaries could specifically recall BRDG and, unprompted, name personnel involved, despite the modest size of the support each individual beneficiary received.

\(^{300}\) Interview 55, 5/29/13; interview 58, 6/10/13; interview 57, 6/5/13; interview 62, 6/10/13; interview 65, 5/31/13. The general impression was that it was the Government equivalent of a Ph.D. student’s 300 dollar travel grant; appreciated, but modest in impact.

\(^{301}\) Interview 57, 6/5/13

\(^{302}\) Interview 58, 6/10/13

\(^{303}\) Interview 64, 6/19/13; interview 60, 6/27/13; interview 41, 6/10/13. That is, while many of the beneficiaries were Government ministries and actors, the project only coordinated with beneficiaries individually rather than with some focal government actors who might have assisted in allocating what (overall) was a fixed envelope of funding. This
Systematic Effects

The end of CISCAB – that is, the lack of a follow on project - caught many actors by surprise, and the trajectory of civil service reform suffered substantially as a result.\textsuperscript{304} As Allen, then CSA Director General, felt it "was like they pulled the carpet out from under us." Oblayon Neymah, then the head of the Civil Service Reform Directorate,\textsuperscript{305} said that post-CISCAB "We became the orphan in the process." As one person involved in the sector put it, "Ever since they [DFID/CISCAB] left, huge gap to trying to get back to the track that we were on. We were on a fast track to attaining our goals."\textsuperscript{306} Another said “what CISCAB was doing for CSA – to have to replace that – was really kind of difficult, to be frank... it put the entire reform of the CSA back for some time.”\textsuperscript{307}

The decision not to renew was driven by a broader DFID decision to shift focus back to Sierra Leone, closing the Liberia office and investing in Liberia only through multidonor

\textsuperscript{304} Interview 50, 6/5/13; interview 43, 6/25/13; interview 56, 6/4/13; interview 55, 5/29/13

\textsuperscript{305} The Civil Service Reform Directorate (CSRD) was the unit inside CSA set up by CISCAB to help drive the reform efforts. The unit still exists, though in substantially weakened form.

\textsuperscript{306} Interview 46, 6/11/13

\textsuperscript{307} Interview 39, 6/12/13
mechanisms such as the Health Pooled Fund. Indeed, DFID continued to see the project as rather successful, giving it an 85% success rating overall and giving no sub-component less than the 2nd-highest possible rating (on a 5 point scale) on the project’s completion report. (DFID 2010)

The Liberian government, in a move that suggests the Government found CISCAB to be quite valuable, fought for DFID to continue funding CISCAB, a rather unusual move regarding such a modestly sized intervention. On 13 November 2009 (before the program closed in March 2010, but after the announcement that it would not be continued) President Ellen Johnson Sirleaf wrote then-Prime Minister Gordon Brown, asking for the program to continue. Even this, however, was insufficient pressure for the funding to be continued; DFID had decided to refocus their attention on Sierra Leone, and on one account Brown was briefed that any monies heading to Liberia would detract from the Brown’s efforts to support free health care in Sierra Leone.

---


309 Interview 55, 5/29/13; interview 56, 6/4/13

310 Interview 56, 6/4/13
While ASI and DFID took action to attempt to secure continuing funding for civil service reform from other donors, these efforts bore little fruit.\textsuperscript{311} CISCAB left behind a roadmap for civil service reform and some momentum, but the engine driving the reform had been lost as well as the donor who had previously been supplying the fuel. This drop in reform energies was referred to by virtually every interviewee,\textsuperscript{312} and more than one international actor suggested that there was some level of irresponsibility associated with DFID’s withdrawal.\textsuperscript{313} DFID’s pullout also contributed to a widespread view that the project had been viewed by DFID as unsatisfactory,\textsuperscript{314} which itself strengthened those who might have opposed reform while simultaneously discouraging other donors from tackling civil service reform.

The end of BRDG also marked OTI’s departure from Liberia, and BRDG had no straightforward successor; this met much less resistance than did CISCAB’s close.\textsuperscript{315} The diffuse, multisectoral nature of BRDG meant that no stakeholders were as ill affected by BRDG’s departure and the lack of a clear follow on program as was the case regarding CISCAB. It also makes it harder to evaluate the continuity of BRDG’s efforts after the

\textsuperscript{311} Interview 67, 6/24/13; interview 55, 5/29/13

\textsuperscript{312} e.g. interview 50, 6/5/13; interview 65, 5/31/13; interview 55, 5/29/13; interview 66, 6/7/13; interview 59, 6/12/13; interview 39, 6/12/13

\textsuperscript{313} Interview 40, 6/4/13; interview 56, 6/4/13; interview 67, 6/24/13

\textsuperscript{314} Interview 65, 5/31/13; interview 63, 6/25/13; interview 49, 5/30/13

\textsuperscript{315} Interview 64, 6/19/13; interview 70, 5/31/13; interview 68, 6/19/13
program’s departure, as those efforts were varied and in most specific cases relatively small.

For some of the civil society projects in particular the end of BRDG funding led the activities supported to cease.\textsuperscript{316} Some share of the physical infrastructure support of Government- including the President’s radio studio and computers and office supplies in various agencies – is still in use; less clear is the impact from analytic work BRDG supported.\textsuperscript{317} The only example of a BRDG program or process (rather than equipment or supplies) that, once set in motion, continued to grow following OTI’s departure is the Songhai agricultural center, which was initially supported by BRDG funding.

On one view, this is an unfair question; one interviewee argued that BRDG’s mission was not sustainable activities, but rather quick impact “pre-sustainable development” support.\textsuperscript{318} However, others – including those from USAID and DAI – seem not to have seen it that way, instead seeing BRDG as a bridge to broader capacity building efforts to follow; e.g. multiple interviewees expressed concern that BRDG’s lack of formal transition or any plan for other USAID/US Embassy efforts to build upon what OTI had begun likely had deleterious effects on the longer term impact of those efforts.\textsuperscript{319}

\textsuperscript{316} Interview 52, 6/4/13; interview 45, 6/10/13; interview 44, 6/6/13

\textsuperscript{317} Interview 57, 6/25/13; interview 55, 6/12/13; interview 40, 6/5/13

\textsuperscript{318} Interview 68, 6/19/13

\textsuperscript{319} Interview 70, 5/31/13; interview 60, 6/27/13; interview 64, 6/19/13
So what did these projects accomplish? While both projects were well regarded by their beneficiaries and had short term impacts, substantial progress – even if incomplete – on civil service reform has left more substantial positive impacts on Liberia than the potpourri of activities supported by BRDG. The departure of both projects was done without any real continuity to ensure that the activities initiated were to continue; however the vocal opposition to CISCAB’s departure and the extent to which actors clearly felt its absence speak to the greater value it had for beneficiaries.

*Barriers to Success*

BRDG was designed to be broad and thus less deep than was CISCAB. It however accomplished what it set out to do; it is hard to find substantial fault in how BRDG was implemented. As discussed in greater detail below, motivated staff, great working relationships, and a flexible and adaptive approach clearly served BRDG well. It is striking how many interviewees mentioned how exceptional this was among USAID projects with which they have interfaced; how unlike other USAID capacity building efforts BRDG was.\(^\text{320}\) This was not because it was administratively easy to implement BRDG; far from it. The BRDG contracting modality was in fact much more administratively complex than those normally used by OTI; this added time and cost to the project.\(^\text{321}\)

---

\(^\text{320}\) Interview 44, 6/6/13; interview 57, 6/5/13; interview 62, 6/10/13; interview 58, 6/10/13

\(^\text{321}\) Interview 70, 6/12/13; interview 60, 6/3/13. BRDG was a time & materials contract rather than the SWIFT (cost + fixed fees) modality normally used by OTI. BRDG is a modality normally used by other USAID units; that OTI was able to get substantial flexibility with a contracting modality not specifically designed for same underscores that the orientation and quality of the OTI implementers was consequential here.
CISCAB, on the other hand, had substantial implementation challenges. The core implementing team of ASI and Subah Belleh personnel had substantial tensions that seem to have had their roots in the management style of the ASI team leader.\textsuperscript{322} These tensions eventually convinced ASI management that it was time for a change of ASI team leadership; ASI then discussed the matter with DFID and encouraged the team leader to exit, which she did.\textsuperscript{323} There was also substantial tension between CISCAB beneficiaries, tension that ultimately led the project to mostly exclude the GRC.\textsuperscript{324}

There were substantial external obstacles for CISCAB as well; ministries resisted reforms that would circumscribe their control over political appointments in favor of a more professional, routinized civil service. Despite repeated endorsement of the strategy and push for reform from President Ellen Johnson Sirleaf, Ministers still managed to resist change in the absence of any process that, given the weakness of core of government follow

\textsuperscript{322} Interview 38, 6/7/13; interview 67, 6/24/13; interview 49, 5/30/13; interview 53, 10/3/13

\textsuperscript{323} Interview 43, 6/25/13; interview 67, 6/24/13; interview 38, 6/7/13; interview 42, 5/31/13; interview 53, 10/3/13. The team leader left of her own accord, and in such a way that it is not clear she was fully aware that this was by popular consensus. Interview 67, 6/24/13; interview 54, 6/5/13. It should also be noted that the team leader was much beloved by Dr. Allen and many of the Liberian counterparts for her energy, engagement, and focus on getting things done in practical ways; she was described as “a force” (intended positively) by numerous Liberian government actors (Interview 55/ 6/12/13; interview 50, 6/5/13; interview 65, 5/31/13; interview 59, 6/12/13; interview 61, 6/7/13)

\textsuperscript{324} Interview 43, 6/25/13; interview 67, 6/24/13; interview 61, 6/7/13; interview 51, 6/11/13
up, might have helped to make them accountable;\textsuperscript{325} this reform also required legislative action to be implemented.\textsuperscript{326} Reforming appointment power was one small part of a larger civil service reform strategy; it is, however, far and away the element of CISCAB that seems to have been most salient to actors at the time, and is certainly among the most recalled elements of the project.\textsuperscript{327} One might question the decision to focus on this difficult reform rather than the myriad technical reforms that might not have met with such unified political opposition. BRDG suffered no such resistance, in large part because it did not attempt to tackle entrenched resistance, instead providing support only when conditions were aligned for that support to be well received.

**Differences in Navigation Strategies: BRDG and CISCAB**

What led to the differential performance of BRDG and CISCAB? Did USAID/OTI and DFID engage in substantially different navigation strategies? To what extent did OTI’s unique status within USAID allow it to engage in navigation strategies different than those of most USAID projects?

\textsuperscript{325} Interview 39, 6/12/13; interview 42, 5/31/13; interview 53, 10/3/13; interview 59, 6/12/13

\textsuperscript{326} Interview 61, 6/7/13

\textsuperscript{327} Interview 59, 6/12/13; interview 61, 6/7/13; interview 38, 6/7/13; interview 50, 6/5/13; interview 66, 6/7/13; interview 51, 6/11/13
Design

Both USAID and DFID designed projects that built on relatively long country- and context-specific experiences; that of DFID staff in the years before CISCAB launched with EC projects, that of OTI in previous in-country projects.

The roots of DFID’s CISCAB went back to 2004, when DFID personnel participated in the design of the EU’s institutional support program; a seconded DFID employee, Chris Gabelle, went on to manage the EU’s program.328 Seeing problems in Sierra Leone as growing out of Liberia,329 DFID staff began to conceive of supporting reform of the civil service in Liberia as they were doing in Sierra Leone; DFID also saw civil service reform as an area of comparative advantage.330 Design work began in earnest after the election of 2005 led by Gabelle and a senior DFID governance advisor (Garth Glentworth OBE). This was a long process, involving what both government and DFID actors saw as extensive consultations; Gabelle’s substantial Liberia experience played a significant role in the project design.331 An initial focus on capacity building was broadened into a more

328 Interview 56, 6/4/13; interview 43, 6/25/13

329 A view eventually echoed by the Special Court for Sierra Leone, which in 2012 found former Liberian president Charles Taylor guilty of war crimes for his role in the conflict in Sierra Leone.

330 Interview 56, 6/4/13; interview 43, 6/25/13; on at least one view Marcus Manuel, the person with whom Health Minister Gwenigale had critically interacted in pushing forward support for the OFM (DFID’s health project, to be discussed later in this chapter) was also critical in backing CISCAB.

331 Interview 56, 6/4/13; interview 67, 6/24/13; interview 43, 6/25/13; interview 54, 6/5/13
systematic reform program, a transformation that may have been initiated by GoL actors who felt more comprehensive reform needed.\textsuperscript{332}

DFID decided to implement through a contractor, and bids were solicited; Allen and Dr. Harold Monger (the head of LIPA) flew to London to be part of the selection process.\textsuperscript{333} Adam Smith International (ASI) was one of the final bidders; by common account ASI had the stronger bid, strengthened by a stronger local Liberian partner in the bid (Subah Belleh Associates) and the dynamism of the proposed ASI CISCAB team leader, Nellie Mayshak.\textsuperscript{334} Following the award to ASI there was a 3-month inception period in which the CISCAB core team\textsuperscript{335} revisited the project design in conversation with the main project beneficiaries. Their initial assessment suggested that there was much more that needed to be done in terms of basic equipment than DFID had initially envisaged; that to do work in this environment a focus on infrastructure was needed to complement the focus on policy, structure, and skills.\textsuperscript{336} While DFID does not normally support infrastructure in public service reform programs, DFID agreed to do so in the case of CISCAB.\textsuperscript{337} The design of CISCAB was based primarily on the judgment of DFID’s field level staff, then adjusted

\textsuperscript{332} Interview 55, 5/29/13

\textsuperscript{333} Interview 55, 5/29/13; interview 43, 6/25/13

\textsuperscript{334} Interview 43, 6/24/13; interview 55, 5/29/13; interview 54, 6/5/13; interview 42, 5/31/13; interview 67, 6/24/13

\textsuperscript{335} The core team is the implementation team formed by Adam Smith International and Subah Belleh, the local partner, to implement the project.

\textsuperscript{336} Interview 54, 6/5/13

\textsuperscript{337} Interview 54, 6/5/13
before implementation began in earnest as a result of the judgment of the CISCAB core team.

The seed for BRDG came – unlike in any USAID project thus far – from outside USAID. Liberian Minister of Finance Amara Konneh, then a staff member in President Sirleaf’s office, asked USAID for assistance with a communication strategy. USAID turned to OTI in an effort to respond as quickly as possible to the request. The initial plan for BRDG was developed by USAID in collaboration with Minister Konneh, with a focus on supporting critical government institutions with both materials and skills on demand.

While the project was initially to focus exclusively on supporting the Government of Liberia, the scope was broadened shortly before implementation began to allow for civil society and media groups to receive support as well. (Development Alternatives Inc. 2006)

The specific activities funded under BRDG were determined during implementation. Some activities grew out of requests (both solicited and unsolicited), including from Government actors; some stemmed from the DAI and OTI teams’ reading of public messages from Liberian government officials or issues raised in the popular press; and some were identified by project staff themselves, who felt very free to propose projects that were

---

338 Interview 68, 6/19/13; interview 70, 5/31/13; interview 60, 6/27/13

339 Interview 70, 5/31/13

340 Interview 60, 6/27/13; interview 70, 5/31/13

341 Interview 52, 6/4/13

342 Interview 70, 5/31/13; interview 60, 6/27/13
generated from their (soft information) understanding of Liberia’s needs.\textsuperscript{343} As one BRDG staffer put it, “We are Liberians. We know exactly what is going on and what will be helpful to society.”\textsuperscript{344} There was, as noted above, no central Government hand in deciding precisely what to support, though some BRDG activities were the result of specific ministry requests.\textsuperscript{345} Each activity was designed so as to have clear, monitorable outputs (e.g. consultants’ reports, computers supplied) – on the view of some BRDG staff members this monitorability was critical to getting a project approved.\textsuperscript{346} BRDG was thus – unlike any other USAID project examined thus far - very much driven by Navigation by Judgment in the choice of where to direct funds. Each individual project, however, required a substantial dose of Navigation by Measurement.

\textit{Implementation}

The CISCAB implementation team was given substantial scope to act in the manner they saw best, including modifying plans and activities when necessary, with DFID approval.\textsuperscript{347} Implementation decisions were left very much to the core team’s discretion; the core team in turn relied substantially on Government actors for implementation guidance.\textsuperscript{348} The selection of consultants for short term pieces of work involved

\begin{itemize}
\item \textsuperscript{343} Interview 52, 6/4/13; interview 47, 6/4/13; interview 41, 6/10/13
\item \textsuperscript{344} Interview 47, 6/4/13
\item \textsuperscript{345} Interview 52, 6/10/13; interview 57, 6/25/13
\item \textsuperscript{346} Interview 45, 6/25/13; interview 41, 6/19/13
\item \textsuperscript{347} Interview 54, 6/5/13; Interview 38, 6/7/13; interview 56, 6/4/13
\item \textsuperscript{348} Interview 67, 6/2/13; interview 54, 6/5/13
\end{itemize}
Government actors on the selection committee, who were generally seen as quite effective.\textsuperscript{349} That said, there were some less successful consultants; the consultants who were meant to lead the actual drafting of the strategy, for example, proved quite disappointing; as such the core team led the actual drafting process.\textsuperscript{350} There was a clear distinction between DFID and their contractor, ASI; ASI implemented with substantial discretion while DFID supervised.

BRDG exhibited a very different working relationship between USAID and the contractor (in this case, DAI) than CISCAB or, for that matter, other USAID projects examined in this work. In general DAI and OTI seemed to be operating as a nearly unified, relatively seamless BRDG administrative team; this “positive, collaborative, and open” working relationship is also noted in the externally contracted final project evaluation for BRDG. (Nicolls and Kupperstein 2007, p.27) Following identification and approval of a project idea by the DAI chief of party (Barb Lauer) and OTI, DAI would develop a budget to be discussed with the beneficiary and approved by OTI, with the grant then managed by DAI staff.\textsuperscript{351} DAI and OTI staff met weekly and collaborated on activity development and review;\textsuperscript{352} when consultants departed their exit briefing was with both DAI and OTI staff.\textsuperscript{353}

\textsuperscript{349} Interview 55, 5/29/13; interview 38, 6/7/13
\textsuperscript{350} Interview 42, 5/31/13; interview 38, 6/7/13
\textsuperscript{351} Interview 48, 6/12/13; interview 60, 6/27/13; interview 47, 6/5/13
\textsuperscript{352} Interview 52, 6/4/13; interview 47, 6/4/13; interview 41, 6/10/13; interview 64, 6/19/13
\textsuperscript{353} Interview 45, 6/10/13; interview 41, 6/10/13
A number of Liberians on both the staff and the beneficiary side noted how differently interacting with OTI was from other USAID projects with which they have engaged. One government beneficiary praised what he described as the continuing “mid-level bureaucratic contact” with BRDG; as he put it, “Normally with USAID we’re involved with the RFP [Request for Proposals], but after that involvement is at the whim of implementers.... They [implementers on projects other than BRDG] knew they had to get [Government’s] OK, and if they got it that was it.” Other responses on the difference between BRDG and OTI more generally to other USAID programs focused on differences in BRDG’s internal processes. These comments focused on the greater power given to staff than true in other USAID projects and the speed of BRDG. Three Liberian respondents without specific prompt mentioned the greater flexibility of BRDG as a significant comparative strength; this was also a strength of OTI mentioned by international staff.

---

354 Interview 57, 6/5/13; this echoes the views expressed in interview 51, 6/11/13, where USAID was described as normally being run “in an enclave context, with very minimal capacity building on a long term basis”.

355 Interview 45, 6/10/13

356 Interview 58, 6/10/13; interview 41, 6/10/13

357 Interview 58, 6/10/13; interview 48, 6/12/13; interview 44, 6/6/13

358 Interview 60, 6/27/13; interview 70, 5/31/13; interview 64, 6/19/13. This is not to suggest that implementation was absolutely seamless – an incident of procurement corruption was reported by one interviewee, where someone buying goods on behalf of the project was soliciting bribes. (Interview 60, 6/3/13) While there were no other projects in which corruption was revealed by interviewees, my suspicion is that this has more to do with a lack of open-ness and less to do with actual greater levels of corruption. Far more indicative is, to me, the response of project staff to the employee accused of malfeasance
Revision

The objectives of BRDG were revised, if modestly, throughout the project period; before implementation began the objectives were broadened to include “other key actors” rather than merely Government beneficiaries. (Development Alternatives Inc. 2006) Then in June 2007 objective 1 shifted to emphasize communication to the exclusion of other areas (e.g. planning & budgeting) in supporting the Government and other key actors. (Clemens 2007) This revision seems to have been quite easy, as was the extension of the project duration.\textsuperscript{359} Where most USAID projects might revise their objectives only with a substantial amount of effort, for OTI such revisions are \textit{de rigeur}; as one person with knowledge of many OTI projects put it, “I would guess never did [OTI] have a programme that the goals at the beginning are the same as the goals at the end. They are specifically designed to be adapted as the situation changes.”\textsuperscript{360}

CISCAB had arguably more major revisions, including (but not limited to) the change of direction from CISCAB to CISREP and the change of personnel on the

\textsuperscript{359} Interview 70, 5/31/13; interview 60, 6/27/13; interview 64, 6/19/13

\textsuperscript{360} Interview 60, 6/27/13
implementation team. The general view was of CISCAB modifying plans and activities when necessary with DFID approval.

Seen through the lens of navigation strategies, both these projects Navigated by Judgment as regards revisions; DFID’s CISCAB had the more substantial revisions but was also the longer project. CISCAB’s revisions were driven by the judgment of contractors’ field level staff – to whom DFID allowed a great deal of autonomy and discretion. By contrast, BRDG was revised via OTI’s judgments – that is, by the judgment of USAID personnel.

_Differential Capacity Building Navigation in Sum_

There are a range of ways one could describe the differential ways USAID and DFID responded to essentially similar insights, environments, and needs assessments. USAID’s program was more explicitly political in the sense that pleasing Government of Liberia officials in a country of strategic importance was seen as desirable. This may well have

---

361 Both noted above with relevant interview/document cites.

362 Interview 54, 6/5/13; Interview 38, 6/7/13; interview 56, 6/4/13

363 Interview 70, 5/31/13; implied, not stated explicitly. The M&E guide for the project does get fairly explicit on this point; “It is important to keep in mind that OTI’s programs are put in place to achieve “political’ results as defined by US Foreign Policy objectives. While this is increasingly true for all US foreign assistance programs, for OTI, the key challenge is to support the US interest in helping nation-states recover from the chaos of violent internal conflict, and in making the transition to a stable, democratic and just government and society. Unlike other parts of USAID, OTI’s role is to be fast moving, opportunistic, flexible and responsive to changing situations in the short to medium term. It uses the small in-kind grant mechanism and, in Liberia, the Short Term Technical Assistance consultancy, to move money and other resources quickly to where they can demonstrate US commitment, support local agencies public and private to keep the peace.
played a role in the decision to spread beneficiaries widely in the design phase, and as such BRDG’s lower level of long term impact than CISCAB may be as intended, in some sense. That said, it’s also notable that BRDG focused exclusively on small projects with quite visible outcomes; every activity yielded some visible result, be it a successfully held event, a piece of technology in use, or a consultant’s report.

CISCAB, on the other hand, took on a riskier, more difficult, and less externally legible strategy. Indeed, the CISCAB Project Memorandum and Framework – the DFID-prepared document laying out the plan for CISCAB prior to implementation – says “The fact has to be faced that post-conflict conditions in Liberia do provide a high level of risks to CISCAB… however if no attempt is made, conditions will never improve.” (DFID 2006b, p.27) As one DFID official with knowledge of both CISCAB and civil service reform efforts more broadly put it, “[CISCAB was] the unsung stuff – civil service reform is the nasty, gritty, doesn’t get much honors, doesn’t get the plaudits, but it gets things done because it changes minds and it changes ways of doing things.”

OTI showed a substantial degree of Navigation by Judgment; the statements of Liberian actors regarding the exceptional manner of BRDG’s engagement relative to other USAID projects to which they have been exposed suggest that this is not simply a difference between USAID’s manner of engagement in Liberia relative to South Africa but rather a

---

364 Interview 56, 6/4/13
result of OTI's different manner of engagement and level of autonomy within USAID. That said, BRDG still focused on discrete projects tractable to Navigation by Measurement in contrast to CISCAB’s focus on a much less externally observable and monitorable process. DFID’s tackling of a high risk but high reward project was critical to its comparative success. The general political authorizing environment insecurity reported by USAID actors seems related to the less ambitious, more measurable project chosen in this sector.

5.3: Health

In the previous chapter it appeared that USAID’s Navigation from the Top – with a heavy dose of Navigation by Measurement – proved superior in the South African health sector to DFID’s Navigation by Judgment and focus on working through Government institutions. Did DFID and USAID engage in parallel navigation strategies in Liberia? Did Navigation by Measurement still prove the superior strategy in health?

In 2014 Ebola ravaged Liberia’s health system; to read some of the news coverage of the outbreak one would think Liberia’s health system saw little improvement from Ellen Johnson Sirleaf’s inauguration in 2006 until Ebola’s arrival in early 2014.\footnote{The first cases of Ebola in Liberia were confirmed in late March 2014.} This is not the case; health was broadly seen as one of the most rapidly improving parts of the Liberian government prior to the crisis.\footnote{Interview 24, 6/4/13; interview 23, 6/11/13} Health data support this view; between 2007 and 2013 infant mortality fell by 24%, prenatal care by a skilled provider increased by over 20%, and
vaccination rates increased across the board. (LISGIS 2014)\textsuperscript{367} While Liberia’s health system was indeed overwhelmed by Ebola, that does not preclude a discussion of what contributed to progress in the Liberian health system between 2006 and 2013.\textsuperscript{368}

Both USAID and DFID had funded NGOs in the health sector prior to 2006, with both Governments tending to fund NGOs from their own countries.\textsuperscript{369} DFID and USAID began from similar institutional trajectories, and had similar – largely absent – legacies of working on health systems, or Government capacity building, on which to build.\textsuperscript{370} The election of Ellen Johnson Sirleaf prompted both IDOs to shift towards systems building in the health sector with projects aimed at building systems. USAID began project work in PMTCT via Liberia’s National AIDS Control Program (NACP) immediately after Ellen Johnson Sirleaf’s inauguration; the vehicle via which this support was delivered was Action For West Africa Region (AWARe). DFID, on the other hand, initiated its post-election

\textsuperscript{367} Calculations mine. Infant mortality fell from 71 to 54 per 1,000; prenatal care by a skilled provider increased from 79 to 96 percent of the population. These positive figures are not outliers, but rather broadly similar to other comparative data regarding progress between the 2007 and 2013 waves of Liberia’s Demographic and Health Survey (DHS).

\textsuperscript{368} The interviews that undergird this chapter were largely conducted in mid-2013.

\textsuperscript{369} Interview 37, 5/27/13

\textsuperscript{370} There isn’t good evidence that any of DFID and USAID’s NGO support had very much lasting systemic impact, though it may still have provided critical support to individual clinics and patients during the conflict. The consensus view was that there was no major impact of this funding in a systems sense, though the provision of services to Liberians in need was of course laudable. This direct support for NGOs was almost always characterized negatively, with few distinctions if any made between USAID and DFID; the Government was not involved in implementation, was not aware of the budget, and had only the vaguest sense of what activities, where, were occurring. As one individual put it, “You know how it goes – when you do it through NGOs it gets done, but when NGOs leave it all crumbles.” (Interview 30, 6/11/13)
health system support with a focus on systems and strengthening the Ministry of Health’s Office of Financial Management. Both agencies then moved to general, relatively unrestricted support to the health sector.

By 2006 Liberia’s Minister of Health was Dr. Walter Gwenigale. Gwenigale had stayed in Liberia throughout the conflict, treating combatants from all sides in dangerous circumstances. Gwenigale was widely seen as an honest, dedicated reformer, though one who faced a monumental challenge – turning a basically absent system in a country with some 50 doctors (Kay 2014) into a functioning public health infrastructure.

**Intervention Descriptions**

DFID and USAID engaged in somewhat parallel changes during this period – from funding Non-Government Organizations (NGOs) to projects to broader support to Ministry of Health-led efforts. Examining the arc of support from each donor in turn allows for a close examination of the relative constraints each IDO faced and how they navigated the same environment. This comparison is helped by the fact that both IDOs had the same basic strategy in these interventions – that is, building Liberia’s national health systems.

**USAID’s Liberia Health Arc**

The HIV/AIDS-focused Action for West Africa Regional program (AWARE) was a USAID regional intervention run out of Ghana and led by Family Health International

---

371 Interview 24, 6/4/13; interview 29, 6/7/13
372 Interview 36, 6/25/13; interview 11, 6/10/13; interview 21, 6/4/13
(FHI),\textsuperscript{373} a frequent USAID contractor. The USAID mission in Liberia ‘bought in’ to AWARE – that is, the mission put their funds towards an existing regional project in order to expand the project to Liberia.\textsuperscript{374} The state of Liberia’s HIV/AIDS program was typical of the state of dysfunction of public systems as a whole. Of the nearly 50,000 HIV+ Liberians,\textsuperscript{375} something in the range of 600 individuals were on anti-retroviral treatment through government supported programs (and these programs were often short of drugs), with no more (and likely significantly fewer) than a few thousand HIV+ patients receiving treatment via NGOs.\textsuperscript{376} In the five or so years since treatment had come to Liberian institutions,\textsuperscript{377} those few medical personnel treating HIV+ patients faced stigma from fellow medical practitioners and very limited access to drugs. There was little facilities could do beyond test for the disease; a positive result was largely seen, by patients and practitioners alike, as a death sentence.\textsuperscript{378}

\textsuperscript{373} FHI was the lead of 8 partners, “responsible for the overall program management, implementation, and M&E” (FHI 2008b, p. 2). As only FHI was, to my knowledge, involved in the Liberia component of AWARE, the other implementers are not mentioned here.

\textsuperscript{374} Interview 13, 6/1/13

\textsuperscript{375} An estimate based on multiplying the population as estimated by the 2008 census (3.5 million) by the prevalence rate (1.5\%) as estimated by the 2007 DHS.

\textsuperscript{376} Interview 28, 6/12/13; interview 32, 5/28/13

\textsuperscript{377} While NACP was established in 1987, the first case of actual treatment that could be recalled was in 2001, when a Liberian patient returning from overseas had friends who paid for the acquisition and transportation of drugs for her. (Interview 28, 6/12/13) Liberia did not have guidelines on the administration of Anti-Retroviral Therapy (ART) until 2005, when the WHO prepared such a document.

\textsuperscript{378} Interview 28, 6/12/13
Activities in Liberia commenced the week following Ellen Johnson Sirleaf’s inauguration in January 2006. Key components included the training of commercial sex workers as peer counselors, establishing support groups for those living with the disease, and strengthening care and treatment, but would focus on prevention of mother to child transmission (PMTCT). Specifically, the project sought to train health professionals regarding PMTCT, develop new national guidelines for PMTCT, and strengthen NACP’s capacity to monitor and evaluate facilities, collecting data centrally regarding activities nationwide. (NACP 2007) In the recollections of participants from FHI, USAID and MoHSW/NACP, the PMTCT training component is the primary, and in some cases the only, element of AWARE’s support which is recalled.379

The primary mechanism through which AWARE operated was via a sub-agreement with NACP. This in itself is notable – while financial management and all payments remained in the control of FHI staff, this may be the first example in post-war Liberia of a grant formally putting a Liberian government institution in charge of a component of a donor project.380 The 163,800 USD agreement came into force in June 2006, with an initial closing date of May 2007; this agreement was extended due to delayed implementation and

379 Interview 32, 5/28/13; interview 4, 5/30/13; interview 13, 6/1/13

380 As outside donor engagement was relatively novel at the time, this arrangement was perhaps not as surprising to Liberians as it is to me; as one interviewee put it, “it was springtime in Liberia”, with much new money and engagement coming in. Many Liberians in positions of responsibility at the time recall it as a confusing and exciting time; “dizzying” seems a potentially appropriate term.
without any additional funding\(^{381}\) at least four times, until an eventual closing date of March 2008. In 2008 FHI’s agreement with NACP came to an end, as did FHI’s central agreement with USAID West Africa. While there was a follow-on regional program (AWARE II) led by FHI, Liberia did not participate.\(^{382}\)

With the arrival of a new USAID senior health officer in Monrovia in early 2008\(^{383}\) USAID wished to initiate an integrated bilateral project, Rebuilding Basic Health Services (RBHS) – that is, a single project that would focus on the whole of health systems.\(^{384}\) There was certainly logic to transitioning to RBHS, including cost and management efficiencies.\(^{385}\) RBHS built not just on AWARE but also on other USAID health interventions,\(^{386}\) and was

\(^{381}\) In the language of development practitioners this is thus a No-Cost Extension, or NCE; NCEs are \textit{de rigeur} in development, and while I am aware of no formal study of same seem to occur quite frequently within USAID projects in particular.

\(^{382}\) Interview 13, 6/1/13; interview 11, 6/10/13

\(^{383}\) It is possible the arrival was in fact late 2007.

\(^{384}\) Interview 31, 6/2/13; interview 13, 6/1/13

\(^{385}\) Interview 31, 6/2/13; interview 3, 6/17/13

\(^{386}\) Chief among these was Basic Support for Implementing Child Survival (BASICS). BASICS was USAID’s primary mechanism for engagement in health systems strengthening at the central Ministry, and as such is a necessary building block for RBHS and the next generation of USAID support to Liberia’s health sector. Like AWARE, BASICS was an existing project to which USAID Monrovia ‘bought in’. The third of three five year global programs, BASICS III ran from 2004 to 2009 and focused on child health and nutrition in the main, but also on health systems strengthening more broadly. BASICS helped support the development of the National Health Plan and the Basic Package of Health Services. One additional item of support was a trip exposing MoHSW staff to the notion of Performance-Based Financing (PBF). As one senior MoHSW put it, PBF “was sold to us by BASICS.” (Interview 34, 6/13/13)
often referred to as USAID’s “flagship” in Liberia. The 62 million USD, 5-year project aimed to both deliver services directly and strengthen MoHSW’s ability to do so. (JSI 2010) Implemented by John Snow International (JSI), RBHS focused both on direct service delivery via NGOs in 7 of Liberia’s 15 counties (with a particular focus on 5 counties) while also building upon BASICS’ efforts to strengthen body the central MoHSW and the Government’s County Health Teams.

With the arrival of the next USAID health advisor, Randolph Augustin, in Monrovia in late 2010 once again came a shift in focus, in part driven by changes in global strategy at USAID HQ in Washington. Out of this grew a Fixed Asset Reimbursement Agreement (FARA), essentially an agreement to reimburse MoHSW for spending in accordance with a plan agreed between USAID Monrovia and MoHSW. A FARA was signed in September

---

387 Interview 5, 6/6/13; interview 11, 6/10/13

388 Interview 5, 6/6/13; also supported by JSI documents.

389 Via USAID Forward and a shift in USAID strategy initiated under then-new USAID director Raj Shah. Augustin was also, by disposition, much more willing to engage with MoHSW, to collaborate, than was his predecessor; shortly after arrival he initiated conversations regarding the lack of “exit strategy” for RBHS. He became an advocate for Liberia forming a pilot for the expansion of a Fixed Asset Reimbursement Agreement (FARA).

390 While the FARA mechanism was pre-existing, it had traditionally been used for physical works (hence “asset” in the name of the mechanism) such as roads, and had rarely if ever been employed in a context as fragile as that of Liberia.

391 Interview 11, 6/10/13; interview 24, 6/4/13; interview 34, 6/13/13; interview 15, 6/13/13
2011, spanning 5 years and worth 42 million USD. MoHSW now contracted NGOs, with FARA then reimbursing the Government (via monies sent to the Ministry of Finance) based on accomplishment of agreed benchmarks. These benchmarks are largely process, rather than outcome, linked, and thus are things over which MoHSW largely has direct control. RBHS continued, but as a capacity building program only; this of course required quite substantial contract renegotiations both with JSI and other implementers in the health sector.

**DFID’s Liberia Health Arc**

DFID followed a similar strategy of moving from projects to sectoral support. The first DFID health intervention in this new era was support to the Ministry of Health's Office of Financial Management (OFM). At the time the Ministry was unable to generate even simple financial statements. A single individual managed finances, storing the Ministry's financial data on his personal laptop which was itself often not functioning due to computer viruses. The Ministry was using approximately 65 separate bank accounts to manage

---


393 Interview 11, 6/10/13; interview 34, 6/13/13; interview 23, 6/11/13

394 Interview 25, 6/6/13; interview 11, 6/10/13; interview 5, 6/6/13. Some contracts that had not yet been awarded were canceled outright.

395 Interview 29, 6/7/13; interview 24, 6/4/13; interview 23, 6/11/13

396 Interview 6, 6/7/13. While it is certainly possible that corruption played some part here – that opacity was in the strategic interest of particular actors – it seems more likely that a simple lack of capacity was the primary cause.
funds, making reconciliation alone more work than a single individual could do were he to have the capacity and disposition to do so; the last financial report that could be located was almost 20 years old.\footnote{397 Interview 29, 6/7/13}

The OFM project commenced in June 2007, just about the same time that DFID’s country representative arrived. The two-year project had an initial ceiling of 745,000 GBP, which rose to approximately 2.25 million GBP (including extensions in scope to include the pooled fund – discussed below – as well as extension from a 2 to 3 year project) before coming to a close in June 2010. Price Waterhouse Coopers (PwC) was contracted to implement the project, working primarily through a Ghanaian expert, Alex Nartey, placed inside MoHSW to establish the unit. Nartey worked alongside a Liberian counterpart, Toagee Karzon, who had previously been the chief accountant for the UN Development Program (UNDP) in Liberia.\footnote{398 Interview 6, 6/7/13; interview 22, 6/11/13; interview 24, 6/4/13} Nartey served as the OFM’s head, with Karzon supporting.\footnote{399 Interview 24, 6/4/13; interview 22, 6/11/13; interview 29, 6/7/13; interview 6, 6/7/13} At its close in 2010 DFID did not renew the OFM project; however the Pooled Fund that DFID had played a lead role in establishing (to be discussed below) continued to fund support to the OFM. UNICEF took DFID’s supervision role and Ernst & Young won the tender and replaced PwC, but retained the actual implementing personnel; Nartey moved from implementation to support, while Karzon began leading day-to-day functioning of the
This support to OFM ended in 2012 and the OFM is now functioning, by all accounts quite successfully, without any external support, with staff now contracted by the Ministry directly.\footnote{Interview 6, 6/7/13; interview 22, 6/11/13}

One smaller component of PwC’s OFM contract was to explore possibilities for a Health Sector Pooled Fund (PF) that would fund activities in support of the newly created Government National Health Plan.\footnote{Interview 24, 6/4/13; interview 6, 6/7/13; interview 22, 6/11/13; interview 23, 6/11/13} Pooled funds – multidonor efforts to pool money in a particular sector – are a preferred modality of DFID’s worldwide. Given DFID’s smaller profile in Liberia, DFID saw one potential ‘comparative advantage’ as its ability to promote joint donor efforts such as pooled funds.\footnote{Interview 2, 6/25/13}

In implementation, it quickly became clear that the PF was as much a focus of DFID as was the OFM.\footnote{Interview 2, 6/25/13; interview 24, 6/4/13} The initial 3-month PwC position to explore the possibility of a PF was extended to nine months, then in extended another two years to actually manage the start up of the PF; management of the PF became part of PwC’s contract in March 2008, which is also when the fund received its first funding – 8 million USD from DFID.\footnote{Interview 24, 6/4/13} The pooled fund
gained additional donor participation from IrishAid in December 2008, followed by UNICEF and UNHCR.406

The primary purpose of the PF from the standpoint of MoHSW was to increase control over donor funding and ensure it was put in service of the Government’s wishes.407 Donors contribute to the fund, which is managed by a pooled fund manager (initially employed by PwC, then Ernest & Young, and since 2012 Hughes Development International).408 Proposals are made to a steering committee composed of donors and the government (including both MoHSW and the Ministry of Finance); once agreed, funds flow from the pooled fund accounts through normal government channels (including the OFM).409

**Health Intervention Results**

Having laid out these narratives, and before getting to the “why”, it seems reasonable to ask: Is there a significant gap in performance between DFID and USAID? If so, which organization performed better? Aiding this comparison is the similar behavior by USAID and DFID prior to the period examined here (funding NGOs and humanitarian

406 Interview 2, 6/25/13; interview 24, 6/4/13; interview 29, 6/7/13

407 Interview 29, 6/7/13; interview 23, 6/11/13; interview 34, 6/13/13

408 Interview 29, 6/7/13. Hughes Development International (HDI) is a firm set up by Jacob Hughes, the PwC consultant who assisted in setting up the Pooled Fund. By way of full disclosure, I have previously worked on an HDI contract (though in South Sudan, not Liberia).

409 Interview 24, 6/4/13; interview 29, 6/7/13
endeavors) and the same ‘exogenous shock’ – the end of conflict and the election of Ellen Johnson Sirleaf – prompting a shift in behavior. For both organizations this shift prompted a move from NGO funding to projects aimed at building government health systems, then from project-based support to broad sectoral funding mechanisms. Table 5.2 provides a sense of these parallel trajectories, summarizing USAID and DFID intervention milestones described above on a common time scale.

Table 5.2: Timeline of Key Activities in the Liberian Health Sector, USAID & DFID

<table>
<thead>
<tr>
<th></th>
<th>Pre 2005</th>
<th>2006 (EJS Election)</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DFID</strong></td>
<td>Funds NGOS</td>
<td>Agrees to fund OFM</td>
<td>Opens office in Monrovia; OFM project begins</td>
<td>OFM project expands to include PF, PF receives 1st contribution (from DFID)</td>
<td>OFM/PF project extended by one year</td>
<td>OFM project ends</td>
<td>PF contribution continues</td>
</tr>
<tr>
<td><strong>USAID</strong></td>
<td>Funds NGOS</td>
<td>AWARE, BASICS buy-in begins</td>
<td>AWARE ends; RBHS begins</td>
<td>BASICS ends</td>
<td></td>
<td></td>
<td>FAR A initiated; RBHS altered</td>
</tr>
</tbody>
</table>

Since the broad goals of each of these projects aimed at health system strengthening the focus in making performance comparisons will be on the effects of interventions on systems while projects were underway and, in the case of completed projects, after the project departed. In making this comparison the section below breaks from the structure of the comparative performance sections that have preceded it, instead grouping USAID and DFID interventions by type in comparing across interventions.410

410 This is done for the purposes of simplifying the comparison; there are just too many interventions here to do anything like a systematic job in examining all of them in combination while maintaining narrative clarity.
Projects: DFID's Support to OFM & PF Preparation; USAID’s AWARE & RBHF

Both USAID & DFID attempted via early projects to work through, and strengthen the systems of, Government in addressing PMTCT (USAID’s AWARE) and MoHSW financial management (OFM).

The OFM continues today as a competent, widely praised unit without external support; and while external support to the OFM continued beyond the DFID-funded project, the DFID project did a great deal of the hard work of getting the OFM up and running and pointed in the right direction.  

Many actors noted that the “cleaning house” inside the Ministry's finance unit which this project accomplished is a type of task that has often been undertaken with donor support but rarely been accomplished. The President singled out MoHSW as the Ministry to have advanced furthest in improving financial management systems in her January 2009 annual message (the rough equivalent of the US State of the Union), also explicitly linking this improved financial management to the success of the Pooled Fund, saying “The Ministry of Health has made the most progress in the establishment of financial management systems. In recognition of this achievement, the Pool Fund established last year for the sector attracted the support and the praise of several of our development partners.” (Sirleaf 2009)

---

411 E.g. interview 11, 6/10/13; interview 34, 6/13/13; interview 24, 6/4/13; interview 23, 6/11/13; interview 22, 6/11/13

412 e.g. interview 21, 6/4/13
One actor close to the OFM project described it as “one of the more effective projects that I’m happy to have been part of”, a view broadly representative of actors from DFID, PwC, and MoHSW alike. These improved financial systems also made routing funds through government systems more attractive. By one estimate, assets under management by MoHSW rose from under 10 million USD in 2007 to nearly 70 million by 2012, with only a small fraction of that coming from Government coffers.

By contrast, very little, if any, of AWARE’s efforts continued once funding ended; trainings ceased, and equipment purchased by the project became contentious. The exception is national PMTCT guidelines that were developed under AWARE in collaboration with the United Nations Children’s Fund (UNICEF). While not implemented initially, these guidelines later guided national PMTCT policy. Staff whom had been involved in AWARE departed for NGOs and other institutions, and it seemed that there was a general demotivation regarding PMTCT. As one individual involved from the donor side put it, “If you go after 2 years and you suddenly stop when you are about to scale up, have the momentum to go to the next level... you will not achieve all that you set out to do. Wouldn’t achieve everything that we wanted to achieve through that means... Something

---

413 Interview 6, 6/7/13

414 Interview 24, 6/4/13

415 There are many actors with continued hard feelings about this; e.g. interview 32, 5/28/13. The dispute was between the hospital where trainings took place (Catholic Hospital) and the Government National AIDS Control Program (NACP).

416 Interview 4, 5/30/13; interview 16, 5/31/13

417 Interview 4, 5/30/13; interview 9, 5/30/13; interview 10, 5/29/13
got lost when the program ended in the way it did."\textsuperscript{418} Lwopu Bruce, then the NACP deputy head, reported that the decline in activities was noticed by beneficiaries; “[Clients] would say what happened to you people? We don’t hear from you people.”

RBHS was notably more impactful; it accomplished much in terms of service delivery, as well as pioneering innovations such as performance based financing that have since become national policy.\textsuperscript{419} That said, in an echo of PEPFAR in South Africa this effective structure was outside Government, with (in the views of many inside and outside USAID) detrimental implications for RBHS’ long term ability to sustainably improve health systems, the project’s primary goal.\textsuperscript{420} RBHS has produced a significant number of reports, strategies, and guidelines. (JSI 2009, 2010) This effort was rarely mentioned by interviewees, who almost always focused on the direct service delivery components of RBHS.\textsuperscript{421}

In response to open-ended questions regarding RBHS it was striking how often individuals framed the project as, in the words of one individual, “basically running a parallel health system.”\textsuperscript{422} This parallel system had a significantly different salary scale to that of the Government. While MoHSW salaries average 85 USD/month, drivers and

\textsuperscript{418} Interview 13, 6/1/13

\textsuperscript{419} Interview 5, 6/6/13; interview 35, 6/11/13

\textsuperscript{420} Interview 11, 6/10/13; interview 24, 6/4/13; interview 23, 6/11/13

\textsuperscript{421} e.g. interview 8, 6/9/13; interview 24, 6/4/13; interview 29, 6/7/13

\textsuperscript{422} Interview 24, 6/4/13
janitors for RBHS were by at least one account making three times as much;\textsuperscript{423} and even were that to overstate the differential, there are multiple cases of staff leaving MoHSW for RBHS employment.

The general sense was that RBHS was a well designed, well executed project run by earnest, caring individuals at JSI.\textsuperscript{424} JSI/RBHS made significant efforts to include the Ministry in critical decisions to the extent the project structure allowed.\textsuperscript{425} It was just separate from the government, using a different operational approach – efficient but not necessarily transferrable to Government, and thus possibly requiring perpetual external support. Little financial information flowed back to the Ministry, making it difficult to know precisely what was being accomplished or to get a holistic picture of spending in Liberia.\textsuperscript{426} DFID’s OFM project was significantly more successful than USAID’s AWARE and RBHS, certainly on a value for money basis.\textsuperscript{427}

\textit{Sectoral Funding: PF & FARA}

USAID and DFID then shifted to support via broader mechanisms. For DFID this was via the Pooled Fund (PF) whose development DFID supported as part of the OFM project, \hfill

\textsuperscript{423} Interview 29, 6/7/13

\textsuperscript{424} Interview 34, 6/13/13; interview 8, 6/9/13

\textsuperscript{425} Interview 5, 6/6/13; interview 25, 6/6/13; interview 34, 6/13/13

\textsuperscript{426} Interview 37, 5/27/13

\textsuperscript{427} RBHS is approximately ten times the size of the OFM support project.
then made the first contributions towards. For USAID this was via a Fixed Assets Reimbursement Agreement (FARA), via which USAID reimburses the ministry based on explicitly measured performance in accomplishing discrete tasks.

The PF was a marked success. As one study put it, prior to the Pooled Fund “Multiple donors funded multiple NGOs within a few geographic areas, which resulted in no clear sense of which organization was responsible for providing what services to what population.” (Hughes, Glassman, and Gwenigale 2012, p.7) The shift to the PF – with NGOs being contracted but out of central ministry revenues – initially improved this coordination problem but with a decline in health delivery, as MoHSW was unable to fill the supervision role previously played by donors.\textsuperscript{428} The PF Steering Committee responded by funding monitoring and evaluation from the PF’s resources, and performance improved.\textsuperscript{429} NGOs have increasingly been held to account, with performance compared to workplans which are designed in consultation with MoHSW; performance is by all accounts now well above levels in the period prior to PF financing.\textsuperscript{430}

Liberian actors were universal in their support for the PF as an important step in gaining control over the sector and in improving coordination; as Liberia’s Chief Medical Officer,\textsuperscript{431} Dr. Eunice Dahn, put it “From the word ‘go’, from the time we established the
\footnotesize
\begin{footnotes}
\item[428] Interview 34, 6/13/13; interview 24, 6/4/13
\item[429] Interview 24, 6/4/13; interview 29, 6/7/13
\item[430] Interview 34, 6/4/13
\item[431] In early 2015 Dahn became Minister of Health.
\end{footnotes}
Pooled Funding, the Ministry of Health has been in the driver’s seat.” Beginning in 2011 the PF also piloted what Liberian health officials call “contracting in” – paying the government’s county health team rather than an NGO to implement services; this pilot is broadly seen as a success, and as of late 2013 was due to expand.\textsuperscript{432} The PF remains a critical, and widely positively regarded, institutional feature of the Liberian health sector.\textsuperscript{433}

DFID’s advocacy and stewardship was critical to the Pooled Fund’s establishment; the PF has improved allocative efficiency and government ownership over the sector. As DFID put it in a publication regarding lessons learned from joint donor funding mechanisms:

In a context where major donors by volume do not participate in pooled funds arrangements, the Liberia pool fund raised the relevance and effectiveness of multiple small donors while increasing the stewardship role of the Ministry of Health and Social Welfare. As the fund is managed from within the MOHWS, uses country systems, and the steering committee is chaired by the MOHWS, the fund has been successful at increasing the MOHWS’s ability to coordinate health sector donors and increasing the MOHWS’s fiscal decision making-space. (DFID 2013b, p. 15)

FARA is also seen as a move in the right direction by a wide variety of actors as compared to RBHS and other USAID health interventions.\textsuperscript{434} The performance of facilities transitioning from RBHS to FARA declined initially but not precipitously in part due to

\textsuperscript{432} Interview 34, 6/13/13; interview 24, 6/4/13

\textsuperscript{433} Interview 23, 6/11/13; interview 34, 6/4/13; interview 24, 6/4/13

\textsuperscript{434} Interview 11, 6/10/13; interview 34, 6/13/13; interview 23, 6/11/13; interview 37, 5/27/13
delays in MoHSW processes as the Ministry adjusted to the additional administrative burden; one observer characterized this as similar to the decline one might expect e.g. if there were a change in contractors between phases of a large USAID project.\footnote{Interview 11, 6/10/13; interview 34, 6/13/13 (characterization in interview 11)} Performance has continued to improve since, and FARA is largely seen as successful in delivering services.\footnote{Interview 34, 6/13/13; interview 11, 6/10/13}

While FARA is largely seen as a success, it has been constrained by the pre-financing requirement; that is, the need for MoHSW to spend only to be reimbursed by USAID ex-post.\footnote{Interview 11, 6/10/13; interview 19, 6/11/13; interview 37, 5/27/13} In Liberia’s constrained fiscal environment, finding the money to finance operations has not always been easy. One observer noted it was “quite optimistic to think a poor country is going to be able to pre-finance a large USAID health portfolio.”\footnote{Interview 24, 6/4/13}

FARA also places a quite significant administrative burden on the Ministry, and a weekly meeting is held between USAID, RBHS, MoHSW’s FARA coordinator (a dedicated position), and senior MoHSW personnel.\footnote{Interview 15, 6/13/13; interview 11, 6/10/13; interview 34, 6/13/13} One senior MoHSW official described USAID as “Guiding the process so that we can’t make mistakes.”\footnote{Interview 34, 6/13/13} By all accounts these meetings are

\footnote{Interview 11, 6/10/13; interview 34, 6/13/13 (characterization in interview 11)}

\footnote{Interview 34, 6/13/13; interview 11, 6/10/13}

\footnote{Interview 11, 6/10/13; interview 19, 6/11/13; interview 37, 5/27/13}

\footnote{Interview 24, 6/4/13}

\footnote{Interview 15, 6/13/13; interview 11, 6/10/13; interview 34, 6/13/13}

\footnote{Interview 34, 6/13/13}
critical to FARA’s success; however, it is hard to imagine that MoHSW would be able to have a similar meeting with every donor were that to be required for fund distribution.441

FARA increased government ownership of USAID funds, though those funds were still not as flexible or helpful as those in the Pooled Fund.442 While PF resources can be deployed (and revised) as the steering committee sees fit, revisions to FARA require a formal negotiation involving USAID HQ (revisions have indeed occurred, though the process for doing so has been somewhat lengthy).443 In addition, unlike contributors to the pooled fund, in the case of FARA risk is borne by the Ministry, rather than the contributing donors. Minister of Health Gwenigale spoke at length regarding both the superiority of FARA to RBHS and its continued weaknesses as compared to the PF:

People who put the money in the Pooled Fund, they are taking their risk. They put it there knowing what we will do with it... even though I have a good management system put in place to manage it in the Pooled Fund, they still are taking the risk they put their money there. The Americans are not taking any risk; because what they are doing is ‘use your money, if we want to we will pay you back or we will not pay you back if we don’t accept what you have done’.... They are not taking any risk by saying use your own money and we’ll replace it. The benefit for me, however, is I can control how much overhead they get from my money. If I use 10 or 14 million dollars for health, and they replace it, even if I was taking that money to contract someone to do

441 That said, FARA is a quite large sum of money; were the administrative burden to be scaled by contribution size, that might be more sustainable administratively.

442 Interview 23, 6/11/13; interview 34, 6/13/13

443 Interview 15, 6/13/13; interview 11, 6/10/13
something for me, I would not be giving an overhead as big as what John Snow gets when they get this contract signed in Washington; because when you sign in Washington for that money [JSI] will be using some for Boston where John Snow has their office... Here, when I use my own money to be reimbursed, I’m not going to be giving that kind of overhead. So there is a benefit to me to use my money... If [USAID] give it to a group in the US, they will be using it every time they fly in and so forth, so that overhead will be much larger, and [there] will be less to do the things I decided. – Dr. Walter Gwenigale, Liberian Minister of Health

Once again, DFID’s intervention – the Pooled Fund, now wholly owned and managed by Government, with DFID continuing to contribute – seems the more successful intervention as judged at the time of interviews (mid-2013).

Viewing these arcs of support in aggregate, USAID delivered less than DFID even without adjusting for DFID’s smaller level of funding. It is striking that DFID focused on interventions that more greatly empowered a recipient Government both USAID and DFID were interested in strengthening. DFID’s interventions also were delivered in a way that naturally led to their continuation (including working more closely through MoHSW).

**Barriers to Success**

The various interventions discussed here had a variety of issues in implementation; AWARE was particularly beset by conflict, though issues were mentioned regarding a

---

444 Interview 20, 8/30/13; interview 16, 5/31/13; interview 28, 6/12/13; interview 32, 5/28/13
number of the projects discussed above. These difficulties are unconvincing in explaining the differences between USAID and DFID’s success in the health sector, however. DFID benefited from responding more quickly and fully to the Liberian Government, in so doing empowering the reformers that both DFID and USAID sought to support. The clear question, then, is what prevented USAID from responding in the same manner as DFID. This is all the more striking in that USAID had exactly the same opportunities as did DFID.

Indeed, the OFM project – a product of the thinking of Minister Gwenigale and his senior advisor, Bill Martin – was first presented to USAID for funding.\textsuperscript{445} In April 2006 MoHSW developed a proposal for donor support to create and staff an office of financial management with the goals of improving accountability and management as well as attracting additional funding; this plan included a proposal for a ‘basket fund’ which was to evolve into the Pooled Fund.\textsuperscript{446} The proposal was first discussed with USAID, who seemed by some accounts “not enthused”\textsuperscript{447} by the idea; an approach was then made to the US Ambassador to Liberia, Donald Booth, directly by Minister Gwenigale.\textsuperscript{448} Ambassador Booth was receptive and pushed USAID to engage; in conversation with USAID it was estimated it would take over a year for funding to get an OFM up and running.\textsuperscript{449} While these discussions were ongoing (in approximately July 2006) Marcus Manuel, DFID’s

\textsuperscript{445} Interview 29, 6/7/13; interview 23, 6/11/13; interview 24, 6/4/13

\textsuperscript{446} Interview 29, 6/7/13

\textsuperscript{447} Interview 29, 6/7/13

\textsuperscript{448} Interview 29, 6/7/13; interview 23, 6/11/13

\textsuperscript{449} Interview 29, 6/7/13; interview 23, 6/11/13
Deputy Director for West Africa, visited the region. Gwenigale and Manuel spoke at a cocktail party, where Gwenigale requested support; DFID agreed to fund it quite quickly and the project moved forward.\textsuperscript{450} Similarly, MoHSW tried to push USAID to contribute to the Pooled Fund, but was told repeatedly that Congressional authorization made this impossible.\textsuperscript{451} Key USAID actors also wished, at least initially, to contribute to the pooled fund; FARA was in some ways an attempt to approximate the benefits of the PF within USAID’s rules.\textsuperscript{452}

In the OFM project DFID was implementing a project USAID had proved less responsive to; in the PF DFID had catalyzed a fund at least some USAID actors would have preferred to their own mechanisms, if it had been possible for USAID to implement in this manner. USAID organizational constraints certainly limited the exercise of agent judgment in the latter (PF) case, and may have done so in the OFM case as well.

It is intriguing to entertain what might have transpired had the OFM reform indeed been funded by USAID. My view is that it is hard to imagine USAID effectively navigating the politically difficult, soft information-demanding OFM project. I also have trouble imagining the Pooled Fund emerging in quite the same way, or quite as quickly, from the original proposal’s “basket fund” idea if the OFM project had been under USAID’s

\textsuperscript{450} Interview 21, 6/4/13; interview 23, 6/11/13; interview 29, 6/7/13; interview 36, 6/25/13

\textsuperscript{451} Interview 23, 6/11/13; interview 34, 6/13/13

\textsuperscript{452} Interview 11, 6/10/13
stewardship. DFID’s initial PF investment clearly aligned DFID’s reputation with the success of the Pooled Fund, something that would have been impossible for USAID, which due to congressional restriction could not have contributed to a pooled fund. It seems somewhat unlikely that USAID would have fought to create a pooled fund to which it could not contribute, though of course it is impossible to know this with certainty.

**Differences in Navigation Strategies: USAID and DFID Health Support Arcs**

As touched on briefly above, constraints from above seem to have constrained USAID field actors’ decisions in at least one case – contributing to the pooled fund. Was this restriction indicative of broader Navigation from the Top? Additionally, how did DFID navigate its way through this support arc?

*Design*

The decision to buy in to AWARE was made by the USAID mission in Liberia based largely by remote (DC-based) staff.\(^{453}\) The specific design of the intervention was also done on some interviewee’s views largely within USAID and AWARE’s implementer, FHI;\(^ {454}\) an unusual choice for a project whose aim was improved government systems.

---

\(^{453}\) Interview 13, 6/1/13; while the decision was made in-country, it seems it was substantially influenced by DC supervisors.

\(^{454}\) Interview 32, 5/28/13; interview 17, 5/28/13; interview 10, 5/29/13
RBHS seems to have been motivated first and foremost by the judgment of a new USAID health advisor that an integrated program was the right next step. This design shows Navigation by Judgment, aided by lobbying in DC that created the space for this decision and helped make the advisor’s advocacy of such a position career enhancing.

The move to FARA was similarly driven in substantial part by the judgment of a newly arriving health advisor, albeit one in consonance with changing views at headquarters. In 2010 USAID launched USAID Forward, which included new ways of operating, including a move towards working via developing country’s systems. In a fairly substantial shift, the percentage of mission funds flowing through local systems would now be one of the eight “topline” (that is, primary) indicators for progress on implementing USAID Forward. In this environment RBHS was now out of step; what had

---

455 Interview 31, 6/2/13

456 Interview 25, 6/6/13; this lobbying may have included JSI, the firm that was eventually to win the competitive tender to implement RBHS. Career enhancing from interview 7b, 6/12/13

457 Interview 11, 6/10/13; also supported by the USAID forward website, www.usaid.gov/usaidforward. USAID had historically been one of the donors perceived as most reticent to work through country systems – to have funding actually flow via government procurement and budgeting systems.

458 Those indicators are: Number of country development cooperation strategies (CDCSs) completed; Number of completed high quality evaluations; Number of emerging foreign service officer (FSO) and foreign service national (FSN) leaders participating in formal mentoring programs; Mission program funds per year implemented through local systems; U.S. staff work plans with at least one direct engagement performance measure; Total value of public-private partnerships with a minimum of 1:1 leverage; Percent of program funds dedicated to leveraging commercial private capital through the Development Credit Authority; and Number of missions actively supporting mobile money initiatives. These indicators are, on the view of this work, troubling in their potential for distortion.
once seemed innovative now smacked of the parallel systems USAID was newly seeking to avoid.

On the DFID side, the re-engagement with Liberia that opened up the space for DFID programming was very much driven by high politics. As one person close to the decision put it, in response to Ellen Johnson Sirleaf’s call for engagement DFID felt they “should be seen to be responding.” DFID began preparations towards the launching of a field office in early 2006, undertaking a series of trips to Liberia to attempt to determine priorities and begin projects even before the arrival of a Monrovia-based country representative.

In contrast to this Navigation from the Top the decision to support the OFM project in specific was largely a judgment call, albeit one made by a senior London-based official (Marcus Manuel) while in Liberia. In making this decision, however, Manuel relied on the judgments of DFID Liberia field staff – namely, the view that Gwenigale and the MoHSW leadership were reformers whose hand ought be strengthened if possible. This was one of the key drivers of DFID’s internal support for this unusual project. This highlights one critical element of Navigation by Judgment – the way in which an autonomous, secure organization has the opportunity to base critical decisions on what might be called

459 Interview 2, 6/25/13
460 Interview 2, 6/25/13; interview 21, 6/4/13
461 Interview 2, 6/25/13; interview 21, 6/4/13; interview 36, 6/25/13
462 Interview 36, 6/25/13; interview 21, 6/4/13. Mr. Manuel's direct conversation with the Minister while visiting Monrovia played a role as well in this decision.
“derogated judgment”. An HQ official deferring in part to the judgment of field staff whom the HQ official himself trusts is just such an example, and precisely the sort of thing that the literature\textsuperscript{463} would predict is likely to encourage field staff to gather soft information.

Presence may be necessary, but does not seem to be sufficient for successful design. In the case of AWARE the USAID’s mission presence seems not to have counted for very much, with minimal involvement of MoHSW actors. In the case of RBHS there was a much more concerted attempt to incorporate local views and knowledge; however, there were constraints with regards to the degree to which those views could be incorporated imposed by the structure of the vehicle (an integrated project) via which support was being delivered. While FARA is very innovative by USAID terms, and certainly laudable, there were constraints as to the possible scope of the project.

Both USAID and DFID’s designs involved some degree of Navigation by Judgment. The most striking difference between USAID and DFID regarding project design is that in each USAID case the idea for the project came primarily from inside USAID, though informed by events going on in-country; in each DFID case the idea for the project came from Liberia itself. In addition, USAID actors (particularly as regards FARA) seem to have been constrained regarding how far their judgment could push – that is, the range of options inside the ‘envelope’ in which their judgment, their discretion, was operable.

\textsuperscript{463} e.g. Aghion Tirole, discussed in chapter 2.
Implementation

The implementation arrangements of both USAID and DFID support in the health sector both used contractors for their projects. FHI (in AWARE) and to a lesser extent JSI (in RBHS) both worked with relative autonomy from their USAID supervisors. While USAID personnel were present in Liberia, they were not actively involved in day-to-day activities in the health sector. This is not to say that either organization just did as they saw fit; both FHI and JSI worked under regimes of reporting and documentation that imposed significant *de facto* constraints on their behavior. AWARE reported on measures such as the number of people trained; the sub-agreement between AWARE and NACP had process and output indicators against which NACP was reporting. RBHS was also very focused on meeting quantitative targets, and encouraged facilities that it funded to do similarly, in part via performance based contracts where providers were rewarded for meeting both output and administrative indicators.

Within whatever space was left available, the contractors (or, in the case of AWARE, NACP as-delegated by the contractor) had control. FARA in some sense deepens this approach, with the measurable targets the explicit focus of attention, the means to achieve them explicitly left in Government’s hands (though *de facto* with USAID rather involved to

---

464 Interview 20, 8/30/13; interview 25, 6/6/13

465 Interview 11, 6/10/13; interview 5, 6/6/13

466 Interview 20, 8/30/13; also supported by NACP-AWARE agreement (NACP 2006)

467 Interview 25, 6/6/13. An example of an output indicator used is % of pregnant women seen having 4 prenatal visits; an administrative indicator, submitting reports on time. This focus on the quantitative is also supported by RBHS annual reports.
ensure targets are met, as noted above). This is quite similar to USAID’s method of
navigation in South Africa’s health sector – Navigation by Measurement seems, as predicted
by chapter 2, to have given implementers discretion over the means to achieve those
measures.

DFID’s small office in Monrovia was technically reporting to HQ via the much larger
office in Sierra Leone; however neither Freetown nor London played an active role, and
DFID personnel on the ground had a relatively free hand to manage as they saw fit.\textsuperscript{468} For
PwC (in supporting OFM and developing the PF) the situation was quite different than for
JSI or FHI. While the formal contracting mechanism of PwC was similar to that of USAID
contractors, and there were also targets against which to report, the contract and its
targets seemed less binding, less constraining to PwC.\textsuperscript{469} PwC also did not have the same
zone of practical autonomy as did the USAID contractors. DFID personnel seem to have
been frequently involved in PwC’s day-to-day activities, in a manner not quite \textit{as} hands on
as USAID OTI regarding BRDG, but moving in that direction.\textsuperscript{470} Indeed, what was effectively
the envelope of discretion available for USAID implementers – albeit subject to USAID
reporting requirements – was effectively exercised by DFID rather than contractors.

Underscoring USAID’s propensity for Navigation by Measurement, the USAID arc of
support also encouraged the use of performance based financing \textit{within} MoHSW – that is,

\textsuperscript{468} Interview 21, 6/4/13; interview 2, 6/25/13

\textsuperscript{469} Interview 27, 6/13/13

\textsuperscript{470} Interview 21, 6/4/13; interview 27, 6/13/13
paying individuals or teams based on measured results. One of the ways in which RBHS’ systems differed from MoHSW’s was via the use of performance based contracting of implementing NGOs, or the paying of incentives based on performance. The use of performance based contracting was in fact MoHSW’s policy, though MoHSW was not implementing it; (MoHSW 2008) it had become MoHSW’s policy via the exposure and advocacy of BASICS, USAID’s previous project.471 RBHS was quite proud of its performance based modality, and published work highlighting same. (Brennan et al. 2010) The FARA-reimburseable contracts signed by MoHSW with NGOs were the first MoHSW contracts to truly operationalize performance-based financing.472 A performance-based financing office was established in MoHSW with FARA support and RBHS technical assistance, and a facility-level bonus scheme based largely on facility outputs was devised.473 This performance scheme has been expanded beyond the counties directly funded by FARA, with intentions to make it nationwide.474

471 Interview 34, 6/13/13

472 Some Pooled Fund contracts signed beginning in 2009 had claimed to be performance-based, but had no standard for performance assessment and paid no bonuses. (MoHSW 2013b, p. 8)

473 Interview 30, 6/11/13; interview 35, 6/11/13

474 The tight Navigation by Measurement which is FARA’s core operating modality has driven attention towards FARA targets; and it is possible that FARA will have an effect somewhat parallel to PEPFAR in the long term, with this Navigation by Measurement improving health performance. I view this as unlikely; unlike in South Africa’s PEPFAR, FARA’s targets are – by design – focused on process rather than outcomes so as to ensure these targets are fully within Government’s control. (Interview 15, 6/13/13; interview 11, 6/10/13) These process indicators are, unlike the outcome indicators of PEPFAR, very much susceptible to Goodhart’s Law.
Revision

DFID’s OFM was not – on the view of DFID actors – initially on its way to the success it ultimately enjoyed; the initial PwC staffer assigned to support OFM was seen as not up to the job.\textsuperscript{475} DFID actors – in a clear example of Navigation by Judgment – pushed for the individual’s replacement. This push came quickly enough that it would have been impossible for distant supervisors to see measures of poor performance. This revision proved critical in leading to the arrival of Nartey, who was ultimately to lead the project to the success described above.

The expansion of the Pooled Fund component also relied on DFID staff’s judgments about the viability of the fund.\textsuperscript{476} The investment in the Pooled Fund’s development also facilitated DFID’s exit strategy; in 2011 DFID removed its Liberia office, continuing to contribute to the Pooled Fund but with no active program to manage.\textsuperscript{477} As such in this case the judgment of DFID’s field staff was nicely aligned with the interests of DFID HQ, meaning that some degree of field staff judgment oriented towards career concerns rather than achieving maximal impact cannot be dismissed.

For USAID by far the most significant change in this period was RBHS’ shift away from direct service provision to a focus solely on capacity building with the initiation of the FARA. Indeed, the FARA is a sort of revision in its entirety, as it required a significant

\textsuperscript{475} Interview 29, 6/7/13; interview 24, 6/4/13; interview 21, 6/4/13

\textsuperscript{476} Interview 2, 6/25/13

\textsuperscript{477} This exit is also what led to CISCAB’s close, as discussed in the capacity building section.
reprogramming of USAID’s preexisting commitments in the health sector. Advocacy of USAID field-level personnel, in particular USAID’s health advisor Augustin, played a substantial role in this revision;\textsuperscript{478} while it may well have been Augustin’s judgment that a move towards FARA was a good idea, it was one very much in keeping with corporate changes from the top of USAID. This suggests that individuals and their judgments are consequential, at least sometimes – that it is not the case that, even at USAID, organizational rules dictate all outcomes.

The massive RBHS revision shows that given enough political will from above, major revision can occur. This is the exception, however; a number of actors characterized USAID as much less flexible and adaptable, making statements to this effect when asked to compare specific USAID and DFID projects or USAID and DFID more generally.\textsuperscript{479}

\textit{Differential Health Navigation in Sum}

In the Liberian health sector USAID’s Navigation from the Top performed less well than DFID’s Navigation by Judgment, the opposite relationship as that seen in the South African health sector. This is as predicted in chapter 2, with Navigation from the Top performing less well in Liberia (the more uncertain & opaque environment) than South Africa. The cases here offers a few examples in which Liberia’s uncertainty played a critical role in the difficulties faced by Navigation from the Top or Navigation by Measurement;

\textsuperscript{478} Interview 25, 6/6/13; interview 11, 6/10/13

\textsuperscript{479} Interview 55, 5/29/13; interview 54, 6/5/13; interview 69, 6/3/13; interview 54, 6/5/13
arguably AWARE’s failings were related to the difficulties of ex-ante planning and external observability, as is the need for revision in FARA.

It is striking that – in parallel to the South African health sector, and to a lesser extent all other USAID projects - all the USAID interventions were designed so as to be highly externally monitorable. FARA, RBHS, and AWARE all focused on observed outputs whereas the OFM project and Pooled Fund financing did not. Indeed, it seems that the design of the OFM project - like CISCAB before it - could only have succeeded with an IDO that was comfortable with high degrees of opacity and did not require externally observable progress (e.g. via measurement) in the short term. DFID’s projects had elements of Navigation from the Top; USAID also had some ability to Navigate by Judgment. Nonetheless, DFID had much greater propensity to Navigate by Judgment, whereas USAID tended to Navigate by Measurement. Even USAID Forward – USAID’s reform effort – engaged in Navigation by Measurement, with the ‘topline’ indicators focuses of attention by USAID staff.\textsuperscript{480}

The navigation methods employed in the health sector seem quite related to political authorizing environments and organizational autonomy. As was the case in South Africa, Congress cast a long shadow in Liberia’s health sector. The word “Congress” was mentioned at least eight times in interviews in the context of constraints on design and management, with the majority of those mentions not from USAID personnel. One USAID official quipped, in response to a question about a tight monitoring and reporting regime,

\textsuperscript{480} Interview 7b, 6/12/13
“Congress doesn’t trust us [USAID]. Thinks we’re all a bunch of hippies.” By contrast, Parliament was mentioned precisely zero times by any actor.

DFID operates with more autonomy relative to its political authorizing environment than does USAID. The constraints that authorizing environment places on USAID have an effect on project design. DFID’s greater organizational autonomy relative to their political authorizing environment also arguably played a role in the advocacy of collective funding mechanisms, and pooled funds, more generally; indeed, one need look no further than the counterfactual of USAID’s inability to engage in similar mechanisms and the explicit linking of this to Congress to demonstrate this. It is plausible that DFID’s organizational autonomy is linked to their corporate position favoring pooled funds. The increased difficulty of directly attributing contributions to results and loss of direct control over funds necessitated by contributions to pooled funds may make it more difficult for a more constrained IDO like USAID to make similar contributions to pooled funds.

5.4: The Impacts of Constraint & Navigation Strategy

The previous sections focused on the different navigation strategies employed by USAID and DFID as well as the relationships between these navigation strategies and success. Chapter 2 has some specific hypotheses via which Navigation from the Top and Navigation by Measurement are anticipated to affect project success – namely, through the

481 Interview 7a, 5/23/13
ability to utilize soft information in decision-making and through the motivation and quality of staff. This section focuses on evidence from both the capacity building and health cases in investigating these channels.

**Soft Information**

In these cases we have multiple examples of soft information proving critical to project success. The decision to fund the OFM project is a clear judgment call based on DFID field actors’ soft information; then in implementation the replacement of the initial project lead – twice – in search of someone with the skills to succeed required soft information. In this latter case in particular there was some cost to DFID staff; changing controller was direct involvement in an agreed contract, and to push for the individual’s replacement required some confrontational statements by DFID to PwC.

The assessment that the initial PwC OFM controller was inadequate is precisely the kind of observable but non-verifiable data – something field-level actors can observe and agree on but it would be difficult to prove to a dubious supervisor if required to justify – that chapter 2 suggests will be excluded from a production process when actors do not feel, or in fact are not, sufficiently autonomous. This case fits the theory well – field-level actors, confident in their judgments and in control of the project from the point of view of their organization pushed for a change not in *de jure* project rules but in *de facto* performance via the replacement of the key project personnel assigned by the contractor. This was done
at field level without the substantial involvement of DFID HQ; it is hard to imagine a USAID officer taking a similar decision without significant consultation with Washington.

OTI’s administration of BRDG also allows an examination of whether OTI’s increased envelope for independent action and involvement in project management and implementation (what Lawson 2009 describes as OTI’s “hands on approach”) translated into BRDG’s greater use of soft information. That BRDG was more empowered is clear – as noted above, interviewees mentioned the greater power BRDG staff had than their experiences in other USAID projects. One of the ways this power was reflected was in staff’s ability to recommend specific activities for support and play a role in deciding what projects would receive support. This allowed BRDG to take advantage of soft information from staff that “Know exactly what is going on and what will be helpful to society.”482 Soft information played a critical role in BRDG’s relative Navigation by Judgment and consequent success, one enabled by OTI’s relative flexibility and autonomy.483 It is interesting that OTI was able to achieve flexibility and use soft information, but was not able to engage in activities which were opaque – which were not completely legible, comprehensible, to those outside the context.

482 Interview 47, 6/4/13; quote repeated from above.

483 Nonetheless BRDG proved less successful than DFID’s CISCAB, which also demonstrated the benefits of soft information in navigating a complex political space, albeit incompletely.
Staff Satisfaction & Agent Quality Revisited

OTI’s greater autonomy and empowerment of Navigation by Judgment also allows an examination of whether, consistent with Chapter 2’s theory, OTI staff are more motivated and/or satisfied than their USAID peers – whether autonomy has workplace effects. The scant empirical data suggests this is the case; OTI scores higher on morale and satisfaction in annual Federal Employee Viewpoint surveys than the rest of USAID by significant margins.\textsuperscript{484} A 2009 Congressional Research Service report notes the unusual (for USAID) “high staff morale at OTI” (Lawson 2009, p. 9).

BRDG provides data consistent with these broader dynamics; to say job satisfaction amongst Liberian BRDG employees was high is an understatement. While by many accounts there was quite a bit of work at BRDG,\textsuperscript{485} many staff, when asked to compare BRDG to other aid project on which they have worked, remarked how uniquely satisfying the experience of working with BRDG was. One team member described it as “one in a million”, the kind of coworkers and working environment that comes “once in a lifetime.”\textsuperscript{486} More than one individual described it as the project they were proudest to have been involved in in their careers.\textsuperscript{487} At the project’s close the staff prepared a “yearbook” – resembling an American high school yearbook – of project activities and staff; multiple staff

\textsuperscript{484} Interview 56, 6/4/13

\textsuperscript{485} Interview 44, 6/6/13; interview 45, 6/10/13

\textsuperscript{486} Interview 45, 6/10/13

\textsuperscript{487} Interview 44, 6/6/13; interview 48, 6/12/13
members still had this yearbook easily accessible more than five years after the close of the project.

This great work environment had, in the opinion of some BRDG employees, much to do with the ability of staff to Navigate by Judgment. Multiple interviewees felt BRDG provided the ability to contribute – to feel one’s work was consequential and could impact program direction. As one team member put it, there was “room for thinking on your own, for taking decisions.” Speaking directly to whether this satisfaction and ability to Navigate by Judgment crowded in higher quality staff, multiple interviewees noted the exceptional quality of their colleagues as compared to other projects on which they have worked.

5.5: Liberia Conclusion – Similar Objectives, Different Methods & Results

DFID’s Navigation by Judgment proved the superior strategy both in the relatively opaque domain of capacity building and the relatively observable domain of health delivery. This proved the superior strategy in part because of the better ability of DFID to

488 Interview 45, 6/10/13; interview 44, 6/6/13

489 Interview 41, 6/10/13; interview 48, 6/12/13; interview 44, 6/6/13; interview 45, 6/10/13; interview 64, 6/19/13

490 Interview 41, 6/10/13

491 Interview 45, 6/10/13; interview 44, 6/6/13; interview 48, 6/12/13
take advantage of greater autonomy to make design, implementation, and revision decisions that Navigated by Judgment.

The in-country authority to make substantively significant implementation decisions was a critical part of success in these projects; that this was the case seems related to the difficulty of understanding Liberia remotely. As Raymond Muhula, a World Bank Public Sector Reform specialist who was not in Liberia at the time of most of these projects, put it regarding his transition from headquarters to Monrovia “Engaging in a fragile state with the ambitions of HQ staff is a driver of failure; when you sit in HQ and when you sit here you see things vastly differently”. Relatedly, DFID felt the loss of its in-country office; as one DFID actor put it:

There is a big difference [to being in-country]... any modifications or changes to implementing projects, it's easier to do so, and to understand why you're doing so, and to respond to the client requests in a more responsive manner if you're sitting in a country office and you're involved in the day-to-day context of what's going on. If you're remote, or if you're sitting in HQ, or you don't really have that personal relationship or feel for the day-to-day implementation; or the difference between being a governance specialist, for example, or an administrator, you see it from a technical point of view as opposed to ‘oh well, I can't do it from an administrative point of view.”

These statements speak to the importance of soft information given the opacity (rather than the frequently changing environments) of fragile states.

This is not to say that organizational autonomy and navigation are monolithic; and indeed these cases show changes both over time and across units. Over time, FARA demonstrates significantly more autonomy than its predecessors; across units, the relative

---

492 Interview 21, 6/4/13
autonomy of OTI is quite strongly linked to the flexibility of BRDG relative to other USAID projects. While OTI is the exception for USAID at the organizational level, it is in some sense the exception that proves the rule; the OTI project examined here was still influenced in design by the need to focus on things that could be quickly observed and reported upon.

One thing that comes to the fore in these cases is that a similar mechanism – a multi-year implementation contract – does not seem to have the same result in imposing inflexibility for DFID as it does for USAID. One long-serving USAID staff member summarized the effects of such contracts on USAID in saying “Rigidities of the [contracting] mechanism affect the way we operate, the way we design, etc.” This makes sense – knowing you will contract will change both the scope of what is designed and how it is measured, articulated. One could imagine contracting as a modality has the natural result of focusing agencies towards less complicated, and ex-ante predictable, tasks. DFID, however, seems not to have the same issues though they have similar structures – that is, similar mechanisms did not seem to imply similar rigidities. One USAID senior official stated “the real world requires a lot of flexibility; the world of procurement requires a lot of certainty.” By contrast, one veteran of numerous development agencies, said “[With DFID] nothing is set in stone. DFID – of the development partners I’ve worked for –is the most flexible to respond to changes, tweaks.” The “world of procurement” seems to be

493 Interview 7, 6/12/13
494 Interview 7, 5/23/13
495 Interview 21, 6/4/13. While this individual had collaborated with USAID on many occasions, he had never worked for USAID, and as such this quotation on its most natural reading ought not be read as directly comparing the two organizations.
very different depending on which IDO you examine. This suggests that contracting form may matter less, or at least be mediated by, informal understandings and working procedures.

Another striking difference between USAID and DFID is as regards continued activities after the end of projects. One senior Liberian government official, speaking about the lack of continuation of activities begun during USAID’s AWARE, said:

In most of these projects sustainability is talked about, it’s flagged as a priority, but it’s really not. It’s one of those things that makes sense to talk it, but the reality’s quite different. People know that it’s [the project] not sustainable; the institutional arrangements are not worked out, especially the funding dries up, and everyone goes back to square zero.496

This insight applies to many of the USAID projects here, but less so DFID’s interventions. CISCAB’s end may have led to a loss of momentum, but by no means a return to “square zero”. The OFM project and Pooled Fund have continued to progress after the end of the project. If sustainability is the maintenance of the level reached during the project, in DFID’s Liberia interventions have clearly surpassed the sustainability standard, with the project acting as a catalyst, a spark to launch continued future progress.

USAID’s relative Navigation from the Top does not mean there is no Navigation by Judgment present in its programs, and these cases provide evidence of same. The reverse is also true; DFID is also not a fully autonomous agency with no Navigation from the Top. DFID’s move into Liberia was prompted by optics and (implicitly) legitimacy concerns,

496 Interview 10, 5/29/13
DFID’s departure in 2011 related seemingly to changing political winds in favor of a greater focus on Sierra Leone. By summer 2013 DFID had reversed course once again, placing an individual back in Monrovia as a permanent DFID representative. Some saw this as linked to President Johnson Sirleaf’s appointment by UN Secretary General Ban Ki-moon alongside UK Prime Minister David Cameron as co-chairs497 of a panel to develop the post-2015 Millennium Development Goals in May 2012.498 Optics may have prompted reengagement once again.

These cases call into question some of this work’s assumptions about environmental uncertainty and country context. Liberia is very much an aid dependent state; this certainly changed the bargaining relationship between IDO and recipient, as did Liberia’s relative lack of development (as compared to South Africa). However the Liberian government was still able to ‘push back’ in places (e.g. pushing for the pooled fund and influencing the transition of RBHF to FARA). In both South Africa and Liberia, then, there have been some more confident and capable government units (NDOH in South Africa, MoHSW in Liberia) and some less so (NACP in Liberia, DPLG/COGTA in South Africa).

497 Along with Indonesian President Susilo Bambang Yudhoyono.

498 The implication being that this prompted PM Cameron to want to seem responsive to President Sirleaf and Liberia. Data seems supportive of this argument; in February 2012 an independent analysis of the UK Aid Budget noted that Liberia was estimated to receive a total of 10 million GBP from 2010-2014, though it noted that the figure had not yet been revised following Liberia’s November 2011 presidential election; in May 2012 an updated operational plan for Liberia gave a 2011-2014 estimate of 24 million GBP. By June 2013, little had changed save the MDG appointment and DFID’s apparent subsequent re-focus on Liberia; DFID’s updated operational plan now suggested 29 million GBP in spending from 2011-2014, or a 20% increase. (Fitzsimons, Rogger, and Stoye 2012; DFID 2012, 2013)
There is an argument for deconstructing recipient country in thinking about the strength of bargaining and the implications of same.

These cases also make clear the way political authorizing environments constrains IDO autonomy, influencing the available envelope of navigation strategies. These political constraints then necessarily limit the scope of possible action, the ‘envelope’ in which field-level staff can maneuver. It may not be the case that field-level staff always wishes to reach a point outside their envelope of maneuverability; put another way, the constraint is not always binding on behavior. But sometimes the teeth do bite, and avoiding them leads to agreements other than what would have been agreed in their absence.

5.6: Comparing Across Countries: Bringing Liberia and South Africa Cases Together

Both this chapter and its predecessor have grouped projects by country, then by task domain. While this allows for the most direct comparison of USAID and DFID performance and navigation by country and task, it makes it more difficult to see comparisons that cut across country. As such this section briefly explores comparisons broken down by task domain (Capacity Building and Health) and IDO (USAID and DFID) to gain additional analytic leverage on these data.
Per Chapter 2, capacity building is the domain in which Navigation by Judgment should prove most helpful, Navigation by Measurement the least effective. There is evidence to support this; while DFID’s more Navigation by Judgment-prone project proved the superior performer to USAID’s in each country, LGSP – USAID’s South Africa effort – was most effective precisely where it relied most on judgment (with the placing of advisors). BRDG was also helped by the relative autonomy of OTI and the flexibility and Navigation by Judgment it enabled. While projects were hampered by Navigation from the Top, there were also costs to Navigation by Judgment, places (e.g. CMTP’s capture by Deloitte, CISCAB’s losing direction following the launch of the Civil Service Reform Strategy) where the empowered judgment makers made decisions that seemed to lessen the success of the project from either error (CISCAB) or due to a mis-alignment of the interests of contracted agents and supervisors (Deloitte in CMTP).
The theory would also predict that the returns to soft information should be higher in Liberia than in South Africa. It is not immediately clear from these cases that this is true; in the projects involved here task has to a large extent trumped broader national environment. Indeed, it might be fairer to say that the relevant ‘environmental uncertainty’ proved to be task specific – South African municipal environments proved to be opaque, complex, unpredictable, and difficult to monitor and steer from the outside. As such DFID’s greater focus on, and effective project leadership by, hands on advisors proved more effective than USAID’s externally monitorable delivery.

Both USAID’s capacity building efforts focused on externally monitorable delivery – either trainings (LGSP) or supplies and short-term consultancies with policy outputs (BRDG). In some sense BRDG’s arguably superior performance to LGSP is a credit to OTI & DAI’s better ability to incorporate soft information; it also surely is because BRDG, unlike LGSP, did not attempt to tackle politically difficult problems, instead engaging in more superficial activities where these were likely to be welcome.

None of these projects were an unmitigated success; in each of these projects there were elements that didn’t work. In some ways these cases compare the guaranteed mediocrity of Navigation from the Top with the sometimes-success of Navigation by Judgment. Capacity building may just be a difficult thing to get right, one that ex-ante requires risk. Balanced against this is the incredible gains to be had from success; none of the capacity building projects was very expensive (certainly as compared to the health
interventions), and when impactful these projects clearly had substantial long term impacts.

**Health**

Table 5.4: Health Projects Examined

<table>
<thead>
<tr>
<th></th>
<th><strong>USAID</strong></th>
<th><strong>DFID</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rebuilding Basic Health Services <em>(RBHS)</em>, 2008-present: Implemented by JSI</td>
<td>Implemented by Price Waterhouse Coopers</td>
</tr>
<tr>
<td></td>
<td>Fixed Asset Reimbursement Agreement <em>(FARA)</em>, 2011-present</td>
<td>Health Sector Pooled Fund <em>(PF)</em>, 2008-present</td>
</tr>
<tr>
<td><strong>South Africa</strong></td>
<td>President’s Emergency Plan For AIDS Relief <em>(PEPFAR)</em> as administered by USAID &amp; CDC, 2004-present; various implementers</td>
<td>HIV/AIDS Multisectoral Programme <em>(MSP)</em>, 2003-2008; implemented initially by Futures Group, succeeded by HLSP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rapid Response Health Fund <em>(RRHF)</em>, 2008-2009; implemented by HLSP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Strengthening South Africa’s Response to HIV and Health <em>(SARRAH)</em>, 2010-2015; implemented by HLSP</td>
</tr>
</tbody>
</table>

The health sector is where this work’s theory predicted Navigation by Measurement was most likely to succeed – particularly when Navigation by Measurement interacted with a relatively certain environment, as was predicted to be the case in South Africa. The findings on success are just what theory predicted – Navigation by Measurement fared better in health than in capacity building, broadly speaking; it also fared better in South Africa than Liberia, and was in fact associated with the more successful intervention in the South African health sector.
This simple “correlational” confirmation of the theory is in some ways obfuscatory. First, it ignores the differences – most notably in terms of objectives - between the South African and Liberian Navigation from the Top attempts. RBHS’ success in running a parallel health system – arguably the most inherently externally observable and contractible part of the Liberian health effort - gets less credit in this chapter than PEPFAR’s did in the previous chapter, simply because RBHS was explicitly aimed at building national systems (thus making an effective parallel system a fact partially demonstrative of impact, partially warranting of criticism). It is possible that Navigation by Measurement could, and would, work for service delivery even in Liberia; just that it is an inferior strategy to follow in building government systems, be it via FARA, PEPFAR, or in the context of a USAID project. Whether FARA’s performance based contracts with NGOs and (now) with government county health teams succeeds will provide useful new data in this regard.499

It does seem that USAID uses Navigation by Measurement both in health tasks where this work’s theory suggests it is likely appropriate to do so (PMTCT delivery in South Africa; possibly performance based contracts for implementers under contract to the national government in Liberia) and in those where the nature of indicators chosen and project goals make success less likely (FARA’s process-based indicators, which are both

499 It is possible, then, that Navigation by Measurement may be superior in delivery of simple, externally monitorable interventions with clear outcomes (e.g. PMTCT) irrespective of environmental uncertainty; possibly useful where monitoring is tricky but possible (e.g. general health systems), with environmental uncertainty critical in the analysis; and almost certainly inferior when the task is less tractable to measurement (e.g. reforming a health financing office). More work is needed to investigate this hypothesis.
subject to Goodhart’s Law problems and need frequent revision due to Liberia’s uncertainty).

DFID’s Navigation by Judgment in health worked in each of these cases in close collaboration with national governments. When those Governments were in a mode of newly implementing reforms with earnestness and zeal (in Liberia and during Hogan’s regime in South Africa) they saw impressive success, accomplishing great things with modest expenditures. Where DFID encountered more resistant or apathetic governments, their health efforts floundered. To what extent this is due to Navigation by Judgment itself and to what extent a function of DFID’s propensity to work through Government is difficult to disentangle; it certainly seems DFID’s results were more contingent on counterparties than USAID’s.
In some of these cases DFID may well have benefited from less Navigation by Judgment; in others more may have been appropriate. The near-universality of the strategy is itself notable, and somewhat contrary to expectations. The theory would suggest that DFID had the ability, but not the need, to Navigate by Judgment – that greater autonomy is sufficient but not necessary to pursue a more soft information-utilizing navigation strategy. This would also imply that we should see DFID as having substantial internal variation in navigation strategy, in effect using judgment to make decisions regarding the use of Navigation by Judgment as the nature of the case warrants.
It’s certainly the case that – looking at the DFID cases as a whole – we see substantial variation in whose judgment leads the *de facto* navigation of interventions. Starting from the field and moving ‘up’:

- **CMTP**: The ISFs - individual field advisors - directs DFID’s involvement where the project most succeeds.
- **CISCAB**: Contractor’s field team with some input from DFID staff.
- **OFM and MSP**: DFID staff’s judgment is front and center, often trumping that of contractors.
- **RRHF**: DFID and Government both make critical judgments, to the benefit of the project.
- **SARRAH**: Government interlocutors may often direct the project to its detriment.
- **Pooled Fund**: After DFID-directed set up, the PF steering committee (composed of Government and many donors) is now making judgments in implementation (including about DFID’s contributions).

While we see remarkable variation in who the judge is, we see less variation in the decision to judge. This confirms part of the theory; DFID seems to be engaging in nuanced, if not always fully conscious, variation about whose judgment counts in ways plausibly driven by the nature and goals of the project. The two cases where the level of judgment is closest to the field – CMTP and CISCAB – are in the task domain where returns to soft information are predicted to be highest, and thus judgment being conducted as close to the ‘action’ as possible most appropriate.
In these cases DFID never fails to Navigate by Judgment to a substantial degree. This may be because all these projects were perceived to benefit from judgment; it also may be the case that there is a degree of bureaucratic capture reminiscent of Barnett & Finnemore. Empowered bureaucrats may believe Navigation by Judgment the right solution even where an honest ex-ante reading of the environment (the MSP during denialism seeming the best case in point) would suggest otherwise.

An unexamined faith in judgment and/or in Government collaboration may not be optimal, a failing borne of organizational autonomy rather than political authorizing constraint. It also may be the case that the greater use of contractors – driven on the views of some by political changes in the UK and/or the need to spend a rapidly expanding budget – means roles DFID staff could once successfully Navigate by Judgment are now in the hands of contractors, leading to a mis-alignment of interests. Robin Gorna, a former DFID Health advisor, essentially argued just this, saying

I think health advisors are just becoming more kind of overseeing contracts with others who are delivering the fun work, for me this is less of an issue that others in that I don’t have a long, long DFID career, but I certainly know from my colleagues that the often feel some frustration that historically they were able to really get in there and spot the good projects and get the money behind them and make things move forward, whereas now the project design is much, you’re designing much bigger, longer, more expensive programmes... it often doesn’t play to the strengths of the health advisor who often has a great deal of personal technical skill that they want to bring to the conversation. So even though there remains this distinction between the advisor and the programme manager, advisors are getting much more pulled into the contracting process and the programme management, and doing less of the getting involved with the technical skills which play to their strengths. And I think this is just the inevitability of DFID spending more and more money, means that it becomes more of an accountability procuring role than a technical role. And certainly I've heard from a number
of colleagues that they’ve found that frustrating and there has been quite an exodus of quality DFID advisors in the last few years.

This suggests that optimal navigation strategy interacts with not just the technical design of the project but also the organizational plan for implementation. It also suggests that getting this mix wrong may have agent quality and motivation implications similar to those USAID may experience due to its low level of operating autonomy.

**USAID**

Table 5.6: USAID Projects Examined

<table>
<thead>
<tr>
<th>Capacity Building</th>
<th>Health</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Liberia</strong></td>
<td></td>
</tr>
<tr>
<td>Rebuilding Basic Health Services (RBHS), 2008-present: Implemented by JSI</td>
<td>Fixed Asset Reimbursement Agreement (FARA), 2011-present</td>
</tr>
<tr>
<td><strong>South Africa</strong></td>
<td></td>
</tr>
<tr>
<td>Local Government Support Programme (LGSP) phase 2, 2004-2010; implemented by Research Triangle Institute (RTI)</td>
<td>President’s Emergency Plan For AIDS Relief (PEPFAR) as administered by USAID &amp; CDC, 2004-present; various implementers</td>
</tr>
</tbody>
</table>

Turning, finally, to USAID: USAID is remarkably consistent across these contexts, certainly when OTI’s project (BRDG) is removed from consideration. USAID proper sees much less variation in navigation strategy across the countries, task domains, and projects examined here than does DFID.\(^5\)\(^0\) This persistent Navigation from the Top – including a heavy does of Navigation by Measurement – is as we might expect to see from a

\(^5\)\(^0\) With the notable exception of OTI, chosen so as to investigate the potential for variation.
constrained agency worried about legitimacy in the face of a rather insecure political authorizing environment.

USAID extends Navigation by Measurement to areas where it is unlikely to work well – LGSP, in which USAID’s own prior project specifically warns of the difficulties to be had from pursuing an external training strategy is the clearest case where USAID ought, and arguably did, know it was pursuing a strategy unlikely to succeed. USAID’s use of Navigation by Measurement in PMTCT training (in AWARE) was much less effective than when applying measures towards PMTCT drug delivery (in PEPFAR). FARA strongly suggests that USAID’s Navigation by Measurement is effectively imposed from above rather than chosen from below. That CDC approaches PEPFAR differently from USAID suggests that this imposition is not merely a function of the political authorizer (i.e. Congress), but of the dyadic relationship between that authorizer and the organization. This does not mean USAID always gets it wrong; USAID’s navigation proved superior in the South African health sector, and might well prove superior in other tasks and environments tractable to Navigation by Measurement – particularly if DFID were to attempt to Navigate by Judgment in those domains.

USAID’s constraints are also not absolute; to say they have diminished freedom of action relative to DFID is not to suggest that there is no such envelope of discretion. OTI’s existence, with its more flexible and Navigation by Judgment-operating procedures, is testament to that. USAID’s decision to use OTI precisely in the intervention where my theory suggests soft information and Navigation by Judgment would see the highest
returns – a capacity building project in an uncertain environment – suggests that USAID may be deploying OTI with a view to trying to use appropriate navigation where possible within the available envelope.

Comparing Across Countries in Sum

This section has allowed for the integration of Liberia and South African cases in an effort to get a better sense of task domain and intra-organizational variation. USAID’s relative consistency of navigation strategy is as anticipated by theory; DFID’s relative consistency of navigation strategy is more intriguing. Also intriguing is the way the interaction of task domain and national-level environmental uncertainty appears to be the beginning, not the end, of the relevant set of facts regarding optimal navigation. This augurs for a need to consider something like the effective micro-environment of an intervention in the context of its goals and the ‘local’ uncertainty and monitorability (e.g. South African municipal governments having little connection to South Africa’s general environment; RBHS’ focus on systems vs. PEPFAR’s focus on delivery) in choosing the right navigation strategy for the job.
CHAPTER 6: CONCLUSION

Words are but the vague shadows of the volumes we mean.
― Theodore Dreiser (1927, p. 6)

In 1856 Henry Bessemer wanted to monetize his new process for producing steel; he attempted to do so by selling licenses to use his patented method, and found a number of willing buyers. The buyers were none too pleased by the results; following the patent to the letter, they produced merely useless lumps. Bessemer changed course, instead opening up his own factory and simply producing Bessemer steel himself.\footnote{This account is from Gordon 1991} Bessemer knew more than he could write down; his steelmaking process required tacit knowledge. It also required a skilled operator, one who could observe the process of production and make adjustments appropriately – that is, someone who could Navigate by Judgment based on soft information.

Bessemer wasn't exceptional – far from it. Ask a skilled machine operator or craftsman how they perform their trade; the answer will often sound quite a bit like Navigation by Judgment. We all Navigate by Judgment when we get the timing right on a joke in a social gathering, decide who's trustworthy, or recognize our child’s cry in a crowded nursery.

Judgments can be wrong, and often are. Navigation from the Top, and Navigation by Measurement, reduce the impact of errors in judgment by reducing the discretion of those...
who might reach erroneous conclusions. Less discretion also reduces the possibility of the ‘chickens running the henhouse’, bureaucratic capture from below by autonomous actors. This work has argued that in contrast to the classic principal-agent model this ‘cure’ is often worse than the disease; in many circumstances this solution has negative net effects as compared with the counterfactual.

6.1: The Argument Restated

Empowering field level agents and allowing them to make relatively unconstrained judgments that lead the design, implementation, and revision of activities is the superior organizational navigation strategy in many international development interventions. This is because Navigation by Judgment allows the incorporation of soft information, observable non-verifiable features of the environment. It also allows agents to make use of tacit knowledge, things agents know how to do but cannot codify. The returns to Navigation by Judgment, then, are increasing the less externally monitorable the task and the more opaque and uncertain the environment. Political authorizing environments prevent IDOs from choosing optimal navigation strategies. As such, more autonomous organizations engage in Navigation by Judgment more readily; they consequently perform better in more unstable environments and in tasks that are harder to monitor externally. The inverse is also true; the more certain the environment and observable the task, the more Navigation from the Top – and, in the case of projects oriented towards targets, Navigation by Measurement – is the appropriate organizational strategy. Less autonomous IDOs will also
see organizational consequences via the exit of quality staff and reduced organizational learning.

6.2: The Evidence

Chapter 3 examined the performance of nine IDOs of varying levels of autonomy across task and environment using the world’s largest aid project performance dataset. The nature of the data limits the analysis to comparative performance within organizations; however, we saw that more autonomous IDOs see much less of a decline in performance in more fragile environments than do less autonomous IDOs. This effect is clustered in tasks that are more difficult to externally monitor and does not seem to be driven by differential evaluator bias across IDOs or the selection of IDOs into countries or projects. This effect also holds when looking within recipients – when only considering how performance changes as recipients become more, or less, fragile over time. This result also holds when only the political authorizing environment instability sub-scale of the autonomy measure is employed, suggesting that these performance differences relate to political authorizing environments.

This quantitative analysis provides prima facie evidence but leaves many open questions, including many core to the argument. Are more autonomous IDOs actually performing better in harder to monitor tasks and more uncertain environments, or only more consistently? Is the better performance of more autonomous IDOs due to different
navigation strategies? Chapters 4 and 5 employed case studies in one relatively stable and certain environment (South Africa) and one much less so (Liberia). In each country a relatively politically insecure and constrained IDO (the US’s USAID) was compared with a much more autonomous organization (the UK’s DFID). The USAID-DFID comparison in each country involved four cases, one in a relatively externally observable and contractible task domain (health), the other in a relatively difficult to monitor task domain (capacity building). These cases investigated the core argument of the work, looking at comparative performance as well as whether USAID and DFID responded to similar problems with different navigation strategies. They also allow a more nuanced examination of the relationship between political authorizing environments, organizational autonomy, agent autonomy, and project performance.

Chapters 4 and 5 provide strong support for the argument; DFID’s more autonomous agents often Navigate by Judgment or empower others to do so. USAID focuses on measurement across the board, biasing projects towards the observable. This Navigation by Measurement has deleterious effects, seen in this work particularly in the capacity building projects where process indicators employed had little to do with actual human capital improvements. Failing to Navigate by Judgment constrains USAID’s ability (relative to DFID’s) to make use of soft information in planning, implementing, and revising projects. The greater constraints faced by USAID at both the organizational and field agent levels stem clearly from political authorizing environments in a number of instances (e.g. FARA), as does the focus on meeting targets for Congressional reporting (e.g. LGSP). USAID’s lower levels of autonomy seem to be associated with lower levels of workplace
satisfaction and greater propensity to exit – indeed, the exception in autonomy terms within USAID (OTI) is also the exception in workplace satisfaction. Navigation by Judgment proves the superior strategy in three of the four case ‘pairs’ – the exception, South Africa’s health sector, is the case pair where the returns to Navigation by Judgment were hypothesized to be most modest and Navigation from the Top most likely to succeed.

Evidence in chapters 3, 4, and 5 all points to the legibility, the opacity, of recipient country environments as the critical environmental factor. In chapter 3, there was suggestive evidence that rapid changes (e.g. disasters) did not differentially effect performance, but measures of the difficulty of understanding environments from a distance did. Chapter 4 provided an (unanticipated) example of an opaque environment – while South Africa may be relatively stable and monitorable, South African municipal governance is an environment that is not terribly legible. In this environment Navigation by Judgment proved the superior strategy. The Liberia of Chapter 5 was changing rapidly; nonetheless, the superiority of Navigation by Judgment as employed by DFID seemed linked to the selection and successful implementation and revision of opaque tasks (CISCAB and the OFM support).

The case studies suggest that task and environment together create a level of legibility or monitorability that is not as simple as the sector and country. In the Liberian health sector reforming the OFM was unmonitorable but delivering PMTCT was not. Recipient country “environment” may be task specific, with a need to consider the legibility
and measurability of particular tasks in particular sectors in particular countries in determining the appropriate navigation strategy.

The evidence in chapters 4 and 5 also adds nuance to the complex relationship between measurement and navigation strategy articulated in chapter 2. Within USAID projects, measures are associated with significant autonomy over means to achieve those ends, coupled with very little flexibility in revising targets or pursuing other useful elements of intervention. This underscores that Navigation by Measurement may well be the best strategy when one or a group of measures act as effective summary statistics for an intervention (e.g. PMTCT treatment), environments are certain, and activities are observable.

We also see evidence that Navigation by Judgment has costs. Agents and contractors can point projects in the wrong direction, particularly when – as the classic principal-agent model would predict – agents have goals (e.g. profit) that are not in the interests of principals. In the projects examined here, these costs were broadly smaller than the benefits to be reaped from Navigation by Judgment. This suggests, however, that there is likely to be such a thing as too much Navigation by Judgment, even in the tasks where it is most likely to be the better strategy.\footnote{As noted earlier, DFID’s use of Navigation by Judgment across the board – rather than increasing Navigation by Judgment markedly in the most difficult to observe projects – suggests that there may be some over-use of the technique as well, a Barnett & Finnemore-esque capture of the organization by its agents.} It also speaks to the complexities around the increasing use of contractors, who may have different incentives than agency
staff. Contractors may also have different abilities than IDO staff, as highlighted by Robin Gorna’s remarks on talent spotting and the difficulties when the organization tries to “contract out the intelligence in the same way”.503

Organizational learning was touched upon in Chapter 2; both the quantitative and qualitative data provide hints of this organizational learning channel. USAID’s lack of incorporation of lessons learned – and articulated – in LGSP is striking, and not the only piece of suggestive evidence that the more constrained IDO is less able to learn from experience. Multiple Liberian officials commented on their surprise that USAID seemed initially unable to cope with the complexity of post-conflict Liberia; that USAID seemed to have no institutional memory to draw on from other contexts, instead muddling through and getting better over time.504 It is striking that USAID – an agency so focused on targets and Navigation by Measurement - doesn’t assess the holistic success of projects ex-post, and thus is not part of the dataset examined in Chapter 3.

Examining the universe of IDOs, USAID is typical in this regard - the organizations that collect, and were willing to share, results data tend to be those with more autonomy.505 USAID personnel have noted that, due to the potential for it to be asked for/seen by Congress, USAID very rarely codifies its failures lest they be called to task for having

503 The full remarks are in Chapter 4.

504 This point was made most directly in interview 69, 6/3/13

505 As assessed by the scale employed in chapter 3; see Table 3.17 for details. I have not separately systematically evaluated which IDOs fail to collect ex-post data rather than collect such data but were unwilling to make such data public.
wasted taxpayer dollars. Relatively, USAID interviewees were much more likely to ask to be made anonymous than their DFID counterparts; this may well be because they believed USAID would be less tolerant of dissent, of criticism. There is suggestive evidence here that the more insecure agency (USAID) engaged in less codification of failure even as it focused on measures with which to justify itself to political authorizers. An organization that cannot formally admit failure will have difficulty learning from same; more work is warranted to further explore this channel. By contrast, in addition to allowing for codification of failure and thus learning, Navigation by Judgment allows for heterogeneity of implementations; it lets more flowers bloom, and thus has a richer variety of practice from which the organization can learn.

6.3: Alternative Explanations of IDO Success

The narrative in this work – particularly in the case study chapters – has touched on elements of project success and failure beyond those articulated in the core argument. This section will briefly address how other explanations of international development success frequently mooted by practitioners and those who comment on foreign aid relate to the data presented here.

506 Outside the context of these interviews and this academic work.
Country Ownership & Political Will

Many argue that the more a particular project is owned by the recipient – the greater the extent that it is adopted at the request of, and is led by, senior actors in the developing country – the more likely it is to be successful.

If indeed country ownership contributes significantly to performance, the question remains how one gets country ownership – how one ensures that projects are “owned”. What is often framed as a feature of recipient countries is, I would argue, fundamentally about the interaction between donor and recipient. The OFM’s being ‘shopped’ to USAID and only subsequently to DFID is suggestive of realized country ownership across development organizations being at least partially endogenous to donor organizational features, with USAID less able (or willing, or interested) in engaging in this fully owned project.

DFID, the more autonomous organization, allowed Liberia’s MoHSW to guide strategic direction more than did USAID. It makes sense that more autonomous organizations and personnel, with more independent scope of action, would be less constrained by plans developed remotely and thus more able to seize opportunities when they arise, to truly develop plans in concert (or in response to) recipient country actors rather than impose plans when they arise. As one MoHSW senior official put it, “USAID consults on strategy, but it’s their strategy”.507

---

507 Interview 17, 5/28/13
Political will – roughly, the desire of senior political actors for reforms to occur and the amount of capital (human, political, reputational, financial) recipient country political leaders are willing to put forward in an effort to implement – is often put forward as a cause of intervention success or failure. What’s true of country ownership is also true of political will – the same actors can have different realized political will towards different projects. In Liberia MoHSW seemed to have more will to fix the OFM than to improve the functioning of the NACP; by responding to Government’s priorities DFID’s project thus had more political will. To a lesser extent, DFID’s CMTP seemed more important to South Africa’s DPLG than did USAID’s LGSP.

Political will can also push in different directions. CISCAB was very important to Liberia’s CSA and President, so much so that they were able to expend substantial political capital in an effort (albeit an unsuccessful one) to continue the project. Pushing against CISCAB were Liberian senior politicians who did not want to lose appointment powers. In South Africa political will pushed strongly against HIV/AIDS treatment; USAID was able to succeed by working outside the National Department of Health, while DFID was not. That said, DFID’s network inside Government put them in a better position to exploit opportunities when political winds, and will, changed in South Africa.

---

508 This also speaks to the unhelpfulness of talking about political will at a national level. Political will is an abstraction that may suffer from meaninglessness borne from failure to decompose whose will, and for what, is being discussed.
Country ownership and political will are often an important part of success; these factors seem to predict well what interventions will continue to be championed after funding ebbs or donor political winds change, and thus is a harbinger of continuity if not necessarily quality. Political will and country ownership are not randomly distributed among interventions. They depend critically on the realized set of interventions in a given country and thus on features of IDOs that influence the realized intervention set. IDO autonomy and Navigation by Judgment is likely to increase the proportion of projects that are country owned and regarding which recipient country political leaders demonstrate political will.

**Corruption**

While there were instances of small-scale malfeasance (e.g. in BRDG) or questionable actions by contractors (e.g. Deloitte in CMTP) in the cases examined here, none of these projects experienced massive corruption and fraud. It is possible that the lessons drawn from these cases (though less so the quantitative empirics) would not be robust to significantly corrupt environments, particularly in the form of recipient country ministries. However, this begs questions of project selectivity, country ownership, and political will explored above. One would think in a corrupt environment an autonomous agency engaging in Navigation by Judgment would be best able to identify and support activities which were likely to include political will for positive reform, or were it to be the case that there were no such projects either decide to fund nothing or choose the least-bad options.
Being In-country

While autonomy is relatively novel in the development discussion, the notion of formal devolution – of placing offices in developing countries so as to be closer to clients/recipients – is not. The logic for locating decision-making staff in the country parallels in some ways the theory on which the autonomy hypothesis rests; by being ‘closer to the action’, the organization will perform better.

Interviewees – particularly in Chapter 5 – highlighted the returns to ‘being there’, the way it changes views and deepens understanding. But while the most successful examples of Navigation by Judgment involve an empowered, experienced in-country decision maker, there is evidence in this work that being in country is neither necessary nor sufficient.

On sufficiency, Chapter 3’s quantitative analysis suggests that having an in-country office is useful, but not equally so; offices only positively impact project success for more autonomous IDOs. Chapters 4 and 5 also provide a number of examples where the presence of in-country staff did not lead to the benefits of presence being reaped by a project – LGSP in South Africa is a particularly strong illustration of this. This makes sense; if the organization cannot make use of soft information, being in a position to collect it has minimal impact for the production process.
By the same token, DFID’s CISCAB in Liberia Navigated by Judgment despite the absence of a substantial DFID supervisory presence in country; while there was a DFID country representative present, supervision of the project largely involved actors based in Sierra Leone and London with frequent visits to Liberia. In this case DFID provided the space for the (in-country) contractor to Navigate by Judgment. This suggests that a substantial in-country presence by an IDO is not strictly necessary for Navigation by Judgment. That DFID had a lighter ‘footprint’ in-country in both South Africa and Liberia than USAID suggests that at the very least a larger office does not seem associated with better use of soft information across these agencies.\(^{509}\)

Without some in-country presence or delegation to in country contractors, it is very difficult to gather soft information. Actual presence in the county does not guarantee that soft information will be used, however; even \textit{de jure} formal devolution of authority is the beginning, not the end, of the story. It is critical to soft information incorporation that actors have the \textit{de facto} space to Navigate by Judgment.

**Alternative Explanations in Sum**

Many things contribute to the success and failure of development interventions. Increased Navigation by Judgment has the potential to move many of these ‘needles’ - e.g. realized political will and country ownership, realized corruption – in the right direction,

\(^{509}\) It might still be the case that office size matters within agency – that a larger office might have returns for more autonomous IDOs. Future research would be needed to investigate this.
making success more likely. There is a compelling argument for focusing on organizational features of IDOs beyond the impact any of these features may have on corruption or political will: an IDO’s organizational features differ from these items, in that they are wholly controlled by those providing the funds. Organizational design is the “low-hanging fruit” of international development, the factor in development outcomes arguably most changeable by Western governments and polities.

6.4: Navigation by Judgment and the Literature

Bureaucratic Politics

Carpenter argues that “bureaucratic autonomy is politically forged.” (Carpenter 2001, p. 353) IDOs are no exception. IDOs certainly exist in a political space, one in which political authorizing environments are critical. The notion that political authorizing environments relate to bureaucratic function is not a new one, or one unique to the world of foreign aid. McCubbins, Noll, & Weingast (1989) are foundational to a small forest of publications focusing on how the enacting coalition of bureaucracies affects the administrative procedures of organizations. Yet this – and, indeed, the vast majority of political science and economics work on management in public agencies – starts from a principal-agent model, in which political principals are actively involved in managing organizations and agents in service of the principal’s goals. This work takes a different view; the principal may get more from agents by controlling them less. In so doing it builds on Mansbridge (2009) and Bohnet, Frey, & Huck (2001) in suggesting that attempts to
control may sometimes be self-defeating; following Mansbridge, it argues that the paucity of effective *media res* controls ought focus great attention on agent quality.

One might worry that the failure of less autonomous IDOs is a feature, not a bug, for political authorizers. On this view perhaps it is the case that, per the structural politics school (e.g. Moe 1990), the ineffectiveness of IDOs is precisely what political authorizers intend – that e.g. USAID has been designed to fail because this is in the interests of political authorizers. It is very difficult to prove that this is not the case – but it is very hard to imagine how it could be. By the time a development project is implementing the political spoils have already been divided; any benefits to constituents in the home country have already been awarded at the moment the contract is signed. An ineffective development project simply fails to generate as much value for the recipient country while generating no clear additional value for political authorizers. As such, it is hard to imagine how a political authorizer who had any preference for helping others – or, indeed, was indifferent to the welfare of others – might gain from an ineffective development agency.

I have argued that whether the principal might indeed accomplish more with less control depends critically on the nature of the task, building on a thread in the bureaucracy literature that goes back at least to Wilson’s (1989) typology of organizations. Wilson, in fact, goes so far as to suggest the distortionary power of measurement and its interaction with task observability and autonomy:

> In coping organizations as in procedural, management will have a strong incentive to focus their efforts on the most easily measured (and thus most easily controlled) activities of their operators. They cannot evaluate or often even see outcomes, and so
only the brave manager will be inclined to give much freedom of action to subordinates. (p. 171)

What distinguishes this dissertation is an articulation that connects what Wilson might have called the “propensity for bravery” – that is, the likelihood of subordinates being given freedom of action – all the way ‘up’ to differential political authorizing environments. In doing so it brings together the “low politics” of Lipsky’s (1980) street level bureaucrats with the “high politics” of rational choice models. Political authorizing environments, organizational structure, organizational strategies (what I call Navigation strategies), and results are contingently related – certain kinds of political authorizing environments and thus organizational strategies are more likely to fail and succeed depending on the nature of the task and environment.

Management Studies

Beyond international development, this work invokes a long, if somewhat dormant, thread in the management literature. That organizational strategy is contingent on features of the environment, with autonomy a key lever, has a long history in Organizational Behavior. (Lawrence and Lorsch 1967; Thompson 1967). Key in this theoretical literature is uncertainty, with greater environmental volatility (and thus less potential for task routinization) associated with a higher optimal level of agent discretion and autonomy (March and Simon 1958; Perrow 1967; Thompson 1967).

This work builds upon these themes in the context of international development while bringing them into conversation with a discussion of task observability and
contractibility that I believe to be novel in the organizational behavior literature. In doing so these findings connect two emerging threads in the management literature: that goals and targets sometimes go too far, and that more opacity and less observation of agents can sometimes yield better results. (Bernstein 2012; Ordóñez et al. 2009) The evidence found here in support of the interaction between task, environment, organizational strategy, autonomy, and organizational resources is ripe for further exploration, as is the crowding in (or out) of soft information and tacit knowledge as mechanisms underlying these effects.

This work has focused on international development not just because of the advantages of IDOs as a context of study, but because it matters substantively. It matters not just as an abstract concern regarding efficiency but tangibly to the lives of literally hundreds of millions of people. However, the dimensions along which Navigation by Judgment and Navigation from the Top augur for better or worse organizational performance—argued here to be task observability and environmental unpredictability—matter well beyond the confines of IDOs, with potential relevance to a range of organizations, particularly those that often work in novel and complex contexts or task domains.

6.5: THE WAY FORWARD

Limitations & Further Work

IDO$s$—and the aid industry more broadly—offer scholars of bureaucratic politics, development studies, public management, and organizational behavior the prospect of a
relatively unexplored area where one might expect large effect sizes, novel contexts in which to generate theory or explore its boundaries, and substantively significant potential impacts for research findings. Potential margins of future research, some of which have already been suggested, include exploring variation in intra-IDO autonomy between projects, and countries; hiring and staff review processes and incentives; performance measurement (both in the human resources sense and in the organizational/project performance sense); staff rotation practices; the role of staff quality, including the feedback loop between staff quality and work environment; and organizational learning, among others.

While the empirical strategy has been designed so as to maximize internal validity – to ensure that conclusions drawn from these data accurately reflect the data – it is nonetheless possible that additional data would lead to different conclusions. Including the data of additional IDOs (if further data became available) might alter the findings of the quantitative analysis. On the qualitative side, the complexity of projects and the number of margins on which they can differ means that a similar analysis of a wider range of projects and IDOs would almost surely produce a fuller picture of IDO behaviors. It is certainly possible that such an exploration could lead to substantively different conclusions than those articulated here; only future research would indicate whether this is the case.

510 One recent paper from the World Bank research department finds that who supervises a project—that is, individual-level fixed effects—plays a greater role in project success than any other feature of the project (Denizer, Kaufmann, and Kraay 2013).
As to external validity within the universe of IDOs, while there is every reason to believe these findings extendable to the universe of aid organizations, more investigation is needed to confirm this is the case. More work is also needed on the strength and weaknesses of working through government, and on the potential for ‘division of labor’ – for IDOs to focus on the things their political authorizing environments make viable – as a potential route towards greater aggregate impact. The growth of contracting by IDOs also provides an opportunity for further work.

This work has only begun exploring the rich relationship between navigation strategy and performance. It is possible that Navigation by Measurement may be superior in delivery of simple, externally monitorable interventions with clear outcomes (e.g. PMTCT) irrespective of environmental legibility; possibly useful where monitoring is tricky but possible (e.g. general health systems), with environmental legibility critical in the analysis; and almost certainly inferior when the task is less tractable to measurement (e.g. reforming a health financing office). More work is needed to investigate this hypothesis. Future research could also articulate the variety of intermediate positions between these two extremes, or other dimensions of navigation left unexplored by this work.

There is also an open question as to how far these findings can generalize beyond aid organizations. More work is need as to what extent tight management crowds out soft information and tacit knowledge utilization in contexts where the observability challenge is substantially less than that of foreign aid.
Implications for IDOs

It is striking that the implication of this work’s findings points in the opposite direction from that IDOs seem to be moving. USAID Forward seems focused on indicators and target setting; DFID has lowered the threshold for discretionary spending and moved towards greater ex-ante articulation of plans and justification of same via “Business Cases”, a practice likely to crowd out tacit knowledge and soft information.

The growth of contracting is a (possibly related) trend pointed to by many DFID staff as a negative. Ken Sigrist, a long time DFID contractor,\(^{511}\) argued:

This whole process of packaging up technical assistance and contracting it out in large lumps actually takes away all of the flexibility, all of that opportunity to pivot quickly and to take advantage of circumstances, you can’t build that into the bloody contract.... I want to be able to say to DFID, as your advisor on the ground, I consider us to be stuck for these reasons, I consider the way to break out of this situation is the following, this is the strategy I propose, pros and cons and all the rest of it, I want to be able to have that conversation and that is more difficult now.

Sigrist’s views echoes those of Gorna and other DFID actors quoted in this work.

One possibility is that USAID and DFID’s supervisory styles and relationships with contractors are both optimal responses to the relationship between the political authorizing environment and the agency itself. In the case of USAID, the only way to achieve de facto flexibility may be via opacity; via turning over day-to-day implementation to contractors in such a way as to become unaware, and thus not have to account directly to

\(^{511}\) Sigrist was the head of CISCAB in its last phases, and thus is listed as an interviewee in the Liberia capacity building section of Appendix I.
Congress for day-to-day decisions. On the other hand DFID’s move towards contracting may reduce effective Navigation by Judgment just as USAID’s contracting opens up a window for Navigation by Judgment. That the answer to the Williamsonian “make or buy” question moves net navigation in opposite directions depending on political authorizing environment is an additional avenue where the study of IDOs might yield results of more general theoretic interest.

Changes within IDOs over time underscore the fundamentally mutable nature of the key variables discussed here; Carpenter’s “politically forged” agency autonomy is not static. There has been a move in the halls of Government – particularly in the UK – towards a policy of something like earned autonomy; of greater delegation to those who perform well. This work suggests that such a policy may in many cases be the inverse of what might work best – that better performance is likely to be a product of greater autonomy.

The focus on comparative IDO performance naturally points towards thinking of the ecosystem of aid organizations as a whole. If individual IDOs find it difficult to change it may still be possible for the aid system as a whole to serve common goals more effectively. This work provides fairly strong evidence that constrained IDOs would be best focusing on tasks and recipient country environments tractable to measurement – e.g. road construction or drug delivery. This would allow organizations that can Navigate by Judgment to focus on tasks where they have a comparative (and absolute) advantage. In the context of the cases examined here, one can imagine that USAID may have struggled
with CISCAB as it did with LGSP; DFID may have struggled to implement PEPFAR PMTCT treatment as effectively.

There is little evidence that this kind of specialization is occurring at present. The project selection data presented in Chapter 3 strongly suggests that IDOs are not focusing on the projects where they are most likely to succeed. In some cases – e.g. that of USAID – this appears to be due to earmarks, to Congressional requirement that money be spent in particular projects or with particular focus. These kinds of restrictions – and, indeed, requirements about the sectoral or country focus of funds – restricts Navigation by Judgment almost by definition. What kinds of such restrictions augur for the best possible organizational performance is an open question, one future research might fruitfully address.

**IDO Field Agents as Street Level Bureaucrats**

If the quality of IDO staff is not an important part of producing results, then many recipient country Government actors are substantially wrong. As one senior Liberian official put it in connecting staff quality to delivery and the importance of soft information:

---

Earmarks were frequently raised by USAID personnel as a reason for inefficiency. One case in point from a USAID staffer: “So there’s an earmark for South Africa saying well five hundred thousand dollars must go towards achieving basic education objectives, but we wouldn’t have education funding having been budgeted or allocated, but you still get the earmark so then you go out and you find out how do you use say EG [Economic Growth] funding and arguably justify it for use in an education project... you need to be able to use your EG funding for EG purposes and you’ve got to meet earmark so one can do that you’ve obviously got to justify from an economic growth point of view why working in basic education or certain projects which have an education flavour also have an economic growth outcome.” (Interview 78, 7/22/13) Managing earmarks clearly takes a lot of time in addition to introducing rigidities which restrict Navigation by Judgment.
The delivery of programs, delivery of efforts and initiatives are really to a large extent based in the kinds of relationships that [are] developed around personalities of the individuals in leadership positions within the government side and the donor side. There has to be an appreciation of the subtleties on the ground.... [and] a recognition of the genuineness of desire of those government actors and leaders that have the responsibility to carry it, but where you [the donor] can also visibly see their challenges.... You [the donor] then has an appreciation of the reality of what we’re trying to do.\(^{513}\)

It is all the more unfortunate, then, that Navigation from the Top may cause demotivation and the exit of quality staff. This work has provided theoretical foundation and some suggestive evidence\(^{514}\) that the staff who exit are higher quality than those who remain. Even were that not the case, working for an IDO requires job-specific skills and knowledge – tacit knowledge must be built, judgment capacity honed, instincts developed. This takes time, and increases the cost – in the sense of success forgone – of staff turnover. Staff turnover also impedes organizational learning, as a staff member who is learning the job has less ability to provide knowledge back to the center.

We have seen over and over again how IDO field agents and/or the contractors to whom the organization has delegated ‘make’ policy. The project as it exists in memoranda of understanding and program documents certainly has something to do with what is experienced on the ground; but ultimately the project depends on the client orientation, judgment, and quality of IDO street-level bureaucrats. Whether these actors are oriented ‘down’ towards clients and delivery or ‘up’ towards their distant superiors and political

---

\(^{513}\) Interview 1, 6/3/13

\(^{514}\) This is outlined most directly in section 4.4 of chapter 4.
authorizers depends substantially on the institutional environment, which also affects both entry into and exit out of these critical field roles.

Measurement and its Discontents

Where the outcome desired is known, observable, and contractible without distortions, using the outcome as a target is the best strategy. Where the outcome desired is known but contracting on that outcome is impossible – whether that impossibility is due to a lack of outcome observability, the time period in which the data is needed and the outcome emerges, or the susceptibility of the situation to Goodhart’s Law and/or multitask problems – the answer is less clear, depending on features of the case.

In the world of international development it is hard to closely monitor the activities of agents from the center. As such, much of Navigation from the Top in the context of international development comes through targets, through Navigation by Measurement. IDOs respond to these controls – measures are met, targets are reached. These measures do not often focus on clear long term outcomes for a host of reasons.\textsuperscript{515} Goodhart’s Law and multitask problems are very real, made more so by measures focusing on process and output due in part to an effort to generate data in a timely manner. Sometimes this measurement is nonetheless a very good thing; where the measure is closely related to the outcome (e.g. PMTCT drug delivery or road construction) this works wonderfully. As

\textsuperscript{515} Reasons include the unobservability of outcomes, the unattributability of these outcomes to specific projects, or the need for desired outcomes to be mutable ex-ante, discovered by “muddling through”.

302
predicted by theory, where what’s measured has a weaker relationship with outcomes (e.g. municipal governance trainings, health systems strengthening) these measures work decidedly less well.

As quoted above, Wilson (1989) predicted the move towards measurement in the public service, particularly where outputs and outcomes are difficult to observe. He went on to suggest that bureaucrats would not respond uniformly to these measurement schemes, writing:

Some [bureaucrats] will conform their behavior to whatever is being measured (“they want stats, we’ll give ‘em stats”); others will subvert the management strategy by ignoring the measured activities (thus jeopardizing their own chances for advancement) or by generating enough stats to keep management happy while they get on with their own definition of what constitutes good work. (p. 171)

We have seen both the first and the last strategy in this work (presumably those that engage in the middle strategy do so at their peril, and do not last long), but with differences more at the organizational than the individual level. USAID staff tend to “give ‘em stats”; this provides a good deal of autonomy in how one delivers the stats, but little capacity to focus on other things. DFID measures, too; these measures just tend not to constrain, with field staff and contractors seeming to generate enough stats to “keep management happy”. This may be because management sets looser targets; it may be because management cares about more than the numbers. It may be because DFID allows reporting in a way that provides opportunities for field staff to mis-attribute gains to the project. In any case,

516 CMTP mis-reporting in this way was alluded to in Chapter 4. I call this phenomenon “piggybacking” – that is, when an IDO exploits a secular trend to declare success. PwC in the OFM project is perhaps the purest example in these cases. The logical framework on which
DFID’s numbers seem not to bite in the same way as USAID’s; while the measures exist pressure to meet them does not, meaning there are fewer Goodhart’s law distortions of the measures’ relationship with outcome or multitask distortions of agent effort towards the observable and measurable. This is all to say it is not the measure itself that creates problems and crowds out Navigation by Judgment; it is the pressure put upon these measures and targets.

It is often mooted that a unique advantage of control (including measurement) in the public sector is in making agencies accountable to political authorizers and ultimately the public. If indeed it is true that measurement in some circumstances is detrimental to performance, it is unclear what value “accountability” has when it comes along with an actual reduction in the desired performance outcomes. Second, Goodhart’s law suggests that the measures of performance – the vehicles via which accountability is meant to flow – may break down precisely when they are used for accountability practices. In these cases, PwC had to report contains indicators related to the project’s plausible achievements to be sure (e.g. number of personnel trained in financial management), but also indicators that seem only very tangentially related to the (clearly laudable) goal of sound financial management – e.g. reduction of maternal and child mortality rates in Liberia (by 2010 a 10% reduction in maternal mortality and a 5-10% reduction in child mortality relative to 2007 baselines). These are laudable goals to be sure, and there is no doubt that better financial management can (and in the case of the OFM did) play an important role in improving the allocation of funds in the health sector as well as attracting new funds, and as such clearly has a positive impact on real health outcomes. But, of course, Liberia had myriad programs, from a variety of sources (including substantial NGO and RBHS financing) devoted to the same outcomes! It is impossible to estimate the added value of the OFM to these results; and surely this state of affairs was perfectly predictable when the target was conceived. This project “piggybacked” on other projects and secular trends in a way that allows it to demonstrate what look like quantitative results – and indeed would have done so even in the absence of the OFM’s success.
improved performance on measures may not mean more has been accomplished but rather less. It is unclear what the benefit to citizens is of measurement when this measurement serves to obfuscate actual performance and/or reduces performance. Measurement, then, can be in service of accountability or a façade that makes true accountability more difficult depending on the nature of the measure, task, and environment.

**Final Thoughts**

Lurking just out of the frame of this dissertation has been the rapid penetration of technology in all walks of life and the opportunities for increased Navigation from the Top, and Navigation by Measurement, such technologies facilitate. Wilson’s paradigmatic examples of bureaucratic tasks where both outputs and outcomes were unobservable were teachers and police officers. Today teacher measurement via student testing outcomes is an industry; police wear body cameras and the public is full of individuals with camera-equipped phones. This increased monitoring, measurement, and transparency is sometimes very good – as I write these words in April 2015 the US is engaged in a period of introspection regarding racial inequities in policing catalyzed by video of police actions. But it can also have deleterious effects, crowding out judgment, learning, and efforts towards things that cannot be readily measured. Framed this way, the debate on teacher evaluation has much of the flavor of the challenge IDOs face.

In some ways international development is relatively immune from technology’s growth because of lower levels of technological development. In some ways it is not – new
mobile technologies allow the measurement of things that would have been impossible just a decade ago. These technologies have facilitated some of the most impressive gains of the last decade – e.g. those of vaccine delivery. As these technologies expand the tensions highlighted in this work are likely to become more, not less, acute. Political authorizers and IDO managers need to reach a common understanding of where opacity and judgment are in service of political authorizers’ goals, experiment, and adjust accordingly. Doing so involves some risk, and probably increases the likelihood of some – quite politically difficult – costs, like misuse of funds. But to do otherwise is to condemn the enterprise of development aid to failure in some of its most potentially impactful realms of intervention.

Soft information is a delicate thing; for an organization to cultivate it takes time and space, and is not without costs. Navigation by Judgment is not always the best strategy; neither is Navigation from the Top. Choosing the right navigation strategy in international development – and, perhaps, far more broadly – depends on the nature of the task, the technology, the agents, and the environment. There are few ‘one size fits all’ solutions to organizational strategy; and getting it right requires patience, wisdom, and an orientation to the particular case. A move towards more such nuance in IDOs has the potential for great gains, but can only be realized by changes in structure and political authorizing environments that allow for a broader menu of navigation strategies to be available to IDO decision makers.

Where output measurement and tight control by distant principals works well Navigation by Measurement should be used to better deliver vaccines or more efficiently
build electricity transmission infrastructure. But in some of the domains where foreign aid has the potential to make the most difference – in the most fragile states, and in tasks like improving government’s ability to manage – measurement is the least useful, with Navigation by Judgment the optimal strategy. These findings suggest that not only are we not doing all we can to improve aid delivery, the move towards measurement and control across all aid sectors in recent years may actually be making things worse in some sectors. Measurement may lead to the construction of many successful dams but leave recipient countries without the capacity building necessary to manage and maintain those dams or to put the electricity to use. If our drive for results leads us to control aid too tightly, we may end up accomplishing precisely the opposite of what we intend.
APPENDIX I: INTERVIEW SOURCES & METHODS

Case Study Methods

The case studies are drawn from primary documents and a series of semi-structured interviews, the majority of which occurred in-country in spring/summer 2013 (for Liberia, May-June 2013; for South Africa, July-August 2013).

In both Liberia and South Africa I began with a series of contacts regarding each case drawn from personal networks and official (public) documents. While I happily accepted referrals (i.e. ‘snowball sampling’) to other key actors, I made sure to ‘re-seed’ my network several times to minimize the possibility of referral bias; e.g. to ensure I was not only speaking to those Government actors to which I was referred by donor actors (or vice versa).

At the beginning of interviews I made clear to each interviewee that the conversation was confidential by default – that no specific content would be linked to their name without their express permission (which was received in all cases of specific quotation/reference). I also made clear that they could request anonymity at any time, either before, during, or at any time after the interview, and I would list them as anonymous rather than include their proper name and title on the list of interviewees.

In addition to taking notes I made recordings of interviews where context or the preference of the individual (I asked before each interview) did not preclude same. I also
turned off the recording device in the few cases where it became clear that the interviewee was significantly constraining responses due to the recording (e.g. an interviewee frequently glancing at the recorder combined with vague answers containing little content). In the end I recorded over 80% of interviews, and I re-listened to these interviews while writing up full (electronic) notes prior to writing the case studies proper.\textsuperscript{517}

I thought of each project as involving three phases in which autonomy might be relevant – program design, program implementation, and program revision (or lack thereof). For each case I made sure to interview those responsible for leading the project for the donor; those responsible for coordinating day-to-day implementation (sometimes a donor representative, sometimes that of an international contractor, sometimes a government actor – a data point in itself), and the responsible recipient government representative who was the ‘government liason’ on the project. In many cases due to staff rotations multiple people filled a particular role (e.g. lead responsibility for the project on the donor side) over the course of the project; I tried to interview as many of these individuals as possible.

My primary objects of inquiry were descriptions of how the project was formed and why (including their/their agency’s goals), what revisions were contemplated and implemented (and why these and not others etc.), and the process/structure of day to day management – what was possible for various actors to do, and how this translated into

\textsuperscript{517} In a minority of cases I had interviews transcribed (with transcribers signing confidentiality agreements).
project decisions. In general, I was interested in what features of organizations (recipient government bureaus and international development organizations) and the incentives key decisionmakers and implementers faced seemed to have causally influenced outcomes.

At the same time, I tried to keep the door open both to conflicting (e.g. agent autonomy led to worse outcomes) and alternative hypotheses. To what extent was country ownership present, and how did it come about? Did recipient governments have an independent vision of what the project looked like, and if so were they successful in engaging in a dialectical back-and-forth with IDOs? What were the results of said process if so, and in what ways do they differ across donors/projects?

The flow and structure of the interviews themselves were quite dependent on whom the interviewee was; I began each interview by asking the interviewee the open-ended question “How did you first hear of X project?” and then proceeded from there. I always asked about their personal role relative to others, as well as the role of the actor for whom they worked (e.g. the implementing contractor) relative to supervisors. I also always asked (towards the end of the interview) for their evaluation of the success of the project and two open-ended questions, one regarding what could have made it better and one regarding why the project was as successful as it was (both of which opened the door quite widely to alternative hypotheses). Following a few initial interviews in which it proved difficult to get critical responses regarding what might have been different, I began to ask for a project rating on a scale of 1 to 10; I would then ask what would have made the project a 10 rather than the number (e.g. 6, or 8.5) at which they rated the project. Where
relevant I also asked about project reporting and measurement, both at a personal and organizational level – what was being measured, whom it was being reported to, etc., and for a project timeline/narrative regarding the evolution of projects.

I asked questions in as open-ended a way as possible, and tried to ensure I made every response sound equally desired (to avoid implicitly encouraging certain types of answers). I avoided stating my hypotheses at all, and if pressed described the work in the most general and neutral terms (e.g. “I'm interested in the organizational features of development success, and to help others learn from this project”), though in a few cases I became more specific after the end of the interview (my promise to do so having been necessary ex-ante for the interview to proceed). I asked follow-up questions when things were unclear, and additional questions as they were prompted by the project narrative. In general, I was open to the interview flowing in the direction the interviewee wished to take it to a moderate extent, inasmuch as I suspected (as was proven out by events) that many people wanted to express particular positive or negative things regarding the projects in which they had been involved.

In addition to publicly available information, I was also able to access non-public (or in some cases theoretically public, but not publicly available, documents) in concert with interviews or from other individuals. In parallel to the standard employed in interviews, I made clear blanket confidentiality applied to document disclosure. I also filed freedom of information requests with both USAID (US FOIA) and DFID (UK FOI), which led to the further declassification and disclosure of documents. I have verified the veracity of all
referenced documents not from an official source to my satisfaction, and have not incorporated or referred to documents regarding whose accuracy I remain unsure. In a few cases documents were shared on the condition they not be referenced or referred to in any way; these documents are thus not listed in the references section at the end of this work.

Interviewee Data and Numbering Schemes

The following tables indicate all individuals that provided information (mostly by interview, but in a handful of instances by email correspondence) that informs the cases in chapters 4 and 5. Many thanks to all those who were kind enough to give of their time.

A few notes on these tables – first, the number in the left hand column does not correspond to the number in the interview citations (e.g. interview 63, 6/25/13), which are randomized, and not alphabetical, to maintain the promised anonymity to interviewees. The dates of the interviews are omitted from this table, as including them would make it much easier to infer the identity of a given speaker. For the same reason, individuals who contributed in more than one domain – e.g. speaking to South African interventions in both municipal governance and health – are given a new randomly generated number, as to do otherwise would make it exceedingly easy to identify these speakers. As such, though 146 interviews are listed below, there is a degree of overlap; it is small, however. There are more than 135 unique interviews on which these cases draw.
<table>
<thead>
<tr>
<th>#</th>
<th>Surname</th>
<th>First Name</th>
<th>Position</th>
<th>USAID</th>
<th>DFID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Anonymous</td>
<td>Anonymous</td>
<td>COGTA Senior Official</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Bester</td>
<td>Angela</td>
<td>Former DFID staff, then Deloitte; also former DG, Public Service Commission</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Brooks</td>
<td>Frikkie</td>
<td>Head of KZN provincial planning department</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Chipkin</td>
<td>Ivor</td>
<td>Executive Director, PARI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Chrystal</td>
<td>Blake</td>
<td>Supervisory Program Officer, USAID SA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Dei</td>
<td>Colleen</td>
<td>Former USAID SA Mission Chief</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Fortuin</td>
<td>Joe</td>
<td>Director of Aid, COGTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Francis</td>
<td>Virginia</td>
<td>USAID Health Team, former RTI SA staff</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Glasser</td>
<td>Matt</td>
<td>Former USAID advisor in SA on municipal financing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Hackner</td>
<td>Allan</td>
<td>USAID SA Financial Sector Manager (former COR on Municipal)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Harding</td>
<td>Joel</td>
<td>DFID Governance Advisor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Heymans</td>
<td>Chris</td>
<td>Former CMTP chief architect, now WB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Hofmeyr</td>
<td>Beatie</td>
<td>Head of Education and Training Unit, LGSP implementing sub-contractor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Horn</td>
<td>Steve</td>
<td>former ISLGS CoP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Kolker</td>
<td>Joel</td>
<td>Former USAID staff, municipal program</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Konig</td>
<td>Ferdie</td>
<td>CMTP ISF in Phalaborwa, Mpumalanga</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Layte</td>
<td>Michelle</td>
<td>former RTI LGSP CoP (Vaz’s successor)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Madurai</td>
<td>David</td>
<td>Chief Director, Norms, Standards, Policy and Research, COGTA; former Chief Director, Development Planning &amp; Local Economic Development</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Mangokwena</td>
<td>Andries</td>
<td>Advisor in Thulamela under CMTP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Mathivha</td>
<td>Makonde</td>
<td>Municipal Manager, Thulamela, Limpopo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Matomela</td>
<td>Bongani</td>
<td>Former Deputy Project Director, LGSP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Naidoo</td>
<td>Subethri</td>
<td>Former Governance Advisor, DFID; former Local Government sector manager, USAID; former Deloitte program manager on CMTP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Olver</td>
<td>Chippy</td>
<td>Former Deputy Director General DPLG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Powell</td>
<td>Derek</td>
<td>Former Deputy DG, DPLG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Rambulana</td>
<td>Wilson</td>
<td>former LGSP Revenue Enhancement Advisor (trainer)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Sadan</td>
<td>Mastoera</td>
<td>Programme Manager, PSPPD, Office of the SA Presidency</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Surname</td>
<td>First Name</td>
<td>Position</td>
<td>USAID</td>
<td>DFID</td>
</tr>
<tr>
<td>---</td>
<td>---------</td>
<td>------------</td>
<td>--------------------------------------------------------------------------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>27</td>
<td>Savage</td>
<td>David</td>
<td>Former WB staff, now SA Treasury head of Cities Support program</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Snook</td>
<td>Steve</td>
<td>former USAID Democracy and Governance deputy team leader</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Tazewell</td>
<td>Littleton</td>
<td>Deputy Mission Director, USAID South Africa Regional Program</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Thomas</td>
<td>Richard</td>
<td>Former DFID South Africa Governance Advisor on CMTP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Timm</td>
<td>Jeremy</td>
<td>Former CMTP now Treasury muni gov support</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>Toli</td>
<td>Robin</td>
<td>Chief Director, International Development Coordination, SA Treasury</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>TV</td>
<td>Pillay</td>
<td>Head of Municipal Finance, SA Treasury</td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>Vaz</td>
<td>Peter</td>
<td>former RTI LGSP CoP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>Yako</td>
<td>Pam</td>
<td>Former municipal manager, Amathole District; former DG, Environmental Affairs, Water Affairs</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table I.2: South Africa Health Interviews

<table>
<thead>
<tr>
<th>#</th>
<th>Surname</th>
<th>First Name</th>
<th>Position</th>
<th>USAID</th>
<th>DFID</th>
<th>CDC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Agenbag</td>
<td>Rentia</td>
<td>Government and Civil Society Support Manager, SANAC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Anonymous</td>
<td>Anonymous</td>
<td>CDC &amp; USAID PEPFAR Implementer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Anonymous</td>
<td>Anonymous</td>
<td>Senior DC-based PEPFAR official</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Anonymous</td>
<td>Anonymous</td>
<td>Look at title in Liberia chapter (anonymous)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Anonymous</td>
<td>Anonymous</td>
<td>USAID and CDC PEPFAR Implementer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Barker</td>
<td>Pierre</td>
<td>Senior VP, Institute for Health Care Improvement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Barron</td>
<td>Peter</td>
<td>Public health specialist &amp; advisor to DDG Pillay</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Coovadia</td>
<td>Jerry</td>
<td>Director, MaTCH</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Coovadia</td>
<td>Ashraf</td>
<td>Head of pediatric HIV,Rahima Moosa Mother and Child Hospital, Johannesburg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Dei</td>
<td>Colleen</td>
<td>Former USAID SA Mission Chief</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Desmond</td>
<td>Chris</td>
<td>Chief research specialist, Human Sciences Research Council</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Fryatt</td>
<td>Bob</td>
<td>Former DFID Health Advisor, SA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Name</td>
<td>Title</td>
<td>Organization/Position</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>------------</td>
<td>----------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Giddy Janet</td>
<td>Former HIV program coordinator, McCord Hospital, Durban</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Goga Ameena</td>
<td>Senior Specialist Scientist, MRC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Gorna Robin</td>
<td>Former Senior Regional Health and AIDS adviser for DFID Southern Africa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Grant Ken</td>
<td>HLSP Programme Director, SARRAH</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Harding Joel</td>
<td>DFID Governance Advisor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Holst Helga</td>
<td>CEO, McCord Hospital, Durban</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Kok Michelle</td>
<td>Advisor to Precious Robinson, NDOH</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Kumar Smita</td>
<td>USAID PMTCT Lead</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Lesole Lerato</td>
<td>PMTCT Specialist, CDC SA; previous NDOH</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Mahasela Lusanda</td>
<td>Deputy, Research &amp; M&amp;E, Johns Hopkins Health and Education in South Africa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Mazibuko Ntombi</td>
<td>RTC PMTCT Project Manager; former EGPAF &amp; NDOH; former NDOH</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Ngubane Gugu</td>
<td>former HLSP A-Plan Project Manager and Technical Adviser on PMTCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Nkulu Hilary</td>
<td>former DFID SA Programme Manager</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Pattinson Robert</td>
<td>Director, MRC Maternal and Infant Health Care Strategies Unit, University of Pretoria</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Pillay Yogan</td>
<td>NDOH Deputy Director General</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Robinson Precious</td>
<td>NDOH Deputy Director in charge of PMTCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Sanne Ian</td>
<td>CEO, Right To Care</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Schneider Helen</td>
<td>Director, School of Public Health, University of the Western Cape; former SANAC, MRC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Slingers Nevilene</td>
<td>Donor Coordination Manager, SANAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>Taback Rayna</td>
<td>Senior Public Health Advisor, CDC South Africa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#</td>
<td>Surname</td>
<td>First Name</td>
<td>Position</td>
<td>USAID</td>
<td>DFID</td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>---------</td>
<td>------------</td>
<td>--------------------------------------------------------------------------</td>
<td>-------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Anonymous</td>
<td>Anonymous</td>
<td>Former NACP personnel</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Anonymous</td>
<td>Anonymous</td>
<td>Current MoHSW senior personnel</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Anonymous</td>
<td>Anonymous</td>
<td>Senior official, USAID Liberia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Anonymous</td>
<td>Anonymous</td>
<td>USAID Liberia international staff</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Augustin</td>
<td>Randolph</td>
<td>Lead Health Officer, USAID</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Benson</td>
<td>Angela</td>
<td>FARA Coordinator, MoHSW</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Bility</td>
<td>Kalipha</td>
<td>Former Program Coordinator, NACP; now Deputy Minister, Ministry of Agriculture</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Bruce</td>
<td>Lwopu</td>
<td>Head of Blood Safety, MoHSW; former deputy head, NACP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Curran</td>
<td>Desmond</td>
<td>DFID Representative in Liberia 2007-2009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Dahn</td>
<td>Eunice</td>
<td>Chief Medical Officer, MoHSW</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Davis</td>
<td>Natty B</td>
<td>Chairman and CEO, NiC; former Minister without Portfolio and National Coordinator, LRDC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Dolopeh</td>
<td>Dr.Eugene</td>
<td>Former Program Manager, NACP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Duncan</td>
<td>Julie</td>
<td>Commissioner, NAC; former Assistant Minister for Preventive Services, MoHSW</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Dworku</td>
<td>Tanu</td>
<td>Former USAID Health Officer, Former NACP Coordinator</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Dzokoto</td>
<td>Agnes</td>
<td>Senior Technical Officer, AWARE (responsible for Liberia)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Flomo</td>
<td>Matthew</td>
<td>Deputy Minister for Administration,</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table above shows the interviews conducted with various individuals from the Liberian health sector. The interviews were organized by Surname, First Name, Position, and their affiliated agencies, USAID or DFID.
<table>
<thead>
<tr>
<th>#</th>
<th>Surname</th>
<th>First Name</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>Freeman</td>
<td>Josephine</td>
<td>Former PMTCT Coordinator, NACP</td>
</tr>
<tr>
<td>18</td>
<td>Gabelle</td>
<td>Chris</td>
<td>Former lead Liberia Governance Advisor, DFID</td>
</tr>
<tr>
<td>19</td>
<td>Gwenigale</td>
<td>Walter</td>
<td>Minister of Health, MoHSW</td>
</tr>
<tr>
<td>20</td>
<td>Hughes</td>
<td>Jacob</td>
<td>Head of Liberia Health PF Management firm; Former PwC</td>
</tr>
<tr>
<td>21</td>
<td>Jones</td>
<td>Janyaj</td>
<td>M&amp;E Deputy, NACP</td>
</tr>
<tr>
<td>22</td>
<td>Karzon</td>
<td>Toagee</td>
<td>Controller, MoHSW</td>
</tr>
<tr>
<td>23</td>
<td>Lippevald</td>
<td>Theo</td>
<td>RBHS/JSI Deputy CoP</td>
</tr>
<tr>
<td>24</td>
<td>Logan</td>
<td>David</td>
<td>Global Fund Coordinator, MoHSW; former deputy coordinator, NACP</td>
</tr>
<tr>
<td>25</td>
<td>Macaulay</td>
<td>Rose</td>
<td>RBHS/JSI CoP</td>
</tr>
<tr>
<td>26</td>
<td>Manuel</td>
<td>Marcus</td>
<td>Former DFID Deputy Director for West Africa</td>
</tr>
<tr>
<td>27</td>
<td>Mapleh</td>
<td>Louise</td>
<td>PBF Coordinator</td>
</tr>
<tr>
<td>28</td>
<td>Martin</td>
<td>Bill</td>
<td>Former Senior Adviser to the Minister, MoHSW; now PF Manager</td>
</tr>
<tr>
<td>29</td>
<td>McDermott</td>
<td>Chris</td>
<td>Former health lead, USAID</td>
</tr>
<tr>
<td>30</td>
<td>Nartey</td>
<td>Alex</td>
<td>Former lead of PwC team to MoHSW</td>
</tr>
<tr>
<td>31</td>
<td>Niyuhire</td>
<td>Floride</td>
<td>RBHS Advisor on PBF to MoHSW</td>
</tr>
<tr>
<td>32</td>
<td>Nyoweh</td>
<td>Moses</td>
<td>STI Officer, NACP</td>
</tr>
<tr>
<td>33</td>
<td>Sanvee</td>
<td>Dr.Lilly</td>
<td>Head implementer, Catholic Hospital, AWARE</td>
</tr>
<tr>
<td>34</td>
<td>Sieh</td>
<td>Sonpon</td>
<td>Program Coordinator (head), NACP; former M&amp;E on HIV, NACP</td>
</tr>
<tr>
<td>35</td>
<td>Sirleaf</td>
<td>Momolu</td>
<td>Head of Aid Coordination, MoHSW</td>
</tr>
<tr>
<td>36</td>
<td>Subah</td>
<td>Pewu</td>
<td>Head of Project Implementation Unit, MoHSW</td>
</tr>
<tr>
<td>37</td>
<td>Tamattey</td>
<td>Felix</td>
<td>Senior Partner leading PwC Engagement, MoHSW</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table I.4: Liberia Capacity Building Interviews</th>
<th>USAID</th>
<th>DFID</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td>Surname</td>
<td>First Name</td>
</tr>
<tr>
<td>---</td>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td>1</td>
<td>Anonymous</td>
<td>Anonymous</td>
</tr>
<tr>
<td>2</td>
<td>Anonymous</td>
<td>Anonymous</td>
</tr>
<tr>
<td>2</td>
<td>Allen</td>
<td>William</td>
</tr>
<tr>
<td>3</td>
<td>Atuanya</td>
<td>Jenkins</td>
</tr>
<tr>
<td></td>
<td>Name</td>
<td>Designation</td>
</tr>
<tr>
<td>---</td>
<td>---------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>4</td>
<td>Baki Shadi</td>
<td>Head of Biometrics, CSA</td>
</tr>
<tr>
<td>5</td>
<td>Belleh Willie</td>
<td>Partner Subah Belleh Associates; local partner for CISCAB</td>
</tr>
<tr>
<td>6</td>
<td>Callender Elizabeth</td>
<td>Deputy Head, OTI Liberia</td>
</tr>
<tr>
<td>7</td>
<td>Cooper Vicky</td>
<td>Former WB consultant on Civil Service Pay Reform; current Chief of Party, GEMS</td>
</tr>
<tr>
<td>8</td>
<td>Cooper Lloyd</td>
<td>Grants Manager, BRDG</td>
</tr>
<tr>
<td>9</td>
<td>Curran Desmond</td>
<td>DFID Representative in Liberia 2007-2009</td>
</tr>
<tr>
<td>10</td>
<td>Davis Natty B.</td>
<td>Chairman and CEO, NiC; former Minister without Portfolio and National Coordinator, LRDC</td>
</tr>
<tr>
<td>11</td>
<td>Drosaye Alfred</td>
<td>CSAA PAO</td>
</tr>
<tr>
<td>12</td>
<td>Fahnbulle Louise</td>
<td>former OTI staff, Liberia</td>
</tr>
<tr>
<td>13</td>
<td>Fn’Piere Pat</td>
<td>Consultant, BRDG; OTI Field Advisor</td>
</tr>
<tr>
<td>14</td>
<td>Gabelle Chris</td>
<td>Former DFID Governance Advisor in Liberia</td>
</tr>
<tr>
<td>15</td>
<td>Gattorn John</td>
<td>Former Africa Program Manager, OTI</td>
</tr>
<tr>
<td>16</td>
<td>Glentworth Garth</td>
<td>OBE; former senior Governance Advisor, DFID</td>
</tr>
<tr>
<td>17</td>
<td>Hare Sam</td>
<td>Former Deputy Minister, Ministry of Youth and Sports</td>
</tr>
<tr>
<td>18</td>
<td>Hunter Rosslyn</td>
<td>M&amp;E team, BRDG</td>
</tr>
<tr>
<td>19</td>
<td>Johnson Mimi</td>
<td>HR team, BRDG</td>
</tr>
<tr>
<td>20</td>
<td>Kialain David</td>
<td>Former principal deputy, GRC</td>
</tr>
<tr>
<td>21</td>
<td>Lauer Barb</td>
<td>Former CoP for BRDG, DAI</td>
</tr>
<tr>
<td>22</td>
<td>Liberty T. Edward</td>
<td>Director General, LISGIS</td>
</tr>
<tr>
<td>23</td>
<td>Logan James</td>
<td>Former Deputy Minister, Ministry of Agriculture</td>
</tr>
<tr>
<td>24</td>
<td>Mayshak Nellie</td>
<td>Former head, ASI CISCAB team 2007-2009</td>
</tr>
<tr>
<td>25</td>
<td>Muhula Raymond</td>
<td>Public Sector Specialist, World Bank</td>
</tr>
<tr>
<td>26</td>
<td>Neymah Oblayon</td>
<td>Former Reform Directorate CSA; current head of LIPA</td>
</tr>
<tr>
<td>27</td>
<td>O’Neill Dominic</td>
<td>Head of DFID Sierra Leone 2008-2011</td>
</tr>
<tr>
<td>28</td>
<td>Panton Richard</td>
<td>Deputy Director General, Training, LIPA</td>
</tr>
<tr>
<td>29</td>
<td>Patel Jalpa</td>
<td>Former coordinator, ASI CISCAB project, 2009-2010</td>
</tr>
<tr>
<td>30</td>
<td>Sigrist Ken</td>
<td>Former head, ASI CISCAB team 2009-2010</td>
</tr>
<tr>
<td>31</td>
<td>Tarpeh Dominic</td>
<td>Former CISCAB consultant, now with GRC</td>
</tr>
<tr>
<td>32</td>
<td>Thompson James</td>
<td>Subah Belleh staff; former member of CISCAB core team</td>
</tr>
<tr>
<td>33</td>
<td>Wilson Peter</td>
<td>Program Development Officer, BRDG</td>
</tr>
<tr>
<td>34</td>
<td>Wilson Mark</td>
<td>Grants Manager, BRDG</td>
</tr>
</tbody>
</table>
REFERENCES


Barh, Benson, Selam Kebrom, Cecelia Morris, Mbuyi Mutala, and Bruce Grogan. 2011. Rebuilding Basic Health Services (RBHS): Year 2 Assessment. USAID.


Center for Systemic Peace. 2012. “State Fragility Index.”


Development Alternatives Inc. 2007a. BRDG-Liberia subgrants table. Undated, but based on completed activities fairly near to project close.

———. 2007b. BRDG-Liberia Grant Summary Sheets, DAIM044-099


———. 2007g. BRDG Yearbook.


Eyben, Rosalind. 2013. ...-a-framing-paper-for-development-practitioners ( ... Uncovering the Politics of “Evidence” and “Results”. A Framing Paper for Development Practitioners.


John Snow International, Johns Hopkins University Center for Communication Programs, and Management Sciences for Health. 2010. Rebuilding Basic Health Services Project Description.

John Snow International. 2008-2011. Annual and Semi-annual reports to USAID on RBHS.


———. 2006. Office of Financial Management: Office description, organogram, job descriptions, memos to DFID.


———. 2012-2013. Documents related to Performance Based Financing, including policy and actual performance scorecards, results, verification reports.

———. 2013a. Budget Performance Reports.

———. 2013b. Contracting In Guidelines: Background and Context

———. 2013c. Contracting In Guidelines: Summary of Systems, Procedures and Core Competencies for the County Health and Social Welfare Team (CHSWT)


Liberia Ministry of Finance. 2009-2013. Aid Management Reports, Spreadsheets (listing all donor programs reported to Government of Liberia, amounts of financing, etc.)
Liberia National Aids Control Program. 2004-2008. Project documents (various), including sub-grant agreement between NACP and FHI, extensions of same, and project communications, including communications with Catholic Hospital.


———. 2008b. Celebrating Life: Latest PEPFAR Results.


———. 2013. PEPFAR 9th Annual Report to Congress.

———. 2015. PEPFAR South Africa Overview: Partnering to Achieve Epidemic Control in South Africa.


———. 2010b. The Aid Effectiveness Framework for Health in South Africa.


———. PMTCT ART Flowchart.


Sessions, Myra. 2006. HIV/AIDS Monitor: Tracking Aid effectiveness; Center for Global Development Overview of the President’s Emergency Plan for AIDS Relief.


Stickel, Wendy, Sebastian Milardo, and Sarah Scheening. 2012. USAID/Liberia Health Sector FARA: An Assessment of Operational Lessons Learned in Design and Early Implementation. USAID.


———. 2003b. *CMTP Project Logical Framework, 10/2/03.*


———. 2009e. Logical Framework (updated), CISCAB Project.


———. 2009g. Project Completion Report, Support to Primary Health Care Services in Liberia (project 113353)

———. 2009h. Final Report, Support to Primary Health Care Services in Liberia (project 113353)
Various documents discussing, approving funding to NGOs in the health sector.

2010a. Project Completion Report, Office of Financial Management (project 113381)

2010b. CISCAB Project Completion Report.


2013b. Pooled Funding to Support Service Delivery: Lessons of Experience from Fragile and Conflict-Affected States.


undated. Logical Framework (updated), Support to Primary Health Care Services in Liberia (project 113353)

undated, but after June 2010. Project Document, Office of Financial Management (project 113381)

undated. Logical Framework, Office of Financial Management (project 113381)


UNICEF. 2012. Options B and B+: Key Considerations For Countries to Implement an Equity-Focused Approach.


Weijer, Frauke De. 2012. *Rethinking Approaches to Managing Change in Fragile States.*


