



Delegation and Stewardship in International Organizations:

Revisiting the Relationship between Member States and International
Administrations

Inaugural-Dissertation

zur Erlangung des Doktorgrades der Sozialwissenschaftlichen Fakultät
der Ludwig-Maximilians-Universität München

vorgelegt von

Vytautas Jankauskas

München, Februar 2021

Referent: Jun.-Prof. Dr. Steffen Eckhard

Koreferent: Prof. Dr. Christoph Knill

Tag der mündlichen Prüfung: 20. April 2021

Contents

Deutschsprachige Zusammenfassung der Dissertation III

List of Acronyms and Abbreviations IX

Acknowledgements XI

Framework Paper 1

 1 Introduction..... 2

 2 P-A theory and its limitations in studying MS-IPA relationships 7

 2.1 *The dominance of P-A theory* 7

 2.2 *The limitations of P-A theory*..... 9

 3 Probing and advancing our understanding of MS-IPA relationships 14

 3.1 *Common approach and workshare between the core papers*..... 14

 3.2 *Summary of the four core articles* 18

 3.3 *Lessons learned* 26

 4 A new research agenda - Between agency and stewardship in IOs..... 29

References 35

Appendix to Framework Paper..... 46

Paper A..... 51

 Introduction 52

 Beyond agency: Two types of IO delegation relationships..... 54

 Research Design 58

 Empirical Analysis 61

 Discussion 67

 Conclusion..... 69

 References 71

 Appendix to Paper A 75

Paper B 79

 Introduction 80

 Explaining IPAs’ Responses to Member States’ Contestation 81

 Assessing IPAs’ Responses to Contestation by the Trump Administration 85

 Conclusion..... 96

 References 98

 Appendix to Paper B 103

Paper C	133
Introduction	134
Theorizing political interests and the political use of evaluation in IOs	136
Explaining political evaluation use in IOs	140
Research design.....	143
Results	147
Discussion	153
Conclusion and theoretical implications	155
References	157
Appendix to Paper C	162
Paper D	169
Introduction	170
Basic assumptions of Principal-Agent theory	172
Endogenous and exogenous sources of agency slack	174
‘The nature of the beast’: Conceptualizing and Operationalizing IPAs’ Organizational Routines.....	177
‘The leash’: Conceptualizing and Operationalizing IPA Control	179
Empirical Results and Discussion	181
Conclusion.....	189
References	191
Appendix to Paper D	195

Deutschsprachige Zusammenfassung der Dissertation

Im Jahr 1815 wurde in einem der vielen Anhänge zum Schlussdokument des Wiener Kongresses die Zentralkommission für die Rheinschiffahrt gegründet. Ihre Mitglieder waren die Anrainerstaaten des Flusses und ihre Verwaltung bestand aus internationalem Personal. In Anbetracht anderer Errungenschaften des Kongresses schien die Schaffung dieser technischen, regional begrenzten Organisation von geringer Bedeutung zu sein. Und doch markierte sie den Beginn eines neuen Phänomens, das die Art und Weise, wie Weltpolitik betrieben wird, nachhaltig geprägt hat. Sie stellte die Schaffung einer der ersten modernen Internationalen (zwischenstaatlichen) Organisation (IO) dar, in der Nationalstaaten Autorität an eine ständige internationale Verwaltung delegierten.

Seit 1815 und insbesondere nach dem Zweiten Weltkrieg haben Internationale Organisationen ihre Autorität deutlich vertieft (Zürn et al., 2012; Hooghe et al., 2017) und ihre Anzahl ist stark gestiegen. In der Tat sind heute hunderte von Internationalen Organisationen, darunter auch die Rheinkommission, in praktisch allen Bereichen der öffentlichen Politik und unseres Lebens aktiv. Ihre Verwaltungen, die sich aus im Allgemeinen wohl neutralen, multikulturellen, gut ausgebildeten Fachleuten zusammensetzen, haben ihre Ressourcen erweitert (Vaubel et al., 2007) und sich als einflussreiche Akteure in der globalen Weltpolitik etabliert (Eckhard und Ege, 2016).

Während dies wie eine Erfolgsgeschichte für internationale Bürokratien klingen mag, zeichnet die dominierende Perspektive in der IO-Literatur ein anderes Bild. Spätestens seit dem bahnbrechenden Buch „Delegation and Agency in International Organizations“ (Hawkins et al., 2006b) haben sich die Principal-Agent (P-A) Theorie und die Rational-Choice-Perspektive als dominantes Paradigma herauskristallisiert, welches die Basis der Forschung über internationale Delegation bildet (siehe Reinalda, 2013: 17; Tamm und Snidal, 2014). Wissenschaftler in dieser Tradition haben argumentiert, dass die Beziehungen zwischen Mitgliedsstaaten und internationalen Verwaltungen auf Zielkonflikten und Informationsasymmetrien beruhen, so dass die Frage der Kontrolle von zentraler Bedeutung war. Es wird angenommen, dass Agenten (internationale Verwaltungen), wenn sie die Möglichkeit haben, ihren Nutzen auf Kosten ihrer Auftraggeber (Mitgliedsstaaten) maximieren. Daher wurden Agenten im Allgemeinen als opportunistische Akteure betrachtet, die von egoistischem Eigeninteresse angetrieben sind (Williamson, 1985: 30). Daher wurden internationale Verwaltungen größtenteils im Zusammenhang mit Entgehen der Kontrolle der Mitgliedsstaaten (Nielson und Tierney, 2003;

Cortell und Peterson, 2006; Gould, 2006; Conceição-Heldt, 2013), dem Aufblähen eigener Ressourcen (Copelovitch, 2010) oder der Manipulation von Delegationsregeln (Hawkins und Jacoby, 2006) untersucht. Das Hauptanliegen der P-A Literatur war daher, wie man IOs kontrollieren und das opportunistische Verhalten von deren Verwaltungen einschränken kann (Pollack, 1997; Abbott und Snidal, 1998; Nielson und Tierney, 2003; Hawkins et al., 2006b; Copelovitch, 2010; Elsig, 2010; Grigorescu, 2010).

Eine solch negative Sichtweise auf IOs und ihre Verwaltungen ist umso überraschender, wenn man neuere Erkenntnisse berücksichtigt, die der angenommenen Natur des Agenten, d.h. der internationalen Verwaltung, widersprechen. Insbesondere die Erkenntnisse der aufkommenden Forschung zu internationalen Bürokratien deuten darauf hin, dass einige internationale Verwaltungen sich in erster Linie auf die Erfüllung ihrer funktionalen Zwecke konzentrieren, anstatt politische Agenden zu verfolgen (Yi-chong und Weller, 2008; Biermann und Siebenhüner, 2009; Knill et al., 2019; Bayerlein et al., 2020b). Ege (2020: 594) argumentiert daher, dass die „egozentrische Zielorientierung“ von IO-Verwaltungen nicht „als Standardmodus bürokratischen Verhaltens“ angesehen werden sollte.

Obwohl diese Erkenntnisse die Notwendigkeit unterstreichen, konventionelle P-A Annahmen zu überprüfen, werden sie von der klassischen P-A Literatur aufgrund ihres dominanten Fokus auf den Prinzipal bisher ignoriert. Bei der Untersuchung internationaler Delegation konzentrieren sich P-A Wissenschaftler primär auf die Seite des Prinzipals, d.h. der IO-Mitgliedsstaaten, da sie davon ausgehen, dass Agenten antizipieren, wie ihre Prinzipale reagieren und daher ihr Verhalten anpassen, um Sanktionen zu vermeiden. Es sind also die IO-Mitgliedsstaaten – und die Probleme zwischen ihnen –, die IOs antreiben und die es wert sind, untersucht zu werden (Koremenos et al., 2001; Lyne et al., 2006; Pollack, 2006; Tamm und Snidal, 2014: 138). In Anbetracht der widersprüchlichen Befunde zum Agentenverhalten und der Implikation, dass zentrale P-A Annahmen empirisch möglicherweise nicht zutreffen, erscheint eine solche staatszentrierte Perspektive jedoch problematisch.

Vor diesem Hintergrund ist das zentrale Ziel dieser Dissertation, die politisch-administrativen Beziehungen zwischen IO-Mitgliedsstaaten und IO-Verwaltungen neu zu untersuchen. Anstatt eine einseitige Perspektive auf IOs einzunehmen und sich entweder auf die Mitgliedsstaaten oder die IO-Bürokratien zu konzentrieren, argumentiere ich, dass es die Beziehung zwischen diesen Akteuren ist, die die Politikgestaltung und interne Funktionsweise der IOs vorantreibt. Eine IO ist weder ein bloßes Ergebnis der Interessen mächtiger Prinzipale, wie die klassische P-A Theory impliziert (siehe Koremenos et al., 2001; Drezner, 2007; Manulak, 2017), noch

sind Internationale Organisationen in ihrer Gesamtheit als Bürokratien aufzufassen (siehe Barnett und Finnemore, 1999, 2004). Es ist vielmehr die kontinuierliche, tägliche Interaktion innerhalb und zwischen den politischen und administrativen Zweigen, die der Schlüssel zu einer angemessenen Analyse der Arbeitsweise von IOs und der Bedeutung ihrer Bürokratie ist. Ausgehend von dieser Perspektive nähere ich mich in dieser Dissertation den Internationalen Organisationen, indem ich das Verständnis der Beziehungen zwischen Mitgliedsstaaten und IOs von Neuem überprüfe.

Diese Aufgabe ist wichtig, insbesondere wenn man bedenkt, dass man sich bei der Untersuchung der Delegationsbeziehungen in Internationalen Organisationen fast ausschließlich auf die P-A Theorie verlässt. Aber auch im Allgemeinen funktionieren Internationale Organisationen nur, weil sich die beiden Akteure auf eine Zusammenarbeit einigen: Die Mitgliedsstaaten erklären sich bereit, Befugnisse an eine internationale Verwaltung zu delegieren, während sich die internationale Verwaltung bereit erklärt, diese Befugnisse unter Anleitung der Mitgliedsstaaten auszuüben. Wenn die Mitgliedsstaaten beschließen, ihre Autorität zurückzuziehen, verliert die Verwaltung ihre Legitimität und wird obsolet. Wenn die Verwaltung jedoch aufhört, delegierte Aufgaben umzusetzen, verlieren die Mitgliedsstaaten ihre Delegationsvorteile. Letzten Endes ist davon auszugehen, dass wenn die Beziehungen zwischen den Mitgliedsstaaten und der internationalen Verwaltung zusammenbrechen, dann scheitert auch die gesamte Internationale Organisation. Und daher ist es diese politisch-administrative Interaktion, die das Rückgrat aller IOs und womöglich der globalen Politikgestaltung jenseits der Nationalstaaten bildet. Es ist daher sowohl für Praktiker als auch für IO-Wissenschaftler von größter Wichtigkeit, diese Beziehung – ihre Dynamik, Dimensionen, Arten und Bestandteile – besser zu verstehen. Es ist daher das Ziel dieser Arbeit, einen Beitrag zur IO-Forschung zu leisten, indem sie eine ausgewogene, systematische, theoretische und empirische Darstellung der Interaktion von Mitgliedsstaaten mit IO-Administrationen bietet.

Zu diesem Zweck teste und modifiziere ich die konventionellen Annahmen der P-A Theorie, korrigiere ihren verzerrten Fokus auf den Prinzipal und ziehe Erkenntnisse über den Agenten aus der aufkommenden Literatur zur internationalen Bürokratien heran. Insbesondere werden mindestens vier Forschungslücken bei der Untersuchung der Beziehungen zwischen Mitgliedsstaaten und IO-Verwaltungen angesprochen. Erstens befasst sich die Dissertation mit der mangelnden definitorischen Klarheit darüber, wer genau der Prinzipal und der Agent in IOs ist. Sie liefert spezifische Definitionen der beteiligten Akteure und bietet eine mehrdimensionale IO-Delegationsperspektive, die erklärt, welche Dimension am besten zu

welchen Forschungsfragen passt (Jankauskas, 2021). Zweitens problematisiert die Dissertation die bestehende „principal-biased“ Perspektive in Delegationsstudien und die mangelnde Beachtung des tatsächlichen Verhaltens des Agenten. Sie zeigt die Grenzen einer staatszentrierten Sichtweise auf IOs auf und betont die Notwendigkeit, die Charakteristika sowohl der Mitgliedsstaaten als auch der internationalen Verwaltungen in Relation zueinander zu betrachten (Eckhard und Jankauskas, 2019, 2020; Heinkelmann-Wild und Jankauskas, 2019, 2020; Jankauskas, 2019, 2021; Jankauskas und Eckhard, 2019; Jankauskas et al., 2021). Drittens adressiert die Dissertation den Mangel an Wissen über grundlegende P-A Annahmen im IO-Kontext und deren mögliche Variation. Sie testet diese Annahmen empirisch und liefert entsprechend theoretische und konzeptionelle Modifikationen (Jankauskas, 2019, 2021; Eckhard und Jankauskas, 2020; Heinkelmann-Wild und Jankauskas, 2020; Jankauskas et al., 2021). Schließlich erweitert die Dissertation die enge empirische Landschaft der IO-Studien, indem sie eine Vielzahl von IOs untersucht und vergleichende Forschungsdesigns einsetzt.

Als kumulative Arbeit besteht die Dissertation aus vier Kernbeiträgen, die dem Rahmenpapier beigelegt sind und dort ausführlich diskutiert werden (Eckhard und Jankauskas, 2020; Heinkelmann-Wild und Jankauskas, 2020; Jankauskas, 2021; Jankauskas et al., 2021). Darüber hinaus ergänzen sechs „non-core“ Beiträge das Forschungsthema, sind aber nicht dem Rahmenpapier beigelegt (Eckhard und Jankauskas, 2019; Heinkelmann-Wild und Jankauskas, 2019; Jankauskas, 2019; Jankauskas und Eckhard, 2019; Weiss und Jankauskas, 2019; Jankauskas et al., in press). Stattdessen wird auf die Ergebnisse dieser „non-core“ Artikel im Rahmenpapier sowie in den Kernbeiträgen verwiesen. Im Rahmenpapier gebe ich einen systematischen Überblick über die wichtigsten Ergebnisse und Beiträge der Dissertation, diskutiere den aktuellen Stand der Literatur und identifiziere Forschungslücken in der bestehenden Forschung, die in den beigelegten Kernartikeln bearbeitet werden. Außerdem fasst das Rahmenpapier die Ergebnisse der einzelnen Artikel kurz zusammen und diskutiert deren empirische Strategien. Der Hauptfokus liegt auf der Erläuterung der internen Kohärenz der vier Kernartikel und der Erläuterung ihrer Gesamtergebnisse und Beiträge. Abschließend skizziert das Framework Paper Wege für zukünftige Forschung.

Basierend auf den Ergebnissen der Artikel ist die Kernaussage der Dissertation, dass die Beziehungen zwischen Mitgliedsstaaten und internationalen Verwaltungen jenseits der konventionellen P-A Annahmen und der formalen Beschränkungen des Delegationsvertrags stark variieren; sie sind nicht durch hohe Zieldivergenz, Opportunismus und Anfechtung geprägt, wie es die P-A Theorie vorschreibt. Genau wie in unserem Privatleben ist nicht jede Beziehung durch betrügerische Anliegen und das Bestreben, sich gegenseitig zu kontrollieren,

gekennzeichnet. Prinzipale und Agenten können den Modus Operandi ihrer Interaktion steuern und zu einem grundlegend anderen Beziehungstyp übergehen, der auf Vertrauen statt auf Misstrauen basiert und von weicher statt harter Kontrolle sowie von informellen statt formellen Vereinbarungen geprägt ist (Jankauskas, 2019, 2021). Ich nenne einen solchen alternativen, idealtypischen Typus der Delegationsbeziehung „Stewardship“ und theoretisiere über dessen Existenz und Mechanismen in Internationalen Organisationen (ebd.).

In Stewardship-Beziehungen vertrauen die Mitgliedsstaaten der IO-Verwaltung, was eine weichere, informellere Ausübung der Kontrolle ermöglicht. Hierbei versucht die Verwaltung die konkurrierenden Präferenzen der Mitgliedsstaaten auszugleichen und Lösungen im besten Interesse aller Mitglieder zu finden. Sie agiert daher eher als pro-organisatorischer Steward und nicht als eigennütziger, opportunistischer Agent. Dies führt zur Vertrauenswürdigkeit des Managements und ermöglicht damit eine weniger anspruchsvolle, eher informelle, weiche Ausübung der Kontrolle. Solche Stewardship-Beziehungen sind also eher durch Kooperation als durch ständige Anfechtung gekennzeichnet.

Hervorzuheben ist, dass die Dissertation auf der Grundlage umfangreicher Interviewdaten sowohl mit Mitgliedsstaaten als auch mit IO-Verwaltungen sowie einer primären Dokumentenanalyse die Existenz von Stewardship empirisch nachweist – und zwar in einer der wichtigen Organisationen des UN-Systems (World Food Programme) (ebd.). Dies widerspricht P-A-Wissenschaftlern, die argumentieren, dass vertrauenswürdige Agenten „ungewöhnlich“ sowie empirisch selten zu beobachten sind (vgl. Hawkins und Jacoby, 2006: 201).

Außerdem zeigt die Dissertation, dass selbst in IOs, in denen ähnliche institutionelle Strukturen und formale Kontrollmechanismen existieren, deren tatsächliche Nutzung durch die Mitgliedsstaaten stark variiert. Dies hängt dabei vom Ausmaß des Vertrauens sowie davon, ob opportunistische oder vertrauenswürdige Delegationsbeziehungen vorherrschen (Jankauskas, 2021), ab. Dies unterstreicht die informelle Seite der Arbeitsweise von IOs (siehe Stone, 2011) und ruft die Wissenschaft dazu auf, sich mehr auf die Vertrauens-Kontroll-Dynamik zwischen den Prinzipalen und ihren Agenten sowie auf die de facto statt de jure Merkmale der IO zu konzentrieren.

In diesem Zusammenhang lernen wir, dass konventionelle P-A-Annahmen nicht als gegeben angenommen, sondern eher als Variablen behandelt werden sollten. Tatsächlich zeige ich, dass nicht alle IO-Verwaltungen eigennützige Agenten sind, wie die P-A Theorie impliziert (Jankauskas, 2019, 2021; Jankauskas et al., 2021). Ihr Verhalten ist weder nur ein Spiegelbild des formalen Delegationsvertrags (Eckhard und Jankauskas, 2019, 2020) noch ist es ein bloßes

Nebenprodukt der Kontrolle durch die Mitgliedsstaaten, bei der jeglicher Ermessensspielraum für die Verfolgung der eigenen eigennützigen Agenda genutzt wird (Jankauskas, 2021; Jankauskas et al., 2021). Einige Verwaltungen sind in der Tat bestrebt, die verschiedenen Interessen der Mitgliedsstaaten zu managen und ihrem Mandat treu zu bleiben (Heinkelmann-Wild und Jankauskas, 2019, 2020; Jankauskas, 2021; Jankauskas et al., 2021). Während dies die Forschung zur funktionalen Orientierung von internationalen Verwaltungen bestätigt (Yi-chong und Weller, 2008; Biermann und Siebenhüner, 2009; Eckhard und Ege, 2016; Knill et al., 2016; Knill et al., 2019; Bayerlein et al., 2020b; Ege, 2020), behaupte ich nicht, dass alle internationalen Verwaltungen vertrauenswürdige Stewards sind. Aber nur wenn wir die variierende Natur des Agenten berücksichtigen – und diese Variation in Bezug auf den Prinzipal theoretisieren – werden wir in der Lage sein, eine umfassendere Darstellung der Delegationsdynamik auf internationaler Ebene zu erhalten.

Im Wesentlichen zeigt die Dissertation, dass die P-A Theorie für ein umfassendes Verständnis internationaler Delegationsbeziehungen entscheidend ist, aber für sich allein nicht ausreicht. Die eingeführte Stewardship-Theorie ermöglicht es nun, übersehene Stewardship-Dynamiken in rationalistische Delegationsanalysen einzubeziehen, die staatszentrierte Sichtweise aufzugeben und dadurch eine nuanciertere und umfassendere Perspektive auf die interne Funktionsweise von IOs zu bieten.

List of Acronyms and Abbreviations

BRICS – Brazil, Russia, India, China, and South Africa
csQCA – crisp-set Qualitative Comparative Analysis
EU – European Union
FAO – Food and Agriculture Organization
IAEA – International Atomic Energy Agency
IGO – International Governmental Organization
ILO – International Labour Organization
IMF – International Monetary Fund
INGOs – International Non-Governmental Organization
IO – International Organization
IOM – International Organization for Migration
IPA – International Public Administration
IR – International Relations
MDSD – Most Different System Design
MS – Member States
MSSD – Most Similar System Design
NATO – North Atlantic Treaty Organization
NGO – Non-Governmental Organization
OHCHR – Office of the United Nations High Commissioner for Human Rights
P-A – Principal-Agent
SFC – Structured, Focused Comparison
UNDP – United Nations Development Programme
UNEP – United Nations Environment Programme
UNESCO – United Nations Educational, Scientific and Cultural Organization
UNFCCC – United Nations Framework Convention on Climate Change
UNFPA – United Nations Population Fund
UNHCR – United Nations High Commissioner for Refugees
UNICEF – United Nations Children's Fund
UNRWA – United Nations Relief and Works Agency for Palestine Refugees in the Near East
UNSG – Secretary-General of the United Nation
UPU – Universal Postal Union
WB – World Bank
WFP – World Food Programme
WHO – World Health Organization
WTO – World Trade Organization

Acknowledgements

Writing a PhD thesis was quite a journey – with its successes and achievements, but also hardships and disappointments. It has challenged me, sometimes more than I thought it will, but it also allowed me to grow, to learn, and to explore the unique world of academia as well as the fascinating realm of international organizations. Taking this journey, moving forward, and, importantly, bringing it to an end would not have been possible without the support of the many people whom I sincerely thank.

I express my deepest appreciation to my supervisors Steffen Eckhard and Christoph Knill. I would not have started this endeavour without Steffen's trust in me when he asked me to join his research project back in 2017. I would not have finished it without his continuous support and encouragement. I have highly valued the confidence that Steffen put in me; I took it as a commitment to fulfil and preserve. As a supervisor and mentor, he was extremely generous with his time and knowledge, always offering insightful feedback that allowed me to sharpen my thinking and writing. Steffen pushed me when necessary and knew when I should slow down risking biting off more than I can chew. I learned a lot from him – both professionally and personally – and I sincerely thank for his openness and thoughtfulness throughout these years. Also, I am deeply indebted to Christoph who welcomed me with open arms at his Chair and offered me guidance and advice anytime I needed it. Christoph has continued to inspire me with his papers, books, and projects, his relentless drive for new insights and discoveries. I thank Christoph for his generosity and appreciate his time he dedicated to my supervision. In the language of my research, I have been fortunate with my supervisors to operate under stewardship – enjoying mutual trust, soft control and informal interactions, and thereby cooperative and productive working environment – something that I am very grateful for.

Due to the cumulative nature of my dissertation, I have been working on a collection of articles, which means that I had to write papers individually but also worked cooperatively and thereby learned with and from a number of co-authors: my supervisors Steffen Eckhard and Christoph Knill as well as my colleagues Louisa Bayerlein, Jörn Ege, Tim Heinkelmann-Wild, and Moritz Weiss. I sincerely thank them for their kind cooperation and knowledge.

For my research, insights from interviews were essential. I was lucky enough to be able to conduct interviews at the headquarters of major IOs – from the European Commission in Brussels and UN organizations in Geneva, Paris, Rome, and New York to the IMF and World Bank in Washington. I cherish these encounters, both the lessons I learned and the experiences I had: from chatting with Christine Lagarde in an elevator at the IMF to sitting in the office of the UN Under-Secretary-General in New York or talking over an hour with the WHO's Chief de Cabinet – something that would be impossible today, over a year later, as the global COVID-19 pandemic began. Many people helped me with this, more than a hundred IO officials, member state representatives, and experts dedicated their time to explain their work and share their insights, sometimes for hours. I sincerely thank them for this. I also thank Silke Heuser, Timothy Trollope, and Arnoldas Pranckevičius who generously shared their contacts and Leon Apostolopoulos, Marisa Brecht, and Alexa Lenz who assisted in arranging the interviews.

Beside these research trips and my short stays at the University of Konstanz (which location at the Bodensee lakeside I admired), I spent most of my time at the Geschwister-Scholl Institute

of Political Science in Munich. My semi-basement office was located in front of the staff entrance and so I would gladly use my window well to catch up with my GSI colleagues. Much more than that, they offered me the very needed peer support. I acknowledge my colleagues at the Chair: Christian Adam, Louisa Bayerlein, Eva-Maria Euchner, Xavier Fernández-i-Marín, Steffen Hurka, Olivia Mettang, Christina Steinbacher, Yves Steinebach, Dionys Zink. I thank them for their encouragement and feedback on my research, our colloquium sessions and long dinners discussing research ideas and problems. I am also grateful to Annette Ohlenhard for her excellent administrative support. For their mentoring and advice, I sincerely thank Moritz Weiss and Ronny Patz who I first got to know as inspiring teachers during my studies and with whom I was lucky to work during my PhD. Furthermore, from having an afternoon swim at the Eisbach in summer to discussing politics at the Chinese Tower in winter, many of the GSI colleagues became dear friends who I highly appreciate: Felix Biermann, Stefan Jagdhuber, Lisa Kriegmair, Stefan Matern, Eva Ziegler. I especially thank Tim Heinkelmann-Wild who as a friend helped me on so many occasions and as a colleague provided me with motivation and invaluable support by reading and commenting almost all of my papers. Also, many thanks to Alexa Lenz – I could not have asked for a better colleague and friend to share the office.

Similarly, I have greatly benefited from being part of the DFG Research Group “International Public Administration” and working in its two research projects on evaluation and administrative styles. I thank all the group and project colleagues for stimulating and productive research experiences throughout the years as well as the many participants of various workshops, conferences, colloquiums, and meetings for sharing their reflections on my papers. I thank a number of anonymous reviewers who read my work with great caution and provided tough but crucial feedback on my arguments. I also thank DAAD (German Academic Exchange Service) which provided financial support during both my Bachelor and Master studies, allowing my education in Germany and thereby setting the ground for my PhD.

Thanks also to my friends for having my back and for putting up with me being sat in the office for long hours and often weekends. I thank my cousin Augustas who provided with sometimes needed distraction from work; I thank my best friends from Bachelor studies who always inspired me with their enthusiasm and positivity and who supported me ever since: Timothy Carter, Patrick Gutekunst, Aljoscha Nau, Lorenz Oelschlaeger, Stefan Rottler, Timothy Trollope; and I thank my friends from Master studies – Urs Hagen, Tim Heinkelmann-Wild, and Jasper Roskilly – with whom I shared both challenging times and celebrated the good ones.

Finally, and probably most importantly, I truly thank my family – my parents Gailutė and Kęstutis as well as my sister Sigita and her husband Vilius with my (now) six-year-old nephew Dominykas. I thank them for their understanding and unconditional support, as I have spent too many years abroad, not being able to spend more time with them as I wished. They have been an indispensable source of advice, keeping me going and never doubting me. I would not have made it without their love and encouragement. I especially thank my father who has showed so much enthusiasm and interest in my research, shared so many brilliant ideas, and always taught me to never stop learning.

*Vytautas Jankauskas
Munich, February 2021*

Framework Paper

Vytautas Jankauskas

Abstract:

In this Framework Paper, I provide a systematic overview of the (cumulative) dissertation, its main findings and contributions. The Framework Paper discusses the state of the art and identifies research gaps in existing research that are addressed in the attached core articles. It briefly summarizes each paper's findings and discusses employed empirical strategies. Its main focus lies on explaining the internal coherence of the four core articles and elucidating their overall findings and contributions. Finally, the Framework Paper outlines avenues for future research.

1 Introduction

In 1815, in one of the many appendices to the Final Document of the Congress of Vienna, the Central Commission for the Navigation of the Rhine was created. Its members were states bordering the river and its administration was composed of international staff with headquarters in Mainz. Considering other Congress achievements, the creation of this technical, regionally bounded organization appeared to be of a minor significance. Yet, it marked the beginning of a new phenomenon that – as we now know – has fundamentally shaped how the world politics are being conducted. It constituted the creation of a first modern international (governmental) organization (IO) in which nation states delegated authority to a permanent international administration.

Over time, such institutionalized relationships between member states (MS) and international bureaucracies – or international public administrations (IPAs) – increasingly transformed *ad hoc* multilateral conferences into IOs (Claude Jr., 1956: 194). This has constituted “one of the most significant developments in the evolution of the international legal order over the last hundred years” (Collins and White, 2013: 119). Indeed, since 1815, and especially after the Second World War, international organizations have both significantly deepened their authority (Zürn et al., 2012; Hooghe et al., 2017) and rapidly grew in numbers (Figure 1). Today, hundreds of international governmental organizations, including the Rhine Commission, are actively engaged in virtually all aspects of public policy and our lives. Their *administrations*, comprised of arguably neutral, multicultural, well-educated professionals, have expanded their resources (Vaubel et al., 2007) and established themselves as influential actors in global governance (Eckhard and Ege, 2016), helping member states solving issues that require international cooperation. Consider, for instance, the role that the secretariat of the World Trade Organization (WTO) plays in facilitating international trade or the relevance of the World Health Organization and its administration in the global coordination during the COVID-19 pandemic, not to mention the immense importance of such IPAs like the European Commission or the secretariat of the United Nations (UN).

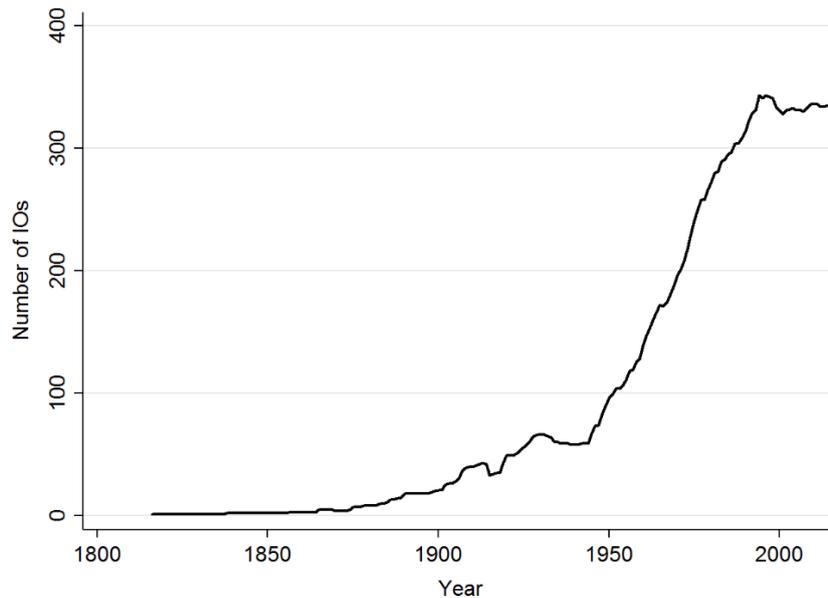


Figure 1: Number of international (governmental) organizations, 1815-2014. Compiled based on data provided by: Pevehouse et al., 2020.

This success story of IPAs addressing global challenges seems at odds with the dominant perspective in IO scholarship. Principal-Agent (P-A) draws a dark picture of IOs as opportunistic agents that seek to betray the member states. At least since the seminal book “Delegation and Agency in International Organizations” (Hawkins et al., 2006b), Principal-Agent (P-A), the rational choice perspective has appeared as a dominant paradigm underlying research on international delegation (see Reinalda, 2013: 17; Tamm and Snidal, 2014). Scholars in this tradition have argued that MS-IPA relationships are based on goal conflict and information asymmetries, so the issue of control has been central. In short, if given an opportunity, agents (IPAs) are thought to maximize their utility at the expense of their principals (member states). Therefore, agents have been generally seen as ‘rogue’ opportunistic actors, driven by ‘self-interest seeking with guile’ (Williamson, 1985: 30). IPAs, too, were largely studied in the context of slacking and escaping member states’ control (Nielson and Tierney, 2003; Cortell and Peterson, 2006; Gould, 2006; Conceição-Heldt, 2013), inflating own resources (Copelovitch, 2010), or manipulating delegation rules (Hawkins and Jacoby, 2006). The key concern for P-A scholars has thus been how to control IOs and constrain IPAs’ opportunistic behaviour (Pollack, 1997; Abbott and Snidal, 1998; Nielson and Tierney, 2003; Hawkins et al., 2006b; Copelovitch, 2010; Elsig, 2010; Grigorescu, 2010).

Such a negative view on IOs and their administrations is even more surprising considering recent findings that contradict the assumed opportunistic, self-seeking nature of the agent, i.e. the IPA. Nascent research on international bureaucracies found that some IPAs focus primarily on fulfilling their functional purposes rather than pursuing political agendas (Yi-chong and Weller, 2008; Biermann and Siebenhüner, 2009; Knill et al., 2019; Bayerlein et al., 2020b). This implies, according to Ege (2020: 594), that “self-centered goal orientation” of IO administrations should not be seen “as the default mode of bureaucratic behavior”.

Although these findings highlight the need to re-examine the conventional assumptions of the dominant P-A perspective, they remained largely ignored due to the principal-bias in the classical P-A literature. When studying international delegation, P-A scholars focus primarily on the principal side, i.e. IO member states, as they assume that agents anticipate how their principals react and therefore adjust their behaviour to avoid punishments. It is thus IO member states – and problems among them – that drive IOs and are worth studying (Koremenos et al., 2001; Lyne et al., 2006; Pollack, 2006; Tamm and Snidal, 2014: 138). However, considering the contradictory findings about the agent’s behaviour and the implication that key P-A assumptions may not hold empirically, such a state-centric perspective appears problematic.

Instead of taking a one-sided perspective on IOs and focusing on *either* member states *or* IO bureaucracies, this dissertation argues that it is the *relationship* between these actors that drives IOs’ policy-making and internal functioning. IOs are no epiphenomenon, shaped by the interests of powerful principals (see Koremenos et al., 2001; Drezner, 2007; Manulak, 2017). IOs are also not just large bureaucracies, insulated from their member states (see Barnett and Finnemore, 1999, 2004). IOs are rather the sum of continuous, daily interactions both within and between the political and administrative branches. This *relational* understanding is key to a proper analysis of how IOs work and how their bureaucracies matter. It is from this angle that I approach IOs in this dissertation, revisiting our understanding of MS-IPA relationships.

This task is important, especially considering the almost exclusive reliance on P-A theory when studying delegation relationships in IOs. But also, in more general terms, IOs function only because two collective actors agree to cooperate: member states agree to delegate authority to an international administration, whereas the international administration agrees to act upon this authority under member states’ guidance. If member states decide to rescind their authority, the IPA loses its legitimacy and becomes obsolete. Yet, if the IPA stops implementing delegated tasks, member states lose their delegation benefits. After all, if MS-IPA relationships collapse, so does the overall international organization. Therefore, it is this political-administrative

interaction that constitutes the backbone of all IOs and, some may argue, of the global policy making beyond nation state. Looking at the *global* nature of key challenges that humanity will face in the 21st century, it is likely that the role of IOs will only increase as nation states will be unable to cope with such issues like climate change, technological disruption, or nuclear war individually. Their relationships with *international* administrations will thus intensify. It is hence of utmost importance for both IO policymakers and IO scholars *to better understand this relationship – its dynamics, dimensions, types, and constituents*. It is the ambition of this thesis to contribute to IO scholarship by offering a balanced, systematic, theoretical and empirical account of MS-IPA interaction.

To this end, I both *test and modify conventional P-A theory assumptions*, correcting its principal bias and drawing on insights on the agent from nascent international bureaucracy literature. In particular, at least four *research gaps* in studying MS-IPA relationships are addressed. First, the dissertation deals with the lack of definitional clarity as to who exactly the principal and the agent in IOs are. It provides specific definitions of involved actors and offers a multidimensional IO delegation perspective, explaining which dimension suits which research questions (Jankauskas, 2021). Second, the dissertation problematizes the existing principal-biased perspective in delegation studies and the lack of attention to the actual behaviour of the agent. It demonstrates the limits of a state-centric view on IOs, highlighting the need to consider the characteristics of both member states and IPAs in relation to each other (Eckhard and Jankauskas, 2020; Heinkelmann-Wild and Jankauskas, 2020; Jankauskas, 2021; Jankauskas et al., 2021; see also Eckhard and Jankauskas, 2019; Heinkelmann-Wild and Jankauskas, 2019; Jankauskas, 2019; Jankauskas and Eckhard, 2019). Third, the dissertation puts so far unquestioned, basic assumptions of P-A theory to an empirical test and develops theoretical and conceptual modifications thereof (Eckhard and Jankauskas, 2020; Heinkelmann-Wild and Jankauskas, 2020; Jankauskas, 2021; Jankauskas et al., 2021; see also Jankauskas, 2019). Finally, while existing knowledge about delegation to IOs rests on a rather small number of IOs and single case studies of the very same IOs, the dissertation broadens the narrow empirical landscape in IO studies by studying a diverse set of IOs and by employing a *comparative* research design.

As a cumulative thesis, the dissertation comprises *four core research articles* which are attached to this Framework Paper and discussed in detail in its remainder (Eckhard and Jankauskas, 2020; Heinkelmann-Wild and Jankauskas, 2020; Jankauskas, 2021; Jankauskas et al., 2021). In addition, *six non-core contributions* add to the research topic, yet are not attached to the Framework Paper (Eckhard and Jankauskas, 2019; Heinkelmann-Wild and Jankauskas, 2019;

Jankauskas, 2019; Jankauskas and Eckhard, 2019; Weiss and Jankauskas, 2019; Jankauskas et al., in press). An overview list of both core and non-core contributions is provided in the Appendix I and II to this Framework Paper, including a list of abstracts of the non-core articles.

While each core paper of this dissertation yields various contributions on its own, they are connected in their objective to overcome the limitations of P-A theory and thereby advance our understanding of MS-IPA relationship. Overall, I find that MS-IPA relationships vary greatly beyond conventional P-A assumptions and formal constraints of the delegation contract; they are not predefined by high goal divergence, opportunism, and conflict as P-A theory prescribes. Just as in our lives, not every relationship is characterized by cheating concerns and efforts to control each other. Principals and agents can actually interact in a way that is fundamentally different from P-A theorizing – in a way which is based on trust rather than distrust, shaped by soft rather than hard control, and informed by informal rather than formal prescriptions (Jankauskas, 2019, 2021). I dub this relationship *stewardship* and theorize on its mechanisms and possibility in international organizations (ibid.). Importantly, drawing on extensive interview data with both member states and IPAs as well as primary document analysis, I show the existence of stewardship empirically – in one of the major UN system organizations (World Food Programme) (ibid.). This contradicts P-A scholars who argue that trustworthy agents are “unusual”, absent and thus not worthy to theorize on (see Hawkins and Jacoby, 2006: 201).

Conventional P-A assumptions should thus not be taken as granted but treated as variables. In fact, I demonstrate that not all IPAs are ‘rogue’, self-serving agents as P-A theory implies (Jankauskas, 2019, 2021; Jankauskas et al., 2021). Their behaviour is neither only a reflection of the formal delegation contract (Eckhard and Jankauskas, 2019, 2020) nor is it a mere by-product of member states’ control, where discretion is *per se* used for pursuing own self-serving agenda (Jankauskas, 2021; Jankauskas et al., 2021). Some administrations are indeed doing their best to manage diverse member states’ interests and stay true to their mandate (Heinkelmann-Wild and Jankauskas, 2019, 2020; Jankauskas, 2021; Jankauskas et al., 2021). While this corroborates research on IPAs’ functional orientation (Yi-chong and Weller, 2008; Biermann and Siebenhüner, 2009; Eckhard and Ege, 2016; Knill et al., 2016; Knill et al., 2019; Bayerlein et al., 2020b; Ege, 2020), I do not claim that all IPAs are trustworthy stewards. But only if we consider the varying nature of the agent – and theorize this variation in relation to the principal – will we be able to get a fuller account of delegation dynamics at the international level.

Overall, this dissertation demonstrates that P-A theory is crucial for a comprehensive understanding of international delegation relationships but is insufficient by itself. The introduced stewardship theory allows incorporating overlooked trust-based dynamics into rationalist delegation frameworks, abandoning principal-biased, state-centric view and thereby providing a more nuanced and comprehensive perspective on IOs' internal functioning.

The remainder of this Framework Paper is structured as follows: first, Section 2 provides a systematic discussion on the state of the art (2.1) and identifies research gaps related to P-A theory and the analysis of MS-IPA relationships (2.2). Then, Section 3 spells out the common denominators of the core articles (3.1), summarizes their main findings and paper-specific contributions (3.2), and provides a summary of overall lessons learned (3.3). Finally, Section 4 concludes by highlighting stewardship theory as a new perspective of advancing our understanding of the MS-IPA relationship in IOs and lays a firm foundation for future research agenda.

2 P-A theory and its limitations in studying MS-IPA relationships

Almost all we know about delegation in IOs, i.e. member states' relationships with IOs in general, and IPAs in specific, essentially stems from studies informed by Principal-Agent theorizing. This section explains how and why P-A theory emerged as a key theoretical perspective underlying research on IO delegation (2.1), thereby clarifying the dissertation's particular attention to this theory. Also, drawing on recent developments in IO research, Section 2.2 identifies and describes in detail four research gaps related to P-A approaches and their understanding of MS-IPA relationship.

2.1 The dominance of P-A theory

Principal-Agent theory has stood out as a fitting approach to study international organizations and specifically their delegation relationships. For one, its rationalist perspective seems to have been widely accepted by IO scholars. P-A theory follows rational choice tradition which invokes the assumption of rationality and strategic action as an explanation of actors' behaviour. Differently put, "given their beliefs about how the world works and their ability to affect it, actors choose their actions in order to best attain their goals" (Tamm and Snidal, 2014: 133). In IO research, too, scholars seem to have agreed that involved parties – both member states and

IO administrations – are indeed rational actors (or at least rationally bounded) (see Tamm and Snidal, 2014).

In addition, IOs constitute typical cases for the application of P-A models (Hawkins et al., 2006b). P-A theory is based on the idea of contractual relationships when the principal delegates authority to an agent. The resulting delegation relationship is characterized by power and information asymmetries: while the ultimate power remains by the principal who can rescind its authority, the expert knowledge favours the agent. As for IOs, delegation of authority lies at the heart of every international organization since member states, as political decision-makers and the ultimate holders of authority, i.e. principals, enable IOs or IO bureaucracies specifically to act on their behalf as policy advisors, implementers, or assistants in intergovernmental bargaining, i.e. as agents. Member states do not have to execute these tasks by themselves and can anytime rescind their authority. Yet, they also know less about the details of these tasks and their implementation as compared to IPAs. Hence, member states-IPA relationships mirror typical structure of principal-agent framework.

Indeed, P-A models have helped scholars to study a variety of aspects related to delegation in IOs. Initially, P-A theory focused on why member states delegate authority to supranational institutions. This question was less interesting for those studying domestic delegation relationships as the fact of delegation seemed to be self-explaining (e.g. McCubbins and Schwartz, 1984; Kiewiet and McCubbins, 1991). Yet, delegation at the international level was perceived as a phenomenon since powerful nation states were thought to be capable of achieving most of the tasks unilaterally and other forms of cooperation were available (see Keohane, 1984). The literature thus examined and identified several benefits of delegating authority to IOs: dispute resolution, credibility enhancement, decision-making facilitation, specialization and policy management, ‘lock-in’ effects, etc. (Abbott and Snidal, 1998; Tallberg, 2002; Pollack, 2003; Hawkins et al., 2006a). More broadly, these factors, embedded into P-A theory, provided an explanation for the increasing internationalization which, as scholars of global governance argue, has shifted the notion of the state “from virtual monopolist to manager of political authority” (Genschel and Zangl, 2014: 337; see also Zürn et al., 2012; Weiss and Jankauskas, 2019).

Subsequently, P-A theory has been widely applied to study how IOs function once delegation took place. In this regard, the key concern for IO scholars has been how to control IO administrations and ensure their faithfulness to member states. According to P-A theory, agents have own preferences which may diverge from those of the principals. Principals therefore have

to control their agents by applying different control mechanisms. The question of control and how member states can avoid agency slack, i.e. “independent action by an agent that is undesired by the principal” (Hawkins et al., 2006a: 8), has therefore been central in the recent decades (Nielson and Tierney, 2003; Hawkins et al., 2006b; Vaubel, 2006; Copelovitch, 2010; Elsig, 2010; Grigorescu, 2010; Stone, 2011; Urpelainen, 2012; Conceição-Heldt, 2013; Dijkstra, 2015; Heldt, 2017; Parížek, 2017).

In sum, it is the simple and straightforward manner in which P-A theory models hierarchical relationships that led it to evolve as the main analytical tool to study delegation in international organizations. From general studies on why states delegate authority to IOs to more specific analyses of *how* member states control IOs after delegation took place, P-A theory has positioned itself at the heart of contemporary IO delegation studies.

2.2 *The limitations of P-A theory*

Albeit its great achievements in increasing our understanding of delegation to IOs, at least four limitations remain when P-A theory is applied to study MS-IPA relationships in IOs: 1) the lack of clarity on the definitions of actors (the principal and the agent); 2) the lack of attention to the actual agent (IPA) behaviour (principal-bias); 3) the lack of knowledge about variation in basic P-A theory assumptions; and 4) the narrow empirical focus on few, prominent IOs.

First, P-A studies of delegation to IOs often *lack clear definitions of the principal and the agent*. There is still a lot of confusion as to who exactly the principal and the agent are. As Ege and Bauer (2013: 143) highlighted, “the major shortcoming in the literature is the failure to properly define the bearer of agency within IOs”.

Regarding the principal, scholars usually refer to member states (see Hawkins et al., 2006b). Yet, it often remains unclear what entities they actually mean. An important difference lies between member state governments, sitting in IOs’ *general* governing bodies (e.g. the World Health Assembly), and their representatives (usually ambassadors) in IOs’ *executive* governing bodies (e.g. the WHO Executive Board). General governing bodies meet rarely in plenary sessions (e.g. once a year) and make ‘macro-level’ decisions such as approving an IO’s budget. Executive governing bodies, by contrast, meet frequently (e.g. monthly) and make ‘micro-level’ decisions such as approving IOs’ programmes or reviewing oversight reports. Elsig (2010) calls these bodies as “proximate principals” as they oversee IO administrations but report to general governing bodies. Depending on the research question, it seems necessary to draw a clear

distinction. For instance, if one asks why states delegate authority to IOs, it is member state governments who make such decisions. Yet, if one asks how states exercise daily control over IO administrations, the scope of analysis should be narrowed down to executive governing bodies.

Regarding the agent, scholars usually refer to IOs in general, thereby conflating their political and administrative branches (see Hawkins et al., 2006b). International organization as such includes both IO staff and the collective of member states represented by IOs' governing bodies. While governing bodies oversee the IO administration, it is the latter that executes delegated tasks and thus is oftentimes the actual agent in the framework. Noteworthy, international bureaucracies were neglected in the early years of IO studies, at that time dominated by International Relations scholars. However, IO research in the recent decades has experienced the so called "public administrative turn" (Trondal et al., 2010: 3) and IPAs are now recognized as distinct actors in their own right (Liese and Weinlich, 2006; Yi-chong and Weller, 2008; Biermann and Siebenhüner, 2009; Tallberg, 2010; Ege and Bauer, 2013; Schroeder, 2014; Weinlich, 2014; Bauer and Ege, 2016; Eckhard and Ege, 2016; Jörgens et al., 2016; Bauer et al., 2017; Knill et al., 2019; Bayerlein et al., 2020b).¹ The lack of clarity regarding the agent may create confusion: For instance, if IOs are blamed for 'agency slack', is it the administration that bears the responsibility of an undesired performance or is the problem related to member states in (executive) governing bodies who fail to agree on taking appropriate action?

In short, there is a clear need of a more precise differentiation between IO member states and IO administrations, i.e. the political and the administrative as well as the principal and the agent (see also Yi-chong and Weller, 2008; Ege and Bauer, 2013).

Second, there still exists *a principal bias* as P-A studies on delegation to IOs predominantly focus on the principal side when explaining IOs' internal functioning. The dominance of a state-centric or principal-biased view relates to the rationalist thinking that agents anticipate how their principals react and therefore adjust their behaviour to avoid punishments. As Gutner and Thompson (2010: 237) write, "IOs in this tradition are member-driven. Although they may have some autonomy, independent behavior is either consciously intended by their state principals or carefully constrained" (see also Abbott and Snidal, 1998). Hence, the intuitive assumption is that it is more important to study member states and their preferences rather than the actual

¹ Few P-A studies explicitly refer to IPAs as agents, see Graham (2014); Hawkins and Jacoby (2006); Johnson (2013); Johnson and Urpelainen (2014).

behaviour of IPAs (Koremenos et al., 2001; Lyne et al., 2006; Pollack, 2006; Tamm and Snidal, 2014: 138).

Accordingly, a mounting body of studies focused on member states and the (in-)effectiveness of various control mechanisms they employ over IO administrations. Vaubel (2006) argued that IOs are in general hard to control due to long delegation chains, severe information asymmetries, and flawed incentives. Nielson and Tierney (2003) applied P-A theory to explain changes in control mechanisms at the World Bank, highlighting the complexity among member states (issues of collective and multiple principals) and its consequences for constraining the Bank's administration. Similarly, Copelovitch (2010: 49) argued that preference heterogeneity among member states provides the opportunity for an IPA (in his case, the IMF staff) to "exploit 'agency slack' and increase its autonomy" (see also Cortell and Peterson, 2006). Urpelainen (2012) highlighted that single states may seek controlling IO administrations unilaterally. Finally, Elsig (2010) stressed the importance of member states' material and social preferences to ensure proper control.

However, as the literature on international organizations has finally acknowledged the distinct actorness of IO administrations, P-A analyses should rebalance their focus on both sides of the equation. Does control by the principal always constrain the agent's opportunism? Is the fear of sanctions the only concern that drives agents' behaviour? These questions can only be answered if both sides of the relationship are equally considered. By treating both sides seriously, we should learn more about how these actors interact with each other and what dynamics drive their behaviour both within and between these actors. A few very recent studies have already demonstrated the benefits of such perspective: Eckhard et al. (2019) showed how the lack of synchronized action between member states and IO administrations can lead to significant institutional failures; Knill et al. (2019) showed how member states' actions and IPAs' perception thereof contribute to the emergence of informal behaviour patterns in IOs; a study by Bayerlein et al. (2020a) showed that the congruence between member states' and IPAs' policy agendas may vary depending on the extent to which member states delegate formal authority to IPAs as well as the extent to which IPAs' administrative cultures foster independent agenda developments. Hence, if such relational dynamics remain ignored, P-A approaches cannot unleash their explanatory power. For instance, by only focusing on one side of the relationship, P-A models become "ill-suited to deal with temporal dynamics of delegation" as they lack a perspective of how principals build and update their view on agents and how agents can influence it (Thatcher and Sweet, 2002: 8).

In short, P-A theory suffers from a principal-bias in its theoretical thinking, thereby missing potential dynamics related to the actual behaviour of the agent. The state-centric perspective on member states-IPA relationship should thus be scrutinized and modified by focusing on both sides of the relationship and their interactions.

Third, and related, many studies that employ P-A theory to analyse delegation to IOs are based on a set of *unquestioned assumptions*. While these are most of the time taken as given, it remains unclear whether these assumptions hold empirically and what happens if they are not met. The key assumptions are:

- *The assumed goal conflict between the principal and the agent, forcing the principal to apply hard control over the agent.* The agent is thought to be a self-interested utility maximizing actor who, if given a chance, will pursue its own agenda deviant to that of the principal. Even though Hawkins et al. (2006a: 7) mention that P-A theory “does not imply any particular assumptions about the preferences of the actors”, the authors themselves assume that goal conflict is given in all delegation relationships: “Since the preferences of the principals and agents are seldom aligned perfectly [...], there is a natural and perhaps inevitable conflict of interest between the parties” (Hawkins et al., 2006a: 24; see also Kiewiet and McCubbins, 1991: 24). Analogously, Cortell and Peterson (2006: 256) write that “PA approaches generally take agents’ preferences as given and assume they will conflict with those of principals”. Preference divergence is central to P-A theory as it creates incentives for the agent to follow own interests and slack. In addition, high information asymmetry creates opportunities for the agent to do so. This explains the so-called ‘agency problem’ when the principal has to choose between costly control or the increased risk of agency slack (see Hawkins et al., 2006a: 24ff.).
- *The assumed anticipatory obedience by the agent.* In case of multiple or collective principals, the agent is thought to focus on the most powerful principal, anticipating the costs of potential escalation with the latter. This implies that IPAs might ally with selected powerful member states to avoid their sanctions, even if that undermines their impartiality towards the rest of the membership (see Urpelainen, 2012).
- *The assumed reflection of principal’s power in the delegation contract.* In line with rational design school (Abbott and Snidal, 1998; Koremenos et al., 2001), P-A theory assumes that the design of delegation generally reflects principals’ interests (Pollack, 2006: 166). Hence, IOs’ institutional policies and practices should primarily mirror the preferences of IO member states rather than IO administrations.

- *The assumed inverse relationship between the principal's control and the agent's opportunistic behaviour.* P-A theory assumes that control by the principal constrains opportunistic behaviour by the agent. In fact, agency slack is treated as an outcome of imperfect or insufficient control. Differently put, P-A theory generally assumes that “the more closely the principal monitors the agent's behavior [...] – the lower the risk of agency slack” (Abbott et al., 2016: 723; Lake and McCubbins, 2006; Kassim and Menon, 2003).

However, we do not really know whether these assumptions hold true empirically and, if not, how does this affect member states-IPA relationships. Do IPAs indeed demonstrate strong goal conflict and information asymmetry vis-à-vis their member states? Does the notion of ‘anticipatory obedience’ really explain how IPAs behave or are there other factors that relate more to the IPA itself rather than the power of member states? Can the IO's institutional design actually shift the power of balance from member states to the IPA, maybe even willingly so by member states? Do IPAs that are controlled less really show more opportunism? Does more control by member states necessarily constrain IPAs' opportunistic behaviour? Can these controls be escaped through informal practices? These questions cannot be answered as long as key P-A assumptions are taken as given and as long as P-A approaches remain principal-biased. In short, there is a need to test basic P-A assumptions empirically and potentially modify them both conceptually and theoretically.

Finally, these unquestioned assumptions might be the result of a fourth limitation of IO studies employing P-A theory, namely their *narrow empirical focus on few, very prominent IOs*. Scholarship on IOs would benefit from widening their empirical landscape and studying a broader variety of IOs as well as using more comparative rather than single case research designs. So far, a small number of IOs dominated scholarly interest. The most prominent examples include the European Union (especially the Commission and the European Court of Justice) (Pollack, 1997, 2003, 2006; Kassim and Menon, 2003; Wonka, 2007; Heldt, 2017) as well as international finance institutions such as the World Bank (Nielson and Tierney, 2003, 2005; Gutner, 2005a, 2005b; Heldt, 2018) and the IMF (Gould, 2003, 2006; Dreher and Vaubel, 2004; Martin, 2006; Vreeland, 2006; Weaver, 2010). In addition, most of these contributions were single case studies.

This focus on the same, small number of IOs and the lack of comparisons that account for similarities and differences between IOs might be one of the reasons for the limitations of P-A studies outlined above. Hence, studying other IOs beyond the ‘usual suspects’ and employing

a comparative perspective seems as a promising strategy for theoretical innovation (see Jankauskas et al., in press).

Overall, to better grasp MS-IPA relationships in IOs, the following conventional P-A assumptions should be tested and modified: 1) the lack of clarity on the definitions of actors (the principal and the agent); 2) the lack of attention to the actual agent (IPA) behaviour (principal-bias); 3) the lack of knowledge about variation in basic P-A theory assumptions; and 4) related empirical drawbacks in P-A studies.

3 Probing and advancing our understanding of MS-IPA relationships

My dissertation addresses the identified limitations of P-A theory. First, I discuss how the four core articles connect with each other and allow me to take a comprehensive perspective on the MS-IPA relationship (3.1). Then, I provide a brief summary of each paper and its specific contributions (3.2). Finally, I summarize how these findings in overall contribute to identified research gaps and advance our understanding of MS-IPA interaction in IOs (3.3).

3.1 Common approach and workshare between the core papers

Figure 2 depicts the multifaceted relationship between member states and IO administrations. It differentiates between its formal and informal dimensions and equally considers behaviour of *both* member states *and* the IPA. To address the identified shortcomings in the existing, P-A based literature, each of the four papers explores a different aspect of the MS-IPA relationship (see Figure 2). Paper A (Jankauskas, 2021) focused on the principal's behaviour vis-à-vis the agent and how the varying nature of the agent affects it. Paper B (Heinkelmann-Wild and Jankauskas, 2020) focused on the agent's behaviour vis-à-vis the principal and how the nature of the principal affects it. Paper C (Eckhard and Jankauskas, 2020) considered the formal delegation contract, whereas Paper D (Jankauskas et al., 2021) accounted for the P-A relationship and its informal practices.

Member states-IPA relationship

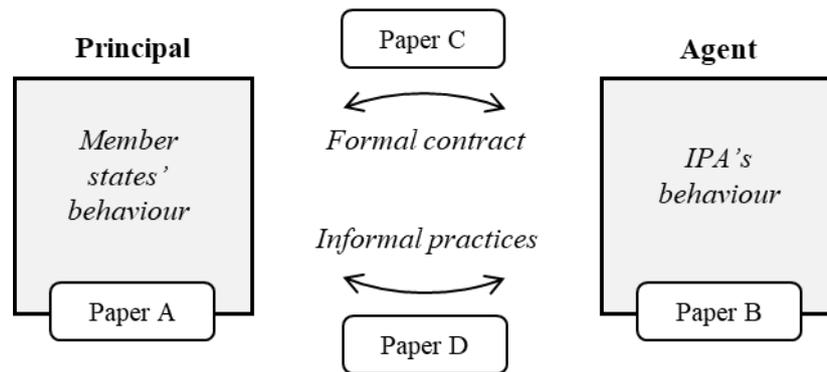


Figure 2: Different angles of member states-IPA relationship covered by core articles.

Approaching the member states-IPA relationship through these four papers allows me to address the identified limitations of P-A theory. First of all, considering the described confusion regarding the *definition of actors in IO delegation relationships*, all four papers offer a precise definition of both the principal and the agent under investigation. Papers B and D focus on member state governments and their relationships with IO administrations, whereas papers A and C apply a narrower focus on the collective of member state representatives in IO executive governing bodies (usually ambassadors) and IPAs' leadership. I also provide a conceptualization of the different dimensions of delegation relationships within IOs, explaining which relationship dynamics are typically studied at which dimension (Jankauskas, 2021).

Furthermore, to correct the described *principal-bias*, the four core articles always consider both the principal and the agent, even if the primary focus was on one of the two actors. Their findings, too, were discussed in the light of existing state-centric perspective in P-A approaches. For example, even if the focus in Paper B was on the agent's behaviour, it investigated how it is affected by the characteristics of the principal.

Moreover, the four papers together test the so far *unquestioned P-A assumptions* (see Table 1 below). In particular, the focus of Paper A allowed me to test the described P-A assumption related to goal conflict and information asymmetries and the related need for member states to apply hard control over the IPA (Jankauskas, 2021). The focus of Paper B allowed me to test the described P-A assumption related to the IPA's anticipatory obedience vis-à-vis selected member states, accounting for varying levels of their power (Heinkelmann-Wild and Jankauskas, 2020). The focus of Paper C allowed me to test the described P-A assumption

related to IOs' institutional design and whether it reflects member states' rather than the IPA's interests (Eckhard and Jankauskas, 2020). Finally, the focus of Paper D allowed me to explore whether informal practices constitute deviations from P-A assumptions, in particular regarding IPA's opportunistic behaviour (Jankauskas et al., 2021).

Core Article	P-A assumption under investigation (see also Section 2.2)
Paper A (<i>Jankauskas, 2021</i>)	The principal (member states) has to apply hard control over the agent (IPA) to reduce inherent goal conflict and information asymmetry
Paper B (<i>Heinkelmann-Wild and Jankauskas, 2020</i>)	Agents' (IPAs') behaviour is constrained by self-serving anticipatory obedience to powerful principals (member states) to avoid punishments.
Paper C (<i>Eckhard and Jankauskas, 2020</i>)	The design of delegation contract (IOs' institutional policies) reflects powerful principals' (member states') interests
Paper D (<i>Jankauskas et al., 2021</i>)	The more the principal (member states) controls the agent (IPA), the less opportunistic the agent (IPA) behaves, and <i>vice versa</i> .

Table 1: Conventional P-A assumptions tested and modified in the dissertation.

Finally, to overcome the *narrow empirical basis* of the IO delegation literature, the four papers broaden our empirical knowledge and approach a diverse set of major IOs and from a *comparative* perspective, utilizing either small-n or medium-n qualitative research designs. Table 2 provides an overview of IOs analysed in each paper.

IO	Core Article			
	A	B	C	D
IAEA		X	X	
ILO			X	X
IMF		X		X
IOM		X	X	X
FAO	X	X		X
NATO		X		X
OHCHR		X		
UNDP			X	
UNEP		X	X	X
UNESCO		X	X	
UNFCC		X		
UNFPA		X		
UNHCR		X	X	X
UNICEF		X	X	
UNRWA		X		
UNSG		X		
UPU		X		
WB		X		
WFP	X	X		
WHO			X	X
WTO		X		

Table 2: Overview of IOs analysed in each paper.

Regarding the sought comparative perspective, Paper A (Jankauskas, 2021) approached the member states-IPA relationship from a small-n perspective as it conducted a comparative case study of two Rome-based IOs (FAO and WFP). Paper B (Heinkelmann-Wild and Jankauskas, 2020) followed a medium-n approach as it conducted a crisp-set Qualitative Comparative Analysis (csQCA) (Rihoux and Ragin, 2009; Schneider and Wagemann, 2013). Although the unit of analysis was not an IO but an IPA's response to contestation by the Trump administration, 32 such cases across 18 IOs were analysed. Furthermore, Paper C (Eckhard and Jankauskas, 2020) and Paper D (Jankauskas et al., 2021) utilized medium-n research strategies as they engaged in a controlled comparison of 9 and 8 IOs respectively (see also Jankauskas et al., in press). In sum, all papers explored member states-IPA relationships from a comparative perspective, across a variety of IOs.

Noteworthy, each of these methods has both its strengths and challenges. For one thing, the use of the csQCA allowed distinguishing relevant constellations of conditions and, as a medium-n technique, enabled higher external validity. Yet, it required to dichotomize complex phenomena

such as dependency on the contesting state to enable appropriate operationalization. Paper B dealt with the issue of potential ‘oversimplification’ by providing full transparency on each operationalization decision in its appendix and running several robustness checks where alternative operationalization decisions were tested. As for comparative case studies in Papers A, C, and D, one of the main challenges was to formulate theoretical categories that are relevant for the specific case, yet, at the same time, are applicable to the overall universe of cases (see Jankauskas et al., in press). Also, selecting cases based on either MSSD or MDSD logics raised the challenge of controlling for all relevant variables. Consequently, the articles discussed the selection criteria and addressed alternative explanations in their research design sections.

Altogether, while the four core papers start from the same understanding of the MS-IPA relationship and all address the limitations of P-A theory, the outlined workshare among them allows this dissertation to efficiently cover different angles of MS-IPA interaction, testing conventional P-A assumptions, and considering both sides of the equation, thereby revisiting this relationship in a theoretically comprehensive and empirically diverse manner.

3.2 Summary of the four core articles

Paper A: Delegation and Stewardship in International Organizations (Jankauskas, 2021)

In this core contribution, I ask how varying levels of goal conflict and information asymmetry between member states and IO administrations affect their interaction, specifically, the principal’s exercise of control over the agent. The relationship dimension is narrowed down to day-to-day interactions between the collective of member states sitting in IOs’ executive governing bodies and IOs’ administrative leadership (the management).

The research question of the article already implies that I relax basic P-A assumptions instead of merely assuming that goal conflict and information asymmetries are inevitable and thus hard control by the principal is the preferable choice. Nevertheless, if we loosen the two assumptions, agency theory loses its explanatory power: “only when there is goal incongruence between the two is agency theory applicable” (Miller, 2005: 205; Tamm and Snidal, 2014). I thus utilize stewardship theory to provide a novel, theory-grounded way of modelling member states-IPA interactions which deviate from conventional P-A assumptions. The article argues that delegation can also take a form of stewardship, where the two actors operate on the basis of low interest divergence and *de facto* reduced information asymmetry.

In *stewardship relationships*, member states trust the IO administration, which enables softer, more informal exercise of control. A steward-like administration seeks to balance the competing preferences among member states and find solutions in the best interest of the membership. Such an IPA acts as a pro-organizational steward rather than a self-interested opportunistic agent. This facilitates the management's trustworthiness and allows less demanding, more informal, soft exercise of control (Table 3). Such stewardship relationships are thus characterized by cooperation rather than continuous contestation.

By contrast, conventional *agency relationships* are based on distrust and so member states as principals indeed are incentivized to apply hard control to reduce agency slack potential. Prevailing high goal divergence and information asymmetry between the principal and the self-serving agent leads to distrust which drives their interaction: executive governing bodies strictly scrutinize the work of the IO management and stick to formal arrangements to reduce the risk of agency slack. Such agency relationships are shaped by contestation as the principal seeks to constrain the agent who, in turn, seeks to escape the principal.

Overall, the newly introduced trust mechanism explains how varying levels of goal conflict and information asymmetry affect the exercise of control in IOs (Table 3). These behavioural patterns are conceived of as relatively stable manifestations of two extreme poles of ideal-typical hierarchical relationships.

Type of relationship	Behavioural premise	Interaction
Agency: (Principal-Agent)	Goal conflict + info. asymmetry <i>High</i>	Exercise of control <i>Hard</i>
Stewardship: (Principal-Steward)	Goal conflict + info. asymmetry <i>Low</i>	Exercise of control <i>Soft</i>

Table 3: Agency and stewardship relationships in IOs.

Empirically, one of the article's main innovations is its novel empirical data drawn from a cross-sectional analysis of FAO and WFP based on a most similar system design, including the analysis of in-depth interviews with both member states and IPA staff. As the study demonstrates, while FAO resembled ideal-typical agency relationship between member states and the IPA, stewardship relationship was found to exist in WFP. In particular, although same formal mechanisms for control exist in both IOs, hard exercise of control was found to be prevalent in FAO. Members states had little confidence in the management, which resulted in

their scrutiny over FAO's internal oversight tools, use of external information sources, and strong tendency to resort to formal procedures when solving issues with the management. By contrast, the WFP Executive Board members were found to rely on the management and accordingly soften their control, relying on WFP's internal oversight and more informal means of interaction. This provides – for the first time – systematic empirical evidence that stewardship relationships exist in IOs, thereby contradicting delegation scholars who argue that loyal subordinates are basically impossible to find and are therefore not worth studying (see Hawkins and Jacoby, 2006: 201).

In addition, by introducing stewardship as an alternative type of relationship (or mode of interaction) between member states and IPAs (see also Jankauskas, 2019), the paper contributes both conceptually and theoretically to delegation and governance literature (Hawkins et al., 2006b; Alter, 2008; Green and Colgan, 2013; Abbott et al., 2016, 2020a, 2020b; Weiss and Jankauskas, 2019). As the theoretical implications of these findings are closely related to the findings of the other core contributions, I discuss them in Section 3.3 on the overall lessons learned. Yet, the paper's specific conceptual contribution relates to its presentation of stewardship, the definition of typical characteristics of both agents and stewards and explanation of their utility functions. I also scrutinize our understanding of the principal's control over the agent by conceptualizing *de facto* rather than *de jure* control characteristics including formal vs. informal arrangements. Finally, the study also speaks to other literatures such as organizational and trust research as it demonstrates how trust affects cooperation between actors in complex organizational settings (Cummings and Bromiley, 1996; Kramer and Tyler, 1996; Kramer and Cook, 2007; Vanneste, 2016). Also, it complements stewardship literature, stemming from management and sociology research, as it shows that insights of stewardship analysis of mostly domestic private actors travel to the analysis of complex member state-IPA relationships at the international level (Donaldson and Davis, 1991; Davis et al., 1997; Arthurs and Busenitz, 2003; Van Sylke, 2006).

Paper B: To Yield or Shield? Comparing International Public Administrations' Responses to Member States' Policy Contestation (Heinkelmann-Wild and Jankauskas, 2020)

The second article of the dissertation focused on member states-IPA relationship in times of disagreement and conflict and analysed how IPAs (as agents) behave in such situations vis-à-vis their principals. It started from the observation that IOs are increasingly contested by member state governments who publicly attack IOs' polity and policies or shift blame to the

administration. IPAs, in turn, have to respond to defend themselves. While some administrations react in a conciliatory way by employing positive, accommodating tone towards the contesting state, other IPAs react adversely by using negative tone, rejecting the accusations, and even shaming the contesting state back. Paper A thus asks what explains this variation in IPAs' communicative responses to member states' contestation.

Referring to the described principal-bias in existing research, the intuitive explanation derived from conventional P-A theory is straightforward: when IPAs are confronted with contestation from a powerful member state, they anticipate potential costs of a further escalation and choose a conciliatory response to avoid costly sanctions. An IPA thus yields to contestation as its dependence on the contesting state constrains an adversarial reaction. If, however, the contesting state is weak, the administration can afford stepping up and shielding its IO. This hypothesis is labelled as the 'dependence condition'.

However, is the notion of 'anticipatory obedience' the only relevant factor that explains how the agent behaves or are there other factors that relate more to the agent itself rather than the power of its principal? The article introduces two further conditions. First, the level of IPA's authority might affect its choice of response. While in some IOs IPAs merely assist member states in intergovernmental bargaining (e.g. the WTO or NATO secretariats), in other IOs they are actively involved in IOs' policy design and implementation (e.g. the IAEA or UNRWA administrations). IPAs which have no policymaking authority have lower stakes in substantive outcomes, are less focal in public, and their public reputation is less important for 'their' IOs' legitimacy. Therefore, such IPAs have lower incentives to adversely defend themselves and can easier afford accepting the blame to end public contestation (conciliatory response). By contrast, IPAs with considerable policymaking authority have higher stakes in their policies, are more focal in public, and their public reputation is highly relevant for their authority as neutral, technocratic experts. Hence, such secretariats are incentivized to respond adversely to avoid undermining own authority and appearing as hypocrites. This hypothesis is called 'authority condition'.

The second additional condition refers to the type of contestation, i.e. how affected IO administrations are. It is easier for an IPA to yield to contestation if it is not targeted directly at the bureaucracy, but, for instance, other member states in the IO's context. By contrast, when member states directly criticize the IPA's performance, necessity, withdraw resources or membership, or threaten to do so, the IPA is incentivized to react adversely and defend itself. It can reject the blame, shift it to others, or attack the contesting state back. Yet, when member

states criticize other member states or IO policies (or not comply with them), the IPA is outside the line of fire and can respond in a conciliatory manner, avoiding becoming the target of the contestation itself. This hypothesis is labelled as the ‘affectedness’ condition.

To test the plausibility of these arguments empirically, a crisp-set Qualitative Comparative Analysis (csQCA) across 32 cases of IPAs’ responses to contestation by the Trump administration (2017-2019) was conducted. After running a number of robustness checks, the study arrived at the following results: IPAs respond conciliatory when they are constrained by the dependence on the contesting state *or* have no incentives to defend adversely. The latter is the case when IPAs have no policymaking authority and are contested indirectly. On the contrary, IPAs respond adversely when they depend less on the contesting state (and are thus unconstrained) *and* are incentivized to respond adversely. The latter is the case when IPAs are targeted directly or enjoy high policymaking authority, so they need to fend off contestation.

Empirically, the paper thus provides first systematic evidence of IPAs’ responses to contestation by the Trump administration. It is indeed remarkable that even if the financial dependency of the investigated IOs on the US varies, in 13 (out of 32) cases IPAs reacted adversely vis-à-vis the US that is both extremely influential in global governance and is known for sanctioning unsatisfactory IOs. Generally, the revealed empirical evidence of numerous IPA responses demonstrate the distinct actorness of IO administrations who – far from being passive technocrats of the past – actively engage in the public discourse (Eckhard and Ege, 2016).

Theoretically, the article demonstrates that IPAs’ (as agents’) behaviour is not merely shaped by anticipatory obedience to powerful member states (cf. Cortell and Peterson, 2006; Lake and McCubbins, 2006). As the csQCA findings show, although the agent’s ‘dependency’ on its principal matters, it is not the only factor that counts: neither the presence nor the absence of any of the three theorized conditions is necessary for a conciliatory or adversarial IPA response. Instead, it is the combination of these conditions that best explains how IPAs behave once contested. As ‘authority condition’ relates to the nature of the IPA and the ‘affectedness’ refers to the context of the situation, these results speak against a principal-biased view when studying IOs and highlight the need to consider the characteristics of both member states and IPAs in relation to each other (see also Heinkelmann-Wild and Jankauskas, 2019). Conceptually, the paper contributes to scholarly debates on IO legitimacy and resilience as the two IPAs’ strategies are already quoted as “essentially two ways for IOs to respond” to external pressures (Debre and Dijkstra, 2020: 4; see also Goritz et al., 2020). Finally, the paper bears implications for IO self-legitimation research as it reveals so far neglected negative-toned communication

strategies (Zaum, 2013; Schmidtke, 2019; Tallberg and Zürn, 2019; Rauh and Zürn, 2020) and contributes to blame and reputation management literature by challenging the widespread assumption that IO member states face no confrontation when shifting blame onto IOs (Vasilopoulou et al., 2014; Schlipphak and Treib, 2017; Sommer, 2020).

Paper C: Explaining the Political Use of Evaluation in International Organizations (Eckhard and Jankauskas, 2020)

The third study of the dissertation delves into political-administrative relationships within IOs' institutional setting, thereby focusing on the formal delegation contract. It specifically focuses on evaluation as an institutional tool that is increasingly used by member states and IO administrations. Traditionally, evaluation is thought to fulfil purely *functional* purposes such as fostering learning or triggering institutional change. Yet, the article asks whether evaluation systematically serves *ex-ante political* interests of either member states or IPAs that go beyond these functional imperatives. The political use of evaluation describes situations in which actors refer to evaluations to realize own political interests in competition with others.

To conceptualize such political evaluation use in IOs, the paper scrutinizes P-A theory and outlines general P-A dynamics both *within* the two actors (among member states and within the IPA) as well as *between* them (member states vis-à-vis the IPA and *vice versa*). As summarized in Table 4, these dynamics generate political interests (in *italics*). Evaluation, in turn, can be useful in achieving them. Member states may use evaluation in intergovernmental bargaining (collective principal dynamics) or in constraining IPAs' bureaucratic influence (control dynamics) and IPAs may use evaluation to justify its initiatives (bureaucratic influence dynamics) or strategically steer the organization internally (collective agent dynamics).

P-A dimension	Principal (MS)	Agent (IPA)
within actor	Collective principal dynamics <i>Justify own interests, prevent other MS' unilateral influence</i>	Collective agent dynamics <i>Contain internal fragmentation</i>
between actors	Control dynamics <i>Control the agent</i>	Bureaucratic influence dynamics <i>Justify own past of future action</i>

Table 4: General P-A dynamics within and between principals and agents and related political interests. MS refers to member states.

Building on this conceptualization, the study theorizes on the conditions under which either member states or IPAs are able to utilize evaluation for their respective political interests. As the argument goes, evaluation should systematically serve ex-ante political interests of one of the two actors depending on who controls evaluation system resources. As demonstrated in another study by Eckhard and Jankauskas (2019), in some IOs it is member states who decide upon IO evaluation units' resources, yet, in other IOs it is the IPA who exercises this control. Following this pattern, evaluation units are expected to orientate towards either member states or the IPA, depending on who dominates these resources. By carefully outlining the causal mechanism of such stakeholder influence, the article argues that evaluation units should accordingly produce evaluations that tend to be used for the respective actor's strategic preferences.

To test the theorized argument empirically, a medium-n controlled comparison of nine IOs, selected on the logic of a MSSD, is conducted. Drawing on IO evaluation policies and 35 research interviews with all relevant actors (member state representatives, IPA staff, and evaluators), the study demonstrates that IO evaluation units systematically oriented to either member states or IPAs as their primary users and sponsors, depending on who controlled their resources. The same pattern was observed in interviewee statements on political evaluation use. In IOs with member state dominated evaluation systems (ILO, UNDP, UNICEF), interviewees mentioned typical member states' interests, i.e. evaluation being primarily used by member states to contain each other's or IPA's influence. In IOs with IPA dominated evaluation systems (IAEA, IOM, UNHCR), respondents referred to typical IPA interests, i.e. evaluation being predominantly used by the IPA to justify its initiatives or avoid internal fragmentation.²

Thereby, Paper C provides first empirical comparative evidence of a political use of evaluation in international organizations. It shows how IOs' institutional policies and practices can influence the balance of power between principals and agents and how IPAs can exert bureaucratic influence through seemingly neutral, value-free technocratic tools.

Theoretically, the study hence demonstrated that – in contrast to state-centric assumptions of rational design theory (Abbott and Snidal, 1998; Koremenos et al., 2001) – IOs' institutional design does not always reflect member states' preferences. Delegation literature should thus equally consider both sides of the relationship as well as the institutional context in which principals and agents interact. In addition, Paper B speaks to the public policy and evaluation

² A more specific, third option was observed in IOs where neither member states nor the IPA had unilateral control over evaluation resources. In such IOs (UNEP, WHO, UNESCO), evaluation unit orientation as well as political evaluation use were also mixed.

literature as it highlights the importance of the political context in which public management tools such as evaluation exist (Hinterleitner et al., 2016; Perl et al., 2018; Fforde, 2019; van Voorst and Mastenbroek, 2019). Neglecting the political context limits our understanding of how organizations and their policies work.

Paper D: Is Control Always Better Than Trust? Determinants of Organizational Agency in Global Public Policy (Jankauskas et al., 2021)

Paper D set out to test and rethink the basic P-A assumptions related to the risk of agency slack, a phenomenon that has been puzzling IO scholars ever since, and potential deviations from the delegation contract through informal practices. So far, the risk of agency slack is seen as a by-product of (lacking) principal's control (see, for instance, Heldt, 2017). This explains why most P-A explanations focused on the principal side rather than the actual agent's behaviour. However, this is also the blind spot of P-A theory: research has demonstrated that IPAs develop a life of their own beyond formal delegation contract and controls that bind them. Hence, Paper D argues that to fully grasp IPA agency (and that of agents more generally) one should consider both its formal discretion determined by the principal and the actual agent's behaviour determined by its (informal) behaviour patterns (which are thereby exogenous to the formal delegation contract).

In particular, the study scrutinized the assumed inverse relationship between principal control and the risk of agency slack: the higher the control over the agent, the more constrained is the agent's opportunism and thus the less likely is the agency slack (and *vice versa*). To test these dynamics, the article conceptualized, operationalized, and measured both formal IPA discretion and its (informal) opportunistic behaviour patterns across 8 major IOs (IMF, IOM, ILO, FAO, WHO, UNHCR, UNEP, NATO). For the former, it presented a novel and encompassing measure of direct and indirect IPA control by member states. For the latter, it measured IPAs' opportunism, proxied by the prevalent informal organizational routines throughout an IOs' policy cycle. Drawing on primary IO documents and in-depth research interviews, the study showed that less control does not necessarily induce the IPA to slack and more control does not necessarily constrain its opportunism. Four short case studies illustrated these findings. On the one hand, the ILO and UNHCR serve as *typical* examples of expected P-A dynamics. As control levels in the ILO are rather high, its secretariat shows low opportunistic behaviour. Analogously, while control levels are rather low in UNHCR, its administration was found to demonstrate highly opportunistic behaviour. Yet, on the other hand, NATO and IMF serve as

deviant cases to P-A theory: although the NATO administration enjoys low control, its opportunistic behaviour remains low. And while the IMF secretariat is put under high control, it nevertheless prevails highly opportunistic orientations. The two factors thus seem to vary across IOs, not necessarily in line with conventional P-A expectations.

Thereby, Paper D conducts a first systematic *empirical* test of classical P-A assumptions related to agency slack potential in IOs. It offers an innovative approach to study IPAs' opportunistic behaviour empirically and reveals such patterns for eight major IOs. It also refines existing measurement of IPAs' control by operationalizing control indicators which 1) show variation across IOs; 2) account for both direct and indirect member states' control; 3) and are sensitive to institutional reporting lines.

Theoretically, Paper D refines our understanding of the risk of agency slack: it depends on the interaction between both the (lack of) control and the patterns of informal agent behaviour as these can both reinforce and mitigate each other. Especially the informal practices in member states-IPA relationships seem to be neglected in conventional P-A scholarship. Even if some P-A scholars sought to consider the nature of the agent more seriously, they focused primarily on its formal side. For instance, Graham (2014) highlighted IOs' complexity, arguing that the more fragmented the IO administration is, the less effective control by member states gets. Yet, and in line with findings of papers A-C, formal IO characteristics do not necessarily explain the agent's behaviour. Indeed, trusting rather than controlling the agent might in some cases be a cheaper and more suitable approach for both the principal and the agent. Finally, the study contributes to research on informal administrative behavior as it demonstrates the utility of applying administrative styles concept to the study of IOs (Richardson, 1982; Knill, 2001; Knill et al., 2019).

3.3 Lessons learned

The analysis of the state of the art highlighted the need to both test and modify conventional P-A assumptions. This has been the core task of the dissertation articles as P-A theory fundamentally shapes our understanding of MS-IPA relationships. The dissertation has demonstrated a number of P-A theory's blind spots and limits when applied to MS-IPA relationships. While each of the core papers yields specific findings and contributions, their common thrust allows to draw general lessons about the member states-IPA relationships and address the identified limitations of classical P-A theory.

First, we learn that *unquestioned P-A assumptions are actually variables*. While suchlike conclusions so far often remained at a conceptual level (see, for instance, Waterman and Meier, 1998), this dissertations put them to empirical tests by combining evidence from both sides of the relationship. I find that Principal-agent or member states-IPA relationships are not necessarily defined by high goal and information asymmetries; not in all IOs are member states incentivized to apply hard control over the administration (Jankauskas, 2019, 2021; cf. Weiss and Jankauskas, 2019). IPAs' as agents' behaviour is not merely shaped by anticipatory obedience to powerful member states; other significant factors related to IPAs' own characteristics also affect how they behave vis-à-vis their principals (Heinkelmann-Wild and Jankauskas, 2019, 2020; Jankauskas et al., 2021). The design of the formal delegation contract and thus of IOs' institutional context does not merely reflect powerful principals' interests (Jankauskas and Eckhard, 2019; Eckhard and Jankauskas, 2020, 2019). Finally, more control does not necessarily constrain IPAs' opportunism and less control does not necessarily induce IPAs to slack (Jankauskas, 2021; Jankauskas et al., 2021).

This, however, would not have been possible without carefully conceptualizing and operationalizing the different aspects of MS-IPA relationship and drawing on a rich basis of about 100 in-depth interviews³ with *both* member state representatives *and* IO administration staff (in addition to document analysis). Interviewing, in particular, proved to be a well-suited tool for data collection and measuring relevant variables or conditions. Interviewing is known as a key technique for “establishing motivations and preferences”, which is “absolutely critical element of any theory” (Rathbun, 2008: 690f.). It allowed me to capture various delegation dynamics by directly speaking with involved actors. To quantify or even reveal such aspects like the extent of goal conflict between member states and the management or the *de facto* exercise of control would otherwise be extremely difficult.

Second, and related to the findings above, we learn that *agents differ in their nature, some being trustworthy stewards* (Jankauskas, 2019, 2021; Jankauskas et al., 2021). This corroborates international bureaucracy research that focuses on the functional orientation of IO staff (Yi-chong and Weller, 2008; Biermann and Siebenhüner, 2009; Eckhard and Ege, 2016; Knill et al., 2016; Jankauskas and Eckhard, 2019; Knill et al., 2019; Bayerlein et al., 2020b; Ege, 2020). P-A scholars should thus shift away from principal-biased perspectives and pre-assumed opportunism by the agent and focus more on the nature of the agent and its actual behaviour.

³ This includes both background interviews that I have conducted throughout the years of the dissertation and the specific interviews utilized in each paper. For the latter, comprehensive interview lists are provided in the Papers' appendices.

Third, we learn that *principal's control does not necessarily affect the agent's behaviour* (Heinkelmann-Wild and Jankauskas, 2020; Jankauskas, 2021; Jankauskas et al., 2021). In fact, IPAs were found to develop own behavioural patterns beyond the delegation contract. This speaks to nascent research on IPAs' informal activities (Enkler et al., 2016; Knill et al., 2016; Bayerlein et al., 2020b). It implies that looking at control alone is not sufficient and might explain why classical P-A scholars struggled to find an answer as to which control types are best to constrain the IPA. Hawkins et al. (2006b) discussed the pros and cons of various control mechanisms that member states employ, Grigorescu (2010) tested when and why new oversight mechanisms are introduced, Heldt argued that rule-based mandate as well as police-patrol and fire-alarm oversight reduce the risk of agency slack (Conceição-Heldt, 2013), whereas centralized monitoring as well as credible sanctions help member states to regain “control of runaway agents after slack has occurred” (Heldt, 2017). However, in some cases the problem might be not the principal's control, but the nature of the agent. Moreover, looking merely at the existence of control is insufficient as well. As I demonstrated, even in IOs with same tools of control in place, these tools can be utilized very differently, depending on the existence of trust between member states and IPAs (Jankauskas, 2021). Hence, these findings call the P-A scholarship to redirect its focus on de facto rather than de jure IO characteristics.

Finally, we learn that *IOs' institutional context may empower both member states and IPAs*. P-A theory would argue that in such cases member states are either unaware of IPAs' ‘institutional advantage’ or are unable to change miscalculated institutional design choices due to collective action problems. However, interview data implied that in some IOs (e.g. the IOM) member states were rather *willing* to shift power towards the IPA despite the potential backlash through its increased discretion (Eckhard and Jankauskas, 2020). And although I did not delve into a systematic analysis of such variation, its findings imply that delegation literature should equally consider both sides of the relationship as well as the institutional context in which principals and agents interact. The implication is also that other factors outside the conventional P-A explanations – such as member states' trust in the IPA – may explain certain design choices or institutional practices in IOs.

In sum, the dissertation advances our knowledge about MS-IPA relationship by drawing the following lessons: 1) the unquestioned P-A assumptions are actually variables; 2) agents (IPAs) actually differ in their nature, some being trustworthy stewards; 3) principal's control does not necessarily affect the agent's behaviour; and 4) institutional context (or the formal delegation contract) may empower not only member states but also IPAs.

4 A new research agenda - Between agency and stewardship in IOs

The central objective of this thesis was to examine political-administrative relationships between IO member states and IO administrations. To this end, I scrutinized P-A theory as the dominant, theoretical perspective to study delegation in IOs and tested its key assumptions, demonstrating variations and deviations from conventional wisdom. While several core papers *modified* the P-A perspective (Eckhard and Jankauskas, 2020; Heinkelmann-Wild and Jankauskas, 2020; Jankauskas et al., 2021), I also *complemented* the delegation relationship based on P-A assumptions with the stewardship relationship that rests on completely diametral assumptions (Jankauskas, 2021; see also Jankauskas, 2019; Jankauskas et al., 2021). The dissertation thereby seeks to move the debate forward by introducing and testing the theory of stewardship (Jankauskas, 2019, 2021) which covers delegation dynamics that P-A theory cannot explain.

The introduced *stewardship theory* considers the nature of the agent arguing that an IPA in its nature may also behave as a loyal steward rather than a self-serving agent and that this affects its relationship with member states. In fact, the flawed assumption among P-A scholars and oftentimes policymakers and the broader public that IPAs *per se* are self-serving bureaucracies might be one of the factors that creates dysfunctional IO designs where steward-like administrations are put under high control. This, in turn, both reduces IPA's competence (see Abbott et al., 2020a) and ignores the nature of the policy problem (see Weiss and Jankauskas, 2019). Stewardship theory might thus help finding fitting control designs and explaining cases where control differences are observed in otherwise similar IOs (Eckhard and Jankauskas, 2019, 2020) or where soft control exists even if that increases the risk of agency slack and alternative P-A explanations can be rejected (Jankauskas, 2019).

Essentially, the introduced trust mechanism implies that the choice of soft control, which puts the principal in a vulnerable position, does not necessarily have to be coercively induced by dysfunctional reasons (control costs or inability to control). Instead, as Paper A shows, it can be member states' trust that makes harder control unnecessary. This logic is what differentiates my stewardship approach from other concepts where scholars refer to trust. For instance, Alter (2006; 2008) writes about delegation to "trustees" and refers to softly controlled international courts (see also Majone, 2001; Abbott et al., 2016; Abbott et al., 2017). However, international courts enjoy low control primarily because hierarchical intervention would diminish their competencies. Thus, member states may question courts' trustworthiness (see Kaltenthaler et al., 2010), yet still have to put up with low control. Soft control in these approaches is thus a

must rather than an outcome of trust. By contrast, theorizing on stewardship where trust is the basis of member states-IPA relationship innovates the literature as IPAs' trustworthiness appear as a mechanism of managing delegation risks. In this regard, stewardship may be a solution to conventional cost-benefit or competence-control trade-offs often highlighted in delegation and governance literature (Hawkins et al., 2006b; Alter, 2008; Green and Colgan, 2013; Abbott et al., 2016, 2020a, 2020b) – if the principal can truly rely on its subordinate, these dilemmas become obsolete (see also Weiss, 2020). By implication, the interaction between member states and IPAs depends primarily on the quality of their relationship rather than structural asymmetries or formal constraints of the delegation contract. Analogous to the famous statement by Alexander Wendt (1992), potential goal conflicts and information asymmetries in delegation relationships is what actors make of it.

To be sure, by this I do not refer to constructivist argumentation or social reading of trust, i.e. emotions or feelings of attachment between the principal and the agent. Instead, member states' decision to exercise soft control over the administration is based on prudential assessment and experience with the administration over time. In line with the rational reading of trust, whether to take the risk on the subordinate's behaviour is a "matter of prudential assessment, not moral choice" (Hardin, 2006: 27). Also, the notion that actors learn from regular interaction and update their actions does not contradict the rational choice perspective (see Stacey and Rittberger, 2003: 869). After all, delegation is not a one-time shot, it is a regular interaction which over time manifests as relatively stable behavioural patterns, leaning towards either agency or stewardship as two ideal-types of delegation relationship. Therefore, stewardship remains compatible with agency theory in terms of assumed rationality, yet, it opens a new frontier in IO research that addresses important blind spots of P-A approaches and thereby advances our understanding of member states-IPA relationships.

In short, I suggest scholars studying international delegation to differentiate between conventional agency relationships and newly introduced stewardship relationships as two extreme poles of ideal-typical member states-IPA interaction. The implication of the dissertation is thus that while P-A theory is useful to explain certain delegation dynamics, it is alone insufficient. Yet, when combined, the two perspectives provide a fuller account of international delegation, able to explain the revealed deviations in classical P-A assumptions.

Several *scope conditions* should be however highlighted with regard to the generalization of these findings and implications. The core articles focused on international *governmental* organizations which consist of three or more member states and have permanent organs

(Pevehouse et al., 2004; Rittberger et al., 2019: 4). Hence, further attention would be required if transitioning my findings to other organizational settings such as non-governmental organizations (NGOs), loose state associations like the G20 or the BRICS, or transnational public-private partnerships. Also, I focused on *major* IOs which have significant resources and are relatively focal in the public. This, however, limits the applicability of findings to minor IOs with less capacities and smaller secretariats as such IPAs may be unable to develop own ‘strategic actorness’ (see Jankauskas and Eckhard, 2019) and preferences.

The stewardship perspective points to three avenues for *future research*: first, while we now know that stewardship relationships in IOs exist, further *mapping of agency vs. stewardship relationships across IOs* is required. For this, following the implemented approach of interviewing both member states and IPAs on their relationship seems to be promising, yet further indicators could potentially be found through IO document analysis (e.g. protocols of executive governing body meetings).

Second, future studies should develop *explanations for the emergence of either agency or stewardship relationships*. While this dissertation did not aim at explaining the emergence of either agency or stewardship, investigating factors that lead to different types of relationship is important. Also, more attention is needed to assess the stability and change of agency and stewardship relationships over time. So far, it seems that researchers and policymakers have focused a lot on how to increase control efficiency and better constrain IPAs’ opportunism and influence. But they rarely considered factors that actually enhance the quality of political-administrative relationships and thus make *hard* control unnecessary in the first place. How can we explain the emergence and ensure the maintenance of stewardship relationships in IOs and beyond? In this regard, based on conducted interviews and dozens of background conversations with IPAs and member states, factors related to IOs’ institutional culture, practices, and structures might be worthy analysing in more depth:

- Features related to *institutional culture* may contribute to stewardship. As for member states, competency of their representatives contributes to professional working environment and facilitates interaction with the management. Also, member states’ propensity to share the burden of involved risks with IO administrations – which usually operate in risky environments – might affect the atmosphere of member states-IPA interactions. As for the IPA, its integrity and openness vis-à-vis the member states may help building trustworthiness: ambassadors often highlighted that they expect the IPA to inform them about issues such as fraud cases immediately rather than reading about

it in the press. Hence, even if performance failures occur, IPAs may save their trustworthiness by approaching member states in advance. Finally, IPAs' ability to maintain impartiality to national interests of single states appears to be crucial.

- Features related to *institutional practices* are also worth considering. As for member states, regular meetings among representatives can help balancing out informational asymmetries among each other. Also, rapid rotation of member state representatives should be scrutinized. It does not only undermine institutional memory, but also creates inefficiencies and competency issues. For instance, representatives who are stationed only for two years can hardly build a comprehensive view of IOs' internal working complexities. As for the IPA, intensive formal and informal communication with IO member states seems to be one of the key mechanisms to reduce information asymmetry and thus suspicion by the principals. Regarding staff rotation on the IPA's side, it seems that in some cases increasing rather than reducing staff rotation might enhance transparency and foster management culture. Interview partners highlighted the rare rotation of FAO's leadership that created negative incentives for the IPA's staff. The average term of the last five Director-Generals was 12.2 years (with two DGs acting for 17 years each), whereas the last five WFP Executive Directors ruled on average 7 years. Finally, the IPA's ability to steer the administration both vertically and horizontally (see Jankauskas and Eckhard, 2019), and thereby strengthen its identity and reputation might also help building stewardship.
- Finally, features related to *institutional structures* may help explaining the existence or absence of stewardship. As for member states, the design and structure of IOs' executive governing bodies could affect their ability to formulate collective preferences and communicate with the IPA management. For instance, the WFP Executive Board has its own Bureau which is elected annually and consists of one member state representative from each regional country list. The Bureau facilitates consensus building among regions and interacts closely with the management. Furthermore, the ways how member states' representations are built, where they are located, how much resources they have, and how much support they receive from the capitals may affect the quality of member states' work in IO governing bodies. Many of the MS representatives are overwhelmed with tasks and information they need to process. And that counts even for most powerful countries. If, say, one ambassador is responsible for eight major IOs in Geneva, how can she or he manage reading hundreds of documents provided by these

IOs weekly? If member states become unable to grasp relevant internal dynamics related to the work of the management, this can negatively affect their relationship. As for the IPA, dedicating a specialized unit for managing relationships with MS might significantly help building mutual trust. For example, the WHO has a unit that is responsible for governing bodies. Its staff briefs and guides incoming Executive Board representatives and fosters informal contacts with the ambassadors. Similarly, reducing internal structural fragmentation inside IO bureaucracies might help the administration to provide a better service to MS (see Graham, 2014; Hanrieder, 2015).

Finally, a third step would be to advance our understanding of *the consequences of stewardship (and agency) on the functioning of IOs*. The way how member states interact with IO administrations may significantly enhance IOs' performance. Organizational and management studies have long argued that trust positively affects organizational processes by reducing transaction costs (Cummings and Bromiley, 1996). It fosters "smooth-running inter-organizational co-operation" and helps in complex decision making processes (Edelenbos and Klijn, 2007). Moreover, given that strict oversight has been observed to negatively affect subordinates' performance efforts (Kramer and Cook, 2007: 112; Abbott et al., 2020a), trust might be seen as "an important alternative to formal governance mechanisms" (Vanneste, 2016: 10). Hence, better understanding how stewardship affects IOs' internal functioning might bear important implications for IO performance research.

To be sure, one should not forget that trust can also be misplaced. As Scharpf (1997: 89) put it, "being able to trust is an advantage – but exploiting trust may even be more advantageous". Thus, the claim is not that all IPAs should be considered as trustworthy stewards. In fact, IPA research has demonstrated the various innovative ways of how IPAs increase their autonomy and bureaucratic influence on IO policy making (Biermann, Siebenhüner 2009; Eckhard, Ege 2016; Eckhard, Jankauskas 2019; Patz, Goetz 2017). Hence, further research is also needed to better identify the 'true nature' of different IO administrations and understand whether and under what conditions do they build trust strategically to exploit delegation asymmetries and pursue own agendas.

After all, whatever key challenges await humanity in the 21st century, they are likely to be of a global nature, ignoring national borders or domestic political systems. As Harrari (2018) wrote, "you cannot build a wall against nuclear winter or against global warming, and no nation can regulate artificial intelligence or bioengineering single-handedly". Indeed, such challenges will require states cooperating with each other rather than turning to individualistic nationalist

agendas, even if the latter might seem to be trending nowadays. But even if nation states will remain the ultimate holders and managers of authority, their ability to effectively deliver their promises to their citizens will require successful cooperation at the international level. As UN Secretary-General António Guterres highlighted, “global problems require global solutions” (UN, 2019).

It thus seems almost inevitable that the level of authority and the scope of tasks delegated to IOs will continue to grow. Of course, IOs are not perfect. As all organizations, they are subjects to inefficiencies, corruption, and performance failures. But it seems that IOs are currently the best platforms we have for multilateral cooperation. Hence, the question of how successful international organizations will be, will define how successfully we will be able to deal with most pressing issues of our times. And for this, member states’ relationships with international administrations will become of even greater importance. It is thus my hope that this dissertation contributes at least partially to a better understanding of how such relationships work.

References

- Abbott KW, Genschel P, Snidal D, et al. (eds) (2015) *International Organizations as Orchestrators*. Cambridge, UK: Cambridge University Press.
- Abbott KW, Genschel P, Snidal D, et al. (2016) Two Logics of Indirect Governance: Delegation and Orchestration. *British Journal of Political Science* 46(4): 719–729.
- Abbott KW, Genschel P, Snidal D, et al. (2020a) Competence versus control: The governor's dilemma. *Regulation & Governance* 14(4): 619–636.
- Abbott KW, Genschel P, Snidal D, et al. (eds) (2020b) *The Governor's Dilemma: Indirect Governance Beyond Principals and Agents*. New York: Oxford University Press.
- Abbott KW, Levi-Faur D and Snidal D (2017) Theorizing Regulatory Intermediaries: The RIT Model. *The ANNALS of the American Academy of Political and Social Science* 670(1): 14–35.
- Abbott KW and Snidal D (1998) Why States Act through Formal International Organizations. *Journal of Conflict Resolution* 42(1): 3–32.
- Alter KJ (2008) Agents or Trustees? International Courts in their Political Context. *European Journal of International Relations* 14(1): 33–63.
- Arthurs JD and Busenitz LW (2003) The Boundaries and Limitations of Agency Theory and Stewardship Theory in the Venture Capitalist/Entrepreneur Relationship. *Entrepreneurship Theory and Practice* 28(2): 145–162.
- Barnett M and Finnemore M (1999) The Politics, Power, and Pathologies of International Organizations. *International Organization* 53(4): 699–732.
- Barnett M and Finnemore M (2004) *Rules for the World: International Organizations in Global Politics*. Ithaca (NY): Cornell University Press.
- Bauer MW and Ege J (2016) Bureaucratic autonomy of international organizations' secretariats. *Journal of European Public Policy* 23(7): 1019–1037.
- Bauer MW, Knill C and Eckhard S (eds) (2017) *International Bureaucracy: Challenges and Lessons for Public Administration Research*. London: Palgrave Macmillan UK.
- Bayerlein L, Kaplaner C, Knill C, et al. (2020a) Singing Together or Apart? Comparing Policy Agenda Dynamics within International Organizations. Online first. *Journal of Comparative Policy Analysis: Research and Practice*: 1–20.
- Bayerlein L, Knill C and Steinebach Y (2020b) *A Matter of Style: Organizational Agency in Global Public Policy*. Cambridge University Press.
- Biermann F and Siebenhüner B (eds) (2009) *Managers of global change: The influence of international environmental bureaucracies*. Cambridge MA: MIT Press.

- Brown RL (2010) Measuring Delegation. *The Review of International Organizations* 5(2): 141–175.
- Claude IL Jr. (1956) *Swords Into Plowshares: The Problems and Progress of International Organization*. New York: Random House.
- Collins R and White ND (2013) Between independence and accountability: Exploring the legal autonomy of international organizations. In: Reinalda B (ed) *Routledge Handbook of International Organization*: London: Routledge, pp. 119–134.
- Conceição-Heldt E da (2013) Do Agents “Run Amok”? A Comparison of Agency Slack in the EU and US Trade Policy in the Doha Round. *Journal of Comparative Policy Analysis: Research and Practice* 15(1): 21–36.
- Copelovitch MS (2010) Master or Servant? Common Agency and the Political Economy of IMF Lending. *International Studies Quarterly* 54(1): 49–77.
- Cortell AP and Peterson S (2006) Dutiful agents, rogue actors, or both? Staffing, voting rules, and slack in the WHO and WTO. In: Hawkins DG, Lake DA, Nielson DL and Tierney MJ (eds) *Delegation and Agency in International Organizations*: Cambridge: Cambridge University Press, pp. 255–280.
- Cummings LL and Bromiley P (1996) The Organizational Trust Inventory (OTI): Development and Validation. In: Kramer RM and Tyler TR (eds) *Trust in organizations: Frontiers of theory and research*. Thousand Oaks, CA, US: Sage Publications, pp. 302–330.
- Davis JH, Schoorman FD and Donaldson L (1997) Toward a Stewardship Theory of Management. *The Academy of Management Review* 22(1): 20–47.
- Debre MJ and Dijkstra H (2020) Institutional design for a post-liberal order: Why some international organizations live longer than others. Online First. *European Journal of International Relations*: 135406612096218.
- Dijkstra H (2015) Shadow Bureaucracies and the Unilateral Control of International Secretariats: Insights from UN Peacekeeping. *Review of International Organizations* 10(1): 23–41.
- Donaldson L and Davis JH (1991) Stewardship Theory or Agency Theory: CEO Governance and Shareholder Returns. *Australian Journal of Management* 16(1): 49–64.
- Dreher A and Vaubel R (2004) The Causes and Consequences of IMF Conditionality. *Emerging Markets Finance and Trade* 40(3): 26–54.
- Drezner DW (2007) *All politics is global: Explaining international regulatory regimes*. Princeton: Princeton University Press.
- Eckhard S and Ege J (2016) International bureaucracies and their influence on policy-making: A review of empirical evidence. *Journal of European Public Policy* 23(7): 960–978.
- Eckhard S and Jankauskas V (2019) The politics of evaluation in international organizations: A comparative study of stakeholder influence potential. *Evaluation* 25(1): 62–79.

- Eckhard S and Jankauskas V (2020) Explaining the political use of evaluation in international organizations. *Policy Sciences* 53(4): 667–695.
- Eckhard S, Patz R and Schmidt S (2019) Reform efforts, synchronization failure, and international bureaucracy: The case of the UNESCO budget crisis. *Journal of European Public Policy* 26(11): 1639–1656.
- Edelenbos J and Klijn EH (2007) Trust in Complex Decision-Making Networks: A theoretical and empirical exploration. *Administration and Society* 39(1): 25–50.
- Ege J (2020) What International Bureaucrats (Really) Want: Administrative Preferences in International Organization Research. *Global Governance: A Review of Multilateralism and International Organizations* 26(4): 577–600.
- Ege J and Bauer MW (2013) International Bureaucracies from a Public Administration and International Relations Perspective. In: Reinalda B (ed) *Routledge Handbook of International Organization*: London: Routledge, pp. 135–148.
- Elsig M (2010) Principal–agent theory and the World Trade Organization: Complex agency and ‘missing delegation’. *European Journal of International Relations* 17(3): 495–517.
- Enkler J, Schmidt S, Eckhard S, et al. (2016) Administrative Styles in the OECD: Bureaucratic Policy-Making beyond Formal Rules. *International Journal of Public Administration* 40(8): 637–648.
- Fforde A (2019) Yes, but what about the authority of policy analysts? A commentary and discussion of Perl et al., ‘Policy-making and truthiness: Can existing models cope with politicized evidence and willful ignorance in a post-fact world?’. *Policy Sciences* 52(1): 153–169.
- Genschel P and Zangl B (2014) State Transformations in OECD Countries. *Annual Review of Political Science* 17(1): 337–354.
- Goritz A, Schuster J, Jörgens H, et al. (2020) International Public Administrations on Twitter: A Comparison of Digital Authority in Global Climate Policy. Online first. *Journal of Comparative Policy Analysis: Research and Practice*: 1–25.
- Gould ER (2003) Money Talks: Supplementary Financiers and International Monetary Fund Conditionality. *International Organization* 57(3): 551–586.
- Gould ER (2006) Delegating IMF conditionality: Understanding variations in control and conformity. In: Hawkins DG, Lake DA, Nielson DL and Tierney MJ (eds) *Delegation and Agency in International Organizations*: Cambridge: Cambridge University Press, pp. 281–311.
- Graham ER (2014) International Organizations as Collective Agents: Fragmentation and the Limits of Principal Control at the World Health Organization. *European Journal of International Relations* 20(2): 366–390.

- Green JF and Colgan J (2013) Protecting Sovereignty, Protecting the Planet: State Delegation to International Organizations and Private Actors in Environmental Politics. *Governance* 26(3): 473–497.
- Grigorescu A (2010) The Spread of Bureaucratic Oversight Mechanisms across Intergovernmental Organizations. *International Studies Quarterly* 54(3): 871–886.
- Gutner T (2005a) Explaining the Gaps between Mandate and Performance: Agency Theory and World Bank Environmental Reform. *Global Environmental Politics* 5(2).
- Gutner T (2005b) World Bank Environmental Reform: Revisiting Lessons from Agency Theory. *International Organization* 59(3): 773–783.
- Gutner T and Thompson A (2010) The politics of IO performance: A framework. *The Review of International Organizations* 5(3): 227–248.
- Hanrieder T (2015) *International Organization in Time: Fragmentation and Reform*. Oxford: Oxford University Press.
- Hardin R (2006) *Trust*. Cambridge, UK: Polity Press.
- Harrari YN (2018) Moving beyond nationalism: Three global problems create a need for loyalty to humankind and to planet. Available at: <https://www.economist.com/the-world-ahead/2018/12/17/moving-beyond-nationalism> (accessed 27 January 2020).
- Hawkins D, Lake DA, Nielson DL, et al. (2006a) Delegation under anarchy: states, international organizations, and principal-agent theory. In: Hawkins DG, Lake DA, Nielson DL and Tierney MJ (eds) *Delegation and Agency in International Organizations*: Cambridge: Cambridge University Press, pp. 3–38.
- Hawkins DG and Jacoby W (2006) How agents matter. In: Hawkins DG, Lake DA, Nielson DL and Tierney MJ (eds) *Delegation and Agency in International Organizations*: Cambridge: Cambridge University Press, pp. 199–228.
- Hawkins DG, Lake DA, Nielson DL, et al. (eds) (2006b) *Delegation and Agency in International Organizations*. Cambridge: Cambridge University Press.
- Heinkelmann-Wild T and Jankauskas V (2019) International Organizations Under Fire: Explaining Communicative Responses to Member State Contestation. Presented at the European Consortium for Political Research (ECPR).
- Heinkelmann-Wild T and Jankauskas V (2020) To Yield or Shield? Comparing International Public Administrations' Responses to Member States' Policy Contestation. Online first. *Journal of Comparative Policy Analysis: Research and Practice*: 1–17.
- Heldt E (2017) Regaining control of errant agents? Agency slack at the European Commission and the World Health Organization. *Cooperation and Conflict* 52(4): 469–484.
- Heldt EC (2018) Lost in internal evaluation? Accountability and insulation at the World Bank. *Contemporary Politics* 48(5): 1–20.

- Hinterleitner M, Sager F and Thomann E (2016) The politics of external approval: Explaining the IMF's evaluation of austerity programmes. *European Journal of Political Research* 55(3): 549–567.
- Honig D (2019) When Reporting Undermines Performance: The Costs of Politically Constrained Organizational Autonomy in Foreign Aid Implementation. *International Organization* 73(1): 171–201.
- Hooghe L and Marks G (2015) Delegation and pooling in international organizations. *The Review of International Organizations* 10(3): 305–328.
- Hooghe L, Marks G, Lenz T, et al. (2017) *Measuring International Authority: A Postfunctionalist Theory of Governance, Volume III*. Oxford: Oxford University Press.
- Huber JD and Shipan CR (2002) *Deliberate Discretion? The Institutional Foundations of Bureaucratic Autonomy*. Cambridge University Press.
- ILO (2016) *Compendium of rules applicable to the Governing Body of the International Labour Office*.
- ILO (2017) ILO Evaluation Policy: GB.331/PFA/8.
- ILO (2018a) Core Voluntary Funding (RBSA) for ILO development cooperation. Available at: https://www.ilo.org/wcmsp5/groups/public/---dgreports/---exrel/documents/publication/wcms_317826.pdf (accessed 16 July 2019).
- ILO (2018b) *Staff Regulations*.
- ILO (2019a) Governing Body. Available at: <https://www.ilo.org/gb/lang--en/index.htm> (accessed 16 July 2019).
- ILO (2019b) Independent Oversight Advisory Committee. Available at: <https://www.ilo.org/public/english/edmas/ioac/index.htm> (accessed 16 July 2019).
- IMF (2015) Terms of Reference for the Independent Evaluation Office of the International Monetary Fund.
- IMF (2017) Distribution of professional and managerial staff by nationality. Available at: <https://www.imf.org/external/pubs/ft/ar/2017/eng/assets/ar17-web-table-3-1.pdf> (accessed 16 July 2019).
- IMF (2018) The External Audit Committee (EAC). Available at: <https://www.imf.org/external/np/eac/eng/eac.htm> (accessed 16 July 2019).
- IMF (2019a) How the IMF Makes Decisions. Available at: <https://www.imf.org/en/About/Factsheets/Sheets/2016/07/27/15/24/How-the-IMF-Makes-Decisions> (accessed 16 July 2019).
- IMF (2019b) IMF Executive Directors and Voting Power. Available at: <https://www.imf.org/external/np/sec/memdir/eds.aspx> (accessed 16 July 2019).
- IMF (2019c) Rules and Regulations of the International Monetary Fund: N--Staff Regulations. Available at: <https://www.imf.org/external/pubs/ft/bl/rr14.htm> (accessed 16 July 2019).

- IMF (2019d) Where the IMF Gets Its Money. Available at: <https://www.imf.org/en/About/Factsheets/Where-the-IMF-Gets-Its-Money> (accessed 16 July 2019).
- Jankauskas V (2019) Trust-Control Dynamics in International Organizations: Presented at the Workshop of the DFR-Research Unit “International Public Administration”, Konstanz.
- Jankauskas V (2021) Delegation and Stewardship in International Organizations: Accepted for publication. *Journal of European Public Policy*. DOI: 10.1080/13501763.2021.1883721.
- Jankauskas V, Bayerlein L and Knill C (2021) Is Control Always Better Than Trust? Determinants of Organizational Agency in Global Public Policy. Under review (as of January 2021). *International Organization*.
- Jankauskas V and Eckhard S (2019) International Bureaucracies as Strategic Actors: How the Better Regulation Reform Strengthens the European Commission. *Politische Vierteljahresschrift* 60(4): 681–699.
- Jankauskas V, Eckhard S and Ege J (in press) The Method of Structured, Focused Comparison. In: Badache F, Maertens L and Kimber LR (eds) *Introduction to International Organization Research Methods*: Ann Arbor: Michigan University Press.
- Johnson T (2013) Institutional Design and Bureaucrats’ Impact on Political Control. *The Journal of Politics* 75(1): 183–197.
- Johnson T and Urpelainen J (2014) International Bureaucrats and the Formation of Intergovernmental Organizations: Institutional Design Discretion Sweetens the Pot. *International Organization* 68(01): 177–209.
- Jörgens H, Kolleck N and Saerbeck B (2016) Exploring the hidden influence of international treaty secretariats: Using social network analysis to analyse the Twitter debate on the ‘Lima Work Programme on Gender’. *Journal of European Public Policy* 23(7): 979–998.
- Kaltenthaler K, Anderson CJ and Miller WJ (2010) Accountability and Independent Central Banks: Europeans and Distrust of the European Central Bank. *Journal of Common Market Studies* 48(5): 1261–1281.
- Kassim H and Menon A (2003) The principal-agent approach and the study of the European Union: Promise unfulfilled? *Journal of European Public Policy* 10(1): 121–139.
- Keohane RO (1984) *After hegemony: Cooperation and discord in the world political economy*. Princeton University Press.
- Kiewiet DR and McCubbins MD (1991) *The Logic of Delegation: Congressional Parties and the Appropriations Process*. Chicago: University of Chicago Press.
- Knill C (2001) *The Europeanisation of national administrations: Patterns of institutional change and persistence*. Cambridge: Cambridge University Press.
- Knill C, Bayerlein L, Enkler J, et al. (2019) Bureaucratic influence and administrative styles in international organizations. *Review of International Organizations* 14(1): 83–106.

- Knill C, Eckhard S and Grohs S (2016) Administrative styles in the European Commission and the OSCE Secretariat: Striking similarities despite different organizational settings. *Journal of European Public Policy* 23(7): 1057–1076.
- Knill C, Enkler J, Schmidt S, et al. (2017) Administrative Styles of International Organizations: Can we find them, do they matter? In: Bauer MW, Knill C and Eckhard S (eds) *International Bureaucracy: Challenges and Lessons for Public Administration Research*. London: Palgrave Macmillan UK.
- Koremenos B, Lipson C and Snidal D (2001) The Rational Design of International Institutions. *International Organization* 55(4): 761–799.
- Kramer RM and Cook KS (eds) (2007) *Trust and Distrust in Organizations: Dilemmas and Approaches*. New York: Russell Sage Foundation.
- Kramer RM and Tyler TR (eds) (1996) *Trust in organizations: Frontiers of theory and research*. Thousand Oaks, CA, US: Sage Publications.
- Lake DA and McCubbins MD (2006) The logic of delegation to international organizations. In: Hawkins DG, Lake DA, Nielson DL and Tierney MJ (eds) *Delegation and Agency in International Organizations*: Cambridge: Cambridge University Press, pp. 341–368.
- Lall R (2017) Beyond Institutional Design: Explaining the Performance of International Organizations. *International Organization* 71(2): 245–280.
- Liese A and Weinlich S (2006) Die Rolle von Verwaltungsstäben internationaler Organisationen: Lücken, Tücken und Konturen eines (neuen) Forschungsgebiets. In: Bogumil J, Jann W and Nullmeier F (eds) *Politik und Verwaltung*: Wiesbaden: PVS-Sonderheft 37, pp. 491–525.
- Lyne MM, Nielson DL and Tierney MJ (2006) Who delegates? Alternative models of principals in development aid. In: Hawkins DG, Lake DA, Nielson DL and Tierney MJ (eds) *Delegation and Agency in International Organizations*: Cambridge: Cambridge University Press, pp. 41–76.
- Majone G (2001) Two Logics of Delegation: Agency and Fiduciary Relations in EU Governance. *European Union Politics* 2(1): 103–122.
- Manulak MW (2017) Leading by Design: Informal influence and international secretariats. *Review of International Organizations* 12(4): 497–522.
- Martin LL (2006) Distribution, information, and delegation to international organizations: The case of IMF conditionality. In: Hawkins DG, Lake DA, Nielson DL and Tierney MJ (eds) *Delegation and Agency in International Organizations*: Cambridge: Cambridge University Press, pp. 140–164.
- McCubbins MD and Schwartz T (1984) Congressional Oversight Overlooked: Police Patrols versus Fire Alarms. *American Journal of Political Science* 28(1): 165–179.
- Miller GJ (2005) The Political Evolution of Principal-Agent Models. *Annual Review of Political Science* 8(1): 203–225.

- NATO (2016a) NATO Secretary General. Available at:
https://www.nato.int/cps/en/natolive/topics_50094.htm (accessed 16 July 2019).
- NATO (2016b) Resource Policy and Planning Board. Available at:
https://www.nato.int/cps/en/natohq/topics_67653.htm (accessed 16 July 2019).
- NATO (2017) North Atlantic Council. Available at:
https://www.nato.int/cps/ic/natohq/topics_49763.htm (accessed 16 July 2019).
- Nielson DL and Tierney MJ (2003) Delegation to International Organizations: Agency Theory and World Bank Environmental Reform. *International Organization* 57(2): 241–276.
- Nielson DL and Tierney MJ (2005) Theory, Data, and Hypothesis Testing: World Bank Environmental Reform Redux. *International Organization* 59(3): 785–800.
- Parížek M (2017) Control, soft information, and the politics of international organizations staffing. *The Review of International Organizations* 12(4): 559–583.
- Perl A, Howlett M and Ramesh M (2018) Policy-making and truthiness: Can existing policy models cope with politicized evidence and willful ignorance in a “post-fact” world? *Policy Sciences* 51(4): 581–600.
- Pevehouse J, Nordstrom T and Warnke K (2004) The Correlates of War 2 International Governmental Organizations Data Version 2.0. *Conflict Management and Peace Science* 21(2): 101–119.
- Pevehouse JCW, Nordstrom T, McManus RW, et al. (2020) Tracking organizations in the world: The Correlates of War IGO Version 3.0 datasets. *Journal of Peace Research* 57(3): 492–503.
- Pollack MA (1997) Delegation, agency, and agenda setting in the European Community. *International Organization* 51(1): 99–134.
- Pollack MA (2003) *The Engines of European Integration*. Oxford University Press.
- Pollack MA (2006) Delegation and discretion in the European Union. In: Hawkins DG, Lake DA, Nielson DL and Tierney MJ (eds) *Delegation and Agency in International Organizations*: Cambridge: Cambridge University Press, pp. 165–196.
- Rathbun BC (2008) Interviewing and Qualitative Field Methods: Pragmatism and Practicalities. In: Box-Steffensmeier JM, Brady HE and Collier D (eds) *The Oxford Handbook of Political Methodology*: Oxford: Oxford University Press, pp. 685–701.
- Rauh C and Zürn M (2020) Authority, politicization, and alternative justifications: Endogenous legitimation dynamics in global economic governance. *Review of International Political Economy* 27(3): 583–611.
- Reinalda B (2013) International organization as a field of research since 1910. In: Reinalda B (ed) *Routledge Handbook of International Organization*: London: Routledge, pp. 1–23.
- Richardson J (ed) (1982) *Policy Styles in Western Europe*. London: Allen & Unwin.

- Rihoux B and Ragin CC (eds) (2009) *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. Los Angeles, London: Sage.
- Rittberger V, Zangl B, Kruck A, et al. (2019) *International organization*. Houndmills, Basingstoke, Hampshire, UK: Red Globe Press.
- Scharpf FW (1997) *Games real actors play: Actor-centered institutionalism in policy research*. Boulder, Colo: Westview Press.
- Schlipphak B and Treib O (2017) Playing the blame game on Brussels: The domestic political effects of EU interventions against democratic backsliding. *Journal of European Public Policy* 24(3): 352–365.
- Schmidtke H (2019) Elite legitimation and delegitimation of international organizations in the media: Patterns and explanations. *Review of International Organizations* 14: 633–659.
- Schneider CQ and Wagemann C (eds) (2013) *Set-theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis*. Cambridge, New York: Cambridge University Press.
- Schroeder MB (2014) Executive Leadership in the Study of International Organization: A Framework for Analysis. *International Studies Review* 16(3): 339–361.
- Sommer M (2020) Blame-shifting in Times of Permanent Austerity: Evidence from Spain and Greece in the Eurozone Crisis. *Journal of Common Market Studies* 58(3): 562–579.
- Stacey J and Rittberger B (2003) Dynamics of formal and informal institutional change in the EU. *Journal of European Public Policy* 10(6): 858–883.
- Stone RW (2011) *Controlling Institutions*. Cambridge: Cambridge University Press.
- Tallberg J (2002) Delegation to Supranational Institutions: Why, How, and with What Consequences? *West European Politics* 25(1): 23–46.
- Tallberg J (2010) The Power of the Chair: Formal Leadership in International Cooperation. *International Studies Quarterly* 54(1): 241–265.
- Tallberg J and Zürn M (2019) The legitimacy and legitimation of international organizations: Introduction and framework. *The Review of International Organizations* 14(4): 581–606.
- Tamm H and Snidal D (2014) Rational Choice and Principal-Agent Theory. In: Weiss TG and Wilkinson R (eds) *International organization and global governance*: London: Routledge, pp. 132–143.
- Thatcher M and Sweet AS (2002) Theory and Practice of Delegation to Non-Majoritarian Institutions. *West European Politics* 25(1): 1–22.
- Trondal J, Marcussen M, Larsson T, et al. (2010) *Unpacking international organisations: The dynamics of compound bureaucracies*. Manchester: Manchester University Press.
- UN (2019) "Global problems require global solutions": UN Chief on cooperation with League of Arab States. Available at:

- https://www.youtube.com/watch?v=2SrkeIOmK8s&ab_channel=UnitedNations (accessed 27 January 2020).
- UNCHR (2016) UNHCR Policy on Evaluation: UNHCR/HCP/2016/2.
- UNHCR (1950) Statute of the Office of the United Nations High Commissioner for Refugees: General Assembly Resolution 428 (V) of 14 December 1950.
- UNHCR (2016) Rules of procedure: A/AC.96/187/Rev.8. Available at: <https://www.unhcr.org/4201eb512> (accessed 16 July 2019).
- UNHCR (2019a) Applications for Independent Audit and Oversight Committee (IAOC). Available at: <https://www.unhcr.org/5a7337d44> (accessed 16 July 2019).
- UNHCR (2019b) Global Appeal: 2019 Update.
- UNHCR (2019c) Standing Committee meetings. Available at: <https://www.unhcr.org/standing-committee-meetings.html> (accessed 16 July 2019).
- Urpelainen J (2012) Unilateral Influence on International Bureaucrats: An International Delegation Problem. *The Journal of Conflict Resolution* 56(4): 704–735.
- Van Sylke DM (2006) Agents or Stewards: Using Theory to Understand the Government-Nonprofit Social Service Contracting Relationship. *Journal of Public Administration Research and Theory* 17(2): 157–187.
- van Voorst S and Mastenbroek E (2019) Evaluations as a decent knowledge base? Describing and explaining the quality of the European Commission’s ex-post legislative evaluations. *Policy Sciences* 52(4): 625–644.
- Vanneste BS (2016) From interpersonal to interorganisational trust: The role of indirect reciprocity. *Journal of Trust Research* 6(1): 7–36.
- Vasilopoulou S, Halikiopoulou D and Exadaktylos T (2014) Greece in Crisis: Austerity, Populism and the Politics of Blame. *Journal of Common Market Studies* 52(2): 388–402.
- Vaubel R (2006) Principal-agent problems in international organizations. *The Review of International Organizations* 1(2): 125–138.
- Vaubel R, Dreher A and Soylu U (2007) Staff Growth in International Organizations: A Principal-Agent Problem? An Empirical Analysis. *Public Choice* 133(3-4): 275–295.
- Vreeland JR (2006) IMF program compliance: Aggregate index versus policy specific research strategies. *Review of International Organizations* 1(4): 359–378.
- Waterman RW and Meier KJ (1998) Principal-Agent Models: An Expansion? *Journal of Public Admin Research and Theory* 8(2): 173–202.
- Weaver C (2010) The politics of performance evaluation: Independent evaluation at the International Monetary Fund. *Review of International Organizations* 5(3): 365–385.
- Weinlich S (2014) *The UN Secretariat's Influence on the Evolution of Peacekeeping: The Internationalization of Security Politics*. Basingstoke: Palgrave Macmillan.

- Weiss M (2020) Varieties of privatization: Informal networks, trust and state control of the commanding heights. Early view. *Review of International Political Economy*. DOI: 10.1080/09692290.2020.1726791.
- Weiss M and Jankauskas V (2019) Securing cyberspace: How states design governance arrangements. *Governance* 32(2): 259–275.
- Wendt A (1992) Anarchy is what states make of it: The social construction of power politics. *International Organization* 46(2): 391–425.
- Williamson OE (1985) *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. London: Macmillan.
- Wonka A (2007) Technocratic and independent? The appointment of European Commissioners and its policy implications. *Journal of European Public Policy* 14(2): 169–189.
- Yi-chong X and Weller P (2008) 'To be, but not to be seen': Exploring the Impact of International Civil Servants. *Public Administration* 86(1): 35–51.
- Zaum D (ed) (2013) *Legitimizing International Organizations*. Oxford: Oxford University Press.
- Zürn M, Binder M and Ecker-Ehrhardt M (2012) International Authority and its Politicization. *International Theory* 4(01): 69–106.

Appendix to Framework Paper

Appendix I: Overview list of core contributions (attached to the Framework Paper)

1. Paper A

<i>Title:</i>	Delegation and Stewardship in International Organizations
<i>Year:</i>	2021
<i>Authorship:</i>	Single-authored
<i>Publication status:</i>	Accepted (DOI: 10.1080/13501763.2021.1883721)
<i>Journal (impact factor):</i>	Journal of European Public Policy (2019: 4.2)

2. Paper B

<i>Title:</i>	To Yield or Shield? Comparing International Public Administrations' Responses to Member States' Policy Contestation
<i>Year:</i>	2020
<i>Authorship:</i>	Co-authored (alphabetical order) with Tim Heinkelmann-Wild
<i>Publication status:</i>	Published (DOI: 10.1080/13876988.2020.1822144)
<i>Journal (impact factor):</i>	Journal of Comparative Policy Analysis (2019: 1.42)

3. Paper C

<i>Title:</i>	Explaining the Political Use of Evaluation in International Organizations
<i>Year:</i>	2020
<i>Authorship:</i>	Co-authored (alphabetical order) with Steffen Eckhard
<i>Publication status:</i>	Published (DOI: 10.1007/s11077-020-09402-2)
<i>Journal (impact factor):</i>	Policy Sciences (2019: 3.61)

4. Paper D

<i>Title:</i>	Is Control Always Better Than Trust? Determinants of Organizational Agency in Global Public Policy
<i>Year:</i>	2021
<i>Authorship:</i>	Co-authored (first author) with Louisa Bayerlein, Christoph Knill
<i>Publication status:</i>	Submitted, under review
<i>Journal (impact factor):</i>	International Organization (2019: 5.0)

Please note that all information related to the publication status or journal impact factor is as of February 3, 2021. Impact factors refer to 2019 Journal Citation Reports by Clarivate Analytics.

Appendix II: Overview list of non-core contributions (not attached to the Framework Paper), including abstracts

5. Non-core contribution

<i>Title:</i>	International Bureaucracies as Strategic Actors: How the Better Regulation Reform Strengthens the European Commission
<i>Year:</i>	2019
<i>Authorship:</i>	Co-authored (first author) with Steffen Eckhard
<i>Publication status:</i>	Published (DOI: https://doi.org/10.1007/s11615-019-00189-3)
<i>Journal (impact factor):</i>	Politische Vierteljahresschrift (2019: 0.44)

Abstract: The 2015 Better Regulation Reform is recognized as one of the key changes of Juncker’s European Commission, but its political implications remain understudied. Despite its appearance as a seemingly technical evaluation system, we present the reform as a political instrument that enhances the strategic actorness of the Commission, both internally and vis-à-vis the member states. Drawing on primary documents and 16 expert interviews with senior Commission officials, we demonstrate that the Better Regulation Reform enhances the Commission’s ability to act as a unified actor (internal coherence) and contributes to its ability to justify its actions vis-à-vis the member states (external robustness). The article contributes to the literature on international public administration in general and the EU in particular, as it demonstrates how institutional policies may enhance bureaucratic influence. We reveal how an international public administration can conform to member states’ demands for more accountability and transparency yet design the overall evaluation system in a way that contributes to its strategic actorness.

6. Non-core contribution

<i>Title:</i>	The Politics of Evaluation in International Organizations: A comparative Study of Stakeholder Influence Potential
<i>Year:</i>	2019
<i>Authorship:</i>	Co-authored (alphabetical order) with Steffen Eckhard
<i>Publication status:</i>	Published (DOI: https://doi.org/10.1177/2F1356389018803967)
<i>Journal (impact factor):</i>	Evaluation (2019: 1.3)

Abstract: While the political nature of evaluation is widely recognized, few attempts exist to conceptualize and compare these politics. This article develops the concept of evaluation stakeholder influence potential, which builds on four political resources for influence (agenda-setting powers, staff and budgetary resources, access to evaluation results, and access to

evaluators). These resources are measured for both member states and international public administrations in 24 United Nations organizations. We find that the administration—and not member states—have the largest influence potential in almost two-thirds of the international organizations. Our findings allow classifying them into three groups for which we expect differences in political contestation about evaluation use: two extreme-case groups (either member state or administrative dominance) and a group of contested middle cases. This finding of bureaucratic dominance reinforces literature on bureaucrats as powerful evaluation stakeholders in domestic settings and speaks to nascent research on bureaucratic influence in international organizations.

7. Non-core contribution

<i>Title:</i>	Securing cyberspace: How states design governance arrangements
<i>Year:</i>	2019
<i>Authorship:</i>	Co-authored (second author) with Moritz Weiss
<i>Publication status:</i>	Published (https://doi.org/10.1111/gove.12368)
<i>Journal (impact factor):</i>	Governance (2019: 2.9)

Abstract:

The functioning of modern societies increasingly depends on secure cyberspace. Given states' lacking capacities to protect this novel domain, governments draw on a variety of third parties for support. Yet, they face a challenge. While imposed control may limit third parties' competence, the abandonment of hierarchical control contradicts the widespread notion of national security. How do states navigate between these functional and national security imperatives to design governance arrangements? We develop a typology that combines cybersecurity problems (risks vs. threats) with governance modes (delegation vs. orchestration). This helps us to explore more than 100 cybersecurity policies across 15 different states. We find one predominant pattern. Governments delegate authority but maintain elements of hierarchical control, when they respond to threatening attacks. In contrast, governments orchestrate intermediaries by soft inducements to address risks and diffuse vulnerabilities in cyberspace. This contributes to both indirect governance scholarship and the debate on cybersecurity.

8. Non-core contribution

<i>Title:</i>	The Method of Structured, Focused Comparison (Book chapter in: Introduction to International Organization Research Methods, edited by Fanny Badache, Lucile Maertens, Leah R. Kimber)
<i>Year:</i>	2021
<i>Authorship:</i>	Co-authored (first author) with Steffen Eckhard, Jörn Ege
<i>Publication status:</i>	Accepted
<i>Journal (impact factor):</i>	Michigan University Press

Abstract: Structured, focused Comparison (SFC) allows a structured comparison of several cases (e.g. six international organizations), whereby the researcher conducts in-depth analysis within each case based on a standardized set of variables and general questions. The design not only increases the external validity of findings, but also allows for cross-case comparison and a fine-grained theoretical analysis. Typical applications are research questions focused on processes or mechanisms and those that allow for the interplay of several interdependent conditions (causal complexity).

9. Non-core contribution

<i>Title:</i>	Trust-Control Dynamics in International Organizations
<i>Year:</i>	2021
<i>Authorship:</i>	Single-authored
<i>Publication status:</i>	Under revision by the author
<i>Journal (impact factor):</i>	-

Abstract: Although trust is at the root of every delegation relationship, a thorough analysis of how trust fits into the principal-agent (P-A) framework is missing. This article refines traditional P-A approaches by introducing trust as a causal force affecting the principal's choice of control over the agent. The conceptual focus lies on delegation of authority to international organizations (IOs), specifically their bureaucracies—international public administrations (IPAs). A roadmap of indicators allows to identify the emergence, existence, and erosion of member states' trust in IPAs. Theoretically, trust explains why principals may willingly avoid controlling their agents even if that increases the risk of agency loss. Hence, the introduced trust mechanism provides a micro level perception-based explanation of IO control design and its changes over time. Empirically, qualitative content analysis demonstrates that member states' trust in the World Bank correlates with changes in its institutional design across three time periods in the 1980s.

10. Non-core contribution

<i>Title:</i>	International Organizations Under Fire: Explaining IO administrations' responses to member state contestation
<i>Year:</i>	2021
<i>Authorship:</i>	Co-authored (alphabetical order) with Tim Heinkelmann-Wild
<i>Publication status:</i>	Ready for submission
<i>Journal (impact factor):</i>	-

Abstract: Member state governments increasingly contest international organizations (IOs). Yet, we do not know much about how IO bureaucracies respond to such state contestation in public. While some IO administrations *yield* to contestation by conceding to member states, whereas others choose to *avoid* public conflict by ignoring the contestants. Moreover, while some IO bureaucracies react by *blurring* their position, others respond to contestation by *attacking* member states back. To explain this variation, we develop in this paper a theory of IO administrations' responses to member state contestation and argue that these depend on (1) the incentives provided by an IO's intergovernmental or supranational authority and (2) an IPA's constraints shaped by an IO's dependency on the contesting state. We demonstrate the plausibility of our theoretical model in two steps: We first probe by means of a medium-n, cross-case analysis whether the theorized conditions co-vary with IPAs' responses across 36 instances of IO contestation by the Trump Administration. We then conduct a congruence analysis by studying the communicative reactions by the administrations of NATO, WTO, UNHCR, and UNEP. The paper thereby contributes to nascent research on re-legitimation strategies of IOs in the face of contestation.

Please note that all information related to the publication status or journal impact factor is as of January 29, 2021. Impact factors refer to 2019 Journal Citation Reports by Clarivate Analytics.

Paper A

Delegation and Stewardship in International Organizations

This version of the article is accepted for publication (as of January 2021)

In: *Journal of European Public Policy*, single-authored

Abstract: International organizations (IOs) are driven by political-administrative interactions between member states and IO administrations. To model these interactions and understand their outcomes, scholars have predominantly, and almost exclusively, relied on agency theory. Yet, as this paper argues, delegation can also take a form of stewardship, where goal conflict and information asymmetries are low. In stewardship relationships, member states trust the IO administration, which enables softer, more informal exercise of control. Both agency and stewardship relationships are illustrated in a comparative case study of FAO and WFP. As interview data and document analysis show, while FAO exhibits agency, WFP provides an example for stewardship. The findings imply that conventional Principal-Agent assumptions should not be taken as given. Not all IO administrations are self-serving agents. The findings also provide implications on IO control and performance and call for scholarship to redirect its focus on *de facto* rather than *de jure* IO characteristics.

Keywords: agency, control, delegation, international organization, stewardship, trust

Introduction

With the growing authority of international organizations (IOs) (Hooghe et al., 2017), scholars have increasingly shifted their attention to internal IO dynamics. Importantly, IO administrations have been recognized as influential actors in their own right (Barnett & Finnemore, 1999, 2004; Bauer et al., 2017; Biermann & Siebenhüner, 2009; Eckhard & Ege, 2016). Since then, an important question has arisen – how can we model the political-administrative interaction between member states (MS) and IO administrations to better understand how IOs function?

At least since the seminal book ‘Delegation and Agency in International Organizations’ (D. G. Hawkins et al., 2006), agency theory has appeared as a dominant paradigm underlying this research (Reinalda, 2013, p. 17). At its core is the assumption of interest divergence and information asymmetry between principals and their agents. If given an opportunity, agents (IO administrations) maximize their utility at the expense of their principals (MS). Thus, agents have been traditionally pictured as ‘rogue’ opportunistic actors, driven by ‘self-interest seeking with guile’ (Williamson, 1985, p. 30). IO administrations, too, were largely studied in the context of escaping MS’ control (Conceição-Heldt, 2013), expanding resources (Copelovitch, 2010), or engaging in organizational pathologies (Barnett & Finnemore, 1999, 2004). Overall, the key concern has been how to control IOs and ensure their faithfulness (Abbott & Snidal, 1998; Copelovitch, 2010; Elsig, 2010; Grigorescu, 2010; Nielson & Tierney, 2003; Pollack, 1997).

However, empirical findings from growing research on internal IO dynamics imply that not all IOs are shaped by political-administrative contestation. Some administrations were found to willingly follow member states’ orders (Bayerlein et al., 2020) and focus on mandate implementation rather than political self-interest (Biermann & Siebenhüner, 2009; Knill et al., 2019; Yi-chong & Weller, 2008). As Ege (forthcoming, p. 13) highlighted, ‘the prominence of [IO administrations’] responsible behavior... directly challenges self-centered goal orientation as the default mode of bureaucratic behavior’. Also, some MS were found to rely on IO administrations and allow them managing control by themselves (Eckhard & Jankauskas, 2019). Overall, these findings imply that IO delegation relationships are not always defined by high goal divergence and information asymmetries as agency theory prescribes.

Despite these findings, we still lack a systematic analysis of whether and under what conditions does delegation resemble opportunism and contestation or collectivism and cooperation. The distinction is important as it affects IOs' day-to-day business and thus the overall IO performance. To that end, this paper asks how varying levels of goal conflict and information asymmetry between member states and IO administrations affect their interaction, i.e. the exercise of control.

Following Waterman and Meier's (1998, p. 195) advice to refrain from unidimensional theoretical thinking, this article utilizes stewardship theory to introduce an alternative type of delegation relationship which is based on fundamentally different premises. According to stewardship theory – which stems from management and sociology literature (Davis et al., 1997) – the principal's subordinate, the steward, behaves cooperatively and in line with the principal's interests rather than individualistic and self-serving. The steward provides necessary information if needed since more value is placed on cooperation than defection. Hence, in stewardship relationships, goal and information asymmetry is low, so trust, i.e. positive expectations about the other's behavior under the condition of risk (Edelenbos & Klijn, 2007; Luhmann, 1979), drives the interaction. This leads to less demanding, soft exercise of control over the IO administration, relying on internal oversight and informal procedures. By contrast, in agency relationships, goal and information asymmetry is high, so distrust shapes the behavioral premise and leads to contestation between the two actors, i.e. hard exercise of control characterized by scrutiny and formal arrangements. The two relationship types are conceived of as relatively stable modes of interaction, which – in line with rational choice perspective (Stacey & Rittberger, 2003, p. 869) – emerge over time through repeated interaction between MS and IO administrations.

These mechanisms are tested in a comparative case study of the Food and Agriculture Organization (FAO) and the World Food Programme (WFP) based on a most-similar-system design. The IOs share a number of similarities traditionally deemed relevant for the exercise of control, yet they differ in the extent of goal conflict and information asymmetry between MS in the executive governing bodies and IO administration management. As interviews with both actors show, these differences lead to agency in FAO and a stewardship relationship in WFP. While members of the FAO Council exercise hard control over the FAO management, WFP management enjoys soft control by the Executive Board.

The findings demonstrate that IOs' political-administrative relationships are not always contested as conventional agency studies imply. The utilization of stewardship theory offers an alternative view on delegation and provides a number of implications for IOs' day-to-day functioning, performance, and control. This article first defines the specific delegation dimension under analysis and derives the two types of relationships. It then introduces research design, empirical analysis, a discussion of findings, and conclusions.

Beyond agency: Two types of IO delegation relationships

IOs are characterized by political-administrative relationships between member states as political decision-makers and IO administrations as their subordinate task executors. At least three dimensions of such delegation relationships can be distinguished.

The first two prevailed among early studies of IOs. MS governments were treated as a collective principal, while IOs as agents were seen as unitary entities (see (1) in Figure 1) (Abbott & Snidal, 1998). Some studies conceptualized MS as multiple principals (see (2) in Figure 1), focusing on their bilateral relationships with IOs (see (2) in Figure 1) (Nielson & Tierney, 2003). Using a rather rudimentary definition of actors simplified P-A models and allowed analyzing macro-level phenomena such as why states delegate authority to supranational institutions. Yet, it also hindered a more thorough analysis of IOs' day-to-day work.

Only recently, after recognizing IO administrations as distinct actors, scholars started to open up the black box of IOs and their membership (Eckhard & Ege, 2016; Elsig, 2010; Graham, 2014; Grigorescu, 2010). Nevertheless, the definition of actors often remains fuzzy (Yi-chong & Weller, 2008). As Ege and Bauer (2013, p. 143) highlight, 'the major shortcoming in the literature is the failure to properly define the bearer of agency within IOs'.

In line with this criticism, this article narrows down the interaction framework to the third dimension, i.e. the relationship between IOs' executive governing bodies as the principal and IO administration management as the agent (see (3) in Figure 1). This delegation dimension has been largely neglected in the literature but is key to understand IOs' internal functioning (Elsig, 2010, 5f.).

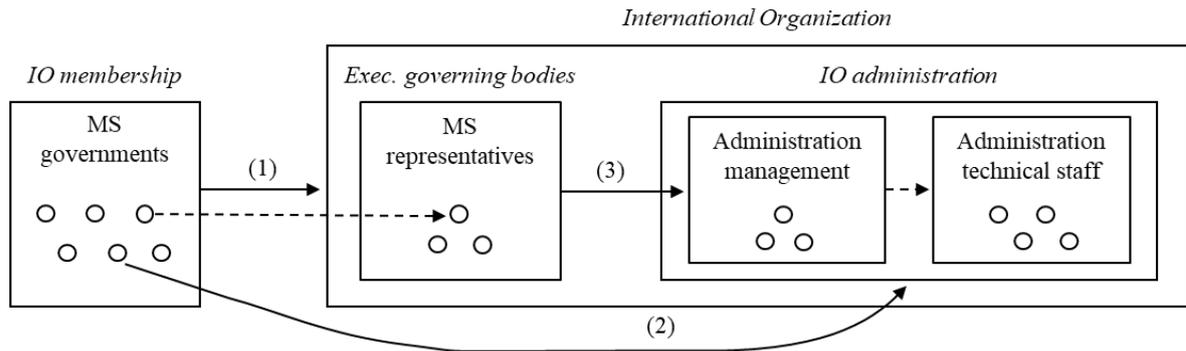


Figure 1: IO delegation dimensions. Bold arrows indicate delegation between principals and subordinates; dashed arrows indicate delegation within the actors. Points indicate single entities within actors, e.g. governments. Author's compilation based on Elsig (2010, p. 499).

Executive governing bodies (EGBs) (e.g. executive boards) consist of (elected) MS which exercise power granted by the overall IO membership.¹ EGBs review and modify IO activities and exercise control over the administration. By contrast, general governing bodies include the whole membership (e.g. the World Health Assembly), but meet rarely, usually follow EGBs' recommendations, and are not involved in daily decision-making. The management of IO administration consists of its senior staff holding leadership positions (e.g. Director-General, (deputy) directors). The management is responsible for the administration's performance and relationship with MS (see Ege & Bauer, 2013, p. 141). It is the regular interaction between these actors that drives IOs' day-to-day functioning and is the focus of this study.

Having defined the two actors under analysis, how can we explain their interaction? Essentially, agency and stewardship theories describe two extreme poles of ideal-typical hierarchical relationships. They fundamentally differ in the behavioral premise, i.e. the assumptions upon which actors act – goal divergence and information asymmetry. Such differences, in turn, affect the behavioral outcome, i.e. the exercise of control by the principal over its subordinate.

Goal conflict and information asymmetry in agency and stewardship relationships

According to agency theory, delegation relationships are based on two assumptions. First, it assumes goal conflict between the principal and the agent. Agents are understood as self-

¹ They are thus sometimes called 'proximate principals' (Elsig, (2010: 499).

interested actors with own preferences; thus, ‘there is a natural and perhaps inevitable conflict of interests between the parties’ (D. Hawkins et al., 2006, p. 24). Agents thereby have an incentive to slack, i.e. to engage in an action undesired by the principals (Cortell & Peterson, 2006; Heldt, 2017; Kiewiet & McCubbins, 1991). For example, IO executive boards may want lower administrative costs, while IO management may wish for ever-more resources.

Second, agency theory assumes high information asymmetry which favors the agent and creates opportunity to slack (D. Hawkins et al., 2006, p. 25; Kiewiet & McCubbins, 1991, p. 25). Agents have more knowledge and expertise than their principals and are thought to use it for their own benefit. For instance, IO management could justify its poor performance by shifting blame to local partners (to which MS rarely have access).

However, existing research implies that goal conflict is not always high and nor is information asymmetry. As Waterman and Meier (1998, p. 197) highlighted, ‘all political-bureaucratic relationships are not a cauldron of conflict’. Yet, their criticism remained at a theoretical level. Findings on IO administrations – as detailed in the introduction – have also implied that interest and information asymmetry varies (Barnett & Finnemore, 1999; Bayerlein et al., 2020; Biermann & Siebenhüner, 2009; Knill et al., 2019; Ege, forthcoming). Yet, if we relax the two assumptions, agency theory loses its power: ‘only when there is goal incongruence between the two is agency theory applicable’ (Miller, 2005, p. 205; Tamm & Snidal, 2014). Thus, an alternative theoretical perspective is needed.

In light of this, stewardship theory, as introduced by early works of Donaldson and Davis (1991) and Davis et al. (1997), is drawn upon to define member states-IO administration (management) relationships that are based on fundamentally different behavioral premises. First, subordinates are assumed to be collective, pro-organizational stewards, motivated ‘to act in the best interests of their principals’ (Davis et al., 1997, p. 24; Van Sylke, 2006). Therefore, goal conflict becomes low – even if there is some disagreement, it does not incentivize the steward to slack. For instance, MS often have divergent interests (Dijkstra, 2015; Nielson & Tierney, 2003). A steward-like administration would seek to balance the competing preferences and find solutions in the best interest of the membership. A self-serving agent-like administration would rather ally with selected states to strengthen its own agenda.

In both cases, actors behave rationally and care about their needs; the two theories are thus compatible in terms of rationality. But the way in which these needs are fulfilled differs:

The steward realizes the trade-off between personal needs and organizational objectives and believes that by working toward organizational, collective ends, personal needs are met. (Davis et al., 1997, p. 25)

Second, in stewardship relationships, risks associated with information asymmetry are also mitigated. The principal does not have to worry that its subordinate would hide important information and use it opportunistically. If needed, the principal can request information and the steward should cooperate. Thus, although the steward maintains its knowledge advantage, de facto information asymmetry is reduced.

Exercise of control in agency and stewardship relationships

The fundamental differences in goal and information divergence affect how actors interact with each other. The focus here is on the exercise of control. IO research usually focuses on the mere existence of control tools (Grigorescu, 2010; D. G. Hawkins et al., 2006; Hooghe et al., 2017). Yet, it is also how they are used. Generally, the exercise of control can be either hard or soft depending on the type of relationship.

Agency relationships are driven by distrust in the agent. Distrust is about one's negative expectations about the other under the condition of risk (Edelenbos & Klijn, 2007; Hoffman, 2002; Luhmann, 1979). Due to goal divergence and high information asymmetry, the principal can never be sure that the self-interested agent will remain loyal if a chance for betrayal occurs. The principal is thus incentivized to apply hard control over the agent to reduce information asymmetry and re-align its goals (D. Hawkins et al., 2006, pp. 26–31; Pollack, 1997, 108f.; Williamson, 1985). This can translate into both additional control mechanisms and/or a more stringent utilization of existing oversight. New control tools might be too costly, so MS may put additional pressure using existing ones. Additionally, hard control is associated with formal procedures along the defined rules. In the absence of trust, actors are incentivized to stick to formal arrangements to keep track of each other's commitments and avoid cheating (Farrell, 2009, p. 106). For example, MS may use primarily formal meetings and written communication to solve issues with the management. Although the principal still has to weigh the benefits of delegation against the costs of control (D. Hawkins et al., 2006, p. 9; Kassim & Menon, 2003; Lake & McCubbins, 2006), agency relationship remains inherently contested: 'there is an ever-present possibility of opportunism, unless it is curbed through controls' (Davis et al., 1997, p. 23).

By contrast, in stewardship relationships, the goal and information conflicts are low, so trust can manifest allowing the principal to be confident about its subordinate’s intentions. Thereby, soft exercise of control becomes feasible. While some scholars argued that trust reduces formal oversight (Hoffman, 2002; M. Weiss, 2020), others observed that trust mitigates the use of formal control (Sitkin & George, 2005). Changing formal mechanisms might be challenging, so MS may put less pressure on existing ones. Also, low goal divergence and information asymmetry cause fewer cheating concerns. This allows solving issues through more informal arrangements (Farrell, 2009). For instance, MS may meet with the management informally to clarify control issues prior to formal sessions.

In sum, agency relationships are based on high goal divergence and information asymmetry between the principal and a self-serving agent (Table 1). Distrust therefore drives the interaction: executive governing bodies are incentivized to exercise hard control over the IO management to reduce the risk of agency slack. Overall, agency relationships are characterized by contestation as the principal seeks to constrain the agent who, in turn, seeks to escape the principal. By contrast, stewardship relationships are based on low interest divergence and information asymmetry. This facilitates the management’s trustworthiness and allows less demanding, soft exercise of control. Overall, such relationships are characterized by cooperation instead of continuous contestation.

Agency relationship: (Principal-Agent)	Goal conflict + info. asymmetry <i>High</i>	→	Exercise of control <i>Hard</i>
Stewardship relationship: (Principal-Steward)	Goal conflict + info. asymmetry <i>Low</i>	→	Exercise of control <i>Soft</i>

Table 1: Differences in agency and stewardship relationships.

Research Design

To test the theorized dynamics empirically, a comparative case study of FAO and WFP is conducted. Case selection is based on a most-similar-system design (MSSD). The principle is to select cases similar on relevant factors but different in the causally relevant variable. In this regard, FAO and WFP share similarities usually deemed relevant for the exercise of control yet differ in the prevailing levels of goal conflict and information asymmetry.

Regarding the principal's characteristics, both IOs as UN organizations have similar governance structures with functioning EGBs – the FAO Council with its sub-committees and the WFP Executive Board. Both bodies hold two to three formal annual meetings for supervision and steering. Regarding decision-making procedures (Cortell & Peterson, 2006, p. 261), both bodies seek for consensus or use unweighted voting otherwise. They consist of 49 (FAO) and 36 (WFP) regionally diverse MS elected for three years and thereby show comparable levels of preference heterogeneity (Conceição-Heldt, 2013, p. 26; Nielson & Tierney, 2003, p. 248).

Regarding the agent's characteristics, FAO and WFP have partially autonomous secretariats acting upon common UN staffing and financial policies. At least formally, both administrations enjoy similarly high policy autonomy (Lall, 2017)² and are capable of developing their own preferences (Cortell & Peterson, 2006, p. 257). Both administrations have around 15.000 employees (Appendix II) and similarly built management teams, thus sharing a similar level of fragmentation at the management level (Graham, 2014, p. 369).

Regarding the principal's ability to control and related costs, the level of formal control applied by the EGBs is similar in both IOs. P-A scholars highlight three types of control: oversight and monitoring, screening and selection, and sanctioning (Conceição-Heldt, 2013; Graham, 2014; D. G. Hawkins et al., 2006; Heldt, 2017; Nielson & Tierney, 2003; Pollack, 1997). These control types include different mechanisms (e.g. audits, evaluations, staff approvals) which – as measured based on primary IO documents – are all in place in FAO and WFP (Appendix I). MS thus have the same tools to constrain the management. In turn, the management in both IOs has a similar level of discretion (Cortell & Peterson, 2006, p. 257).

Regarding the trade-off between control and credibility (Majone, 2001) or competence (Abbott et al., 2020), both IOs operate in the same issue field and represent task-specific IOs (Lenz et al., 2015). FAO and WFP aim at ending hunger, malnutrition and improving food security. WFP delivers humanitarian food assistance, but increasingly engages in development projects (WFP, 2020). FAO provides development assistance to affected countries, but increasingly widens its humanitarian activities, e.g. immediate support to crisis-hit families (FAO, 2020). Consequently, the majority of staff in both administrations work outside the headquarters

² Based on Lall's (2017) measurement of de jure policy autonomy (agenda-setting, decision-making powers, and access to non-state funding), FAO and WFP score 5 and 5,67 accordingly (on the scale from 1 to 6).

(Appendix II). Operating in the same issue field, they thereby face similar needs for political independence.

If we take basic P-A assumptions as given, the described similarities should lead to relatively similar exercise of control in both organizations. Yet, as will be shown in the empirical section, the two organizations fundamentally differ in goal divergence and information asymmetry between EGBs and the IOs' management. While FAO shows high goal conflict and information asymmetry, low values of divergence are observed in WFP.

To assess these values and whether they lead, as theorized, to different exercise of control, 23 in-depth semi-structured interviews were conducted. Both sides in both IOs were interviewed: MS representatives from the EGBs (mostly ambassadors) and administration management staff (e.g. (deputy) directors). The interviewees were asked to describe their relationship with each other; specific questions were raised on goal conflict, information asymmetry, level of trust, exercise of control, and contestation within their IOs. To gain contextual knowledge on political-administrative interaction in the UN, five background interviews with institutions like the UN Secretariat were conducted. Appendix III provides the full list of interviews.

Table 2 summarizes observable implications for each indicator. These observable patterns are conceived of as relatively stable manifestations of ideal-typical relationships, which emerge over time through continuous interaction between MS and IO management. Interviewee statements should thus confirm the continuity of revealed dynamics. Instances of deviant indicator values should remain sporadic and rare.

	Indicator	Value	Observable Implication
Agency relationship	Goal conflict	High	Management follows own interests, even if they conflict with MS interests; management considers selected MS interests
	Information asymmetry	High	Management is hardly accessible to MS; limited information is shared
	Exercise of control	Hard	MS scrutinize internal oversight tools, include external sources; primarily formal arrangements used to solve issues
Stewardship relationship	Goal conflict	Low	Management follows MS interests, even if they conflict with own interests; management considers all MS interests
	Information asymmetry	Low	Management is easily accessible to MS; high amount of information is shared
	Exercise of control	Soft	MS rely on internal oversight, use fewer external sources; primarily informal arrangements used to solve issues

Table 2: *Observable implications.*

Empirical Analysis

First, differences in goal divergence and information asymmetry between FAO and WFP are explored. Second, how these differences affect the political-administrative interaction itself is investigated, looking at the exercise of control in IOs' daily business.

Differences in goal conflict and information asymmetry in FAO and WFP

Starting with FAO, high values of goal divergence and information asymmetry between MS and the IO management were observed.

Regarding goal conflict, the FAO management was found to have a clear agenda along its own interests and continuous clashes with MS were reported. MS in the FAO Council and its sub-committees described the management as 'very opportunistic', pursuing their 'own ideas' rather than following the Council's directions (MS3).³ As one ambassador explained, the management 'does not want to address real issues' and manipulates policymaking, adding that

³ Interviews are quoted using labels (see Online Appendix III). 'MS' indicates interviews with member states, while 'IPA' stands for international public administration, indicating interviews with IO management staff.

(s)he has ‘never seen it as negative as... with FAO’ (MS1). As another ambassador put it, there is ‘polarization between management and... membership’ (MS5). The management itself admitted the discrepancy, saying ‘it is hard to understand what exactly member states’ concerns and interests are’ (IPA5). As one management official stated, the conflict might initially be due to factual errors, yet over time it became an issue of conflicting preferences – primarily ‘priorities of the management, not of the countries’ (IPA2).

Furthermore, the delegation relationship was found to be characterized by high information asymmetry creating tensions between the actors. FAO Council members argued that it is ‘very difficult’ for them ‘to see what is happening [in the administration]’ (MS1, MS8) and for what purposes: ‘is it for the future of the career of their [management’s] number one [Director-General]?’ (MS3). The exchange of information was found to be poor as the management would provide information often too late or in English only, or just send ‘a pile of documents’ (MS3, MS2, MS4, MS8). MS thus argued that they are kept ‘at arm's length’ (MS2, MS4). As one ambassador summarized:

As a Member, you quite often feel only at the surface and you do not know how much they [management] listen... Everything can be manipulated... In FAO, you do not really know, because there are so much instructions coming from the fourth floor [management]... and you always have suspicion... (MS2)

The management was well aware of its informational advantage. It even acknowledged that ‘clearly, the management can influence the debate just by structuring and focusing on what it wants to tell the member countries’ (IPA5). One officer admitted: ‘if it is too controversial... you try to frame it differently...’ (IPA1). For instance, in oversight reports, the management could ‘play up the results which support the direction the management wants to go and downplay the others’ (IPA5). These are exactly the concerns that come with high information asymmetry in agency relationships.

Overall, the observed goal conflict and high information asymmetry in FAO manifests in distrust that drives interaction. Distrust means the lack of MS’ confidence that the management will not harm their interests if it could (Hoffman, 2002). As interview data indicated, MS have ‘suspicion’ (MS2), ‘lack of trust’ (MS3), and ‘trust deficiency’ (MS8) in the FAO management: it is more ‘a Machiavellian approach... than a good nature of trust approach’ (IPA5). This corresponds with opportunism and high risk of agency slack in agency relationships.

The opposite values were observed in WFP. Regarding preference divergence, the management was found to actively accommodate ‘the wishes of the membership’ (MS5) and be responsive to the Executive Board’s requests (MS7). The Board members might be ‘very critical and straightforward’ (MS7), yet, as one ambassador highlighted, ‘even if there are these kind of conflicts, we all know that we are all working for the same purpose’ (MS7). Another ambassador from another continent confirmed that ‘the general relationship is strong’ (MS9). The WFP management also described it as ‘excellent’ (IPA6) and provided several examples where it followed MS’ recommendations despite initially different perspectives (MS5). This, in turn, indicates low interest divergence between the Board and the management.

Regarding information asymmetry, the WFP management was found to be easily accessible to the Board and willing to swiftly provide requested information. As one ambassador illustrated:

Whatever issue we [MS] have, even a technical issue, we can just take the phone, dial the number, and the director will respond, and we can fix an appointment in the day after. (MS7)

Other Board members gave analogous examples (MS9). In contrast to FAO, the WFP management was described as ‘modern, open, democratic leadership’ (MS7). Drawing a direct comparison to FAO, one ambassador said:

It is a completely different story in WFP, where we [MS] are in front and have informal orientations... for all membership to come and to be briefed... In FAO, it is more like you come to a meeting and just a short time before you get a pile of documents. (MS2)

The information exchange between the Board and the management was found to be intense. Representatives of the Board’s five regional lists meet every 4-6 weeks. The management provides information for these meetings and then informs other Board members about the issues discussed to avoid information biases among the membership (IPA6). Thereby, accessible management and intense communication reduces the information asymmetry.

Overall, the WFP case seems to deviate from conventional agency assumptions, showing low values of goal conflict and information asymmetry. This enabled trust to manifest in the relationship. MS expressed their ‘confidence and trust’ (MS7) in the management and praised its integrity (MS9). The management highlighted that interest accommodation and intense information exchange ‘helped to build more trust, because we were open from the get-on...

and this helped throughout the years, this is an evolution' (IPA6). The emphasized continuity also confirms the relatively stable nature of the observed behaviour patterns.

Noteworthy, although interviewees highlighted preference heterogeneity among MS in both organizations, a clear difference was reported on how the management deals with it. In FAO, the latter was claimed to exploit MS' differences: 'FAO leadership is trying to instrumentalize the regional polarization [among membership] for the advancement of its own political purposes' (MS5). The management would engage in the 'manipulation of member states' and 'approach specific representatives to change their minds' (MS6). FAO management itself acknowledged that 'tensions among member states can negatively affect their trust' (IPA3, IPA2). In WFP, MS agreed that bilaterally the management shows more attention to donors (MS5), yet in the Executive Board, it seeks to treat the membership equally. The management was aware of potential impartiality issues and chose to 'to reach, support everyone equally, give [Board members] the same information' (IPA6). One management officer would intervene if other administration officials ignored Board members from non-donor lists (IPA6). This shows that despite MS heterogeneity, goal conflict with the management should not be taken as given. MS may distrust each other but show confidence in the management's pro-organizational rather than individualistic nature.

Differences in exercise of control in FAO and WFP

How the observed differences in goal and information divergence affect the interaction itself? Since same control tools exist in both IOs (Appendix I), the focus lies on how these tools are utilized. Hard exercise of control was found to be prevalent in FAO, which was linked to the described distrust in the management and the resulting suspicion of FAO's internal oversight (MS1, MS2, MS3, MS4). For instance, talking about evaluation and audit, MS argued that produced reports 'are not sufficiently neutral or critical because they are influenced by the senior management' (MS7). The latter itself acknowledged that MS oftentimes pushed for more stringent control 'to reduce management influence' (IPA1). This translated into a more intense and careful reading of oversight reports, longer discussions on findings and recommendations, and a more active use of external information. As one ambassador summarized:

We really make sure that these oversight issues are looked at carefully. We are in permanent exchange with oversight organs, we ask the questions that have to be asked and which are sometimes painful for the management... The oversight role that membership

plays is very attentive, very active, and if you wish also very tough... sometimes oversight is the only possible instrument we have to contain or confront them [management]. (MS5)

Another ambassador claimed that MS often 'seek for other resources within the organization' to 'get all the stories, make own evaluation of evaluation' (MS1); for example, they would focus more on external audits (MS4).

FAO management also noted that 'the pressure on the management is always there' (IPA3) and MS 'ask very specific questions and confront management responses' (IPA4). Discussions of oversight reports often lead to 'a very long follow up process' with MS 'using information they got from other sources' (IPA2). MS would put the management 'on a podium' (IPA5) and engage in 'very technical discussions about validity of data, use of indicators' (IPA4). This did not seem to be isolated incidents, but rather a stable mode of control: 'in FAO, there is a history with member states' as they have a 'habit' in reading and scrutinizing oversight reports (IPA2).

In addition to this, formal procedures were most prevalent when dealing with control-related issues. The informal interaction was found to be low. One ambassador even said that in the recent past 'senior management members were not authorized to have direct, private interaction with Permanent Representatives' (MS7). The management, on its part, argued that MS were 'not frequently seeking out' consultation (IPA5). Several examples illustrated these dynamics. Recently, Director-General implemented a management's restructuring reform without ex ante MS' approval. This infuriated the Council who had reform-related concerns. Instead of clarifying these issues informally, the management called the Council to formally discuss whether approval is necessary (MS5, MS6, IPA6). In another case, MS complained about delays in the management's reports, so they set a deadline and threatened to otherwise take respective items off the agenda (MS1). Such formal procedures were said to be 'very intense' and less efficient (MS4): 'the degree of redundancy, and duplication and triplication of discussions is a nightmare' (MS5). Overall, control-related interaction with the management reflected the contested nature of agency relationships: 'sometimes we [MS] manage to catch the management off-guard and we actually enjoy that' (MS4).

Shifting to WFP, a softer exercise of control was observed. Control processes were found to be 'very smooth' with issues 'addressed immediately' (MS6). MS expressed their confidence in the 'independent analysis' of internal oversight (MS9) and 'comprehensive control... and transparent communication' by the management (MS5). For example, one ambassador recalled incidents with food quality and delivery. The management informed the Board and offered 'to

contact the investigating unit' for more information (MS5). This ensured MS that the undesired performance (agency slack) is taken care of, so no additional pressure was needed: 'when it comes to things that should not happen, they [management] play it open' (MS5, MS4). The COVID-19 situation provides another example. Due to the pandemic, staff was unable to conduct usual oversight. Yet, the Board demonstrated its understanding:

We are in a permanent discussion with the WFP [management] and... the membership is willing to give more leeway and discounts... We have to give the organization a bit more flexibility in these challenging times. (MS5)

The ambassador then drew a direct comparison to FAO, illustrating the theorized mechanism when goal conflict and high information asymmetry lead to distrust affecting the exercise of control:

There is a clear difference in how WFP and FAO are dealing with issues... that affects the tolerance that member states have towards them... if you have to be afraid that they [management] try to hide something from you and that if you are not super attentive you might overlook something, you are a bit distrustful from the very beginning. (MS5)

WFP management also agreed that MS 'rely quite a lot' on internal oversight, so the Board expressed less of a need for external information (IPA9). The latter was described as 'very actively engaged' and committed 'across all geographical lines' which helped to make oversight more efficient (IPA9, IPA7, IPA8).

Moreover, more informal ways of dealing with oversight prevail in WFP. The management was found to have 'a much stronger practice of having informal contact with member states' (MS2, MS9). Their interaction is 'less formalized [than in FAO]' (MS5) and 'very easy' (MS7), meaning that agreements can be reached prior to formal meetings: 'you could see what is coming, you could digest it' (MS2). As one MS representative clarified:

Before every annual session, the Executive Director would meet with list countries informally to discuss agenda items and solve potential issues. This interaction makes things run very smoothly. (MS6)

A more specific experience was shared by another ambassador from another continent:

When it came to the issue of the WFP strategic plan... before bringing it to the Executive Board for a formal discussion and adoption, we had seven or eight rounds of informal

consultations... We had a good impression and a positive feeling that our view counts, we had the confidence and trust... Later, that plan was officially presented at the Board and adopted with some slight comments only... This is the way how WFP operates. (MS7)

WFP management also highlighted the intensity of informal communication (IPA8) and gave examples of informal practices (e.g. weekly briefings or off-the-record discussions) that help dealing with MS (IPA8, IPA6, IPA7). Different to FAO, the established trust in the relationship allowed smoother, less formalized interaction:

Trust by member states makes life easier. There is less discussion time, less hashing out the issues, less stress for both sides – less negative communication. (IPA6)

While this by no means made control in WFP unnecessary, low goal and information asymmetry allowed softer exercise by the Executive Board, relying on internal oversight, restraining from external information, and using more informal arrangements.

Discussion

Summarizing the findings above, the political-administrative interaction in FAO reflects a continuous contestation between the two actors. Due to high interest divergence and information asymmetry, members states have little confidence in the management and thus exercise hard control over the latter. They scrutinize FAO's internal oversight tools, refer to external sources, and rely primarily on formal procedures to solve on-going issues with the management. This resembles agency relationship, where MS in the Council seek to constrain the FAO management who seeks to escape the MS. The opposite is true of WFP which demonstrates a stewardship relationship. Given low goal conflict and de facto reduced information asymmetry, the Board members rely on the management as a steward and accordingly soften their control, relying on internal oversight and more informal means of interaction.

Hence, in line with the theorized mechanism, the different patterns in goal conflict and information asymmetry led to different modes of control exercise in the two IOs, which otherwise share a number of similarities. Using qualitative interview data allowed observing this link directly. A number of 'smoking gun' evidence statements were reported; for instance, in FAO:

Basically, there is a lack of trust and so we [MS] read the reports word by word, line by line, punctuation point by punctuation point. (MS6)

Overall, the level of agreement among interviewed MS and management officials within the two IOs is remarkably strong. Also, both IOs are based in Rome, so MS representatives often sit in both governing bodies. This allowed most interviewees to compare both IOs directly, allowing to control for respondents' personal characteristics. Furthermore, one might have expected the management to overstate its 'good' relationships with MS, yet FAO administration was aware of the prevailing premises of distrust. Importantly, most interviewees were longstanding officials, thus able to provide insights on long-term relationship dynamics rather than a snapshot of the situation. They confirmed the relative stable nature of these relationships, independent from staff rotation or nationality. As one ambassador described it for WFP:

We often say that leadership is closely linked to a person... [yet] I know WFP for more than 20 years and I knew a number of Executive Directors and they had different cultural, political backgrounds, personalities, but one thing was common – they all wanted to run the organization in a transparent and democratic way. (MS7)

Since interviewing all MS or management officials was not feasible, generalizations had to be made on statements by a limited number of interviewees. While this is common in qualitative research, the validity of data was increased by reaching out to MS from different regions and speaking with both senior- and middle-level management officials (see Appendix III).

One could argue that WFP's exclusive reliance on voluntary contributions⁴ constitutes hard control and the management shows accommodation out of opportunistic self-interest to avoid funding cuts (cf. Miller, 2005, p. 209). If that was the case, we would observe the management strategically focusing on main donors in the Executive Board and complaints of discrimination by the recipient states. However, interviewees from different country lists (see Appendix III) including both recipient and donor states consistently confirmed low goal conflict and information asymmetry with the management. As one recipient country stated, 'management reacts well to [all] member states... otherwise there will be a lack of confidence in them' (MS4). While differences in voluntary contributions might affect bilateral management-donor

⁴ For FAO, voluntary contributions constitute 61% of its total budget (Online Appendix II).

relationships, they are of less significance for the management's relationship with the EGB as a collective principal.

Can these findings be transferred to other IOs or other contexts? So far, the generalizability is limited to (larger) IOs with EGBs and administrations with distinct authority. However, these are also the 'usual suspects' in IO research. As the general models of agency and stewardship relationships are in principle applicable to all types of hierarchical relationships, future research could tailor the theorized modes of interaction to study other governance dimensions such as bilateral donor-agency relationships or IO-NGO partnerships.

Conclusion

Starting from the premise that not all IO delegation is driven by opportunism and contestation, the article differentiated between two ideal-types of relationships between member states and IO administration. In conventional agency relationships, high goal divergence and information asymmetry were showed to manifest as distrust in the IO management and lead to hard exercise of control by the IO's executive governing body (the FAO case). Yet, the study also introduced stewardship as a so far neglected, alternative delegation relationship that is based on low goal divergence and information asymmetry. As demonstrated in the WFP case, this enabled member states' trust in the management and led to a softer exercise of control.

The findings imply that basic agency assumptions should indeed be treated as variables rather than constants (cf. Waterman & Meier, 1998) and IO scholars should refrain from treating IO administrations as 'rogue' agents per se. Instead of exclusively relying on agency theory, scholars should investigate the conditions under which different delegation relationships emerge and what further effects on IO control and performance they have.

To be sure, this paper did not aim at explaining why the two IOs developed different levels of preference and information divergence. WFP management efforts to reduce information asymmetry and signal interest accommodation appear as deliberate trust-building strategies which could be employed by other administrations too. Research on IO administrations, management culture and reputation (Bauer et al., 2017; Knill et al., 2019) could, thus, further explore how stewardship relationships emerge.

This is important considering that different interaction modes seem to be associated with IOs' overall performance. Softer control is considered to enhance the subordinate's competences (Abbott et al., 2020; Lall, 2017) and more informal procedures increase decision-making efficiency (cf. Farrell, 2009, p. 106). Indeed, scholars oftentimes highlight FAO's poor performance and praise WFP for its efficiency (Lall, 2017, p. 248). However, conventional agency approaches remain skeptical since lower control is associated with higher risk of agency slack (Kassim & Menon, 2003; Lake & McCubbins, 2006). Yet, once stewardship relationships are established, the negatively inversed dilemma between (1) less control and higher agent competence (but high risk of slack) or (2) more control and less competence (but low risk of slack) disappears (cf. Abbott et al., 2020).

Moreover, IO performance failures are usually associated with either issues among member states (Nielsen & Tierney, 2003) or bureaucratic pathologies and rent-seeking behavior by IO administrations (Barnett & Finnemore, 1999). Yet, goal conflict is unlikely to be reduced if MS are unable to communicate their interests to the management but also if the management is unwilling to accommodate these interests. To better understand IO policy-making, future research should thus focus on the political-administrative interaction dynamics rather than looking at either the principal or the agent.

Finally, regarding IO oversight, the findings imply that the same tools for control are used differently depending on the member states' relationship with IO administration. Which control tools exist may depend on different factors, some even unrelated to the administration's performance. For instance, control tools might be introduced due to states' domestic accountability needs, norm diffusion, or issue linkages (see Grigorescu, 2010). Also, once control mechanisms are established, they are hard to revoke. This might explain why WFP has the same control tools in place as FAO despite its trustworthiness. Scholars studying delegation should therefore focus less on the mere existence of formal control (e.g. Grigorescu, 2010; Heldt, 2017; Hooghe & Marks, 2015), but more on the actual exercise of control and – in line with Lall (2017, p. 276) – 'de facto rather than de jure characteristics of IOs'.

References

- Abbott, K. W., Genschel, P., Snidal, D., & Zangl, B. (2020). Competence versus control: The governor's dilemma. *Regulation & Governance*, 14(4), 619–636.
- Abbott, K. W., & Snidal, D. (1998). Why States Act through Formal International Organizations. *Journal of Conflict Resolution*, 42(1), 3–32.
- Barnett, M., & Finnemore, M. (1999). The Politics, Power, and Pathologies of International Organizations. *International Organization*, 53(4), 699–732.
- Barnett, M., & Finnemore, M. (2004). *Rules for the World: International Organizations in Global Politics*. Cornell University Press.
- Bauer, M. W., Knill, C., & Eckhard, S. (Eds.). (2017). *International Bureaucracy: Challenges and Lessons for Public Administration Research*. Palgrave Macmillan UK.
- Bayerlein, L., Knill, C., & Steinebach, Y. (2020). *A Matter of Style: Organizational Agency in Global Public Policy*. Cambridge University Press.
- Biermann, F., & Siebenhüner, B. (Eds.). (2009). *Managers of global change: The influence of international environmental bureaucracies*. MIT Press.
- Conceição-Heldt, E. da (2013). Do Agents “Run Amok”? A Comparison of Agency Slack in the EU and US Trade Policy in the Doha Round. *Journal of Comparative Policy Analysis: Research and Practice*, 15(1), 21–36.
- Copelovitch, M. S. (2010). Master or Servant? Common Agency and the Political Economy of IMF Lending. *International Studies Quarterly*, 54(1), 49–77.
- Cortell, A. P., & Peterson, S. (2006). Dutiful agents, rogue actors, or both? In D. G. Hawkins, D. A. Lake, D. L. Nielson, & M. J. Tierney (Eds.), *Delegation and Agency in International Organizations* (pp. 255–280). Cambridge University Press.
- Davis, J. H., Schoorman, F. D., & Donaldson, L. (1997). Toward a Stewardship Theory of Management. *The Academy of Management Review*, 22(1), 20–47.
- Dijkstra, H. (2015). Shadow Bureaucracies and the Unilateral Control of International Secretariats: Insights from UN Peacekeeping. *Review of International Organizations*, 10(1), 23–41.
- Donaldson, L., & Davis, J. H. (1991). Stewardship Theory or Agency Theory: CEO Governance and Shareholder Returns. *Australian Journal of Management*, 16(1), 49–64. h
- Eckhard, S., & Ege, J. (2016). International bureaucracies and their influence on policy-making: A review of empirical evidence. *Journal of European Public Policy*, 23(7), 960–978.
- Eckhard, S., & Jankauskas, V. (2019). The politics of evaluation in international organizations: A comparative study of stakeholder influence potential. *Evaluation*, 25(1), 62–79.

- Edelenbos, J., & Klijn, E. H. (2007). Trust in Complex Decision-Making Networks: A theoretical and empirical exploration. *Administration and Society*, 39(1), 25–50.
- Ege, J. (forthcoming). What do international bureaucrats (really) want? *Global Governance: A Review of Multilateralism and International Organizations*.
- Ege, J., & Bauer, M. W. (2013). International Bureaucracies from a Public Administration and International Relations Perspective. In B. Reinalda (Ed.), *Routledge Handbook of International Organization* (pp. 135–148). Routledge.
- Elsig, M. (2010). Principal–agent theory and the World Trade Organization: Complex agency and ‘missing delegation’. *European Journal of International Relations*, 17(3), 495–517.
- FAO. (2020). Structure and finance. Food and Agriculture Organization of the United Nations. <http://www.fao.org/about/who-we-are/en/>
- Farrell, H. (2009). *The political economy of trust: Institutions, interests and inter-firm cooperation in Italy and Germany*. Cambridge University Press.
- Graham, E. R. (2014). International Organizations as Collective Agents: Fragmentation and the Limits of Principal Control at the World Health Organization. *European Journal of International Relations*, 20(2), 366–390.
- Grigorescu, A. (2010). The Spread of Bureaucratic Oversight Mechanisms across Intergovernmental Organizations. *International Studies Quarterly*, 54(3), 871–886.
- Hawkins, D., Lake, D. A., Nielson, D. L., & Tierney, M. J. (2006). Delegation under anarchy: states, international organizations, and principal-agent theory. In D. G. Hawkins, D. A. Lake, D. L. Nielson, & M. J. Tierney (Eds.), *Delegation and Agency in International Organizations* (pp. 3–38). Cambridge University Press.
- Hawkins, D. G., Lake, D. A., Nielson, D. L., & Tierney, M. J. (Eds.). (2006). *Delegation and Agency in International Organizations*. Cambridge University Press.
- Heldt, E. (2017). Regaining control of errant agents? Agency slack at the European Commission and the World Health Organization. *Cooperation and Conflict*, 52(4), 469–484.
- Hoffman, A. M. (2002). A Conceptualization of Trust in International Relations. *European Journal of International Relations*, 8(3), 375–401.
- Hooghe, L., & Marks, G. (2015). Delegation and Pooling in International Organizations. *The Review of International Organizations*, 10(3), 305–328.
- Hooghe, L., Marks, G., Lenz, T., Bezuijen, J., Ceka, B., & Derderyan, S. (2017). *Measuring International Authority: A Postfunctionalist Theory of Governance*. Oxford University Press.
- Kassim, H., & Menon, A. (2003). The principal-agent approach and the study of the European Union: Promise unfulfilled? *Journal of European Public Policy*, 10(1), 121–139.

- Kiewiet, D. R., & McCubbins, M. D. (1991). *The Logic of Delegation: Congressional Parties and the Appropriations Process*. University of Chicago Press.
- Knill, C., Bayerlein, L., Enkler, J., & Grohs, S. (2019). Bureaucratic influence and administrative styles in international organizations. *Review of International Organizations*, 14(1), 83–106.
- Lake, D. A., & McCubbins, M. D. (2006). The logic of delegation to international organizations. In D. G. Hawkins, D. A. Lake, D. L. Nielson, & M. J. Tierney (Eds.), *Delegation and Agency in International Organizations* (pp. 341–368). Cambridge University Press.
- Lall, R. (2017). Beyond Institutional Design: Explaining the Performance of International Organizations. *International Organization*, 71(2), 245–280.
- Lenz, T., Bezuijen, J., Hooghe, L., & Marks, G. (2015). Patterns of International Organization: Task Specific vs. General Purpose. *Politische Vierteljahresschrift*, 49, 131–156.
- Luhmann, N. (1979). *Trust and power*. John Wiley & Sons.
- Majone, G. (2001). Two Logics of Delegation: Agency and Fiduciary Relations in EU Governance. *European Union Politics*, 2(1), 103–122.
- Miller, G. J. (2005). The Political Evolution of Principal-Agent Models. *Annual Review of Political Science*, 8(1), 203–225.
- Nielson, D. L., & Tierney, M. J. (2003). Delegation to International Organizations: Agency Theory and World Bank Environmental Reform. *International Organization*, 57(2), 241–276.
- Pollack, M. A. (1997). Delegation, agency, and agenda setting in the European Community. *International Organization*, 51(1), 99–134.
- Reinalda, B. (Ed.). (2013). *Routledge Handbook of International Organization*. Routledge.
- Sitkin, S. B., & George, E. (2005). Managerial trust-building through the use of legitimating formal and informal control mechanisms. *International Sociology*, 20(3), 307–338.
- Stacey, J., & Rittberger, B. (2003). Dynamics of formal and informal institutional change in the EU. *Journal of European Public Policy*, 10(6), 858–883.
- Tamm, H., & Snidal, D. (2014). Rational Choice and Principal-Agent Theory. In T. G. Weiss & R. Wilkinson (Eds.), *International organization and global governance* (pp. 132–143). Routledge.
- Van Sylke, D. M. (2006). Agents or Stewards: Using Theory to Understand the Government-Nonprofit Social Service Contracting Relationship. *Journal of Public Administration Research and Theory*, 17(2), 157–187.
- Waterman, R. W., & Meier, K. J. (1998). Principal-Agent Models: An Expansion? *Journal of Public Admin Research and Theory*, 8(2), 173–202.

- Weiss, M. (2020). Varieties of privatization: Informal networks, trust and state control of the commanding heights. *Review of International Political Economy*. Early view.
- WFP. (2020). Overview. World Food Programme. <https://www.wfp.org/overview>
- Williamson, O. E. (1985). *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. Macmillan.
- Yi-chong, X., & Weller, P. (2008). 'To be, but not to be seen': Exploring the Impact of International Civil Servants. *Public Administration*, 86(1), 35–51.

Appendix to Paper A

Appendix I: Measurement of formal control in FAO and WFP

Type of control	Control Mechanism	Existence (1=yes; 0=no)	
		FAO	WFP
Oversight and Monitoring	Internal Audit	1 Office of the Inspector General	1 Inspector General & Oversight Office
	External Audit	1 FAO External Auditor	1 WFP External Auditor
	Internal Evaluation	1 Office of Evaluation (OED)	1 Office of Evaluation
	External Evaluation	1 UNEG, MOPAN, JIU	1 UNEG, MOPAN, JIU
	Inspection	1 Office of the Inspector General	1 Inspector General & Oversight Office
	Investigation	1 Office of the Inspector General	1 Inspector General & Oversight Office
	Ombudsman / Conflict Resolution	1 Ombudsman and Ethics Office	1 Ethic's Office
	Ethic / Integrity Assurance	1 Ombudsman and Ethics Office	1 Office of the Ombudsman
	Staff policies	1 Whistle-blower, conflict of interest, financial disclosure	1 Whistle-blower, conflict of interest, financial disclosure
	Advisory oversight to exec. governing body	1 FAO Oversight Advisory Committee	1 Advisory Committee on Administrative and Budgetary Questions
Screening and Selection⁵	Inspector General appointed in consultation with member states	1 On advice of Oversight Advisory Committee and after consultation with Finance Committee	1 On advice of Audit Committee and consent of the Executive Board
	Evaluation Director appointed in consultation with member states	1 Member states recommend, joint panel appoints	1 The Board approves the appointment made by Executive Director
Sanctioning	Majority voting in executive governing body	1 Unweighted majority voting	1 Unweighted majority voting (if consensus not possible)

⁵ In general, the administrative heads of FAO (Director General) and WFP (Executive Director) are responsible for their organizations' staff.

Table 1.A: Formal control mechanisms in FAO and WFP. Own compilation based on IOs' primary documents and official websites (as of 2020).

The chosen three types of control are most common to standard IO delegation studies. Oversight and monitoring as well as screening refer to ex ante control mechanisms available to member states, whereas sanctioning corresponds to ex post control mechanisms (McCubbins and Schwartz, 1984).

Appendix II: General IO characteristics

IO	Mandate	Staff		Budget	
		Total	Outside HQ (%)	Total budget (USD)	Voluntary contributions (%)
FAO	Ending hunger and malnutrition	17,000	68%	2.6 billion	61%
WFP	Ending hunger and malnutrition	11,500	90%	7.2 billion	100%

Table 2.A: FAO and WFP general characteristics (as of 2019/2020). Source: FAO (2020) and WFP (2020).

Appendix III: Interview list

Nr.	Label	IO	Group ⁶	Position	Date	Place
IO interviews						
1	MS1	FAO, WFP	Member states	Ambassador	May 2018	Rome
2	MS2	FAO, WFP	Member states	Ambassador	May 2018	Rome
3	MS3	FAO, WFP	Member states	Deputy Permanent Representative	May 2018	Rome
4	MS4	FAO, WFP	Member states	First Secretary	July 2020	Phone Interview
5	MS5	FAO, WFP	Member states	Ambassador	August 2020	Phone Interview
6	MS6	FAO, WFP	Member states	Minister	August 2020	Phone Interview

⁶ On the member state side, countries' representatives (usually ambassadors) in the executive governing bodies were interviewed. On the IO management side, senior administrations' officials with leadership roles (e.g. directors, deputy directors, senior officers such as division or team leaders) were interviewed.

7	MS7	FAO, WFP	Member states	Ambassador	August 2020	Phone Interview
8	MS8	FAO	Member states	Deputy Permanent Representative	August 2020	Phone Interview
9	MS9	WFP	Member states	Deputy Permanent Representative	August 2020	Phone Interview
10	IPA1	FAO	IO Management	Senior Officer	May 2018	Rome
11	IPA2	FAO	IO Management	Senior Officer	May 2018	Rome
12	IPA3	FAO	IO Management	Deputy Director	July 2020	Phone Interview
13	IPA4	FAO	IO Management	Senior Officer	May 2018	Rome
14	IPA5	FAO	IO Management	Senior Officer	May 2018	Rome
15	IPA6	WFP	IO Management	Director	August 2020	Phone Interview
16	IPA7	WFP	IO Management	Director	March 2020	Phone Interview
17	IPA8	WFP	IO Management	Head of Unit	March 2020	Phone Interview
18	IPA9	WFP	IO Management	Director	March 2020	Phone Interview
Background interviews						
19	UN Secretariat			Under-Secretary-General	February 2019	New York
20	UN Evaluation Group			Specialist	May 2018	Rome
21	UN OIOS			Director	February 2019	New York
22	MOPAN			Advisor	July 2018	Paris
23	UN Joint Inspection Unit			Inspector	November 2017	Phone Interview

Table 3.A: List of conducted interviews. In many cases, same member state representatives covered both organizations. Such interviews were longer in length as questions on both IOs were raised.

Interviews with member states				
Label	IO	Country list/region*	Position	Term in the governing body
MS1	FAO	Europe	Ambassador	2017-2019
MS2	FAO	Europe	Ambassador	2015-2018
MS3	FAO	Europe	Deputy Permanent Representative	2015-2019
MS4	FAO	Latin America and the Caribbean	First Secretary	2019-2022
MS5	FAO	Europe	Ambassador	2017-2020
MS6	FAO	Asia	Minister	2017-2020
MS7	FAO	Europe	Ambassador	2014-2020
MS8	FAO	North America	Deputy Permanent Representative	2019-2022
MS1	WFP	D	Permanent Representative	2015-2021
MS2	WFP	D	Permanent Representative	2015-2016, 2018-2019
MS3	WFP	D	Deputy Permanent Representative	2017-2021
MS4	WFP	C	First Secretary	2017, 2019
MS5	WFP	D	Permanent Representative	2018-2022
MS6	WFP	B	Minister	2012-2018, 2020-2021
MS7	WFP	E	Permanent Representative	2014-2020
MS9	WFP	D	Deputy Permanent Representative	2013-2022

*Table 4.A: List of interviews comparing regional diversion among member states. *The Executive Board of WFP consists of 5 lists of regionally diverse countries (from A to E). FAO Council members are listed according to their regions.*

Paper B

To Yield or Shield?

Comparing International Public Administrations' Responses to Member States' Policy Contestation

This is the published version of the article

In: *Journal of Comparative Policy Analysis: Research and Practice* (2020), Online First, co-authored with Tim Heinkelmann-Wild

Available at: <https://doi.org/10.1080/13876988.2020.1822144> (last accessed 12.01.2021)

Abstract: When member states contest policymaking in international organizations, some international public administrations (IPAs) react in a conciliatory way while others are adversarial. This article argues that IPAs' dependence on contesting states, their policymaking authority, and affectedness from contestation shape communicative responses. A Qualitative Comparative Analysis of 32 cases of contestation by the Trump administration indicates that IPAs yield when they are constrained by dependence on the United States or have no incentive to defend themselves. IPAs fend off contestation when they are unconstrained and incentivized by attacks on an international organization's polity, the bureaucracy, or policies in whose making they were substantially involved.

Keywords: legitimacy; international organizations; international public administrations; policy contestation; qualitative comparative analysis; public communication

Introduction

The policymaking of international organizations (IOs) is increasingly politicized. The deepening of their competences and broadening of their policy portfolios have rendered IOs increasingly prominent and contested in the public (Zürn et al. 2012; Goritz et al. 2020). Attacks on major IOs are not confined to the long-standing critics from the realms of civil society. Member state governments, too, increasingly utter their criticism publicly rather than behind closed doors. Public contestation offers governments not only the opportunity to avoid responsibility for negative outcomes by shifting blame onto them, but also to serve – or even mobilize – nationalist sentiments within their constituencies (Gerhards et al. 2009; Vasilopoulou et al. 2014; Dellmuth and Tallberg 2017; Schlipphak and Treib 2017; Daßler et al. 2019; Heinkelmann-Wild and Zangl 2019; Sommer 2020; Traber et al. 2019; Heinkelmann-Wild et al. 2020). President Donald Trump’s attempt to shift the responsibility for the COVID-19 crisis in the United States onto the World Health Organization (WHO) is illustrative.

A fast-growing literature strand found that, faced with such threats to their legitimacy, IOs professionalized their public communication (Ecker-Ehrhardt 2018b) and engaged in public self-legitimation (Zaum 2013b; Gronau and Schmidtke 2016; Schmidtke 2018; Rauh and Zürn 2020; Tallberg and Zürn 2019; Dingwerth et al. 2019a, 2019b). Far from being the passive, unpolitical and technocratic actors of the past, IOs are increasingly able and willing to strategically influence public discourse as “managers of (de)politicization and legitimacy” (Ecker-Ehrhardt 2018a, p. 23). This literature suggests that IOs accommodate contestation by presenting themselves in a positive light and broadening their justification narratives to include, inter alia, democratic values.

We agree with this stand of literature and assume that IO secretariats – international public administrations (IPAs) – are capable and willing to communicate strategically in public vis-à-vis contesting member states. However, we observe substantial variation in the substance of their responses. Sometimes, IPAs indeed adopt a *conciliatory* communicative approach vis-à-vis contesting states that aims at accommodating their criticism. For instance, even when the US froze their budgetary contributions to the WHO, its Director-General responded diplomatically without directly naming and shaming the US. And when Sweden criticized the Joint United Nations Programme on HIV and AIDS (UNAIDS) for sexual assault as well as mismanagement and threatened to cut its budget, the latter responded by accepting the accusations and signaling accommodation. At other times, however, IPAs opt for *adversarial* communicative responses. For example, when the Trump administration criticized the UN

Population Fund (UNFPA) for its population planning programs in China, its secretariat directly rejected these claims as erroneous. And when the US then threatened to cut its funding, UNFPA shamed the US for the humanitarian consequences. Moreover, after Ukraine blamed the Organisation for Security and Co-operation in Europe (OSCE) for partiality, the organization's management responded by rejecting the accusations and shaming the contesting state. The question thus is: under which conditions do IPAs respond *conciliatory* or *adversely* to member state contestation?

We suggest that an IPA's choice between a conciliatory and an adversarial communicative reaction to member state contestation depends on constraints, posed by the IPA's dependence on the contesting state(s), as well as incentives, posed by the IPA's policymaking authority and its affectedness from contestation. The paper proceeds as follows. First, we draw on the growing literature on blame and reputation management to develop our theory of IPAs' communicative responses to member state contestation. Second, we assess the plausibility of our theory through a crisp-set Qualitative Comparative Analysis (csQCA) of 32 cases of IPA responses to contestation by the Trump administration. The paper concludes by summarizing the theoretical and normative implications of our findings.

Explaining IPAs' Responses to Member States' Contestation

When IO member state governments contest IOs and their policies, this may affect their reputation and overall legitimacy and thus their authority (Schlippak and Treib 2017). This is vital for IOs as their authority depends on the public's belief in their legitimacy to execute delegated tasks (Tallberg and Zürn 2019). As member state contestation has legitimacy costs, IPAs prefer to respond publicly in order to preserve the authority of "their" IOs.¹ Like other political actors that are publicly contested, IPAs try to avoid blame and its consequences, such as reputational costs and potential sanctions (Heinkelmann-Wild and Zangl 2019, 2020, see also Carpenter and Krause 2011; Maor et al. 2013; Busuioc and Lodge 2017). They will thus employ blame and reputation management strategies in public (Heinkelmann-Wild and Zangl 2019, 2020; see also Hood 2011; Hinterleitner and Sager 2017; Maor 2020). Ideal-typically, IPAs can choose between two responses that follow distinct logics:

¹ We use the term IPA to differentiate IO administrations from member states in IO governing bodies. IPAs are "hierarchically organized group[s] of international civil servants with a given mandate, resources, identifiable boundaries, and a set of formal rules of procedures" (Biermann and Siebenhüner 2009, p. 37).

- **Conciliatory response:** IPAs can seek to avoid (more) legitimacy losses by enclosing contestation and avoiding further escalation – and thus potential sanctions – by pleasing the contesting states. To that end, IPAs might opt for adopting a positive-toned communication, complimenting the contesting state(s), admitting alleged failures and even their responsibility for them, or taking ownership for the solution. Following the motto “if you can’t beat them, join them”, IPAs thus yield vis-à-vis the contesting state(s).
- **Adversarial response:** IPAs can also try to avoid (more) legitimacy costs by denying an alleged grievance or even de-legitimizing the contesting state(s). To that end, IPAs might opt for adopting a negative-toned or even hostile communication, shaming the contesting state(s), rejecting alleged failures or at least its responsibility for them. Following the motto “offense is the best defense”, IPAs thus shield “their” IOs vis-à-vis the contesting state(s).

These two public responses are incompatible as their general messages tend to contradict and undermine each other. In contrast to secret diplomacy behind closed doors, in public, an IPA cannot, at the same time, accept and reject a grievance put forward by the contesting state(s). When an IPA opts for an adversarial strategy, and thus refuses to accept blame, it risks further escalation by the contesting state(s). And when an IPA adopts a conciliatory response, it might satisfy the contesting member state(s) but, at the same time, (implicitly or explicitly) accepts blame for the alleged grievance in public. An IPA’s conciliatory response might even risk a backlash from stakeholders – be it other member states, the transnational civil society, or the broader public – who disagree with the thrust of contestation. Satisfying one subset of the public might thus upset another part of the audience (Zaum 2013a, p. 19; see also Carpenter and Krause 2011, p. 29). Hence, while the two response types might come in different variants (see e.g. Gilad et al. 2013; Hood et al. 2016, p. 545), they are eventually mutually exclusive.

When faced with member state contestation, IPAs – as boundedly rational actors with strategic agency – must therefore decide whether to put their efforts into pleasing the contesting state(s) but accepting the blame (conciliatory response); or to openly defend their reputation but bear potential sanctions by the contesting state(s) (adversarial response). To answer this question, we draw on insights from research on blame and reputation management. This scholarship highlights the intensity of contestation (Maor et al. 2013; Hood et al. 2016; Busuioc and Lodge 2017; Traber et al. 2019; Kruck et al. 2020), the contested actor’s tasks and responsibilities (Carpenter and Krause 2011; Heinkelmann-Wild and Zangl 2019, 2020), and its relationship

with contesting actors (Maor et al. 2013, p. 587; Hinterleitner and Sager 2019, p. 137). We therefore suggest that an IPA's choice depends on the constraints, posed by the IPA's dependence on the contesting state(s), as well as the incentives posed by the IPA's policymaking authority and its affectedness from contestation.

Dependence Condition

An IPA's dependence on the contesting member state(s) constrains its response as it affects the harm a potential escalation of the conflict could cause it (cf. Maor et al. 2013, p. 587; Heinkelmann-Wild and Zangl 2019, p. 4; Hinterleitner and Sager 2019, p. 137). While IPAs generally depend on IO member states, they might do so to a varying extent (Lyne et al. 2006, p. 58). IPA representatives will thus anticipate the harm of sanctions on the IO's functionality in case of an escalation of the conflict with the contesting state(s).

If an IPA is independent from the contesting state(s), it is rather unconstrained by their potential sanctions: even if the contesting state(s) escalate the conflict and employ sanctions, the IPA will not severely suffer from them. Hence, if an IPA is independent from the contesting member state(s), it can afford an adversarial response. By contrast, if an IO secretariat depends on the contesting state(s), its response choice is constrained by the harm of potential sanctions. After all, if the contesting state(s) escalate the conflict and employ sanctions, the IO will severely suffer from them. If an IO is vulnerable vis-à-vis contesting member state(s), its administrations cannot afford to attack the contesting state(s), but instead will respond by conceding to the criticism.²

We thus expect: *if an IPA is dependent on the contesting state(s), it is constrained by the potentially harmful escalation and will respond in a conciliatory way; and if an IPA is not dependent on the contesting state(s) it is rather unconstrained in its response.*

Authority Condition

An IPA's level of authority provides it with incentives for a conciliatory or an adversarial response. After all, IPAs' involvement in IO policymaking shapes their public visibility as well

² To be sure, non-contesting states might be dissatisfied about an IPA's conciliatory response to the contesting member state(s), triggering contestation on their part. However, this threat is rather diffuse and unlikely to immediately escalate into material sanctions. By contrast, the threat posed by contesting states is concrete as they are already one step "ahead" in the escalation process.

as their stakes in the polices (cf. Carpenter and Krause 2011; Heinkelmann-Wild and Zangl 2019, 2020). IOs vary in the extent to which member states have delegated authority to IPAs (Zürn et al. 2012; Hooghe and Marks 2015; Busch et al. 2020). In some IOs, member states design and implement policies largely on their own while IPAs coordinate intergovernmental bargaining. Such IO administrations have lower incentives to respond adversely but can afford conceding or accepting blame to end public contestation. When they are not substantively involved in IO policymaking, IPAs have lower stakes in substantive outcomes. And as they mainly assist member states, their public reputation is less important for “their” IOs’ legitimacy. After all, IPAs without policymaking authority are generally less focal in public (Rittberger et al. 2017; Heinkelmann- Wild and Zangl 2019, 2020).

In other IOs, member states have delegated considerable authority for the design and/ or implementation of IO policies to IO administrations (Eckhard and Parížek 2020). IPAs with policymaking authority have strong incentives to respond to member state contestation adversely and not in a conciliatory way. As they are substantially involved in IO policymaking, by conceding and thereby (implicitly or explicitly) accepting blame, they would undermine their own expertise and might appear hypocritical (Knill et al. 2018). After all, not only the legitimacy of IO policies but also such IPAs’ authority depends on their public recognition as neutral, technical experts (Barnett and Finnemore 2004). Their reputation is even more at stake as IPAs’ policymaking authority renders them focal in public (Rittberger et al. 2017; Heinkelmann-Wild and Zangl 2019, 2020).

Hence, we hypothesize: *if an IPA has policymaking authority, its incentives are strong to respond adversely to member state contestation; and if an IPA only assists member states in policymaking, it is incentivized to respond in a conciliatory way.*

Affectedness Condition

IPAs’ affectedness by contestation, specifically whether the contesting member state(s) directly target(s) the IO’s polity or administration, also shapes their incentives to respond in a conciliatory way or adversely (cf. Maor et al. 2013; Hood et al. 2016; Busuioc and Lodge 2017; Traber et al. 2019; Kruck et al. 2020). After all, it is easier for an IPA to concede to contestation that does not target the IO’s polity or the bureaucracy itself but relates to policies or member states.

All else being equal, in cases of direct contestation by member state(s), where an IPA or an IO's overall polity is assaulted directly – by criticizing its performance or even necessity, or by withdrawing resources or membership – the IPA is incentivized to respond adversely. As conceding to direct contestation equals blame acceptance, IPAs will defend their reputation through, for instance, rejecting the blame or launching a counter-attack against the contesting state(s).

In instances of indirect contestation, where not the IPA or an IO's polity but IO member states or policies are criticized or not complied with, the IPA is, all else being equal, tempted to respond in a conciliatory rather than adversarial manner. As the IPA is not in the line of fire, it does not have to bear high reputational costs when conceding to the contesting state(s). An adversarial response, by contrast, could even attract contestation targeting the IPA directly.

We thus expect: *if an IO's polity or bureaucracy is not affected directly by contestation, its IPA has less incentive to respond adversely, while if contested directly, the IPA is incentivized to respond adversely.*

Assessing IPAs' Responses to Contestation by the Trump Administration

To evaluate the plausibility of our theoretical argument, we conducted a crisp-set Qualitative Comparative Analysis (csQCA) across 32 cases of IPAs' responses to contestation by the Trump administration. QCA is increasingly used in Public Administration research as it allows for a context-sensitive analysis of complex social phenomena (Thomann and Ege 2020). As a set-theoretic, truth table-based technique, QCA enables us to arrive at parsimonious solutions of necessity and sufficiency based on patterns of medium-*n* cases (Rihoux and Ragin 2009; Schneider and Wagemann 2013).

Case Selection and Sample Construction

We examined major IO administrations' responses to contestation by the Trump administration. Our focus is on major, authoritative IOs as these meet the assumed *scope conditions* of our theoretical argument: major IOs are prominent and powerful enough to incentivize member states to shift blame onto them or mobilize nationalist publics against them. And they are not

only able to communicate to the public but are also incentivized to publicly respond to their delegitimation as they have much authority to lose (see Heinkelmann-Wild and Zangl 2020).

We opted to focus on contestation by the US under the Trump administration as characteristics of this country and government, as well as the sheer number of instances, provide us with the empirical variation required to assess whether (combinations of) the theorized conditions shape IPAs' responses. We thus do not claim to test our argument in a sample that is representative of all (major) IOs, but rather adopt the logic of a mostsimilar- case design (Przeworski and Teune 1982, pp. 32–33). We thereby necessarily make sacrifices regarding the external validity and thus generalizability of our findings. At the same time, this design supports the internal validity of our results as our case selection allowed us to control for potential confounding factors while probing whether the theorized conditions matter.

The focus on contestation by the US under the Trump administration enabled us to *control* for country characteristics and individual properties of its government officials, including their communication style and foreign policy outlook. In addition, the number of contesting states is a constant as the Trump administration was rather isolated in its criticism of IOs.³ Moreover, focusing on the three-year period from Trump's inauguration in January 2017 until December 2019 allows us to control for potentially conflating, exogenous developments, such as the COVID-19 pandemic.

Our case selection, at the same time, allowed us to observe *variation* in the theorized conditions. As the US is generally powerful but possesses varying power within IOs, focusing on the US allowed us to observe (relative) differences in IPAs' dependence on the very same state. And as the Trump administration contested numerous IOs, the cases in our sample also show variation regarding IPAs' policymaking authority and affectedness.

To identify the responses of major IO administrations to contestation by the Trump administration, we first compiled a list of major IOs. As a starting point, we drew on the dataset of major IOs compiled by Hooghe and Marks (2015). We disaggregated the UN into its individual agencies and funds to include both UN Secretariat and such large organizations as the UN Educational, Scientific and Cultural Organization (UNESCO) or the United Nations Children's Fund (UNICEF). We then excluded all IOs where the US was not a member or whose secretariat does not meet our definition of IPA. For the period ranging from Donald Trump's inauguration in January 2017 to December 2019, we then searched in IPAs' press

³ While Israel supported US contestation in a few cases, including Israel would not have changed the values of the respective cases regarding the respective IOs' dependence on the contesting states.

releases as well as the media database Factiva for accounts referring to the Trump administration and the respective IO. This allowed us to identify instances of contestation as well as IPAs' responses.⁴ Our final sample comprised 32 cases across 18 IOs. Table S.4 in the Supplementary Material lists all cases, including the Trump administration's contestation and examples for IPAs' responses.

Operationalization and Data

To calculate the csQCA, we had to operationalize the outcome as well as the conditions in dichotomous terms.⁵ For each case, we assessed an *IPA's response*, based on a qualitative content analysis of its official statements as well as interviews in and coverage by the media. To measure the two types of IPA responses, we relied on the tone and content of their public statements: (1) When an IPA opts for a conciliatory response, it adopts a positive tone and signals its accommodation of the contesting state through praising it, accepting the accusations, or even neglecting contested IO norms. (2) In the case of an adversarial response, the IPA adopts a negative tone and signals its objection to the contestation through defending contested IO norms, rejecting the accusations or blaming the contesting state. The IPAs' response was adversarial in 13 cases and conciliatory in 19 cases.

To assess the *dependence condition*, we used the share of an IO's total budget provided by the US. Existing literature suggests that resources are a good proxy for an IPA's dependence on member states (Lyne et al. 2006, p. 58; Bauer and Ege 2016; Eckhard and Jankauskas 2019). We opted for the share of US contributions in 2016/2017 as this level constitutes a plausible baseline for IPAs' expectations. We used 15 percent as a cut-off point. When the US contributed more than 15 percent to the total IO budget (including both core and voluntary contributions), the IPA was coded as dependent; if less, we consider the IPA independent. While thresholds are always somewhat arbitrary, we hold it plausible since IOs receiving less than 15 percent, such as the Universal Postal Union (UPU) (6 percent) or the UN Environment Programme (UNEP) (4 percent), can clearly be considered less dependent on the US than IOs such as the North Atlantic Treaty Organization (NATO) (22 percent) or the UN Refugee Agency (UNHCR) (37 percent). Moreover, in most IOs receiving more than 15 percent of funding from the US, the latter is also the single largest donor. Finally, the distribution of US contributions across

⁴ When the same IO was contested several times, but the object of contestation changed, we treated them as separate cases.

⁵ See Supplementary Material, Section 2, for detailed coding rules, examples, and raw data.

IOs confirms that these cluster around the 15 percent threshold (see Supplementary Material, Figure S.1).⁶ In 23 out of the 32 cases, the IPA is dependent on the US.

To assess the *authority condition*, and thus an IPA's involvement in IO policymaking, we referred to its primary tasks (cf. Tallberg et al. 2013). Based on the IO's formal mandate, we coded an IPA's policymaking authority depending on whether its task is (1) to implement policies or provide policy advice to member states; or (2) to assist member states in bargaining. In 24 of the 32 cases in our sample, the IPA possesses policymaking authority.

To assess the *affectedness condition*, we had to determine the substantial thrust of contestation. We differentiated instances of contestation between two categories: (1) direct contestation, i.e. when Trump administration officials criticized the bureaucratic processes explicitly, questioned the overall necessity of the IO, or proclaimed its exit or budget cut; and (2) indirect contestation, i.e. when the Trump administration criticized the behavior of other member states or questioned the policies of an IO without explicitly naming it. Table 1 shows the dichotomized values for the outcome as well as the conditions.

⁶ In cases where the Trump administration's contestation comprised withdrawing membership (e.g. UNESCO) or the complete funding (e.g. UNFPA), we also coded the IPAs as independent as they cannot expect to obtain the same funding level as under the Obama administration.

CaseID	Dependence	Authority	Affectedness	Adversarial
IOM	1	1	0	0
UNHCR1	1	1	0	0
UNHCR2	1	1	0	0
UNHCR3	1	1	0	0
WFP1	1	1	1	0
WFP2	0	1	1	1
NATO1	1	0	1	0
NATO2	1	0	0	0
NATO3	1	0	0	0
FAO	1	1	1	0
UNSG1	1	1	1	0
UNSG2	1	1	1	0
UNFCC1	1	0	1	0
UNFCC2	1	0	0	0
WB1	1	1	0	0
WB2	1	1	0	0
IMF1	1	1	0	0
IMF2	1	1	1	0
UPU	0	0	0	0
UNEP	0	1	0	1
IAEA1	1	1	1	1
IAEA2	1	1	0	1
IAEA3	1	1	1	1
WTO1	0	0	1	1
WTO2	0	0	1	1
UNFPA1	0	1	1	1
UNFPA2	0	1	1	1
UNRWA1	1	1	1	0
UNRWA2	0	1	1	1
UNESCO	0	1	1	1
UNHRC	0	1	0	1
UNICEF	1	1	0	1

Table 1: Cases and coded values (dichotomized).

Results

To identify (combinations of) conditions associated with a conciliatory or an adversarial response, we constructed a *truth table* of our 32 cases (Table 2). Out of eight logically possible configurations of the three conditions, six are clearly associated with one outcome. However, the first two rows are inconsistent: while the majority of cases shows a conciliatory outcome, IPAs in two cases per each row responded adversely. These responses stem from the International Atomic Energy Agency (IAEA) and UNICEF. This inconsistency is not trivial considering that 17 out of 32 cases show these configurations. Indeed, “contradictory

configurations are one of the most challenging features of csQCA” (Yamasaki and Rihoux 2009, p. 135). After probing the feasibility of various approaches to resolve inconsistent configurations laid out by the literature (Rihoux and Ragin 2009, pp. 49–50; Schneider and Wagemann 2013, pp. 120–123),⁷ we eventually used the frequency criterion advocated by Schneider and Wagemann (2013, p. 122): As the majority of cases in both rows shows a conciliatory response with at least 0.75 consistency, we treated these configurations as showing a conciliatory response (see Rihoux and Meur 2009, p. 49; Schneider and Wagemann 2013, pp. 122, 227–228). To be sure, this is a limitation of our analysis which is also reflected in the interpretation of our findings (see discussion in Section 3.4). However, we accept this limitation as the prime objective of this paper is not to account exhaustively for the full empirical diversity in our cases, but to introduce a so far neglected phenomenon – IPAs’ adversarial responses to member states’ contestation – and to offer a first parsimonious model for its explanation.

	Dependence	Authority	Affectedness	Adversarial
WFP1, FAO, UNSG1, UNSG2, IMF2, UNRWA1, IAEA1, IAEA3	1	1	1	0 (0.75)
IOM, UNHCR1, UNHCR2, UNHCR3, WB1, WB2, IMF1, IAEA2, UNICEF	1	1	0	0 (0.78)
NATO1, UNFCC1	1	0	1	0
NATO2, NATO3, UNFCC2	1	0	0	0
UNFPA1, UNFPA2, WFP2, UNRWA2, UNESCO	0	1	1	1
UNEP, UNHRC	0	1	0	1
WTO1, WTO2	0	0	1	1
UPU	0	0	0	0

Table 2: Truth Table.

To identify whether there are necessary or sufficient conditions associated with the two outcomes, we then calculated the csQCA using the application TOSMANA (Cronqvist 2019).

⁷ In the Supplementary Material, we checked whether the inconsistencies stem from an omitted condition by calculating two further model specifications: we probed whether experience with earlier contestation by the same government or earlier sanctions by the contesting state are also associated with IPAs’ responses. However, the results do not substantially differ from our original solution and the inconsistencies are not resolved (Section 4). We also reexamined the operationalization of the outcome and conditions by testing alternative cut-off points (Section 3.2), yet the obtained solutions were all subsets of our original solution. Finally, we ran a model that excluded all 17 contradictory cases (Section 3.3). Its results led to a sharp decrease of the solution’s coverage. When we allowed the algorithm to make assumptions for the lack of configurations, the results equaled our original solution. In the end, we opted to not sacrifice coverage and parsimony for consistency.

We find that neither the presence nor the absence of any of the three conditions is necessary for a conciliatory or adversarial response.⁸

Turning to sufficiency, for *conciliatory responses*, our analysis indicates that either an IPA's dependence on the contesting state (DEPENDENCE) or the indirect contestation and an IPA assisting member states in policymaking (authority*affectedness) are sufficient conditions. The solutions are formalized in Boolean terms whereby “*” signifies a logical “and”, while “+” denotes a logical “or”. Each constellation is defined by a condition's presence (upper-case letters) or absence (lower-case letters). Hence, “adversarial” indicates a conciliatory response, while “ADVERSARIAL” indicates an adversarial response. Table 3 depicts for both sufficient conditions the parameters of consistency, i.e. “the degree to which the empirical information deviates from a perfect subset relation” (Schneider and Wagemann 2013, p. 129), and coverage, i.e. “how much of the outcome is covered (explained) by the condition in question” (Schneider and Wagemann 2013, p. 139). Taken together, the complete solution for a conciliatory response covers all cases of conciliatory IPA responses (and thus has a coverage of 1.00) and a comparatively high consistency of 0.82.

Solution	DEPENDENCE	+	authority*affectedness	→	adversarial
Consistency	0.82		1.00		
Raw coverage	0.95		0.21		
Unique coverage	0.79		0.05		

Table 3: Causal paths leading to a conciliatory response.

The first sufficient condition “DEPENDENCE” covers almost all cases (0.95). Take NATO1 as an example. Its administration depended on the US, which provided 22 percent of NATO's funding. In early 2017, Trump directly contested the organization by criticizing its issue prioritization and questioning its overall necessity (BBC 2017). In line with our expectation, the response by the NATO Secretary-General was conciliatory: he welcomed President Trump's “very strong message on defense spending” and admitted that NATO has “to step up and do more” on fighting terrorism.⁹ Another typical example is UNHCR2. When the US tightened its asylum and migration policies in clear violation of UNHCR standards (indirect contestation), the UNHCR remained friendly towards the US and declared that it stands “ready

⁸ The dependence condition comes closest. For the outcome conciliatory response, we obtain a consistency score of 0.95. Yet, due to the deviant cases IAEA1, 2, 3 and UNICEF, its coverage of 0.82 is below the 0.9 threshold.

⁹ Jens Stoltenberg, Secretary General of NATO, cited by McCaskill and Lima (2017).

at all times to support the United States” (UNHCR 2018). Further corroborating our theory, the conciliatory response was not without costs. In fact, the organization faced harsh public criticism for its “outrageous inability to challenge the Trump Administration in areas that are basically UNHCR’s bread and butter”.¹⁰

The consistency of the “DEPENDENCE” condition is 0.82 and thus clearly above the 0.75 threshold. The inconsistency stems from the four deviant cases: IAEA1, 2, 3 and UNICEF. Although these two IPAs are amongst the most dependent IOs – the US contributed 31 percent to the IAEA’s and 24 percent to UNICEF’s budget–, their responses were always adversarial. The IAEA strongly rejected any blame for its partiality and incompetence, whereas UNICEF shamed the US for its “extremely worrisome”¹¹ policy to separate children from their parents at the border.

The second sufficient condition for a conciliatory response “authority*affectedness” covers fewer cases (0.21) and only one case, UPU, uniquely. UPU’s administration was not dependent on the US as it contributes only 6 percent to its budget. UPU assists member states in policymaking as a “forum for cooperation between postal sector players” (UPU 2020). The US did not contest UPU’s polity or secretariat directly but criticized China for misusing its postal scheme. As expected, UPU’s Director-General responded in a conciliatory way by signaling accommodation and diplomatically requesting a meeting with the US government to “further discuss the matter” (UPU 2018). This condition is completely consistent as it covers no case showing an adversarial response.

Turing to *adversarial responses*, the results imply that an IPA’s independence from the contesting state combined with either policymaking authority (dependence*AUTHORITY) or direct contestation (dependence*AFFECTEDNESS) are both sufficient conditions. Table 4 shows that the consistency of both sufficient conditions assumes the maximum of 1.00 and both have a coverage of 0.54 (as they do not cover IAEA1, 2, 3 and UNICEF). The complete solution for an adversarial response has a coverage of 0.69 and a consistency of 1.00.

¹⁰ Joel Charny, executive director of Norwegian Refugee Council USA, cited by Welsh (2018).

¹¹ Caryl Stern, President of UNICEF USA, cited by Hart (2018).

Solution	dependence*AUTHORIT Y	+	dependence*AFFECTEDNE SS	→	ADVERSARIAL
Consistency	1.00		1.00		
Raw coverage	0.54		0.54		
Unique coverage	0.15		0.15		

Table 4. Causal paths leading to an adversarial response

The first sufficient condition “dependence*AUTHORITY” covers UNEP and UNHRC uniquely. Consider the administration of UNEP, which is independent from the US as it contributes only 4 percent to the UNEP’s budget. It was indirectly contested by President Trump when he decided to leave the Paris Agreement and issued environmentally harmful policies, yet without attacking UNEP directly. UNEP’s bureaucracy has considerable policymaking authority as an expert body for environmental politics that provides policy advice to member states. As it depends on its recognition as neutral, technical expert, we would expect the IPA to see its reputation severely threatened even when contestation does not target the administration directly but rather its domain of expertise. Accordingly, UNEP’s Executive Director responded adversely by stating that Trump “clearly made a mistake”.¹²

The second sufficient condition for an adversarial response “dependence*AFFECTEDNESS” uniquely covers the two WTO cases. Its IPA is comparatively independent from the US, which contributes 11 percent to its budget. When the Trump administration directly blamed the WTO for being unfair and threatened to exit, its Director-General questioned Trump’s trade policies and adversely claimed that the US challenge does “not mean the end of the multilateral trading system”.¹³

As the complete solutions of sufficiency for an adversarial and a conciliatory IPA response are logical complements, we can depict them together in Figure 1. The upper-right quadrant shows typical cases for a conciliatory response, while the lower-left quadrant shows typical cases for an adversarial response. The lower-right quadrant comprises the four deviant cases.

¹² Erik Solheim, Executive Director of UNEP, cited by Agence Francaise-Press (2017).

¹³ Roberto Azevedo, Director-General of WTO, cited by Swanson (2019).

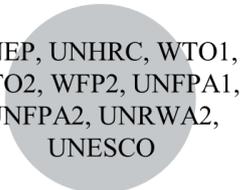
Conciliatory response	
Adversarial response	 
	<p>dependence * (AUTHORITY + AFFECTEDNESS)</p> <p>DEPENDENCE + authority * affectedness</p>

Figure 1: Complete solutions for conciliatory and adversarial responses

Several *robustness checks* increase our confidence in the findings (see Supplementary Material, Section 3). Firstly, we tested whether our findings hold up to methodological triangulation (see Skaaning 2011). We conducted a statistical analysis of contingency tables to test the null hypothesis that there is no relationship between the obtained solution and the outcome. The obtained chi-square value assures us that the identified patterns are not random, with the probability being 1 percent of rejecting the null hypothesis when it is true. Moreover, we probed the robustness of our analytical decisions during the QCA regarding calibration, raw consistency and row frequency. Our results remain largely robust even when ignoring truth table rows with only one case as well as changing the cut-off point of the dependency condition.

Discussion

As all theorized conditions are part of the solutions for IPAs' conciliatory or adversarial responses to contestation by the Trump administration, the analysis lends support to the

plausibility of our theoretical framework. Moreover, as the QCA identified alternative pathways and combinations of conditions leading to the two outcomes, we can further specify our theory: IPAs respond in a conciliatory way when they are constrained by their dependence on the contesting state *or* have no incentives to defend themselves because they are not directly affected or have no policymaking authority. IPAs react adversely when they depend less on the contesting state and are thus unconstrained *and* need to fend off contestation – as it either targets them directly or touches upon policies in whose making they were substantially involved.

At the beginning of the analysis, we discussed the problem of contradictory cases. Comparing our solution with the solution when excluding all contradictory cases from the first two truth table rows (see Supplementary Material, Section 3.3) allows us to nuance our findings regarding conciliatory IPA responses (see Yamasaki and Rihoux 2009, pp. 133–134). Being part of both solutions, we can hold that the combination of no policymaking authority and indirect contestation *strictly* leads to a conciliatory IPA response. By contrast, when excluding inconsistent cases, we find that dependency leads to the outcome only when combined with no policymaking authority whereas our original solution included dependency as sufficient condition on its own. Hence, dependency on the contesting member state *possibly* leads to conciliatory IPA responses, while it *strictly* does so only when the IPA lacks policymaking authority.

Can these results be transferred to IPA responses to member state contestation beyond the Trump administration? As our research design aimed at internal – and not external – validity, the generalizability to other cases is limited. Nevertheless, our analysis demonstrated that IPAs do not only respond in a conciliatory way but also adversely vis-à-vis member states. And not just vis-à-vis any state, but the US, the most influential state in global governance, and, specifically, the Trump administration that showed a record of sanctioning unsatisfactory IOs. The fact that IPAs even under these circumstances opted for an adversarial response suggests that this phenomenon should be more widespread than acknowledged in the state of the art.

Our results bear implications for at least three strands of literature. First and foremost, our finding that IPAs do not confine their public communication in the face of contestation to positive self-representations and accommodation, but also engage in negative, adversarial responses, contrasts a common assumption in the *IO self-legitimation literature* (Zaum 2013b; Gronau and Schmidtke 2016; Schmidtke 2018; Dingwerth et al. 2019a, 2019b; Rauh and Zürn 2020; Tallberg and Zürn 2019). This research strand could profit from broadening their

perspective on IO self-legitimation in the face of contestation to also include adversarial responses.

Secondly, our analysis demonstrated that insights of the *blame and reputation management literature* travel to the analysis of IPAs' responses to public contestation (cf. Hinterleitner and Sager 2017; Maor 2020). This finding challenges the widespread assumption that IO member states are unconstrained in shifting blame onto IOs (Gerhards et al. 2009; Vasilopoulou et al. 2014; Schlipphak and Treib 2017; Sommer 2020). By contrast, our analysis indicates that IPAs will fight back when attacked directly by member states as long as they do not have to fear severe sanctions. Our paper thereby ties in with recent research that emphasizes that IO bureaucracies are no passive scapegoats but active blame avoiders (Heinkelmann-Wild and Zangl 2020).

Finally, and relatedly, our findings qualify the *IO agency literature* which usually treats IOs as unitary actors which are bound to their member states (Nielson and Tierney 2003; Hawkins et al. 2006). Our results advise against a state-centric, monolithic view of IOs and highlight the strategic actorness of IO bureaucracies. IPAs are not always servants of their member states' power, but they are distinct authorities that carefully defend their legitimacy. Scholars analyzing IOs' behavior should thus differentiate between member states (and their governing bodies) and IPAs as actors in their own right (see also Bayerlein et al. 2020; Goritz et al. 2020).

This paper can, however, only constitute a first step towards a better understanding of how and why IO administrations position themselves vis-à-vis contesting member states. Future research should systematically test whether our findings hold beyond contestation by the Trump administration and further specify our theory. For instance, when do the constraints of dependency on the contesting state(s) trump the incentives for an adversarial response posed by policymaking authority, and vice versa? More generally, does the dependency condition hold for contestation by other powerful states, such as China or the United Kingdom, as well as for weaker states? How does the number of contesting states impact IPAs' responses? And, finally, how does an IO's overall legitimacy shape its IPA's response to contestation?

Conclusion

In this paper, we examined under which conditions IO administrations respond in a conciliatory or adverse way to member state contestation. A QCA of 32 cases of contestation by the Trump

administration indicated that IPAs yield when they are constrained by dependence on the US or have no incentive to defend themselves. By contrast, IPAs fend off contestation when they are unconstrained and incentivized by attacks at an IO's polity, the bureaucracy, or at policies in whose making they were substantially involved.

Our findings have ambivalent normative implications for IPAs' accountability and IO legitimacy in general. The analysis implies that IPAs are self-interested actors whose public communication is driven by reputational and legitimacy concerns. The normative evaluation of IPAs' responses to member state contestation ultimately hinges on the "truthfulness" of contestation and the legitimacy of the status quo. After all, while an adversarial response might sometimes be – as often in cases of contestation by the Trump administration – a justified defense of the rule-based, liberal international order, it can also aim at avoiding accountability. And while a conciliatory response might sometimes advance accountability, at other times it means yielding to illegitimate state interests. As often in cases of contestation by the Trump administration, conceding vis-à-vis powerful member states might seem beneficial for IPAs in the short term. However, by neglecting the principles of the liberal international order in favor of nationalist-populist governments, IPAs risk undermining IO legitimacy – and thereby their own authority – in the long run.

References

- Barnett, M. and Finnemore, M., 2004, *Rules for the World: International Organizations in Global Politics* (NY: Cornell University Press).
- Bauer, M. W. and Ege, J., 2016, Bureaucratic autonomy of international organizations' secretariats. *Journal of European Public Policy*, 23(7), pp. 1019–1037.
- Bayerlein, L., Kaplaner, C., Knill, C., and Steinebach, Y., 2020, Singing together or apart? Comparing policy agenda dynamics within international organizations. *Journal of Comparative Policy Analysis*, pp. 1–20.
- BBC, 2017, *Trump Worries Nato with 'Obsolete' Comment*. Available at <https://www.bbc.com/news/world-uscanada-38635181> (accessed 2 January 2020).
- Biermann, F. and Siebenhüner, B. (eds.), 2009, *Managers of Global Change: The Influence of International Environmental Bureaucracies* (Cambridge MA: MIT Press).
- Busch, P.-O., Heinzl, M., Kempken, M., and Liese, A., 2020, Mind the gap? Comparing de facto and de jure Expert Authority of International Public Administrations in financial and agricultural policy. *Journal of Comparative Policy Analysis*.
- Busuioac, M. and Lodge, M., 2017, Reputation and accountability relationships: Managing accountability expectations through reputation. *Public Administration Review*, 77(1), pp. 91–100.
- Carpenter, D. P. and Krause, G. A., 2011, Reputation and public administration. *Public Administration Review*, 72(1), pp. 26–32.
- Cronqvist, L., 2019, Tosmana [Version 1.61]. Available at <https://www.tosmana.net>.
- Daßler, B., Heinkelmann-Wild, T., and Kruck, A., 2019, When does the barking dog bite? Hegemonic decline, the Trump administration and varieties of contestation of multilateral institutions. Paper presented at the Annual Meeting of the Political Science Section of the German Society for American Studies, Heidelberg.
- Dellmuth, L. M. and Tallberg, J., 2017, Elite communication and popular legitimacy in global governance. Working Paper.
- Dingwerth, K., Schmidtke, H., and Weise, T., 2019a, The rise of democratic legitimation: Why international organizations speak the language of democracy. *European Journal of International Relations*.
- Dingwerth, K., Witt, A., Lehmann, I., Reichel, E., and Weise, T., 2019b, *International Organizations under Pressure: Legitimizing Global Governance in Challenging Times*, First ed (Oxford: Oxford University Press).
- Ecker-Ehrhardt, M., 2018a, IO public communication and discursive inclusion: How the UN reported the Arms Trade Treaty process to a global audience. *Journal of International Relations and Development*.

- Ecker-Ehrhardt, M., 2018b, Self-legitimation in the face of politicization: Why international organizations centralized public communication. *The Review of International Organizations*, 13(4), pp. 519–546.
- Eckhard, S. and Jankauskas, V., 2019, The politics of evaluation in international organizations: A comparative study of stakeholder influence potential. *Evaluation*, 25(1), pp. 62–79.
- Eckhard, S. and Parížek, M., 2020, Policy implementation by international organizations: A comparative analysis of strengths and weaknesses of national and international staff. *Journal of Comparative Policy Analysis*.
- Francaise-Presse, A., 2017, United nations environment programme chief urges China to do more on climate. Available at <https://www.ndtv.com/world-news/united-nations-environment-programme-chief-urges-china-to-do-more-on-climate-1735504>.
- Gerhards, J., Offerhaus, A., and Roose, J., 2009, Wer ist verantwortlich? Die Europäische Union, ihre Nationalstaaten und die massenmediale Attribution von Verantwortung für Erfolge und Misserfolge, in: F. Marcinkowski and B. Pfetsch (Eds) *Politik in der Mediendemokratie*, 1st ed (Wiesbaden: VS Verlag für Sozialwissenschaften), pp. 529–558.
- Gilad, S., Maor, M., and Bloom, P. B.-N., 2013, Organizational reputation, the content of public allegations, and regulatory communication. *Public Administration*, 25(2), pp. 451–478.
- Goritz, A., Schuster, J., Jörgens, H., and Kolleck, N., 2020, International public administrations on Twitter: A comparison of digital authority in global climate policy. *Journal of Comparative Policy Analysis*.
- Gronau, J. and Schmidtke, H., 2016, The quest for legitimacy in world politics – International institutions' legitimation strategies. *Review of International Studies*, 42(3), pp. 535–557.
- Hart, M., 2018, Why detaining children is harmful. *UNICEF USA*. Available at <https://www.unicefusa.org/stories/why-detaining-children-harmful/34488> (accessed 2 June 2020).
- Hawkins, D. G., Lake, D. A., Nielson, D. L., and Tierney, M. J. (eds.), 2006, *Delegation and Agency in International Organizations* (Cambridge: Cambridge University Press).
- Heinkelmann-Wild, T., Kriegmair, L., and Rittberger, B., 2020, The EU multi-level system and the Europeanization of domestic blame games. *Politics and Governance*, 8(1), pp. 85–94.
- Heinkelmann-Wild, T. and Zangl, B., 2019, Multilevel blame games: Blame-shifting in the European Union. *Governance*.
- Heinkelmann-Wild, T. and Zangl, B., 2020, Nicht nur wehrlose Sündenböcke!: Schuldvermeidungsstrategien internationaler Organisationen. *Politische Vierteljahresschrift*.
- Hinterleitner, M. and Sager, F., 2017, Anticipatory and reactive forms of blame avoidance: Of foxes and lions. *European Political Science Review*, 9(4), pp. 587–606.
- Hinterleitner, M. and Sager, F., 2019, Blame, reputation, and organizational responses to a politicized climate, in: T. Bach and K. Wegrich (Eds) *The Blind Spots of Public*

- Bureaucracy and the Politics of Non-Coordination*, Vol. 21 (Cham: Springer International Publishing), pp. 133–150.
- Hood, C., 2011, *The Blame Game: Spin, Bureaucracy, and Self-Preservation in Government* (Princeton: Princeton University Press).
- Hood, C., Jennings, W., and Copeland, P., 2016, Blame avoidance in comparative perspective: Reactivity, staged retreat and efficacy. *Public Administration*, 94(2), pp. 542–562.
- Hooghe, L. and Marks, G., 2015, Delegation and pooling in international organizations. *The Review of International Organizations*, 10(3), pp. 305–328.
- Knill, C., Steinebach, Y., and Fernández-i-Marín, X., 2018, Hypocrisy as a crisis response? Assessing changes in talk, decisions, and actions of the European Commission in EU environmental policy. *Public Administration*, 40(3), pp. 193.
- Kruck, A., Hobbach, R., Heinkelmann-Wild, T., and Daßler, B., 2020, The Trump Administration and IOs: Varying frames of institutional contestation. Paper presented at the conference of the Section “International Relations” of the German Association for Political Science (DVPW), Freiburg.
- Lyne, M. M., Nielson, D. L., and Tierney, M. J., 2006, Who delegates? Alternative models of principals in development aid, in: D. G. Hawkins, D. A. Lake, D. L. Nielson, and M. J. Tierney (Eds) *Delegation and Agency in International Organizations* (Cambridge: Cambridge University Press), pp. 41–76.
- Maor, M., 2020, Strategic communication by regulatory agencies as a form of reputation management: A strategic agenda. Early view. *Public Administration*.
- Maor, M., Gilad, S., and Bloom, P. B.-N., 2013, Organizational reputation, regulatory talk, and strategic silence. *Public Administration*, 23(3), pp. 581–608.
- McCaskill, N. D. and Lima, C., 2017, Trump reverses on NATO: ‘It’s no longer obsolete’. *Politico*. Available at <https://www.politico.com/story/2017/04/donald-trump-nato-not-obsolete-237166> (accessed 2 March 2020).
- Nielson, D. L. and Tierney, M. J., 2003, Delegation to international organizations: Agency theory and World Bank environmental reform. *International Organization*, 57(2), pp. 241–276.
- Przeworski, A. and Teune, H., 1982, *The Logic of Comparative Social Inquiry* (Malabar, Florida: R.E. Krieger Publishing Company).
- Rauh, C. and Zürn, M., 2020, Authority, politicization, and alternative justifications: Endogenous legitimation dynamics in global economic governance. *Review of International Political Economy*, 27(3), pp. 583–611.
- Rihoux, B. and Meur, G. D., 2009, Crisp-Set Qualitative Comparative Analysis (csQCA), in: B. Rihoux and C. C. Ragin (Eds) *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques* (Los Angeles, London: Sage), pp. 33–68.

- Rihoux, B. and Ragin, C. C. (eds.), 2009, *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques* (Los Angeles, London: Sage).
- Rittberger, B., Schwarzenbeck, H., and Zangl, B., 2017, Where does the buck stop? Responsibility attributions in complex international institutions. *Journal of Common Market Studies*, 55(4), pp. 909–924.
- Schlipphak, B. and Treib, O., 2017, Playing the blame game on Brussels: The domestic political effects of EU interventions against democratic backsliding. *Journal of European Public Policy*, 24(3), pp. 352–365.
- Schmidtke, H., 2018, Elite legitimation and delegitimation of international organizations in the media: Patterns and explanations. *The Review of International Organizations*, 42(1), pp. 3.
- Schneider, C. Q. and Wagemann, C., 2013, *Set-theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis* (Cambridge, New York: Cambridge University Press).
- Skaaning, S.-E., 2011, Assessing the robustness of crisp-set and fuzzy-set QCA results. *Sociological Methods & Research*, 40(2), pp. 391–408.
- Sommer, M., 2020, Blame-shifting in times of permanent austerity: Evidence from Spain and Greece in the Eurozone crisis. *JCMS: Journal of Common Market Studies*, 58(3), pp. 562–579.
- Swanson, A., 2019, Trump cripples W.T.O. as trade war rages. *The New York Times*. Tallberg, J., Sommerer, T., Squatrito, T., and Jonsson, C., 2013, *The Opening up of International Organizations* (Cambridge: Cambridge University Press).
- Tallberg, J. and Zürn, M., 2019, The legitimacy and legitimation of international organizations: Introduction and framework. *The Review of International Organizations*.
- Thomann, E. and Ege, J., 2020, Qualitative Comparative Analysis (QCA) in public administration, in: B. G. Peters and I. Thynne (Eds) *Oxford Research Encyclopedia of Public Administration* (Oxford: Oxford Univ. Press).
- Traber, D., Schoonvelde, M., and Schumacher, G., 2019, Errors have been made, others will be blamed: Issue engagement and blame shifting in Prime Minister speeches during the economic crisis in Europe. *European Journal of Political Research*, 59(1), pp. 45–67.
- UNHCR, 2018, UNHCR statement on new US regulation on asylum. Available at <https://www.unhcr.org/news/press/2018/11/5be5bfac4/unhcr-statement-new-regulation-asylum.html> (accessed 2 March 2020).
- UPU, 2018, *Statement on the Decision of the United States of America to Withdraw from the Universal Postal Union Treaties*. Available at http://news.upu.int/no_cache/nd/statement-on-the-decision-of-the-united-states-of-america-to-withdraw-from-the-universal-postal-union-treaties/ (accessed 2 March 2020).
- UPU, 2020, The UPU. Available at <http://www.upu.int/en/the-upu/the-upu.html>. (accessed 2 March 2020).

- Vasilopoulou, S., Halikiopoulou, D., and Exadaktylos, T., 2014, Greece in crisis: Austerity, populism and the politics of blame. *Journal of Common Market Studies*, 52(2), pp. 388–402.
- Welsh, T., 2018, UNHCR balances interests in response to Trump asylum policy. Available at <https://www.devex.com/news/unhcr-balances-interests-in-response-to-trump-asylum-policy-93849> (accessed 2 March 2020).
- Yamasaki, S. and Rihoux, B., 2009, A commented review of applications, in: B. Rihoux and C. C. Ragin (Eds) *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques* (Los Angeles, London: Sage), pp. 123–145.
- Zaum, D., 2013a, International organizations, legitimacy, and legitimation, in: D. Zaum (Ed.) *Legitimizing International Organizations* (Oxford: Oxford University Press), pp. 3–25.
- Zaum, D. (ed.), 2013b, *Legitimizing International Organizations* (Oxford: Oxford University Press).
- Zürn, M., Binder, M., and Ecker-Ehrhardt, M., 2012, International authority and its politicization. *International Theory*, 4(1), pp. 69–106.

Appendix to Paper B

1 Constructing the sample and assessing IPAs' responses

In the paper, we analyse public communication by the international public administrations (IPAs)¹⁴ of major IOs in response to contestation by the Trump Administration. To construct our sample, we started with the sample of major IOs compiled by Hooghe and Marks (2015). We disaggregated the UN into its individual agencies and funds to include both UN Secretariat and such large organizations as UNESCO or UNICEF. We then excluded all IOs where the US was not a member or whose secretariat does not meet our definition of IPA. For the period ranging from Donald Trump's inauguration in January 2017 to December 2019, we then searched in IPAs' press releases as well as the media database Factiva for accounts referring to the Trump Administration and the respective IO. This allowed us to identify instances of contestation as well as IPAs' responses.

By means of keyword search¹⁵ in the media database Factiva as well as IOs' press releases, we identified accounts which refer to the Trump Administration or the US in general as well as the respective IO from the election of Donald Trump in November 9, 2016 through December 31, 2019. This allowed us to identify not only instances of direct contestation, but also instances of indirect contestation through IPAs' responses. To qualify as a case and make it into our sample, two attributes had to be present:

- **Contestation by the Trump Administration:** includes both direct and indirect contestation. *Direct contestation* includes cases when President Trump or his administration members publicly name the IO and criticize its mandate, procedures, performance, or question its overall necessity. It is also direct contestation when the Trump Administration decides (or threatens) to withdraw resources from IO budget or to end its membership. *Indirect contestation* includes cases where President Trump or his administration members publicly criticize other IO member states, question the IOs' social purpose and underlying norms (e.g. asylum rights) or act against it, for instance, by issuing policies that go against the norms and standards the IO represents (e.g. in the field of environmental protection).

¹⁴ IPAs are defined as “hierarchically organized group[s] of international civil servants with a given mandate, resources, identifiable boundaries, and a set of formal rules of procedures” (Biermann and Siebenhüner (2009: 37); see also Bauer et al. (2017)). For the purpose of this paper, our focus is on IPAs' leading representatives, including senior management and press officers. In our analysis, we only included responses by these actors but not those of individual member states or governing bodies.

¹⁵ Keywords included IO name and the name of the contesting country (the US or Trump).

- **A response by the respective IO administration:** a press release, public statement, or media interview by the senior IO officials where the issue of contestation is addressed.

Within these responses, we specifically looked at both response object and sender:

- **Response object**, i.e. the case-specific instance of member state contestation which is addressed in the response. The coding of the object was based on the observed instances of contestation (see section 2 below).
- **Response sender**, i.e. an IPA official who responds to member state contestation. In our analysis, we included statements by the heads of IOs or their main departments as well as communication officers.

In some cases, the IO administration did not mention the contesting state specifically. Yet, if the context allows attribution of the IPA communication to the contestation, it should be included in the analysis (for instance, when a journalist asks the head of IO to respond to President Trump, his answer is a public response to contestation regardless whether he or she mentions Trump or the US or not).

This procedure resulted in 32 cases of IPA responses to contestation across 18 major IOs.

2 *Operationalization of the outcome and the conditions*

2.1 *Operationalizing the outcome: coding IPAs' responses*

In order to empirically identify and distinguish the two different IPAs response approaches, we relied on two indicators: the tone and the type of a response statement.

- **Tone of response statement:** when coding the *tone* of a response statement, we differentiated between (1) positive; and (2) negative sentiment (cf. Schmidtke 2018; Tallberg and Zürn 2019; Rauh et al. 2020).
- **Type of response statement:** when coding the type of a response statement, we differentiated between (1) statements signalling accommodation of the contesting state through neglecting contested IO norms, accepting the accusations, and praising the contesting state; and (2) statements signalling the objection to the contesting state through defending contested IO norms, rejecting the accusations, shaming the contesting state (cf. Greuter 2014: 74; Gronau and Schmidtke 2016: 542; Hood et al. 2016; Heinkelmann-Wild and Zangl 2019).

The specific constellations of these observable manifestations for each strategy are briefly presented below and are summarized in Table S.1:

- **Conciliatory response:** When an IPA opts for a conciliatory response, we expect it to adopt a positive tone and signal its accommodation of the contesting state through neglecting contested IO norms, accepting the accusations, and praising the contesting state.
- **Adversarial response:** In the case of an adversarial response, we expect the IPA to adopt a negative tone and signal its objection to contestation through defending contested IO norms, rejecting the accusations, and shaming the contesting state.

In our content analysis, the tone and type of IPA response statements scored unambiguously as conciliatory or adversarial – we thus did not identify any statements that showed mixed indicators of the two response modes (e.g. negative tone combined with signaling accommodation). To ensure that the coding of IPA response statements can be reliably reproduced, in a first step, each statement was coded independently by the authors. In a second step, coding decisions were compared. When disagreement occurred, these were discussed based on the coding rules and resolved consensually (Kuckartz 2016).

Since our research interest is not on the level of individual statements but IPAs' overall response, we aggregated IPAs' response statements for each case. In the rare case that not all statements had the same outcome value (i.e. conciliatory or adversarial response), we assigned the case to the outcome showed by the majority of statements.

Instances of direct contestation where no IPA response was found, were excluded from the sample. This concerned, for instance, UN Women, which was contested indirectly but did not issue an official statement. When the same IO was contested several times, but the object of contestation has changed, cases were coded separately. For instance, UNHCR has been indirectly contested three times: (1) In 2017, President Trump issued a refugee order, challenging humane migration and refugee protection (UNHCR1); (2) President Trump tightened rules on asylum seekers in 2018 (UNHCR2); and (3) the US issued controversial migration policies in 2019 (UNHCR3). Following these rules, we arrived at 32 IPA responses – and thus cases – in our sample, 13 of them adversarial and 19 conciliatory. For an overview of IPA responses, including examples for each case, see Table S.4.

	Conciliatory response	Adversarial response
Tone	<p>Positive tone communication: e.g. “agree”, “support”, “compliment”, “leadership”, “cooperate”, “welcome”</p>	<p>Negative tone communication: e.g. “reject”, “impossible”, “unfortunately”, “disappointment”, “mistake”, “demand”</p>
Type	<p>Signaling accommodation by</p> <p>(i) Accepting the accusations: “I welcome his [Trump’s] very strong message on defense spending, on burden sharing, and on NATO’s role in fighting terrorism, that we have to step up and do more.” (McCaskill and Lima 2017)</p> <p>(ii) Praising the contesting state: “[IOM and UNHCR] hope that the US will continue its strong leadership role and long tradition of protecting those who are fleeing conflict.” (IOM and UNHCR 2017)</p> <p>(iii) Neglecting contested IO norms: “Decisions on which refugees will be resettled and where they go are made solely by the [...] the United States.” (Reuters 2018)</p>	<p>Signaling objection by:</p> <p>(i) Rejecting the accusations: “We [UNRWA] reject in the strongest possible terms the criticism that UNRWA’s schools, health centers, and emergency assistance programs are ‘irredeemably flawed.’” (UNRWA 2018b)</p> <p>(ii) Shaming the contesting state: President Trump “clearly made a mistake” in withdrawing from the Paris Agreement (Agence Francaise-Pressé 2017).</p> <p>(iii) Defending contested IO norms: The US child separation policy is “unconscionable” and an “abuse” which may cause “irreparable harm.” (Toolsi 2018)</p>

Table S.1: Observable implications of IPA’s communicative responses to state contestation.

2.2 Operationalizing the dependence condition

To assess the *dependence condition*, we determined the relative IPA’s dependence on the US by using the share of an IO’s total budget provided by the US as a proxy (Lyne et al. 2006: 58). We opted for the US share of total contributions in 2016/2017 since we assume that this level functions as a baseline for IPAs’ calculations vis-à-vis the Trump Administration.¹⁶ We derived information about US budget contributions from IOs’ internet presence (e.g. UN Chief Executives Board for Coordination 2020) as well as dataset compiled by McArthur and Rasmussen (2017).

¹⁶ While we acknowledge that the share of contributions under the Trump Administration, which came into office in February 2017, might have changed (also as part of the contestation) as compared to the previous Obama administration, we assume that IPAs formulate their response considering the level of contributions they were used to in earlier years.

To decide whether an IPA is dependent on the US or not, we used 15% of total contributions (including both core and voluntary contributions) as a cut-off point. While every cut-off point is somewhat arbitrary, our choice was based on the following considerations: First, we hold that this decision is empirically plausible in a sense that IOs like the UNHCR with the US funding reaching around 37% can be certainly considered as dependent on the US while IOs such as UNEP, whose funding by the US reaches around 6%, are clearly less dependent. The 15% cut-off thus offers a plausible middle point. Second, in the large majority of IOs that receive more than 15% of their funding from the US, the US is also the single largest donor. Third, a visual assessment of the distribution of relative US contributions in Figure S.1 shows that there are three different clusters: IOs with US contributions over 30% (UNHCR, IOM, WFP, UNRWA, IAEA), between 30% and 15% (UNICEF, NATO, UNSG, UNFCC, FAO, IMF, WB), and below 15% (UPU, UNEP, WTO). In section 4.2 we also conduct robustness checks for using 30% as a cut-off point. The 15% threshold thus constitutes the cluster of IOs with the lowest contributions.

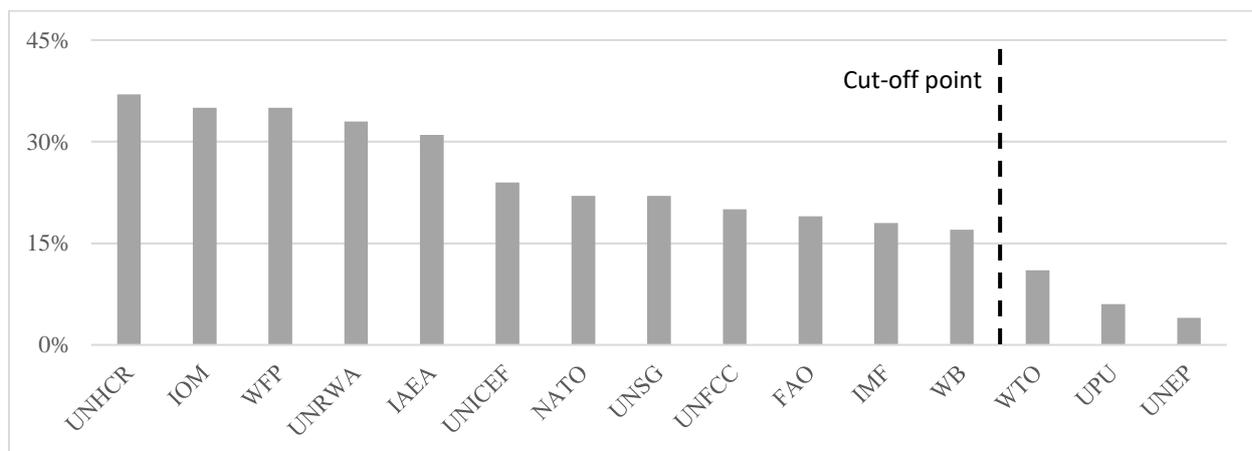


Figure S.1: US share of contributions to total budget.

Finally, we considered including weighted voting rights as an additional measure for an IPA's dependence from a contesting member state. Weighted voting rights are "voting practices where some members of the institution have greater voting power than others, giving the former greater influence over IGO decisions" (Blake and Payton 2015: 382). Weighted voting renders IOs asymmetric since states possessing unequal voting rights are in a privileged position to both foster and prevent outcomes. However, the only two IOs with weighted voting in our sample are the IMF and WB where the asymmetric voting share of the US in fact corresponds to US predominant resource contributions slightly over 15%.

In five cases, the Trump Administration's contestation comprised either immediate exit from an IO (UNESCO, UNRWA2, UNHRC) or a complete funding cut (UNFPA1,2). Since these IPAs cannot expect to obtain the same funding level as under the Obama Administration (even when responding conciliatory), they were also coded as low dependent. In section 4.2 we also conduct robustness checks to test whether our findings hold when treating only those IOs as independent which the US exited during contestation.

Taken together, we arrive at the following coding rule:

- **Dependence on contesting state:** The IPA was coded as dependent on the Trump Administration when the US contributed more than 15% to an IO's total budget (as of 2016/2017).
- **Independence from contesting state:** We coded an IPA as independent from Trump Administration when the US contributed less than 15% to an IO's total budget (as of 2016/2017) or contestation in the respective case took the form of a complete funding cut or membership termination.

Please note that the WFP2 case deviates in its dependence from the WFP1 case. This coding decision was based on the following consideration: A closer look at the WFP2 case revealed that not WFP in general but its Palestine program specifically was contested by the US and the response also stemmed from the Palestine country director. As US contributions to the Palestine program are considerably smaller (3%), than to the WFP in general, we consequently changed the dependence for the WFP2 case from "1" to "0". The alternative of dropping the case and focusing only on the communication of bureaucrats at an IOs' center would not have changed our overall findings but only decrease the coverage of our solution for adversarial IPA responses. In the sample, in 23 out of the 32 cases the IPA was highly dependent on the US, while the dependence of the IPAs in the remaining 9 cases was low.

2.3 Operationalizing the authority condition

To assess IPAs' *policymaking authority*, we referred to the primary tasks the IPA conducts. These were coded based on IOs' formal mandates as well as self-description at their online appearances. The following rules were applied:

- **IPAs without policymaking authority:** We coded an IPA as lacking policymaking authority when its main task is to assist member states in bargaining.

- **IPA with policymaking authority:** We coded an IPA as possessing policymaking authority when its main task is to provide policy advice to the member states or implement policies.

Consider the following two examples: the NATO secretariat, consisting of International Staff, steers the “process of consultation and decision-making” among the member states (NATO 2017, 2018b). It is merely tasked with managing member states’ relationships and assisting intergovernmental bargaining. By contrast, UNEP’s administration has considerable policymaking authority. Its declared mission “is to provide leadership and encourage partnership in caring for the environment” (UNEP 2019). The organization is active in seven thematic issue areas: climate change, disasters and conflicts, ecosystem management, environmental governance, chemicals and waste, resource efficiency, and environment under review (UNEP 2019). The nature of UNEP work is primarily policy advice and expertise. In 24 of the 32 cases, the IPA possessed policymaking authority.

2.4 Operationalizing the affectedness condition

To assess the *type of contestation*, we determined the thrust of the Trump Administration’s contestation. To identify instances of contestation, we again relied on a content analysis of the media coverage (see section 1). The covered period begins with the election of Donald Trump Starting at November 9, 2016 and ends in December 31, 2019. Thus, contestation by the president-elect is included in our sample. The following coding rules were applied to differentiate between direct and indirect contestation:

- **Direct contestation:** We coded that an IPA is contested directly when the Trump Administration assaulted an IO’s bureaucracy by specifically naming the IO and criticizing its mandate, procedures, performance, or even the overall necessity in general; and when the US government withdrew resources or its membership from an IO, or threatened to do so.
- **Indirect contestation:** The Trump Administration’s contestation was coded as indirect when the US government representatives criticized the behavior of other member states, questioned the IOs’ social purpose or acted against it (i.e. non-compliance).

Take as an example the Trump Administration claiming that the structure and fiscal outlays of the UNRWA are “unsustainable” and operations “irredeemably flawed” (Staff 2018). This constitutes a direct contestation to the IO and its bureaucracy, since the object of blame is

clearly defined. By contrast, the Trump Administration did not directly criticize the work of UNEP but contested its environmental policies as well as its domain of expertise by neglecting (human-made) climate change.

As the coding rules imply, we covered contestation by both words and deeds (Deitelhoff and Zimmermann 2018; Tallberg and Zürn 2019). Table S.2 summarizes types of contestation and their observable manifestations covered in our analysis:

	Direct	Indirect
Words	<ul style="list-style-type: none"> • Criticizing IPA • Threat of funding cut • Threat of exit 	<ul style="list-style-type: none"> • Criticizing IO member states • Criticizing IO norms
Deeds	<ul style="list-style-type: none"> • Terminating funding • Terminating membership 	<ul style="list-style-type: none"> • Non-compliance with IO norms

Table S.2: Observable implications of contestation by member states.

In our sample, the Trump Administration contested IPAs directly in 14 of the 30 cases. Please note that we could have coded an additional UNHRC case of direct contestation as the US exited the organization. However, as the withdrawal decision is linked and immediately occurred after the first instance of contestation, where it was indirectly contested by Trump's policy to separate children from their parents at the border, we consider it as one case. To be sure, including a second UNHRC case of direct contestation would not change the solution obtained in the QCA. Please note that we refer to UNHRC as an equivalent to the overall organization, i.e. Office of the United Nations High Commissioner for Human Rights (OHCHR).

Table S.3 summarizes the empirical observations for all three conditions and the outcome and Table S.4 provides a more substantive overview for each IO through a short summary of the Trump Administration's contestation as well as examples of IPAs' responses.

Case	Dependence	Authority	Affectedness	Outcome
IOM	High (35%)	High	Indirect	Conciliatory
UNHCR1	High (37%)	High	Indirect	Conciliatory
UNHCR2	High (37%)	High	Indirect	Conciliatory
UNHCR3	High (37%)	High	Indirect	Conciliatory
WFP1	High (35%)	High	Direct	Conciliatory
WFP2	Low (3%)	High	Direct	Adversarial
NATO1	High (22%)	Low	Direct	Conciliatory
NATO2	High (22%)	Low	Indirect	Conciliatory
NATO3	High (22%)	Low	Indirect	Conciliatory
FAO	High (19%)	High	Direct	Conciliatory
UNSG1	High (22%)	High	Direct	Conciliatory
UNSG2	High (22%)	High	Direct	Conciliatory
UNFCC1	High (20%)	Low	Direct	Conciliatory
UNFCC2	High (20%)	Low	Indirect	Conciliatory
WB1	High (17%)	High	Indirect	Conciliatory
WB2	High (17%)	High	Indirect	Conciliatory
IMF1	High (18%)	High	Indirect	Conciliatory
IMF2	High (18%)	High	Direct	Conciliatory
UPU	Low (6%)	Low	Indirect	Conciliatory
UNEP	Low (4%)	High	Indirect	Adversarial
IAEA1	High (31%)	High	Direct	Adversarial
IAEA2	High (31%)	High	Indirect	Adversarial
IAEA3	High (31%)	High	Direct	Adversarial
WTO1	Low (11%)	Low	Direct	Adversarial
WTO2	Low (11%)	Low	Direct	Adversarial
UNFPA1	Low (withdrawal)	High	Direct	Adversarial
UNFPA2	Low (withdrawal)	High	Direct	Adversarial
UNRWA1	High (33%)	High	Direct	Conciliatory
UNRWA2	Low (withdrawal)	High	Direct	Adversarial
UNESCO	Low (withdrawal)	High	Direct	Adversarial
UNHRC	Low (withdrawal)	High	Indirect	Adversarial
UNICEF	High (24%)	High	Indirect	Adversarial

Table S.3: Cases and coded values (raw data).

CaseID	Description of contestation	Example for IPA response
IOM	Indirect: Trump Administration issued the 'Muslim Ban', challenging humane migration and refugee protection [27.01.2017] (Executive Order 13769 2017)	Conciliatory: "[...] the US resettlement program is one of the most important in the world. [...] [IOM and UNHCR] hope that the US will continue its strong leadership role and long tradition of protecting those who are fleeing conflict and persecution..." (Press Release) [28.01.2017] (IOM and UNHCR 2017)
UNHCR1	Indirect: Trump Administration issued the 'Muslim Ban', challenging humane migration and refugee protection [27.01.2017] (Executive Order 13769 2017)	Conciliatory: "[...] the US resettlement program is one of the most important in the world. [...] [IOM and UNHCR] hope that the US will continue its strong leadership role and long tradition of protecting those who are fleeing conflict and persecution..." (Press Release) [28.01.2017] (IOM and UNHCR 2017)
UNHCR2	Indirect: Trump Administration tightened rules on asylum seekers [08.11.2017] (Miroff 2018)	Conciliatory: "UNHCR stands ready at all times to support the United States and all governments and civil society partners working to guarantee that any person fleeing life-threatening violence or persecution is able to reach safe ground and is able to have their claim reviewed" (Press Release) [09.11.2018] (UNHCR 2018)
UNHCR3	Indirect: Trump Administration reduced numbers of refugees allowed into the US [16.09.2018] (Reuters 2018)	Conciliatory: "Decisions on which refugees will be resettled and where they go are made solely by the governments of countries that, like the United States, admit refugees for resettlement" (William Spindler, UNHCR spokesman) [18.09.2018] (Reuters 2018)
WFP1	Direct: Trump Administration threatened to cut WFP budget [23.03.2017] (White House 2017a)	Conciliatory: "We have people at the top of their professions in logistics and technology [...] I want America to be great too. And I do believe [...] that America is and can be great because she is good, and that ceasing to be good would make America much less great" (David Beasley, WFP Executive Director) [10.07.2017] (WFP 2017)
WFP2	Direct: Trump Administration cut WFP budget and forced it to reduce operations in Palestinian territories [December 2018] (Farrell and al-Mughrabi 2018)	Adversarial: "WFP has been forced, unfortunately, to make drastic cuts to the number of people that we support across Palestine [...] They [US] have cut funding [...] to the rest of the humanitarian community" (Stephen Kearney, WFP Country Director) [December 2018] (Farrell and al-Mughrabi 2018)
NATO1	Direct: President-elect Trump questioned NATO's necessity, criticized its issue prioritization [January 2017] (BBC 2017)	Conciliatory: "I welcome his [Trump's] very strong message on defense spending, on burden sharing, and on NATO's role in fighting terrorism, that we have to step up and do more" (Jens Stoltenberg, NATO Secretary General) [April 2017] (McCaskill and Lima 2017)
NATO2	Indirect: Trump Administration questioned core NATO's norms (collective defense), blamed other MS for unfair funding, threatened to	Conciliatory: "We understand that this American president is very serious about defence spending. And this is having a clear impact [...] There is a new sense of urgency due

	withdraw US support to NATO [June 2018] (Davis 2018)	to President Trump's strong leadership on defence spending" (Jens Stoltenberg, NATO Secretary General) [12.07.2018] (NATO 2018a)
NATO3	Indirect: Trump Administration blamed NATO member states for not spending 2% for defence [December 2019] (Heath 2019)	Conciliatory: "[Trump's] leadership on defense spending is having a real impact [...] This is unprecedented. This is making NATO stronger." (Jens Stoltenberg, NATO Secretary General) [December 2019] (NATO 2019)
FAO	Direct: Trump Administration criticized FAO management, staff composition, threatened to cut its contributions [February 2017] (Italian Insider 2017)	Conciliatory: "FAO is a proud partner of the US and ready to pursue the US ambitions in agriculture" (Graziano da Silva, FAO Director-General) [July 2017] (Italian Insider 2017)
UNSG1	Direct: Trump Administration criticized UN bureaucracy, threatened to cut its budget [27.02.2017] (White House 2017a)	Conciliatory: "budgetary process in the U.S. is complex and lengthy and it needs to be completed [...] We are indeed very grateful for the support the United States has given to the United Nations over the years as the organization's largest financial contributor." (Stéphane Dujarric, Spokesman for UN Secretary General) [24.05.2017] (Gladstone 2017)
UNSG2	Direct: Trump asked for more effectiveness and accountability of the UN, maintained reduced funding to the UN peacekeeping budget [25.09.2018] (White House 2018)	Conciliatory: "The Secretary-General welcomed the United States' efforts to bring peace and denuclearization in the Korean Peninsula, as well as its support to humanitarian operations around the world" (UN Press Release) [25.09.2018] (UN Secretary-General 2018)
UNFCCC1	Direct: Trump Administration threatened to cut US contributions to the organization [March 2017] (Mathiesen 2017)	Conciliatory: "As Executive Secretary of the UNFCCC I, like many people and organizations around the globe, are watching these developments with interest [...] I have made it clear from the outset, following the change in the US administration, that the secretariat works with all Parties to advance climate action [...]" (Patricia Espinosa, UNFCCC Executive Secretary) [31.03.2017] (UNFCCC 2017a)
UNFCCC2	Indirect: Trump Administration decided to leave the Paris Agreement [01.06.2017] (White House 2017b)	Conciliatory: UNFCCC "regrets the announcement [...] stands ready to engage in dialogue with the United States government regarding the implications of this announcement." (Press Release) [01.06.2017] (UNFCCC 2017b)
WB1	Indirect: Trump Administration accused other WB states of financing too many projects in China [March 2017] (Gallagher 2017)	Conciliatory: World Bank's President Jim Yong Kim said he was "encouraged by his engagement so far with the Trump Administration, including interest from Ivanka Trump in the bank's work on women's issues." (April 2017) (Lawder 2017)
WB2	Indirect: Trump Administration called other WB states to stop lending to China [07.12.2019] (The Guardian 2019)	Conciliatory: "World Bank lending to China has fallen sharply and will continue to reduce as part of our agreement with all our shareholders including the United States." (World Bank Statement) [07.12.2019] (Reuters 2019b)

IMF1	Indirect: Trump Administration promoted protectionist, anti-globalist policies [March 2017] (Stewart 2017)	Conciliatory: “On the US, there is already a great discrepancy between the pre-election rhetoric and the actual steps that have been taken in the last few weeks. My view is to focus on the actual measures, the policy determination, the laws that are implemented – rather than speculate on the possible interpretation of this or that.” (Christine Lagarde, IMF Managing Director) [April 2017] (IMF 2017)
IMF2	Direct: Trump Administration opposed an increase of IMF budget [12.12.2018] (Mayedam Andrew and Mohsin 2018)	Conciliatory: “[...] discussions will continue guided by our members' common interests in ensuring that the IMF remains strong and well-resourced [...] we look forward to continuing to work with the U.S. [...] to ensure [...] that the IMF is sufficiently and efficiently resourced” (Gerry Rice, IMF spokesman) [13.12.2018] (IMF 2018)
UPU	Indirect: Trump Administration blamed China for misusing the UN postal scheme [October 2018] (The Guardian 2018)	Conciliatory: “UPU Director General Bishar A. Hussein regrets the decision and will seek to meet representatives of the Government of the United States of America to further discuss the matter” (Press Release) [17.10.2018] (UPU 2018)
UNEP	Indirect: Trump decided to leave the Paris Agreement, issued anti-environmental policies [June 2017] (Agence Francaise-Pressé 2017)	Adversarial: President Trump “clearly made a mistake” in withdrawing from the Paris Agreement (Erik Solheim, UNEP Executive Director) [July 2017] (Agence Francaise-Pressé 2017)
IAEA1	Direct: Trump Administration criticized IAEA monitoring capabilities, pressured to find evidence against Iran [August 2017] (Sanger and Gladstone 2017)	Adversarial: “What you hear from the [US] state department people is that in the future they have some concerns [about] the rotation or recruitment but this is something that we can manage [...] We are loyal only to the agency [...] We have had access to all the locations that we needed to visit” (Yukiya Amano, IAEA Director-General) [November 2017] (Manson 2017)
IAEA2	Indirect: Trump Administration withdrew from the Iran nuclear deal [08.05.2018] (Landler 2018)	Adversarial: “Iran is subject to the world’s most robust nuclear verification regime under the JCPOA, which is a significant verification gain. As of today, the IAEA can confirm that the nuclear-related commitments are being implemented by Iran” (Yukiya Amano, IAEA Director-General) [09.05.2018] (Iran Watch 2018)
IAEA3	Direct: Trump Administration questioned IAEA expert authority, pressured to change monitoring activities [January 2019] (Reuters 2019a)	Adversarial: “If our credibility is thrown into question and, in particular, if attempts are made to micro-manage or put pressure on the agency in nuclear verification, that is counter-productive and extremely harmful” (Yukiya Amano, IAEA Director-General) [January 2019] (IAEA 2019)
WTO1	Direct: Trump Administration blamed WTO for being unfair to US,	Adversarial: “Whether the shock and awe of Trump Administration trade policy can be

	threatened to exit [30.08.2018] (Micklethwait et al. 2018)	channeled into making the trading system better is an open question” (Alan Wolff, WTO Deputy Director-General) [15.10.2018] (WTO 2018)
WTO2	Direct: Trump Administration criticized WTO for being unfair, blocked appointment of judges, threatened to block future WTO’s budget [December 2019] (Becker 2019)	Adversarial: The US challenge does “not mean the end of the multilateral trading system” (Roberto Azevedo, WTO Director- General) [08.12.2019] (Swanson 2019)
UNFPA1	Direct: Trump Administration criticized UNFPA functioning (esp. programs in China), declared total budget cut [March 2017] (Nichols 2017)	Adversarial: “[The US] decision is based on the erroneous claim [...] UNFPA refutes this claim [...]” (Press Release) [April 2017] (UNFPA 2017)
UNFPA2	Direct: Trump Administration criticized UNFPA functioning (esp. programs in China), renewed total budget cut [March 2018] (UNFPA 2018)	Adversarial: “This unfortunate decision will impede UNFPA’s crucial work to protect the health and lives of hundreds of millions of women and girls around the globe, including in humanitarian setting.” (Press Release) [12.03.2018] (UNFPA 2018)
UNRWA1	Direct: Trump Administration criticized UNRWA functioning and partiality, cut its budget [January 2018] (UNRWA 2018a)	Conciliatory: “The US has consistently been UNRWA’s largest single donor, something we sincerely thank the American people for, and countless American decision-makers – presidents, members of Congress, diplomats and civil servants, who embodied the commitment of assisting a vulnerable people through UNRWA.” (Press Release) [17.01.2018] (UNRWA 2018a)
UNRWA2	Direct: Trump Administration criticized UNRWA functioning and partiality, declared total budget cut [31.08.2018] (Beaumont and Holmes 2018)	Adversarial: “We [UNRWA] reject in the strongest possible terms the criticism that UNRWA’s schools, health centers, and emergency assistance programs are “irredeemably flawed”” (Press Release) [01.09.2018] (UNRWA 2018b)
UNESCO	Direct: Trump Administration accused UNESCO for partiality, criticized its bureaucratic inefficiencies, and withdrew its membership [12.10.2017] (Al Jazeera 2017)	Adversarial: “I obviously regret their [US] departure [...] but this ‘empty chair politics’ is not sustainable because the United States is also affected by everything that UNESCO does.” (Audrey Azoulay, UNESCO Director-General) [10.11.2017] (Adamson 2017)
UNHRC	Indirect: Trump Administration issued controversial child separation policy (when migrant children are forcibly separated from their parents at the border) [May 2018] (Toolsi 2018)	Adversarial: The US policy is “unconscionable” and an “abuse” which may cause “irreparable harm” (Zeid Ra’ad al-Husseini, UN High Commissioner for Human Rights) [18.06.2018] (Toolsi 2018)
UNICEF	Indirect: Trump Administration issued controversial child separation policy (when migrant children are forcibly separated from their parents at the border) [May 18] (Toolsi 2018)	Adversarial: “There is a lot of uncertainty surrounding the separation of children from their parents along the United States’ southwestern border, but one thing is clear: family separation is traumatic for all involved” (UNICEF USA statement) [21.06.2018] (Hart 2018)

Table S.4: Overview of cases.

3. Robustness checks

We conducted four robustness checks. Following the suggestions of Skaaning (2011), we checked the robustness of our findings by means of methodological triangulation, on the one hand, and probing the robustness of our analytical decisions during the QCA regarding calibration, raw consistency, and row frequency, on the other hand (see also, Thomann and Maggetti 2017). The results of these tests increase our confidence in the robustness of our findings.

3.1 Methodological triangulation: chi-square test of independence

We conducted a statistical analysis of contingency tables which allows us to check whether the observed patterns in our sample are only random. Specifically, we test the null hypothesis that there is no relationship between the conditions and the outcome. Table S.5 shows our observations with the complete solution terms and the outcome, contrasted by the expected absolute values for a random distribution (in brackets).¹⁷

	Dependence * (AUTHORITY + AFFECTEDNESS)	DEPENDENCE + (authority + affectedness)	Row totals
Conciliatory response	0 (5)	19 (14)	19
Adversarial response	9 (4)	4 (9)	13
Column totals	9	23	32

Table S.5: Complete solutions for conciliatory and adversarial response – observed values vs. expected values for a random distribution (in brackets).

If the null hypothesis is true, we would expect the overall ratio of IPAs' conciliatory and adversarial responses (i.e. the rightmost column) to correspond to the ratio in the other two columns in Table S.5. Yet, the observed and expected values deviate quite considerably from each other. To evaluate this statistically, we conducted a *chi-square test*. The obtained chi-square value of 18.3 implicates that the null hypothesis can be rejected at the 0.01 level of

¹⁷ The expected value for each cell is calculated by multiplying the row total by the column total, then dividing by the grand total.

significance (99% confidence level). In other words, we are assured that the identified patterns in our sample are not random with the probability being 1% of rejecting the null hypothesis when it is true.

Since this is not yet a proof for the correlation of the conditions and the outcome, we examined the strength of their association by calculating the *contingency coefficient*.¹⁸ The obtained value of 0.60 comes close to the theoretical maximum value of 0.71.¹⁹ This strengthens our confidence that a meaningful relationship between the identified configurations of conditions and the values of the outcome exists.

The remainder of this section turns to the QCA and checks how sensitive its results are to changes regarding calibration, raw consistency, and row frequency (Skaaning 2011: 395–403).

3.2 Calibration

Regarding calibration, we checked whether changes to our cut-off point of 15% for the dependence condition alter our findings by increasing the threshold for dependence to 30% or lowering it to mere membership. While changes in both directions decrease the solutions' coverage (while increasing its consistency), they remain subsets of our original solution.

On the one hand, as discussed in Section 3.1, besides the 6%-points difference between WTO (17%) and the World Bank (11%), a second 7%-point gap is between UNICEF (24%) and IAEA (31%). When using 30% as a higher cut-off point for dependence on the US, the truth table (Table S.6) contains five contradicting rows. The only consistent row shows a conciliatory outcome. The third row has a consistency of 0.8 and can thus be treated as showing a conciliatory outcome, too. Not a single row shows consistently shows an adversarial outcome. Two configurations lack any empirical observations and thus are logical remainders.

¹⁸ The contingency coefficient is a modification of the phi-coefficient (we obtained value of 0.76), and is identical with Cramer's V if one of the variables under examination is binary.

¹⁹ While the contingency coefficient assumes values between 0 and 1, it rarely reaches the maximum value of 1. As its upper limit is a function of the number of columns and rows in the table, we use the theoretical maximum value for its evaluation.

	Dependence	Authority	Affectedness	Adversarial
FAO (0), UNSG1 (0), UNSG2 (0), IMF2 (0), WFP2 (1), UNFPA1 (1), UNFPA2 (1), UNRWA2 (1), UNESCO (1)	0	1	1	C
WB1 (0), WB2 (0), IMF1 (0), UNEP (1), UNHRC (1), UNICEF (1)	0	1	0	C
IOM (0), UNHCR1 (0), UNHCR2 (0), UNHCR3 (0), IAEA2 (1)	1	1	0	C
NATO1 (0), UNFCC1 (0), WTO1 (1), WTO2 (1)	0	0	1	C
WFP1 (0), UNRWA1 (0), IAEA1 (1), IAEA3 (1)	1	1	1	C
NATO2, NATO3, UNFCC2, UPU	0	0	0	0
Logical remainder	1	0	1	?
Logical remainder	1	0	0	?

Table S.6: Truth table using 30% as a threshold for the dependence condition.

We thus cannot find any solution for an adversarial response. For a conciliatory response, the solution (see Table S.7) is a subset of our original solution. Its first part covers NATO2, NATO3, UNFCC2, UPU (at 1.00 consistency) while the second part covers IOM, UNHCR1, UNHCR2, UNHCR3, with IAEA2 being the only deviating case (at 0.80 consistency). The coverage of each part is limited to four out of 19 cases showing a conciliatory response.

Solution	$\text{dependence} * \text{authority} * \text{affectedness}$	+	$\frac{\text{DEPENDENCE} * \text{AUTHORITY} * \text{affectedness}}{\text{SS}}$	→	adversarial
Consistency	1.00		0.80		
Raw coverage	0.21		0.21		
Unique coverage	0.21		0.21		

Table S.7: Solution for conciliatory response using 30% as a threshold for the dependence condition.

The complete solution for a conciliatory response has a coverage of 0.42, while its consistency of 0.89 is close to the maximum of 1. Hence, moving the threshold for dependence from contributing 15% to an IO's budget to 30% leads to a decrease of the complete solution's coverage (0.42 compared to 1) while its consistency improves (0.89 compared to 0.82).

On the other hand, one could argue that the US as the most powerful state in the international system possesses considerable influence in all IOs they are a member of. We therefore checked whether our findings hold when treating only those IOs as independent which the US exited during contestation. In the resulting truth table (Table S.8), only three rows show consistent outcomes, two adversarial and one conciliatory. The three inconsistent rows do not allow for simplification since their consistency is below 0.75. Three configurations constitute logical remainders since they lack any empirical observations.

	Dependence	Authority	Affectedness	Adversarial
UNFPA1, UNFPA2, UNRWA2, UNESCO	0	1	1	1
NATO2, NATO3, UNFCC2, UPU	1	0	0	0
NATO1 (0), UNFCC1 (0), WTO1 (1), WTO2 (1)	1	0	1	C
IOM (0), UNHCR1 (0), UNHCR2 (0), UNHCR3 (0), WB1 (0), WB2 (0), IMF1 (0), UNEP (1), UNICEF (1), IAEA2 (1)	1	1	0	C
WFP1 (0), FAO (0), UNSG1 (0), UNSG2 (0), IMF2 (0), UNRWA1 (0), WFP2 (1), IAEA1 (1), IAEA3 (1)	1	1	1	C
UNHRC	0	1	0	1
Logical remainder	1	1	0	?
Logical remainder	0	0	1	?
Logical remainder	0	0	0	?

Table S.8: Truth table using exit as a threshold for the dependence condition.

The solution for a conciliatory response is a subset of the original solution:

$$\text{DEPENDENCE*authority*affected} \rightarrow \text{adversarial}$$

With only one row showing a consistent conciliatory outcome, four out of 19 cases with a conciliatory response are covered by the solution: NATO2, NATO3, UNFCC2, UPU. The coverage thus only reaches 0.21 (compared to 1 for the original solution), while the consistency of the solution is 1 (compared to 0.82 for the original solution).

With regard to the outcome of an adversarial response, the solution again is a subset of the original solution:

dependence*AUTHORITY → ADVERSARIAL

Five out of 13 cases showing an adversarial response are covered by the solution: UNFPA1, UNFPA2, UNRWA2, UNESCO, UNHRC. The solution thus has a coverage of 0.38 (compared to 0.69 for the original solution). As the original solution, the consistency reaches the maximum value of 1. Hence, when using exit as a threshold for the dependence condition, the consistency remains the same (for a conciliatory response) or even increases (for an adversarial response), while the coverage decreases.

To sum up, our decision to use 15% of budgetary contributions as a threshold for an IO's dependence on the US, compared to the discussed alternatives, allowed us to resolve contradictions in the data and obtain a considerable level of both coverage and consistency. Our solution is robust as the obtained solutions when increasing the threshold for dependence to 30% or lowering it to mere IO membership are subsets of our original solution. Yet, both alternatives come with a substantive decrease in coverage (while consistency increases).

3.3 Raw consistency

We also tested the robustness regarding raw consistency. In our original truth table (see Table 2 in the paper), two rows showed contradictory values for the outcome. Since their consistency was above the threshold of 0.75, we treated both configurations as showing a conciliatory IPA response as an outcome (Rihoux and Meur 2009: 49; Schneider and Wagemann 2013: 122). Two alternative options to treat would have been to exclude the inconsistent configurations from our analysis; or to ignore the contradictory cases and allow the computer to generate assumptions about them

(Rihoux and Meur 2009: 49; Yamasaki and Rihoux 2009: 133; Schneider and Wagemann 2013: 121–122). In the following, we therefore tested how much our results change when we would have opted for one of these options.

When excluding the inconsistent truth table rows in our analysis, the solution for an adversarial IPA remains the same (see Table S.5). And for a conciliatory IPA response, one part of the solution remains identical with the original solution (authority*affectedness), while the second part (DEPENDENCE*authority) is a subset of the original solution (DEPENDENCE) (see Table S.9).

Solution	DEPENDENCE*authority	+	authority*affectedness	→	adversarial
Consistency	1.00		1.00		
Raw coverage	0.26		0.21		
Unique coverage	0.11		0.05		

Table S.9: Causal paths leading to an adversarial response.

The complete solution for conciliatory IPA responses covers six out of 19 cases showing a conciliatory response (NATO1, 2, 3; UNFCC2; UNFCC1; UPU), resulting in a coverage of 0.32. Its consistency reaches the maximum of 1. Hence, excluding the two inconsistent rows in our analysis would lead to a decrease of the coverage of the solution for the outcome of a conciliatory IPA response (0.32 compared to 1 for the original solution), while its consistency improves (1 compared to 0.82 for the original solution).

Alternatively, we allowed the computer to generate assumptions about the two inconsistent configurations and thus treat them as logical remainders for the purpose of the analysis. The solutions for both an adversarial and conciliatory IPA response equal our original solution.

3.4 Row frequency

Finally, we also conducted a robustness test for row frequency to assess whether our results change if we ignore truth table rows showing only one case. In our original truth table (see Table 2 in the paper), there is only row for which we have only one case (UPU). The solution for an adversarial IPA again remains the same. Without accounting for UPU, for a conciliatory IPA response, the solution is a subset of the original solution:

$$\text{DEPENDENCE} \rightarrow \text{adversarial}$$

The solution for the outcome of an adversarial IPA response does not change. Ignoring the row with only one case thus leads to a decrease of the solution's coverage (from 1 to 0.95, since UPU is not covered anymore) while its consistency remains the same.

4. Checking for omitted conditions: the learning and the resilience condition

We also calculated two further model specifications to check whether we omitted a condition. First, we probed whether experience with *earlier contestation by the same member state* shapes IPAs' responses. One could argue that IPAs learn that their initial response was unable to put

an end to contestation and thus react adversarial at the second instance of contestation ('learning condition'). Second, we checked whether *earlier sanctions by the contesting member state* shapes IPAs' responses. One could argue that IOs which experienced a withdrawal of resources or even membership by a contesting state in the past are now better prepared for such instances. They might thus be less constrained to respond in an adversarial manner to new instances of contestation ('resilience condition'). We conducted a QCA for two models, one including the learning condition and one including the resilience condition, below. The results of these tests increase our confidence in our original, more parsimonious solution.

4.1 Adding a learning condition

To check whether past experience with contestation by the Trump Administration shapes IPAs' responses, we conducted a QCA based on our original conditions as well as a 'learning condition'. To assess the 'learning condition', we coded for each case whether it was the first time the Trump Administration contested the respective IO ('0') or not ('1').

The resulting truth table (see Table S.10) comprises nine fully consistent rows, while four rows show contradicting cases. In three of these four rows, the consistency is 0.75 or higher. We therefore treat them as showing a conciliatory IPA response. The contradicting cases showing an adversarial IPA response are IAEA1, 2, 3 and UNICEF. Three configurations lack any empirical observations and thus are logical remainders.

	Resilience	Dependence	Authority	Affectedness	Adversarial
UPU (0)	0	0	0	0	0
WTO1 (1)	0	0	0	1	1
UNEP (1), UNHRC (1)	0	0	1	0	1
UNFPA1 (1), UNESCO (1)	0	0	1	1	1
NATO1 (0)	0	1	0	1	0
IOM (0), UNHCR1 (0), WB1 (0), IMF1 (0), UNICEF (1)	0	1	1	0	0 (0.8)
WFP1 (0), FAO (0), UNSG1 (0), UNRWA1 (0), IAEA1 (1)	0	1	1	1	0 (0.8)
WTO2 (1)	1	0	0	1	1
WFP2 (1), UNFPA2 (1), UNRWA2 (1)	1	0	1	1	1
NATO2 (0), NATO3 (0), UNFCCC2 (0)	1	1	0	0	0
UNFCCC1 (0)	1	1	0	1	0
UNHCR2 (0), UNHCR3 (0), WB2 (0), IAEA2 (1)	1	1	1	0	0 (0.75)
UNSG2 (0), IMF2 (0), IAEA3 (1)	1	1	1	1	0 (0.67)
Logical remainder	1	0	1	0	?
Logical remainder	1	0	0	0	?
Logical remainder	0	1	0	0	?

Table S.10: Truth table for model including 'learning condition'.

When we allow for the inclusion of logical remainders in order to minimize the solutions, we obtain two alternative solutions for a conciliatory IPA response

$$(1) \text{ authority*affectedness} + \text{DEPENDENCE*authority} + \text{DEPENDENCE*learning} + \text{DEPENDENCE*affectedness} \rightarrow \text{adversarial}$$

$$(2) \text{ authority*affectedness} + \text{DEPENDENCE*authority} + \text{DEPENDENCE*learning} + \text{LEARNING*affectedness} \rightarrow \text{adversarial}$$

and one solution for an adversarial IPA response

$$\text{dependence*AFFECTEDNESS} + \text{dependence*AUTHORITY} \rightarrow \text{ADVERSARIAL}$$

The solution for an adversarial IPA response is identical with our original solution. In both of the two alternative solutions for a conciliatory IPA response, one implicant (authority*affectedness) is identical with the original solution. In the first alternative solution, the three remaining implicants are all subsets of the second implicant of our original solution (DEPENDENCE) (see Table S.11).

Solution	authority* affectedness	+ DEPENDENCE *	+ DEPENDENCE *	+ DEPENDENCE *	→ adversarial
	s	authority	learning	affectedness	
Consistency	1.00	1.00	0.81	0.83	
Raw coverage	0.21	0.26	0.47	0.53	
Unique coverage	0.05	0.05	0.21	0.16	

Table S.11: First solution for a conciliatory IPA response for model including 'learning condition'.

In the second alternative solution, two of the three remaining implicants are again subsets of DEPENDENCE, while one implicant is new (LEARNING*affectedness) (see Table S.12). However, this implicant does not cover any of the empirical cases uniquely, but these are all also covered by another implicant of the very same solution (authority*affectedness). Hence, the only implicant that is not a subset of our original solution is driven by the presence of logical remainders when including a fourth condition – and not by empirical observations. In addition, it substantially contradicts the theoretical expectation that experience with contestation by the Trump Administration should be associated with an adversarial, not conciliatory IPA response.

Solution	authority* affectedness	+ DEPENDENCE *	+ DEPENDENCE *	+ LEARNING* affectedness	→ adversarial
	s	authority	learning	affectedness	
Consistency	1.00	1.00	0.81	0.86	
Raw coverage	0.21	0.26	0.47	0.32	
Unique coverage	0.05	0.05	0.42	0.16	

Table S.12: Second solution for a conciliatory IPA response for model including 'learning condition'.

Both alternative solutions have a consistency of 0.85, which is slightly better than the consistency of the original solution (0.82), as IAEA3 is not a deviant case anymore while

IAEA1, 2, and UNICEF remain deviant cases. Yet, both alternative solutions show a decreased coverage of 0.89 compared to 1 in the original solution, as they do not cover UNSG2 and IMF2 anymore. Hence, the new, more complex solution does not considerably fare better than the original one.

To sum up, adding a ‘learning condition’ to our original conditions, we end up with the same solution for an adversarial response and two very similar, but more complex solutions for a conciliatory response. Both are – too large parts or complete – subsets of the original solution, while their coverage and consistency values do not improve significantly compared to the original one’s. This increases our confidence in the original, more parsimonious solution.

4.2 Adding a resilience condition

To check whether past experience with sanctions by the US shapes IPAs’ responses, we conducted a QCA based on our original conditions as well as a ‘resilience condition’. To measure the ‘resilience condition’, we coded for each case whether the US withdrew its funding or membership (‘1’) or not (‘0’) from the administration of George W. Bush took office in 2001 through the end of the Obama Administration.

The resulting truth table (see Table S.13) comprises 11 fully consistent rows, while two rows have a consistency of 0.75. As in both rows only a minority of 2 of 8 cases – IAEA1, 2, 3 and UNICEF – shows an adversarial outcome, we treat them as showing a conciliatory outcome. Three configurations lack any empirical observations and thus are logical remainders.

	Resilience	Dependence	Authority	Affectedness	Adversarial
WTO1, WTO2	0	0	0	1	1
UNHRC	0	0	1	0	1
WFP2, UNRWA2	0	0	1	1	1
NATO2, NATO3	0	1	0	0	0
NATO1	0	1	0	1	0
UNHCR1 (0), UNHCR2 (0), UNHCR3 (0), WB1 (0), WB2 (0), IMF1 (0), IAEA2 (1), UNICEF (1)	0	1	1	0	0 (0.75)
WFP1 (0), FAO (0), UNSG1 (0), UNSG2 (0), IMF2 (0), UNRWA1 (0), IAEA1 (1), IAEA3 (1)	0	1	1	1	0 (0.75)
UPU	1	0	0	0	0
UNEP	1	0	1	0	1
UNFPA1, UNFPA2, UNESCO	1	0	1	1	1
UNFCC2	1	1	0	0	0
UNFCC1	1	1	0	1	0
IOM	1	1	1	0	0
Logical remainder	1	1	1	1	?
Logical remainder	1	0	0	1	?
Logical remainder	0	0	0	0	?

Table S.13: Truth table for model including 'resilience condition'.

When we allow the inclusion of logical remainders in the logical minimization, we end up with two alternative solutions for a conciliatory IPA response

(3) $DEPENDENCE + authority * affectedness \rightarrow adversarial$

(4) $DEPENDENCE + authority * RESILIENCE \rightarrow adversarial$

and two alternative solutions for an adversarial IPA response

(1) $dependence * AUTHORITY + dependence * AFFECTEDNESS \rightarrow ADVERSARIAL$

(2) $dependence * AUTHORITY + dependence * resilience \rightarrow ADVERSARIAL$

The first of the two alternative solutions for each outcome is identical with the solution of our main model. The other two alternative solution are very similar, but the presence of the ‘resilience condition’ logically replaces the absence of the affectedness condition, and vice versa. Overall, our original solution finds support even when including a ‘resilience condition’. This increases our confidence in the original solution.

References of the Appendix of Paper B

- Adamson, Thomas, 2017, New UNESCO chief says US 'empty chair' can't last, reaches out to Israel. *The Times of Israel*.
- Agence Francaise-Presse, 2017, United Nations Environment Programme Chief Urges China To Do More On Climate. <https://www.ndtv.com/world-news/united-nations-environment-programme-chief-urges-china-to-do-more-on-climate-1735504>.
- Al Jazeera, 2017, US and Israel withdraw from UNESCO.
- Bauer, Michael W., Knill, Christoph and Eckhard, Steffen, eds., 2017, *International Bureaucracy: Challenges and Lessons for Public Administration Research* (London: Palgrave Macmillan UK).
- BBC, 2017, Trump worries Nato with 'obsolete' comment.
- Beaumont, Peter and Holmes, Oliver, 2018, US confirms end to funding for UN Palestinian refugees. *The Guardian*.
- Becker, Markus, 2019, Donald Trump Damages America By Defanging WTO. *Spiegel*.
- Biermann, F. and Siebenhüner, Bernd, eds., 2009, *Managers of global change: The influence of international environmental bureaucracies* (Cambridge MA: MIT Press).
- Blake, Daniel J. and Payton, Autumn L., 2015, Balancing design objectives: Analyzing new data on voting rules in intergovernmental organizations. *The Review of International Organizations*, **10**(3), pp. 377–402.
- Davis, Julie H., 2018, Trump Warns NATO Allies to Spend More on Defense, or Else. *The New York Times*.
- Deitelhoff, Nicole and Zimmermann, Lisbeth, 2018, Things We Lost in the Fire: How Different Types of Contestation Affect the Robustness of International Norms. *International Studies Review*, **54**(3), p. 421.
- Executive Order 13769, 2017, Protecting the Nation from Foreign Terrorist Entry into the United States: Signed by Donald Trump on January 27, 2017.
- Farrell, Stephen and al-Mughrabi, Nidal, 2018, U.N. WFP to cut back food aid to Palestinians. *Reuters*.
- Gallagher, Kevin P., 2017, China's influence on global finance grows as US scales back input. *Financial Times*.
- Gladstone, Rick, 2017, Trump's Cuts Would Make Job 'Impossible,' U.N. Says. *The New York Times*.

- Greuter, Nicole, 2014, *Accountability without Election. The Attribution of Responsibility in the Financial Crisis 2007-2010*. Schriftenreihen politischer Kommunikation und demokratischer Öffentlichkeit. Band 10 (Baden-Baden: Nomos).
- Gronau, Jennifer and Schmidtke, Henning, 2016, The quest for legitimacy in world politics – international institutions’ legitimation strategies. *Review of International Studies*, **42**(3), pp. 535–557.
- Hart, Marion, 2018, Why Detaining Children Is Harmful. *UNICEF USA*.
- Heath, Ryan, 2019, Breaking down Trump’s post-NATO hot spots. *Politico*.
- Heinkelmann-Wild, Tim and Zangl, Bernhard, 2019, Multilevel blame games: Blame-shifting in the European Union. *Governance*.
- Hood, Christopher, Jennings, Will and Copeland, Paul, 2016, Blame Avoidance in Comparative Perspective: Reactivity, Staged Retreat and Efficacy. *Public Administration*, **94**(2), pp. 542–562.
- Hooghe, Liesbet and Marks, Gary, 2015, Delegation and pooling in international organizations. *The Review of International Organizations*, **10**(3), pp. 305–328.
- IAEA, 2019, *Remarks by Director General Yukiya Amano at New Year Reception* .
- IMF, 2017, *Transcript of an Interview with IMF Managing Director Christine Lagarde with the Leading European Newspapers Association* .
- IMF, 2018, *Transcript of IMF Press Briefing* .
- IOM and UNHCR, 2017, *Joint IOM - UNHCR Statement on President Trump's Refugee Order* .
- Iran Watch, 2018, *Statement by IAEA Director General Yukiya Amano Following President Trump's Announcement on Iran Deal* .
- Italian Insider, 2017, Washington to cut FAO funding by dlrs 30 million.
- Kuckartz, Udo, 2016, *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*, 3. überarbeitete Auflage (Weinheim, Basel: Beltz Juventa (Grundlagentexte Methoden)).
- Landler, Mark, 2018, Trump Abandons Iran Nuclear Deal He Long Scorned. *The New York Times*.
- Lawder, David, 2017, World Bank's Kim says Trump likes lender's private capital plans. *Reuters*.
- Lyne, Mona M., Nielson, Daniel L. and Tierney, Michael J., 2006, Who delegates? Alternative models of principals in development aid, in: Darren G. Hawkins, David A. Lake, Daniel L. Nielson and Michael J. Tierney (Eds) *Delegation and Agency in International Organizations* (Cambridge: Cambridge University Press), pp. 41–76.

- Manson, Katrina, 2017, UN nuclear agency counters Trump objections to Iran deal. *Financial Times*.
- Mathiesen, Karl, 2017, Trump budget: US to stop funding UN climate process. *Climate Home News*.
- Mayedam Andrew and Mohsin, Saleha, 2018, Trump Snubs Global Order Again as U.S. Rejects IMF Funding Boost. *Bloomberg*.
- McArthur, John W. and Rasmussen, Krista, 2017, Who funds which multilateral organizations?, Global Views No. 8.
- McCaskill, Nolan D. and Lima, Cristiano, 2017, Trump reverses on NATO: 'It's no longer obsolete'. *Politico*.
- Micklethwait, John, Talev, Margaret and Jacobs, Jennifer, 2018, Trump Threatens to Pull U.S. Out of WTO If It Doesn't 'Shape Up'. *Bloomberg*.
- Miroff, Nick, 2018, Trump administration tightens immigration asylum rules as caravans continue to push for U.S. border. *Washington Post*.
- NATO, 2017, International Staff. https://www.nato.int/cps/en/natolive/topics_58110.htm.
- NATO, 2018a, Press conference by NATO Secretary General Jens Stoltenberg at the conclusion of the Brussels Summit. https://www.nato.int/cps/en/natohq/opinions_156738.htm.
- NATO, 2018b, Who's who? https://www.nato.int/cps/en/natohq/who_is_who.htm.
- NATO, 2019, *Press point by NATO Secretary General Jens Stoltenberg and US President Donald Trump* .
- Nichols, Michelle, 2017, Trump budget cut bid would make it 'impossible' for U.N. - spokesman. *Reuters*.
- Rauh, Christian, Bes, Bart J. and Schoonvelde, Martijn, 2020, Undermining, defusing or defending European integration? Assessing public communication of European executives in times of EU politicisation. *European Journal of Political Research*, **59**(1), 397–423.
- Reuters, 2018, U.N. refugee agency gives Trump's latest cuts a weary shrug.
- Reuters, 2019a, U.N. nuclear watchdog warns against pressuring it on inspections.
- Reuters, 2019b, Trump calls for World Bank to stop loaning to China.
- Rihoux, Benoît and Meur, Gisèle de, 2009, Crisp-Set Qualitative Comparative Analysis (csQCA), in: Benoît Rihoux and Charles C. Ragin (Eds) *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques* (Los Angeles, London: Sage), pp. 33–68.
- Sanger, David E. and Gladstone, Rick, 2017, Contradicting Trump, U.N. Monitor Says Iran Complies With Nuclear Deal. *The New York Times*.

- Schmidtke, Henning, 2018, Elite legitimation and delegitimation of international organizations in the media: Patterns and explanations. *The Review of International Organizations*, **42**(1), p. 3.
- Schneider, Carsten Q. and Wagemann, Claudius, 2013, *Set-theoretic methods for the social sciences: A guide to qualitative comparative analysis*, Repr (Cambridge, New York: Cambridge University Press).
- Skaaning, Svend-Erik, 2011, Assessing the Robustness of Crisp-set and Fuzzy-set QCA Results. *Sociological Methods & Research*, **40**(2), pp. 391–408.
- Staff, Toi, 2018, UNRWA laments US decision to cut aid, pushes back against criticism. <https://www.timesofisrael.com/unrwa-laments-us-decision-to-cut-aid-pushes-back-on-criticism/>.
- Stewart, Patrick M., 2017, Trump, the World Bank, and the IMF: Explaining the Dog that Didn't Bark (Yet). *Council on Foreign Relations*.
- Swanson, Ana, 2019, Trump Cripples W.T.O. as Trade War Rages. *The New York Times*.
- Tallberg, Jonas and Zürn, Michael, 2019, The legitimacy and legitimation of international organizations: Introduction and framework. *The Review of International Organizations*.
- The Guardian, 2018, Trump pulls US out of UN postal scheme on China price concerns.
- The Guardian, 2019, Donald Trump calls for World Bank to halt all China lending.
- Thomann, Eva and Maggetti, Martino, 2017, Designing Research With Qualitative Comparative Analysis (QCA). *Sociological Methods & Research*, **66**, 004912411772970.
- Toolsi, Nahal, 2018, UN rights chief blasts Trump's 'unconscionable' child separation policy. *Politico*.
- UN Chief Executives Board for Coordination, 2020, Financial Statistics. <https://www.unsystem.org/content/un-system-financial-statistics>.
- UN Secretary-General, 2018, *Readout of the Secretary-General's meeting with H.E. Mr. Donald J. Trump, President of the United States of America* .
- UNEP, 2019, About UN Environment. <https://www.unenvironment.org/about-un-environment>.
- UNFCCC, 2017a, *Developments in the United States, By Executive Secretary Patricia Espinosa* .
- UNFCCC, 2017b, *UNFCCC Statement on the US Decision to Withdraw from Paris Agreement* .
- UNFPA, 2017, Statement by UNFPA on U.S. Decision to Withhold Funding. <https://www.unfpa.org/press/statement-unfpa-us-decision-withhold-funding>.
- UNFPA, 2018, *Statement on the United States Decision to Again Withhold Funding from UNFPA* .

- UNHCR, 2018, UNHCR statement on new US regulation on asylum. <https://www.unhcr.org/news/press/2018/11/5be5bfac4/unhcr-statement-new-regulation-asylum.html>.
- UNRWA, 2018a, *Statement by UNRWA Commissioner-General Pierre Krähenbühl* .
- UNRWA, 2018b, *Urgent UNRWA Statement*.
- UPU, 2018, *Statement on the decision of the United States of America to withdraw from the Universal Postal Union treaties* .
- WFP, 2017, Massive cuts to food aid would do long-term harm to US national security interests. *World Food Programme Insight*.
- White House, 2017a, America First: A Budget Blueprint to Make America Great Again. Office of Management and Budget. https://www.whitehouse.gov/wp-content/uploads/2017/11/2018_blueprint.pdf.
- White House, 2017b, *Statement by President Trump on the Paris Climate Accord* .
- White House, 2018, Remarks by President Trump to the 73rd Session of the United Nations General Assembly. <https://www.whitehouse.gov/briefings-statements/remarks-president-trump-73rd-session-united-nations-general-assembly-new-york-ny/>.
- WTO, 2018, *DDG Wolff: “There is reason for optimism about the future of the multilateral trading system”* .
- Yamasaki, Sakura and Rihoux, Benoit, 2009, A Commented Review of Applications, in: Benoît Rihoux and Charles C. Ragin (Eds) *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques* (Los Angeles, London: Sage), pp. 123–145.

Paper C

Explaining the Political Use of Evaluation in International Organizations

This is the published version of the article

In: *Policy Sciences* (2020) 53: 667–695, co-authored with Steffen Eckhard

Available at <https://doi.org/10.1007/s11077-020-09402-2> (last accessed 12.01.2021)

Abstract:

Despite a growing literature on the politics of evaluation in international organizations (IOs) and beyond, little is known about whether political or administrative stakeholders indeed realize ex-ante political interests through evaluations. This is, however, especially important considering the booming business of evaluation and the proliferation of institutional assessments both in domestic and international politics. We argue that formally independent IO evaluation units informally orientate towards either member states or the IO administration, depending on who controls the unit's budget, staff, and agenda resources. This should enable either actor to also use evaluation results along pre-defined strategic interests. Interview data gathered among evaluators, secretariat officials, and member state representatives of six IOs support the expected pattern, highlighting striking differences in the orientation of evaluation staff and evaluation use. Findings challenge the technocratic, apolitical image of evaluation, offering practical and theoretical implications for future research.

Keywords: International organization, International public administration, Evaluation, United Nations, Evidence-based policy-making

Introduction

In public policy-making, evaluation is typically perceived as a functional tool in the final phase of a cyclical process (Anderson 1975; Brewer 1974; Howlett et al. 2009; Lasswell 1956). Evaluation is thought to inform policy-makers about past and present experiences, allowing for policy adjustment, learning, and accountability. This essentially follows a functionalist logic, where policy-makers seek effectiveness and efficiency, evaluation being one of the measures they take. Yet, in the political reality of public organizations, evaluation finds itself caught between stakeholder struggles for interests and power (Azzam 2010; Bjornholt and Larsen 2014; Morris and Clark 2013; Pleger et al. 2017; Taylor and Balloch 2005; Wildavsky 1972; Weiss 1998). For instance, evaluation results may be helpful in justifying actors' bargaining positions on policy decisions: "Whenever an evaluation affects the future allocation of resources and, hence, a change in power relationships, it is a political activity" (Wergin 1976, p. 76).

Many authors lament these "politics of evaluation" (e.g. Banner 1974; Taylor and Balloch 2005; Weiss and Jordan 1976), suggesting that evaluation results can become "ammunition in political battles" (Schoenefeld and Jordan 2019, p. 377) or that stakeholders put pressure on evaluators "to misrepresent findings" (Pleger et al. 2017, p. 316). Van Voorst and Mastenbroek (2019) show that the quality of policy evaluation systematically decreases when additional political stakeholders get involved. Despite such insights, there is very little comparative evidence detailing whether policy-makers indeed realize ex-ante political interests through evaluation. Taking it from there, this paper *asks whether evaluation systematically serves ex-ante political interests of policy actors that go beyond the traditional duet of learning and accountability*. Such a political use of evaluation is defined as situations in which actors refer to evaluations (their findings, processes, or recommendations) to realize own political interests in competition with others.

The empirical focus is on international organizations (IOs) in the United Nations (UN) system, which offers an important analytical advantage. Comparative studies of evaluation use in domestic settings are challenging because of idiosyncrasies of individual agencies or political systems. IOs, by contrast, offer a class of comparable cases. UN system IOs all abide by the

same evaluation norms and guidelines of the UN Evaluation Group (UNEG) which rules out a range of confounding factors.¹

To identify political interests that could be realized using evaluation, we apply the classical lenses of principal-agent theory. On the one hand, member states are the principals in IOs, who take all policy decisions (Rittberger et al. 2019). Their political interests relate to internal dynamics among the state collective and their relationship with the IO administration. Evaluation can help states in *intergovernmental bargaining* and in exercising *agency control*. On the other hand, a growing literature emphasizes that IO secretariats (International Public Administrations—IPAs) act as independent actors in international policy processes (Bauer et al. 2017; Biermann and Siebenhüner 2009; Eckhard and Ege 2016; Patz and Goetz 2019; Knill et al. 2019; Johnson and Urpelainen 2014).² Evaluation can be politically useful for IPAs, to *justify past and future action*, or to *steer the organization internally*. For instance, studying the IMF, Hinterleitner, Sager and Thomann (2016, p. 564) found that “evaluations are a relatively obvious way for the IMF to effectively exert indirect influence on member states via its surveillance activities”.

Existing research highlights that powerful stakeholders can use the institutional context surrounding evaluations to exercise their influence (Azzam 2010; Højlund 2014; Raimondo 2018). In IOs, this could be done through the control over IO evaluation systems resources, i.e. evaluation units’ staff, budget, and agenda (see Eckhard and Jankauskas 2019). The main expectation therefore is *that the political use of evaluation should differ systematically between IOs whose evaluation resources are either controlled by the IPA or by IO member states*. Such political use should become visible at two stages: We hypothesize that (1) independent IO evaluation units—as key actors managing centralized evaluations in IOs—should orient to the dominant stakeholder; and (2) the use of evaluation in IO policy-making should systematically reflect the typical political interests of the dominant stakeholder.

Empirically, we realize a small-N comparison. The sample includes nine IOs (IAEA, IOM, ILO, UNICEF, UNDP, UNEP, UNHCR, UNESCO, and WHO), six of which are analysed.³

¹ UNEG defines evaluation as “an assessment, conducted as systematically and impartially as possible, of an activity, project, programme, strategy, policy, topic, theme, sector, operational area or institutional performance” (UNEG 2016).

² IPAs are “bodies with a certain degree of autonomy, staffed by professional and appointed civil servants who are responsible for specific tasks and who work together following the rules and norms of the IO in question” (Bauer et al. 2017, p. 2). For member states, it is primarily their delegated representatives who constitute IO governing bodies.

³ We deliberately exclude three of the IOs from further analysis to ensure interviewees’ anonymity (see section on research design for a detailed explanation). The acronyms refer to the International Atomic Energy Agency (IAEA), International Organization for Migration (IOM), International Labour Organization (ILO), United Nations

They fall into three groups, distinguished according to who controls evaluation system resources (the IPA, member states, or both). We draw on 35 qualitative expert interviews with senior officials in the IO administration, evaluation units, and member state representations. Data demonstrate that differences in the *control over evaluation resources* (iV) link to differences in the *orientation of IO evaluation units* (dV1) to either IPAs or member states and also to differences in the *alignment of evaluation use* (dV2) towards the political interests of either IPAs or member states.

The study is the first to present comparative empirical data on political evaluation use in IOs. It shows that the same evaluation tool, governed by the same UN system-wide evaluation standards, plays a very different role in IO policy-making, depending on who controls evaluation units. While these findings do not claim that evaluation reports as such are biased, we demonstrate that their political use differs systematically. These findings concur with others who argue that the idea of evidence-based policy-making should not be taken for granted (Cooley and Snyder 2015; Hinterleitner et al. 2016; Højlund 2015; Merry 2011; Porter 1995; van Voorst and Mastenbroek 2019). We do not suggest to refrain from using evaluation in public management, but the revealed politics surrounding seemingly functional tools like evaluation should be acknowledged and factored into our understanding of policy-making processes (for a related discussion, see Fforde 2019; Perl et al. 2018). For now, such a political understanding of evaluation appears uncommon among IO practitioners.

In the following sections, we first review the literature on IOs to derive expectations about evaluation stakeholder interests and the associated motives for political evaluation use. We then discuss how these dynamics can be observed empirically, introduce our research design and methods, present the empirical data, and discuss our findings as well as theoretical and practical implications.

Theorizing political interests and the political use of evaluation in IOs

Evaluation researchers have conceptualized the use of evaluation in three ways: “[e]valuations could be used (a) instrumentally, to give direction to policy and practice; (b) politically or symbolically, to justify pre-existing preferences and actions, and (c) conceptually, to provide

Children’s Fund (UNICEF), United Nations Development Programme (UNDP), United Nations Environment Programme (UNEP), United Nations High Commissioner for Refugees (UNHCR), United Nations Educational, Scientific and Cultural Organization (UNESCO), and World Health Organization (WHO).

new generalizations, ideas, or concepts” (Weiss et al. 2005, p. 13). Whereas (a) and (c) follow a functional rationale, the political (or symbolic) evaluation use forms the focus of this article.

In this regard, previous studies indicate that evaluations in IOs are politically contested (Højlund 2015, p. 35; Raimondo 2018, p. 32; van Voorst and Mastenbroek 2019; Weaver 2010, p. 366). If evaluation is indeed used politically, i.e. “in the battle for influence and interest promotion” (Bjornholt and Larsen 2014, p. 407), the first step of the analysis is to systematically unpack the pre-existing political interests of key evaluation stakeholders in IOs and to assess how evaluation may serve them.

Due to their direct embeddedness in decision-making processes, the main evaluation stakeholders in IOs are member states on the one hand and IO administrations on the other.⁴ We apply the classical lenses of principal-agent (P-A) theory to distinguish between member states as the principal and IO bureaucracies (IPAs) as the agent. This way and by drawing on a wide range of literature, we are able to outline general P-A-dynamics referring to dynamics *within* the two actors (among member states and within the IPA) as well as *between* the actors (member states interests vis-à-vis the IPA and vice versa). These P-A-dynamics, in turn, generate political interests which exist independently from issue area or IO characteristics. The following sections describe the different interest dynamics from each of the quadrants (Table 1) and how evaluation can help actors utilizing them.

P-A dynamic	Principal (MS)	Agent (IPA)
within actor	Collective principal dynamics	Collective agent dynamics
between actors	P-A control dynamics (bureaucratic drift)	Bureaucratic influence and justification dynamics

Table 1: Overview of theorized principal-agent dynamics in IOs.

First, IO member states operate as a *collective principal* with heterogeneous political interests (Hawkins et al. 2006). Although states formally resolve their disagreements during voting in the policy-making forum of an IO, they are still constantly fighting over IO policies (Koremenos et al. 2001). In this regard, evaluation might be a strategic tool in states’ negotiations. Its results can be used to strengthen their own argumentation or shift blame to

⁴ This is analogous to key evaluation stakeholders at the domestic level: political decision-makers and bureaucratic servants (Bjornholt and Larsen 2014, p. 408; Leeuw and Furubo 2008; The LSE GV314 Group 2014).

others (Chelimsky 1987). Given that evidence is power (see Botterill and Hindmoor 2012), convincing other state counterparts should become easier when referring to arguably objective findings rather than ideological standpoints. Furthermore, individual member states were observed for seeking influence by manipulating the bureaucratic agent—the IPA (Urpelainen 2012; Dijkstra 2015). For instance, powerful member states would push their unilateral agenda by rewarding or punishing influential staff within the secretariat to serve their interests in policy-making or implementation (Voeten 2008; Streck 2001). Against this backdrop, other states could use evaluation to investigate or prevent such practices. Thus, evaluation could be used as a safeguard to counter unilateral influence within the member states collective.

Second, IPAs have also been recognized as *collective agents*. Graham (2014) argued that IO bureaucracies should be treated as plural actors with structurally fragmented organizational units. Hanrieder (2015) also outlined the concept of fragmentation in the context of bureaucratic complexity, arguing that IO subunits, such as regional or country offices, often have their own interests and power. IO complexity means that “[t]he greater the extent of these subunit authorities, the more fragmented and the less hierarchical is an IO” (Hanrieder 2015, p. 34). It might thus become a fundamental challenge for an IO’s management to keep the competing organizational parts together. Evaluation can be used in this sense as a tool to vertically steer the organization. It provides information about the performance of organizational subunits, allowing senior management to overcome informational asymmetry within the fragmented structure. Just as member states may use evaluation to avoid agency slack from the bureaucracy (see below), the administration can also use evaluation to strengthen itself by preventing drift within its own ranks.

Third, turning to the behavioural dynamics *between* the actors, the extensive principal-agent literature demonstrates that the asymmetric structure of delegation relationships allows the possibility of agency losses, so that *member states have to employ control over IPAs* (Hawkins et al. 2006; Kassim and Menon 2003; Nielson and Tierney 2003). While the principal seeks to control its agent to avoid unwanted behaviour, the latter may have their own preferences and seek to escape the former’s oversight. da Conceição-Heldt (2013, p. 24) described it as “agents’ ability to act independently of their principals and to overreach their delegated authority”. Evaluation may hence serve as an ex-post control instrument for member states to contain bureaucratic influence. Ideally, evaluation reports help to reduce the informational gap by red-flagging unwanted IPA behaviour.

Fourth, evaluation may also serve the exact opposite purpose—for the *IPA* to exercise *bureaucratic influence* if it wants to do less or differently than asked by the member states (Elsig 2011; Vaubel et al. 2007). A number of studies show that IPAs are highly innovative and entrepreneurial in employing various administrative tools and tactics to increase own autonomy and influence policy-making (Patz and Goetz 2019; Knill et al. 2019; Biermann and Siebenhüner 2009; Bauer et al. 2017; Johnson and Urpelainen 2014). Consequently, evaluation might be a tool for IPAs to justify such initiatives. For instance, IPAs may refer to evaluations *vis-à-vis* the member states for resource mobilization strategies, favourable agenda-setting, mandate expansion, or to justify past actions (Easterly 2006; O’Brien et al. 2010; Weiss 1998). This relates to Goffman’s “impression management” (1959), where actors “present different aspects of themselves to suit the particular audiences they are faced with”. For instance, Hayward and colleagues (The LSE GV314 Group 2014, p. 224) demonstrated that UK civil servants put significant efforts to shape evaluation results “making them ‘look good’ or minimizing criticism of their policies”.

Considering the above, evaluation in IOs can be used in many more ways than only for the functional purpose of accountability and learning. Of course, the functional and political imperatives are not necessarily inversely related. We might observe both functional and strategic/political evaluation use in the same IO, even in the same evaluation. However, the aim of this paper is to identify the emergence and existence of the latter. Based on the theoretical identification of different motives for political evaluation use (see Table 2), the following section turns to the explanation of under what conditions which kind of political use should prevail.

Principal-agent dynamic	Actor political interest	Political evaluation use
<u>Member States</u>		
<i>Collective principal</i>	Justifying own interests and preventing unilateral influence	Evaluation used in negotiations among states and to counter unilateral influence
<i>P-A control</i>	Controlling the agent	Evaluation used to contain IPA influence/ agency drift
<u>International Public Administration</u>		
<i>Collective agent</i>	Containing organizational fragmentation	Evaluation used to vertically steer IPA and contain fragmentation
<i>Bureaucratic influence</i>	Justifying past or future action	Evaluation used to justify IPA initiatives and past action

Table 2: Conceptualization of political evaluation use for each actor based on principal-agent dynamics.

Explaining political evaluation use in IOs

For the purpose of this research, we focus on IOs' centralized evaluation function, which covers evaluations produced (or managed) by IO evaluation units. This scope condition excludes decentralized evaluations which are mostly routinized project-level studies, usually conducted by responsible IPA departments. Centralized evaluations are mostly highlevel ex-post assessments that "generally support overall corporate-level policy and strategic decision-making" (JIU 2014, iii). They are discussed both by the management and IO governing bodies and their recommendations are tracked and followed up over time. In 2018, UNDP, for instance, undertook 17 thematic and country programme evaluations (IEO 2018), while ILO conducted 54 centralized evaluations (ILO 2019). The cost of a single report can reach up to one million USD. As an example, the average cost of a thematic evaluation report in FAO is about 500,000 USD (FAO 2016, p. 28). As for the staff working in evaluation units, the numbers can vary from single figures as in the IOM to several dozen as in FAO.

The IO evaluation units are thus at the centre stage of evaluation systems in IOs. According to the UN system-wide evaluation standards, evaluation units are designed as independent actors within the IPA: "Organizational independence requires that the central evaluation function is positioned independently from management functions" (UNEG 2016, p. 11). Whereas centralized evaluation units are *formally* independent, they remain a legal entity of their IO and the IPA.

However, studies at the domestic level suggest that stakeholders can translate their influence *informally* through the institutional context in which evaluation activities are embedded (Azzam 2010; Højlund 2014; Bjornholt and Larsen 2014). We hold that this is also true at the international level. The institutional context is the evaluation system, i.e. the organizational structure and rules that define who controls key resources and procedures of IO evaluation unit, including its budget, staff appointments, reporting lines, overall agenda, etc. (cf. Leeuw and Furubo 2008).

We therefore define *control over evaluation systems* as the independent variable (see Fig. 1). It can be exerted most directly through the evaluation staff, budget, and by means of influencing the evaluation agenda (see Rossi et al. 2004, p. 46; Stockmann et al. 2011). The core expectation is that depending on who controls the resources of evaluation system, we can observe the politics of evaluation at two subsequent stages: through the *orientation of the evaluation unit*

(dependent variable 1); and at the level of *political decision-making* (dependent variable 2). We therefore specify two dependent variables as well as hypotheses linking these variables.

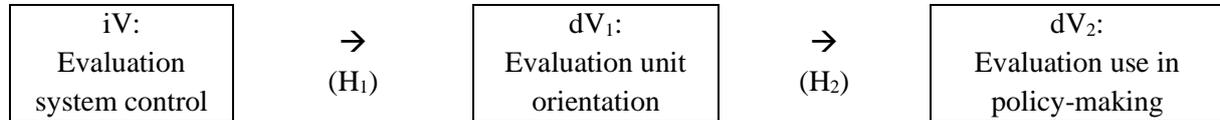


Figure 1: The theorized causal mechanism.

First, despite their formal independence, evaluation units operate as agents in an institutional setting of competing principals (Lyne et al. 2006, p. 44; Schoenefeld and Jordan 2017, p. 277), namely the two stakeholders: the IPA and member states. Given that such formal factors like funding or agenda-setting are valuable resources principals can use for sanctioning or rewarding evaluation units, there should be differences in the *evaluation unit orientation* (dependent variable 1) (Leeuw and Furubo 2008, p. 166; see also O’Brien et al. 2010, p. 432). By evaluation unit orientation, we mean the conscious or unconscious perception by evaluation unit staff of who is the primary sponsor and user of evaluation. This does not necessarily mean that evaluators neglect professional standards and provide biased reports. Yet, according to Weiss (1998, p. 31), it should make a difference for evaluation methodology and conduct, whether the political purpose of evaluation is to generate new ideas to inform IPA policy positions or to ensure, from a member state perspective, that IPAs do not deviate from their mandate. There may also be procedural implications, such as whose comments are primarily (consciously or not) taken into account when drafting terms of references, the report, or recommendations. Overall, the expectation is that the different structural domination by either member states or the IPA determines whose political interest evaluation units will primarily orientate to. On this basis, we derive the following hypothesis on the relationship between evaluation system control and evaluation unit orientation:

H₁: If the evaluation system resources are controlled by member states (or IPA) in an IO, then the evaluation unit orientates primarily towards member states (or IPA) as the sponsor and user of evaluation.

The hypothesis is falsified if we observe no systematic link between the two variables. There might also be cases with mixed stakeholder dominance, for instance, when the IPA allocates

the budget, but member states approve the head of evaluation unit and evaluation agenda. In these cases, we expect evaluation units to balance between the two stakeholders without clearly orientating to either of them.

Second, depending on which stakeholder evaluation units orientate to, i.e. member states or the IPA, they should produce evaluations that tend to serve to and be used for the respective actor's political interests. The reason is, as outlined by principal-agent theory, that "the agent [here evaluation unit] scans the range of principal demands and identifies a point that maximizes the compensation offered by the various principals. Principals with more power and resources thus have a greater impact on agent behavior" (Lyne et al. 2006, p. 58). Previous research on evaluation found that "the more political power or influence stakeholder groups held over evaluation logistical factors (i.e. funding, data access), the more evaluators were willing to modify their design choices to accommodate perceived stakeholder concerns" (Azzam 2010, p. 45; see also Stockmann et al. 2011). Evaluations that serve specific political interests should also tend to be used accordingly, what we term the *evaluation use alignment* with stakeholders' interests (dependent variable 2). In line with Table 2 and depending on the value of the dependent variable 1 (and the independent variable), three patterns of political evaluation use are possible: first, in IOs with evaluation units orientated to member states, we expect evaluation to be primarily used in negotiations among member states to back certain political claims and/or evaluations used to counter unilateral state interests (collective principal dynamics). Also, we expect evaluations in such cases to be used for containing IPA's influence and avoiding bureaucratic drift (agent control). Second, in evaluation systems with units' orientation towards the IPA, we expect evaluation to be used to sanction or steer the behaviour of decentralized IPA units (collective agent) and/or evaluation use for backing or justifying the IPA's own initiatives (bureaucratic influence). The second hypothesis below summarizes our expectations (see also Table 2):

H₂: If evaluation units are orientated to member states (or IPA) as their primary sponsor and evaluation user, evaluation use aligns with typical member state (or IPA) political interests.

The hypothesis is falsified if we observe no systematic link between the evaluation unit orientation and the political use of evaluation results. Again, in cases of mixed evaluation unit orientation, we expect competition between stakeholders and thus mixed political evaluation use aligned to the interests of both member states and the IPA.

Research design

The cause–effect relationship outlined by the two hypotheses implies a causal mechanism, which scholars often test through in-depth case studies along a temporal dimension (Beach and Pedersen 2016; Mayntz 2004). Yet, contrary to a processual approach, which restricts the number of cases that can be studied and thus impairs the external validity of the findings, we apply a comparative most similar systems design (MSSD) according to Mill’s method of difference (Lijphart 1971). Medium-N controlled comparison allows testing theorized claims for a broader set of cases and provides for generalization beyond the selected cases (George and Bennett 2005).

Case selection proceeded in two steps. First, we selected nine IOs which are similar on alternative explanatory dimensions but vary in the independent variable (control of evaluation system resources). IOs in the sample are: the IAEA, IOM, ILO, UNICEF, UNDP, UNEP, UNHCR, UNESCO, and WHO. They all belong to the UN system and as such have a similar membership with governing bodies consisting of member states representatives and a relatively similar IPA structure oriented towards common UN staff policies. As they are all UN organizations, their evaluation activities are defined around the same guidelines, norms, and standards of the UN Evaluation Group. They all have institutionalized evaluation units which are responsible for the centralized evaluation function according to their evaluation policies.⁵ They all conduct evaluations of the IPA’s activities and present annual reports both to the senior management and member states. Finally, evaluation units operate independently from management in formal terms.

Despite these similarities, the nine IOs differ in who formally controls evaluation system resources. Accordingly, IOs were grouped into three clusters. Cluster 1 contains ILO, UNDP, UNICEF whose evaluation system resources are predominantly controlled by member states; cluster 2 includes IAEA, IOM, and UNHCR, where the head of the IPA— and not the member states—decides upon the evaluation system resources; and cluster 3 consists of UNEP, UNESCO, and WHO. In these IOs, none of the resources is approved unilaterally by member states or the IPA.

The operationalization of evaluation system control (independent variable) was based on the following procedure. We focused on evaluation staff, budget, and agenda as these are the key

⁵ These units may conduct evaluations completely by themselves or recruit external consultants. Regardless of which business model is chosen, final reports are under the authority of IO evaluation unit, which makes the distinction between in/house and external consultants less relevant.

evaluation resources identified in the literature (Azzam 2010; Stockmann et al. 2011; Rossi et al. 2004). Drawing on official evaluation policies of all IOs (see “Appendix 1”), we used a scoring system by allocating 1, 0.5, or 0 points to either member states or the IPA. 1.0 point was allocated to the IPA or member states if they unilaterally approved evaluation budget, agenda, or appointment of the evaluation unit head. For instance, according to the UNHCR Policy on Evaluation, the High Commissioner (head of IPA) approves “the annual Work Plan for centralised evaluations and the Evaluation Service Budget” (UNHCR 2016, p. 12). By contrast, in UNDP, the Board “approves the biennial financial appropriation to the Independent Evaluation Office” as well as “the programme of work of the Office” (UNDP 2016, p. 6). 0.5 point was allocated to both stakeholders if they shared the authority to approve these resources. For example, according to WHO Evaluation Policy, the head of evaluation unit is appointed by the Director-General, yet “after consultation with the Executive Board” (WHO 2018). Table 3 summarizes the coding rules.

Resource	Coding rule		
<i>Stakeholder</i>	<i>MS</i>	<i>IPA/MS (both)</i>	<i>IPA</i>
<i>Code</i>	<i>1</i>	<i>0.5</i>	<i>1</i>
Evaluation budget	Member states (IO governing body) approve the evaluation budget	The head of IO administration approves evaluation budget, but after formal consultations with member states	The head of IO administration approves evaluation budget
Evaluation agenda-setting	Member states approve the rolling work plan of evaluation unit	The head of evaluation unit approves the rolling work plan of evaluation unit, but after formal consultations with member states	The head of IO administration or the evaluation unit approves the rolling work plan of evaluation unit
Evaluation staff appointment	Member states (IO governing body) appoint the head of evaluation unit	The head of IO administration appoints the head of evaluation unit, but after formal consultations with member states	The head of IO administration appoints the head of evaluation unit

Table 3: Operationalization of evaluation system control and coding rules.

If one of the stakeholders scored 2.5 or 3 points, the IO was attributed to either cluster 1 (MS control cases) or cluster 2 (IPA control cases) accordingly. If both stakeholders scored more than 0.5 points, the IO was attributed to mixed control cases of cluster 3. See “Appendix 1” for specific references from IOs’ evaluation policies on each resource category and resulting mapping of evaluation system resources.

In the second step of case selection, we randomly selected two out of three IOs per cluster and gathered empirical data on these six IOs. The reason being, we sought to ensure the anonymity of interviewees: the number of evaluation staff per IO is usually low and the topic of this study is sensitive to respondents' professional integrity. Previous studies on similarly contentious topics also anonymized IOs under scrutiny (e.g. Mele et al. 2016). Hence, from this point on, we no longer reveal the identity of each IO but refer only to the clusters (see Table 4).

IO sample	Anonymized label	Control of evaluation resources (iV)	Expected evaluation unit orientation (dV1)	Expected evaluation use alignment (dV2)
ILO, UNDP, UNICEF	Cluster 1 (IO 1, IO 2)	MS	MS	MS
IAEA, IOM, UNHCR	Cluster 2 (IO 3, IO 4)	IPA	IPA	IPA
UNEP, WHO, UNESCO	Cluster 3 (IO 5, IO 6)	MS/IPA	MS/IPA	MS/IPA

Table 4: Expected values to test hypotheses 1 and 2 across 6 IOs clustered in three groups.

The selected six UN organizations are similar on alternative factors that might confound evaluation processes as outlined above. In line with Mill's method of difference, the observation of the variable values as indicated in Table 3 should allow conclusions on the hypotheses. A discussion on alternative explanations is still provided below.

To measure *evaluation unit orientation* (dV1) and *evaluation use alignment* (dV2), we draw on original expert interview data collected for this purpose. We conducted 35 semistructured interviews with officials from three target groups in 2018 and 2019: heads of IO evaluation units, member state representatives (mostly ambassadors) from those governing bodies to which evaluation units report (e.g. executive board, programme committee, executive committee), and senior management officials from respective IOs (e.g. Chef de Cabinet or programme directors). We made sure that all stakeholder groups (evaluators, member states, and IPA) are equally covered (see interview list in "Appendix 2").

To avoid biasing interviewee responses, the questionnaire contained only general questions about the evaluation function in an IO (its system, process, use; see "Appendix 3"). We did not directly ask about orientation and political use of evaluation results. While evaluators or management officials might be aware of the political nature of evaluation in their organization, it is likely that not all would have accurately responded to such questions. Instead, we measure

the key variables indirectly, by assessing what meaning interviewees bring to evaluation and its conduct. This is in line with how Denzin and Lincoln (2000, p. 3) describe the approach: “qualitative researchers study things in their natural settings, attempting to make sense of, or to interpret, phenomena in terms of the meanings people bring to them”. All interviews, each lasting about 60 to 90 min, were recorded and transcribed.

To measure the two dependent variables, we systematically coded interview transcripts using MAXQDA 2018 software. *Evaluation unit orientation* (dV1) was operationalized using two codes. First, all text segments in which evaluation unit respondents referred to member states or IPA as their sponsors (i.e. those who help to fulfil their mandate) or identified themselves with these actors as being part of them were given a code “Evaluation Unit Sponsor”. Depending on whom they referred to, directional sub-codes were attached (either MS or IPA). Second, the same procedure was conducted for text segments where evaluation unit respondents referred to either member states or the IPA as the main users of evaluation results (“Evaluation Unit User”).

Analogously, we followed the same approach to measure *evaluation use alignment* (dV2). Text segments from all interviews where respondents made statements about evaluation use were classified according to the indicators from Table 2: directional codes indicated evaluation use alignment either towards member states’ or an IPA’s political interests (e.g. “MS use/collective principal” or “IPA use/bureaucratic influence”).⁶

In total, 236 text segments (i.e. interviewees’ statements on evaluation unit orientation or political evaluation use) were coded. The interviews were coded by two researchers, focusing on text segments identified as relevant by the first author, and then by resolving any disagreements (see “Appendix 4” for the overview of coded statements per IO and stakeholder group).

Note that our results do not account for variation over time and thus refer to the situation as of 2018/19. Furthermore, we were only able to interview a limited number of stakeholders, which requires us to generalize on the interests of member states or the administration based on statements made by a small number of their representatives. This is, however, a general issue in qualitative studies. Nevertheless, we sought to enhance the representativeness of our interviewees by reaching out primarily to longstanding member state representatives and high-level IPA officials who have a broader view on the overall interest dynamics among their fellow

⁶ Naturally, interviewees also made generic references to learning and accountability. Given that we are interested in the evaluation use along political interests linked to P-A-dynamics, we do not report them.

members. Finally, we do not claim to cover specific particularities of individual evaluation processes; our aim is to reveal general patterns of evaluation use across a number of IOs.

Results

In the following sections, by going from IO cluster 1 to IO cluster 3, we navigate through the two dependent variables and their values for each IO and provide illustrations of the theorized dynamics. The interviews are labelled according to the target group (E for evaluators, MS for member states, IPA for secretariat staff). Note that we also report the share of directional sub-codes for each organization, which serves as an additional indicator for variable values. The aim is not to artificially quantify qualitative data. Instead, we aspire transparency and reliability as to how we decided about the dominant pattern in each IO. It also helps to summarize and illustrate findings.

Variation of evaluation units' orientation (dV1)

Cluster 1 (selected from ILO, UNDP, UNICEF) includes two IOs where member states control both the evaluation budget and agenda. In line with hypothesis 1, we found that in both IO 1 and IO 2 evaluation units tend to clearly orientate themselves towards the member states as their primary principals.

In IO 1, the head of evaluation unit regularly noted that the policy direction “is not set by the bureaucrats, it is set by the member states” and described their interest in evaluation as the “most important question”: one can only make evaluations “as useful as possible by ensuring that you address issues that your constituent and policy-makers are interested in” (E1). This perfectly illustrates what we mean by evaluation unit’s orientation towards one of the stakeholders. Furthermore, the evaluation unit’s director claimed to have the mandate from the governing body to force the IPA to follow his or her requirements, stressing that “if I go back to the governing body and say that there has not been any follow up on the evaluations, there is no one that can stop me”. Such a protection from the member states was implied to be crucial: “unless I have that, I would not be able to hold back the pressure [from the management]” (E1). By contrast, the respondent referred to the IPA mostly in the context of technical issues.

Similarly, in IO 2, the head of evaluation unit highlighted several times that he or she works independently from the administration which is not the evaluation unit’s client (E2). Instead,

the interviewee highlighted consultations with member states as a “critical process” to evaluation and explained that the evaluators’ task was to evaluate—not to aid—the administration (E2). While this sometimes created “a bit of a shock” for the administration, it brought “respect at the [Executive] Board” (E2). The orientation towards member states was illustrated with several examples, where member states supported the evaluation unit vis-à-vis the management: for instance, when IPA staff requested content alterations in evaluation reports, refused to give access to information, or refrained from providing a management response. The interviewee even revealed that such conflictive situations led to attempts by the IPA to alter evaluation policy and intervene into the evaluation unit’s independence. However, due to the support by the member states, such attempts were “lost badly” (E2).

Cluster 2 (selected from IAEA, IOM, UNHCR) includes IOs where the IPA controls both the evaluation budget and agenda. Our empirical data from IO 3 and 4 clearly demonstrate evaluation units’ orientation towards the IPA as a key sponsor and user, which, again, supports hypothesis 1. The head of IO 3 evaluation unit claimed that the IPA and not member states was the key beneficiary of evaluation service. It was stated that “member states get what they want, but they get it not through the evaluation service” (E3). On the one hand, the interviewee stressed the unit’s direct reporting to the head of the administration, while on the other it was noted that member states’ involvement in evaluation activities is rather low and fragmented. This, in turn, shows that the target group of the evaluation unit was the IPA, rather than the governing body. As the head of the unit said, draft results were shared with “the teams that are most associated with the evaluated work” and (s)he would do briefings or “a workshop where we are talking about the findings and the recommendations” (E3).

The evaluation unit of IO 4 also said that the IPA was the key target group for all their evaluations. The interviewee even claimed that it is in agreement with the member states, as they “themselves want to give certain freedom to the Director-General to manage the organization” (E4). Furthermore, although only few statements were made on the indicator of “Evaluation Unit Sponsor”, the interviewee associated him- or herself with the bureaucracy, highlighting the active and flexible internal interaction at different levels, where people know each other and operate “as kind of a family” (E4).

Finally, *cluster 3* (selected from UNEP, WHO, UNESCO) includes IOs where evaluation systems are dominated by both member states and the IPA. Their heads depicted both member states and the IPA as their main sponsors and evaluation users. In IO 5, the unit head emphasized the importance of member states’ push towards a stronger evaluation function and stressed their

recognition of good evaluators' work (E5). It was argued that "member states assign more and more tasks" and that evaluations are often done "because member states wanted us to do it" (E5). Yet, on the other hand, the official regularly referred to the mutual goals with the rest of the IPA and how evaluation helps the organization to improve. Various formats of "internal" discussions within the organization were outlined before proposals would go to member states (E5). Finally, a mixed message regarding the unit's sense of sponsorship was given. The interviewee explained that the IPA's head gave him or her the mandate "to go ahead" and ask everyone to "open everything [he or she] needed to see". At the same time, it was mentioned that member states watch the evaluation work "with a kind of covering hand" and demand it to be independent (E5).

In IO 6, when describing his or her function, the evaluation head claimed to "provide timely advice" to the head of the IPA, implying the administration to be the primary user (E6). However, it was also mentioned that "at the end of the day, we are a member states organization and if they have capacity, interest, and majority to run a certain thing..., there is nothing we can do about it" (E6). Regarding sponsorship, the interviewee said (s)he could openly discuss all matters with the senior management *before* the discussion with the governing body, which seems to be a relevant opportunity for the evaluation unit to ensure management support given that member states tend to pressure the evaluation director to answer very "pedantic" questions. However, further statements were also made on member states' sponsorship in terms of political protection from management's influence saying that "nobody can stop us from writing and sending things to the [Executive] Board... This helps us to build credibility with the member states" (E6). Such contradictions thus indicate a mixed evaluation unit's orientation to both member states and the IPA.

Figure 2 summarizes our findings and illustrates the dominant pattern in the six IOs based on our coding results. A clear trend towards the member states as primary principals can be observed in the responses of interviewees from IO 1 and IO 2, where the overwhelming majority of statements made on evaluation unit orientation were directed to member states (79–93%). In contrast, interviewees from cluster 2 IOs mostly (or even exclusively for IO 4) spoke about the IO administration as a key sponsor and user of evaluation. In IO 5 and 6, statements referred both to the IPA and IO member states, showing a mixed pattern in evaluation units' orientation. All three IO clusters tend to maintain the first hypothesis, allowing us to conclude that evaluations units orientate themselves towards actors who controls the structure of evaluation system. The next question is whether we can also observe a corresponding pattern in the political use of evaluation results.

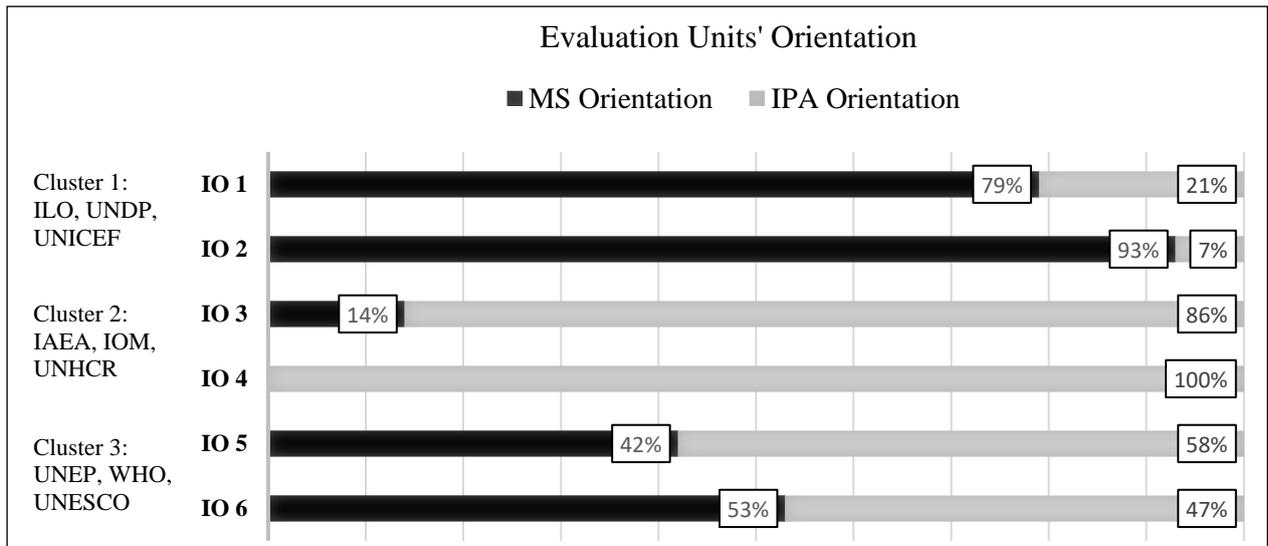


Fig. 2 Illustration of dominant patterns in evaluation unit orientation across six IOs (dV1). The results indicate what percentage of all statements on evaluation unit orientation (made by evaluation unit interviewees) were directed either to member states or the IPA

Variation of political evaluation use alignment (dV2)

Analogously to the previous section, we now turn to the empirical analysis of our second hypothesis. Other than in the previous section, we now include interviews with member state (MS) and administrative representatives (IPA). We report answers corresponding with evaluation use along member states and IPA interest as identified in the theory section (we refer to the dimensions in Table 2 with terms in *italics*).

The findings from *cluster 1* (selected from ILO, UNDP, UNICEF) reveal a pattern of evaluation use predominantly aligned to the interests of member states. In IO 1, when asked about evaluation use, most respondents highlighted the complexity of member states with “conflicting interests among parties”, where evaluation helps to learn about each other’s preferences (MS1; E1) (*collective principal*).⁷ Interviewees also claimed that member states would impose evaluation recommendations as a tool of control if the IPA would not go “in line with the policy direction that they [member states] have in mind” (E1; MS1; IPA1) (*P-A control*). By contrast, only very few statements were made on typical IPA interests. An IPA official said that the administration may occasionally use results as “a negotiation card” within the IPA’s own structure for distinct departments to get support from the senior management (IPA1).

⁷ If several interviews are cited, the first label refers to the direct quotation, whereas further interviews refer to similar statements.

In IO 2, respondents also highlighted member states' heterogeneity, arguing that evaluation is often used in negotiations between states (IPA3; IPA4). Such use was clearly political—due to their “national views”, states would only focus on evaluation findings or topics that suit their purposes, making evaluation a “contested area” (MS2; MS5; MS6). As one IPA official explained, such contestation was especially tangible between developing and developed countries (IPA4) (*collective principal*). Furthermore, interviewees referred to the need to contain the IPA. Member states expressed their concerns that the IO management downplays issues and shows “the good things” only (MS3, MS2). Others suggested that the bureaucracy has its own agenda that is hidden from member states (MS4). In light of this, evaluation was depicted as a tool to control the administration, i.e. to get “an outside view” (MS4), to reveal “where the challenges are” (MS3), and to “send signals” on course correction (MS2). The IPA staff, too, described evaluation as a “top down external mechanism” which member states use to “control and to hold you [IPA] to task” (IPA5) (*P-A control*).

In contrast, cluster 2 (selected from IAEA, IOM, UNHCR) indicates a strong tendency towards evaluation use aligned to IPAs' political interests both regarding internal fragmentation dynamics and in relation to member states. In IO 3, evaluation was described as a tool to gather information on key issues at all levels of organization, from junior staff to senior management, which then allows the administration leadership to set strategic priorities and internally steer the organization by “creating course corrections” (E3; IPA8; MS7; MS8) (*collective agent*). Interviewees also said that the senior management might manipulate member states' positions on certain issues (e.g. country programs) by framing evaluation results “in a politically clever or politically sensitive way” (MS9; IPA7; IPA8). Such a tactic was described as “advocacy on issues” in a dialogue with governments (E3), which perfectly illustrates how evaluation may help the IPA to gain policy-making influence. It was also noted that the IPA may use evaluations to justify performance failures, to show “that things are going better and that it is in their [member states] favor” (IPA7; MS7) (*bureaucratic influence*).

A similar pattern prevails for IO 4. Regarding vertical IO steering, evaluation was said to be helpful to consolidate internal “information and knowledge management” (MS9; E4; IPA10), referring to the IPAs fragmented organizational structures (*collective agent*). In relation to IPA influence and justification, interviewees claimed that evaluation helps the organization to raise additional funding (to “sell projects”) (MS8) and allows the IPA to “promote” its own activities or “justify own mistakes”; for instance, by drawing member states' attention to specific evaluation reports in governing body meetings (E4; IPA9). It was even claimed that evaluation reports would never admit the IPA's own mistakes but would rather point out the external

factors (IPA9) (*bureaucratic influence*). On rare occasions, respondents also noticed that member states would refer to evaluation to oversee the IPA or convince other member states, for instance, in budgetary questions (IPA10).

Finally, a more mixed pattern prevails in *cluster 3* (selected from UNEP, WHO, UNESCO). In IO 5, interviewees again described the diversity of member state interests and claimed that the organization shall “deliver benefit for all the membership and not only for some” (MS11; MS1; MS10; E5) (*collective principal*). Furthermore, evaluation was argued to “detect” issues which might not be reported by the secretariat itself (*P-A control*) (MS11; MS8; IPA12; IPA13). But a significant number of interviewees also pointed out that evaluation served typical IPA interests. For instance, interviewees claimed that evaluation often helps senior management to get a joint understanding about “the different departments doing their own thing” (*collective agent*) (IPA12; IPA14; E5). Interviewees also detailed that evaluation was strategically used by the head of IPA to ask member states for additional funding (MS11), to promote specific programs (MS1), or convince member states if needed: “It makes it much easier to convince people because it is... evidence based. Using evaluation in communication parts is extremely powerful” (*bureaucratic influence*) (IPA14).

In IO 6, interviewees talked about “political struggles” among member states and how evaluation may help to counter single countries trying “to defend their programs... regardless of the results” (*collective principal*) (MS12; MS13; E6). Interviewees also revealed that “the administration is trying to keep the member states out...” (IPA15), whereas evaluation findings provide member states with a baseline for comparison of certain programs (MS12; MS13; IPA16) (*P-A control*). But again, almost half of all statements referred to distinctive IPA interests. For instance, the IO leadership said that they benefited from getting “insight into the processes and the functioning of individual offices where we [senior management] do not have sufficient view on what is going on” (IPA15; IPA16) (*collective agent*). Finally, it was also argued that the IPA would “regularly refer to evaluation reports to defend its own position” *vis-à-vis* the member states, especially when it comes to budgeting or the extension of projects or programs (MS13) (*bureaucratic influence*).

Figure 3 summarizes the findings. The share of interviewee statements referring to either political use alignment with IPA interests, or use alignment with member state interests (or the mix thereof), corresponds with the theorized assumption that the evaluation unit’s orientation determines the type of political evaluation use. Admittedly, evaluation use alignment towards

IPA in IO 5 is more pronounced than we expected (with 63%); however, the pattern is still clearly balanced if taken together with IO 6 and compared to the other two clusters.

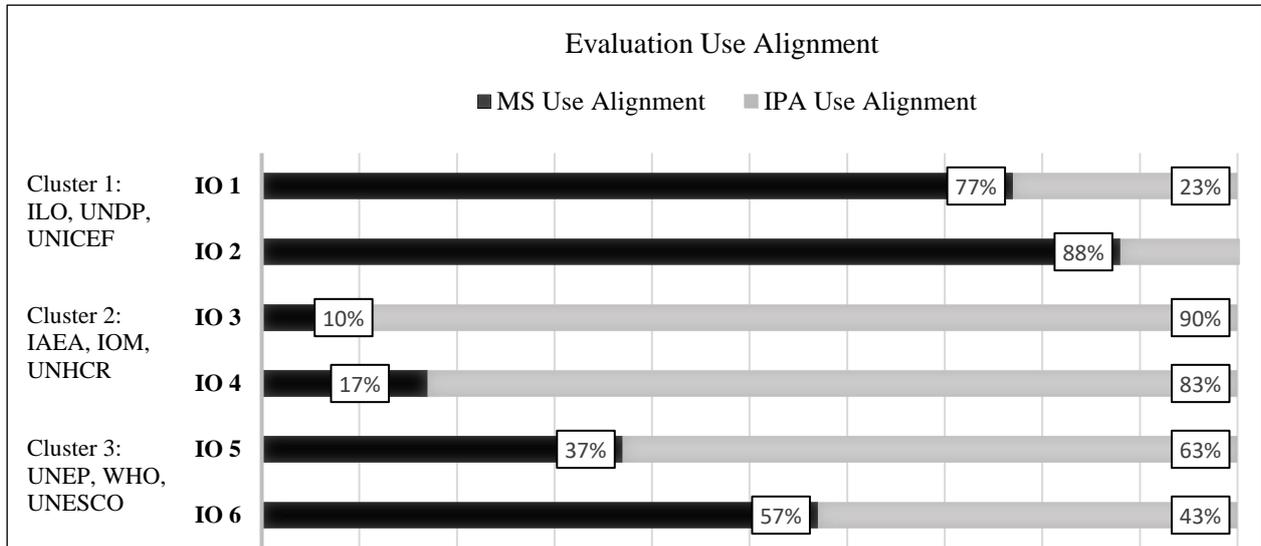


Fig. 3 Illustration of dominant patterns in political evaluation use alignment (dv2). The results indicate what percentage of all statements on political evaluation use (made by all interviewees) were directed either to the interests of the member states or the IPA

Discussion

Overall, the tendency of interviewees—across all three target groups—to frame their statements in the theoretically expected direction is remarkably strong. We find all expectations well confirmed (see summary in Fig. 4). In line with H1, the evaluation units’ orientations were consistent with our expectation that those who control evaluation resources will be perceived as primary sponsors and evaluation users. In line with H2, we found that political evaluation use followed the expected pattern. In cluster 1 (member state dominated IOs), respondent statements on political evaluation use referred to typical member states’ interests (containing unilateral influence; controlling the IPA). In cluster 2 (IPA dominated IOs), respondents predominantly mentioned typical IPA political interests (justification, policy influence and internal steering). In IOs with balanced systems (cluster 3), evaluation unit orientation and use alignment are also mixed.

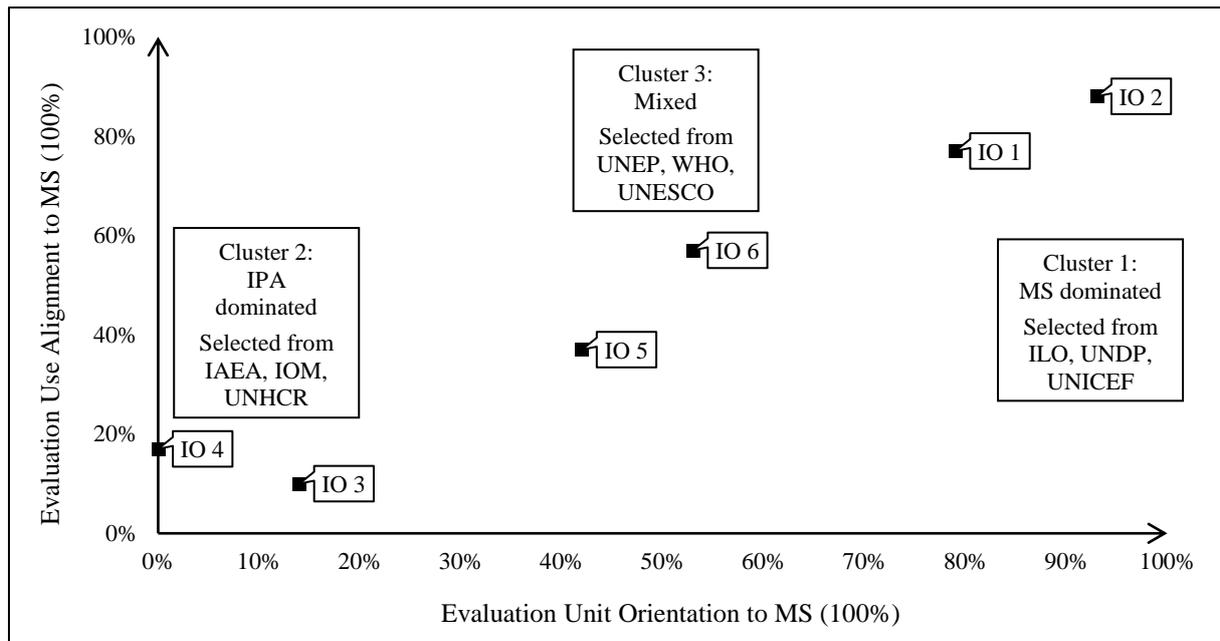


Figure 4: Illustration of empirical findings based on coding results. The results indicate the share of total interviewee references to evaluation unit orientation and evaluation use alignment with 100% meaning that all of the coded statements in the category referred to member states.

The main limitation of this study is the lack of data at the level of evaluation reports and at the level of policy outcomes as our data on political evaluation use consists of statements made by experts who described the use of evaluation results in policy-making processes of their IO in general terms. At the same time, the strength of our approach is that we did not ask about the political use of evaluation directly but measured both dependent variables indirectly by comparing how stakeholders responded differently to similar general questions on evaluation use, depending on who controls evaluation system resources. While findings therefore remain generic, we expect nonetheless that the dynamics theorized and described in this paper should have substantive implications, both for actual evaluation research and reports (see for instance the results by van Voorst and Mastebroek 2019) and for policy decisions (see also the substantive evidence about pressure on evaluators; (cf. Pleger et al. 2017)). Whereas other research designs are necessary to answer detailed questions about policy outcomes, the empirical merit of this paper is in-depth insights from main evaluation stakeholders, particularly evaluation unit experts.

Finally, there is little reason to believe that confounding factors drive our results. On the one hand, the six IOs studied in this paper subscribe to the same evaluation norms and standards, conduct centralized evaluations and report to similar governance structures, which rules out alternative explanations *linked to evaluation policy*. On the other hand, the IOs vary in their

mandate, policy field, and operational profile both within and across the clusters (see “Appendix 5”) This rules out alternative explanations *linked to general IO characteristics*. It is reasonable to conclude, therefore, that evaluation unit orientation and political evaluation use in other comparable IOs should also vary according to the theorized patterns.

Conclusion and theoretical implications

In this paper, we propose a relatively simple, yet novel, theory-founded framework on how to study the political use of evaluation. This is especially useful in times of booming evaluation businesses and the proliferation of institutional assessments both in domestic and international politics (see Cooley and Snyder 2015). Although evaluation is still “widely viewed as the ‘gold standard’ of institutional assessment” (Lall 2017, p. 245), we demonstrate that evaluation systematically serves ex-ante political interests of policy actors, depending on who exerts control over evaluation system resources.

Our findings yield two theoretical implications. The first speaks to the literature on IOs and International Relations. The findings show that institutional IO design matters. At the same time, our results disfavour the *state*-centric view of international institutions which is embedded in rational design theory (Abbott and Snidal 1998; Koremenos et al. 2001). As our findings from cluster 2 and 3 demonstrate, IO *bureaucracies* are able to use IOs’ internal structure (evaluation system) to exert their influence using evaluation as a political tool (Patz and Goetz 2019; Knill et al. 2019; Biermann and Siebenhüner 2009; Bauer et al. 2017; Johnson and Urpelainen 2014). The key theoretical puzzle for future research is to explain the *reasons* for such variation in control over evaluation resources.

Secondly, our results speak to the literature on Public Policy and Evaluation. The findings underline that the purpose, efficiency, and function of public management tools, including evaluation, should always be seen in the context of contested stakeholder interests, especially in such complex organizational environments like IOs. For anthropologists, the expansion of evidence instruments, indicators, and quantification “comes from a political culture that demands more openness and seeks to drive out corruption, prejudice, and the arbitrary power of elites” (Merry 2011, 85; see also Porter 1995). The paper’s findings, however, raise doubts that such a culture can ever be realized in the context of public service organizations. While IOs may be extreme in their internal political contestation (see our literature review), bureaucratic politics and the dichotomy of administration and political actors are also well

known domestically. Neglecting the political nature of evaluation stimulates unintended consequences, when functional tools are harbouring political agendas. In this regard, studies on evaluation should generally pay more attention to the setup of evaluation systems, who controls evaluation resources and how such differences came to be (see Hinterleitner et al. 2016; Fforde 2019; Perl et al. 2018, p. 591; van Voorst and Mastenbroek 2019).

In terms of practical implications, practitioners should consider that political use of evaluation may hinder its functional purposes. As our data implies, an evaluation unit's strong orientation towards one stakeholder leads to disengagement by the other stakeholder. For instance, in IO 2, where member states dominate the evaluation system, IPA officials were sceptical about evaluation's contribution to learning. They criticized the quality of evaluation reports and explained that evaluation is always "at the center of controversy", triggering "defensive behaviors" among staff (IPA6, IPA5). By contrast, in IO 3, where the evaluation unit is controlled by the IPA, member states were less interested in evaluation's benefits and perceived it as an internal "management tool" (MS7).

Future research should thus investigate whether evaluation systems with the mixed control setting (cluster 3 IOs) mitigate politicization and increase evaluation's functional use. After all, the proper answer to evaluation politics is to improve, and not abandon, the evaluation practices. Scholars should also examine the extent to which political interests affect the actual evaluation research and results. Today, evaluation has become a booming industry and it should be relevant whether we are merely looking at politics with other means.

References

- Abbott, K. W., & Snidal, D. (1998). Why states act through formal international organizations. *The Journal of Conflict Resolution*, 42(1), 3–32.
- Anderson, J. E. (1975). *Public policy-making*. New York: Praeger.
- Azzam, T. (2010). Evaluator responsiveness to stakeholders. *American Journal of Evaluation*, 31(1), 45–65.
- Banner, D. K. (1974). The politics of evaluation research. *OMEGA, The International Journal of Management Science*, 2(6), 763–774.
- Bauer, M. W., Knill, C., & Eckhard, S. (Eds.). (2017). *International bureaucracy. Challenges and lessons for public administration research*. London: Palgrave Macmillan.
- Beach, D., & Pedersen, R. B. (2016). Selecting appropriate cases when tracing causal mechanisms. *Sociological Methods & Research*.
- Biermann, F., & Siebenhüner, B. (Eds.). (2009). *Managers of global change. The influence of international environmental bureaucracies*. Cambridge, MA: MIT Press.
- Bjornholt, B., & Larsen, F. (2014). The politics of performance measurement. ‘Evaluation use as mediator for politics’. *Evaluation*, 20(4), 400–411.
- Botterill, L. C., & Hindmoor, A. (2012). Turtles all the way down. Bounded rationality in an evidence-based age. *Policy Studies*, 33(5), 367–379.
- Brewer, G. D. (1974). The policy sciences emerge. To nurture and structure a discipline. *Policy Sciences*, 5(3), 239–244.
- Chelimsky, E. (1987). The politics of program evaluation. *Society*, 25, 24–32.
- Cooley, A., & Snyder, J. (Eds.). (2015). *Ranking the world. Grading states as a tool of global governance*. Cambridge: Cambridge University Press.
- da Conceição-Heldt, E. (2013). Do agents “run amok”? A comparison of agency slack in the EU and US trade policy in the Doha Round. *Journal of Comparative Policy Analysis: Research and Practice*, 15(1), 21–36.
- Denzin, N. K., & Lincoln, Y. S. (2000). *Handbook of qualitative research* (2nd ed.). Thousand Oaks: Sage Publication.
- Dijkstra, H. (2015). Shadow bureaucracies and the unilateral control of international secretariats. Insights from UN peacekeeping. *Review of International Organizations*, 10(1), 23–41.
- Easterly, W. (2006). *The white man’s burden. Why the West’s efforts to aid the rest have done so much ill and so little good*. New York: Penguin Press.
- Eckhard, S., & Ege, J. (2016). International bureaucracies and their influence on policy-making. A review of empirical evidence. *Journal of European Public Policy*, 23(7), 960–978.

- Eckhard, S., & Jankauskas, V. (2019). The politics of evaluation in international organizations. A comparative study of stakeholder influence potential. *Evaluation*, 25(1), 62–79.
- Elsig, M. (2011). Principal–agent theory and the World Trade Organization. Complex agency and ‘missing delegation’. *European Journal of International Relations*, 17(3), 495–517.
- FAO. (2016). *Evaluation of FAO’s evaluation function. PC 120/5*. Rome: Food and Agriculture Organization of the United Nations.
- Fforde, A. (2019). Yes, but what about the authority of policy analysts? A commentary and discussion of Perl et al., ‘Policy-making and truthiness: Can existing models cope with politicized evidence and willful ignorance in a post-fact world?’. *Policy Sciences*, 52(1), 153–169.
- George, A. L., & Bennett, A. (2005). *Case studies and theory development in the social sciences*. Cambridge, MA: MIT Press (BCSIA studies in international security).
- Goffman, E. (1959). *The presentation of the self in everyday life*. New York: Doubleday.
- Graham, E. R. (2014). International organizations as collective agents. Fragmentation and the limits of principal control at the World Health Organization. *European Journal of International Relations*, 20(2), 366–390.
- Hanrieder, T. (2015). *International organization in. Time fragmentation and reform*. Oxford: Oxford University Press.
- Hawkins, D. G., Lake, D. A., Nielson, D. L., & Tierney, M. J. (Eds.). (2006). *Delegation and agency in international organizations*. Cambridge: Cambridge University Press.
- Hinterleitner, M., Sager, F., & Thomann, E. V. A. (2016). The politics of external approval. Explaining the IMF’s evaluation of austerity programmes. *European Journal of Political Research*, 55(3), 549–567.
- Højlund, S. (2014). Evaluation use in the organizational context—Changing focus to improve theory. *Evaluation*, 20(1), 26–43.
- Højlund, S. (2015). Evaluation in the European Commission. For accountability or learning? *European Journal of Risk Regulation*, 6(1), 35–46.
- Howlett, M., Ramesh, M., & Perl, A. (2009). *Studying public policy. Policy cycles and policy subsystems*. Oxford: Oxford University Press.
- IEO. (2018). *Annual report on evaluation. Independent evaluation office*. New York: United Nations Development Programme.
- ILO. (2019). *Annual evaluation report 2018–2019*. Geneva: International Labour Organization.
- JIU. (2014). *Analysis of the Evaluation Function in the United Nations System*. United Nations Joint Inspection Unit. JIU/REP/2014/6. Geneva.
- Johnson, T., & Urpelainen, J. (2014). International bureaucrats and the formation of intergovernmental organizations. Institutional design discretion sweetens the pot. *International Organization*, 68(1), 177–209.

- Kassim, H., & Menon, A. (2003). The principal-agent approach and the study of the European Union. Promise unfulfilled? *Journal of European Public Policy*, *10*(1), 121–139.
- Knill, C., Bayerlein, L., Enkler, J., & Grohs, S. (2019). Bureaucratic influence and administrative styles in international organizations. *Review of International Organizations*, *14*, 83–106.
- Koremenos, B., Lipson, C., & Snidal, D. (2001). The rational design of international institutions. *International Organization*, *55*(4), 761–799.
- Lall, R. (2017). Beyond institutional design. Explaining the performance of international organizations. *International Organization*, *71*(2), 245–280.
- Lasswell, H. D. (1956). *The decision process. Seven categories of functional analysis*. College Park: Bureau of Governmental Research: University of Maryland Press.
- Leeuw, F. L., & Furubo, J.-E. (2008). Evaluation systems. What are they and why study them? *Evaluation*, *14*(2), 157–169.
- Lijphart, A. (1971). Comparative politics and the comparative method. *American Political Science Review*, *65*, 682–693.
- Lyne, M. M., Nielson, D. L., & Tierney, M. J. (2006). Who delegates? Alternative models of principals in development aid. In D. G. Hawkins, D. A. Lake, D. L. Nielson, & M. J. Tierney (Eds.), *Delegation and agency in international organizations* (pp. 41–76). Cambridge: Cambridge University Press.
- Mayntz, R. (2004). Mechanisms in the analysis of social macro-phenomena. *Philosophy of the Social Sciences*, *34*(2), 237–259.
- Mele, V., Anderfuhren-Biget, S., & Varone, F. (2016). Conflicts of interest in International Organizations. Evidence from two United Nations humanitarian agencies. *Public Administration*, *94*(2), 490–508.
- Merry, S. E. (2011). Measuring the world. *Current Anthropology*, *52*(S3), S83–S95.
- Morris, M., & Clark, B. (2013). You want me to do what? Evaluators and the pressure to misrepresent findings. *American Journal of Evaluation*, *34*(1), 57–70.
- Nielson, D. L., & Tierney, M. J. (2003). Delegation to international organizations. Agency theory and World Bank environmental reform. *International Organization*, *57*(2), 241–276.
- O'Brien, T., Payne, S., Nolan, M., & Ingleton, C. (2010). Unpacking the politics of evaluation. A dramaturgical analysis. *Evaluation*, *16*(4), 431–444.
- Patz, R., & Goetz, K. H. (2019). *Managing money and discord in the UN. Budgeting and bureaucracy*. Oxford: Oxford University Press.

- Perl, A., Howlett, M., & Ramesh, M. (2018). Policy-making and truthiness. Can existing policy models cope with politicized evidence and willful ignorance in a “post-fact” world? *Policy Sciences*, 51(4), 581–600.
- Pleger, L., Sager, F., Morris, M., Meyer, W., & Stockmann, R. (2017). Are some countries more prone to pressure evaluators than others? Comparing findings from the United States, United Kingdom, Germany, and Switzerland. *American Journal of Evaluation*, 38(3), 315–328.
- Porter, T. M. (1995). *Trust in numbers. The pursuit of objectivity in science and public life*. Princeton, NJ: Princeton University Press.
- Raimondo, E. (2018). The power and dysfunctions of evaluation systems in international organizations. *Evaluation*, 24(1), 26–41.
- Rittberger, V., Zangl, B., Kruck, A., & Dijkstra, H. (2019). *International organization* (3rd ed.). London: Red Globe Press.
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (Eds.). (2004). *Evaluation. A systematic approach* (7th ed.). Thousand Oaks: Sage Publication.
- Schoenefeld, J., & Jordan, A. (2017). Governing policy evaluation? Towards a new typology. *Evaluation*, 23(3), 274–293.
- Schoenefeld, J., & Jordan, A. (2019). Environmental policy evaluation in the EU. Between learning, accountability, and political opportunities? *Environmental Politics*, 28(2), 365–384.
- Stockmann, R., Meyer, W., & Schenke, H. (2011). Independence of evaluations. *Zeitschrift für Evaluation*, 10(1), 39–67.
- Streck, C. (2001). The global environment facility. A role model for international governance? *Global Environmental Politics*, 1(2), 71–94.
- Taylor, D., & Balloch, S. (Eds.). (2005). *The politics of evaluation. Participation and policy implementation*. Bristol: The Policy Press.
- The LSE GV314 Group. (2014). Evaluation under contract Government pressure and the production of policy research. *Public Administration*, 92(1), 224–239.
- UNDP. (2016). *The UNDP evaluation policy. DP/2016/23*. New York: United Nations Development Programme.
- UNEG. (2016). *Norms and standards for evaluation*. New York: United Nations Evaluation Group.
- UNHCR. (2016). *UNHCR policy on evaluation. UNHCR/HCP/2016/2*. Geneva: United Nations High Commissioner for Refugees.
- Urpelainen, J. (2012). Unilateral influence on international bureaucrats. An international delegation problem. *Journal of Conflict Resolution*, 56(4), 704–735.

- van Voorst, S., & Mastebroek, E. (2019). Evaluations as a decent knowledge base? Describing and explaining the quality of the European Commission's ex-post legislative evaluations. *Policy Sciences*, 52(4), 625–644.
- Vaubel, R., Dreher, A., & Soylu, U. (2007). Staff growth in international organizations. A principal-agent problem? An empirical analysis. *Public Choice*, 133(3), 275–295.
- Voeten, E. (2008). The impartiality of international judges. Evidence from the European Court of Human Rights. *American Political Science Review*, 102(4), 417–433.
- Weaver, C. (2010). The politics of performance evaluation. Independent evaluation at the International Monetary Fund. *Review of International Organizations*, 5(3), 365–385.
- Weiss, C. H. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, 19(1), 21–33.
- Weiss, T. G., & Jordan, R. S. (1976). Bureaucratic politics and the world food conference. The international policy process. *World Politics*, 28(3), 422–439.
- Weiss, C. H., Murphy-Graham, E., & Birkeland, S. (2005). An alternate route to policy influence. How evaluations affect D.A.R.E. *American Journal of Evaluation*, 26(1), 12–30.
- Wergin, J. F. (1976). The evaluation of organizational policy making. A political model. *Review of Educational Research*, 46(1), 75–115.
- WHO. (2018). *Evaluation. Evaluation Policy. EB143/(9)*. Geneva: World Health Organization.
- Wildavsky, A. (1972). The self-evaluating organization. *Public Administration Review*, 32(5), 509.

Appendix to Paper C

Appendix 1: Measurement of stakeholder control over evaluation systems

Primary data sources

IAEA	IAEA Evaluation Policy (2011), Office of Internal Oversight Services (OIOS) OIOS Charter (2014)
ILO	ILO Evaluation Policy (2017), GB.331/PFA/8 ILO policy guidelines for evaluation: Principles, rationale, planning and managing for evaluations (2017), 3rd edition
IOM	IOM Evaluation Policy (2018) Charter of the Office of the Inspector General (OIG) (2015), IN/74 Rev.1 IOM Evaluation Guidelines (2006), OIG
UNICEF	Revised Evaluation Policy of UNICEF (2018)
UNDP	Revised UNDP Evaluation Policy (2019)
UNEP	UNEP Evaluation Policy (2016), Evaluation Office
UNHCR	UNHCR Policy on Evaluation (2016), UNHCR/HCP/2016/2 UNHCR Evaluation Strategy (2018-2022)
UNESCO	UNESCO Administrative Manual (2017), Item 1.6 Internal Oversight UNESCO Evaluation Policy (2014-2021), IOS/EVS/PI/162
WHO	Evaluation: evaluation policy (2018), EB143/9

These documents were used to measure the control over evaluation system resources (independent variable), see below.

Measurement of stakeholder control over evaluation system resources

IO	Evaluation staff appointment	Evaluation budget allocation	Evaluation agenda-setting
ILO	MS&IPA 0.5	MS 1	MS 1
Source: 2017	Not specified in the policy, but: “In terms of our reporting line, I report only to the DG and the governing body”.	The budget approval procedure not specified, but: “The governing body is the decision-making body of the ILO. They have to proof the budget...”; “Budget of EVAL Office decided by	“EVAL will propose to the Governing Body each year... a proposed rolling programme of evaluation work for major independent evaluations”.

ILO Evaluation Policy		constituents in government body meetings” (interview).	
UNDP	IPA/MS 0.5	MS 1	MS 1
<i>Source:</i> 2019 UNDP Revised Evaluation Policy	“The UNDP Administrator appoints the Director of the Office in consultation with the Executive Board, taking into account the advice of the Audit and Evaluation Advisory Committee”.	“The Board approves the biennial financial appropriation to the Independent Evaluation Office”.	“The Board... undertakes periodic reviews and adjustments of such appropriations based on the programme of work of the Office, which the Board also approves”.
UNICEF	MS&IPA 0.5	MS 1	MS 1
<i>Source:</i> 2018 Revised Evaluation Policy	“The Director of Evaluation is appointed by the Executive Director in consultation with the Audit Advisory Committee and the Executive Board, with an external evaluation expert as part of the selection panel”.	“As part of the approval of the integrated budget, the Board approves the budget of the Evaluation Office”.	“The Executive Board... approves the plan for global evaluations”.
IAEA	IPA 1	IPA 1	IPA 1
<i>Source:</i> 2014 IAEA OIOS Charter	Evaluation policy does not specify appointment rules, but: “the Director of OIOS... reports directly to the Director General”.	“The annual work plan is prepared and proposed to the Director General to ensure that the resources required for the internal oversight services are sufficient, appropriate, and effectively deployed to meet the expected outcomes”.	“The annual (OIOS) work plan shall be subject to the approval of the Director General”.
IOM	IPA 1	IPA 1	IPA 1
<i>Source:</i> 2015 Charter of the Office of the Inspector General (OIG)	The head of evaluation unit is appointed by the Inspector General, who “is appointed by the Director General, who shall consult with the AOAC [advisory committee for the DG]”.	“The Director General ensures that OIG is provided with the necessary resources in terms of ...adequate funds...to fulfil its mission and maintain its independence”.	Biennial evaluation plan is submitted to the Director General through the Audit and Oversight Advisory Committee (AOAC).
UNHCR	IPA 1	IPA 1	IPA 1
<i>Source:</i> 2016 UNHCR Policy on Evaluation	“The High Commissioner is responsible for...appointing the Head of the Evaluation Service with the required experience, expertise, profile and qualifications”.	“The High Commissioner is responsible for...with support by the Deputy High Commissioner, approving the annual Work Plan for centralised evaluations”.	“The High Commissioner is responsible for...with support by the Deputy High Commissioner, approving...the Evaluation Service Budget”.
UNEP	IPA 1	MS 1	IPA 1
<i>Source:</i> 2016 UNEP Evaluation	“The Executive Director ensures that adequate and qualified staff are recruited for the effective functioning of the EOU”.	“The UNEA approves the operational budget of the EOU by reviewing and approving the EOU's	“The biennial evaluation work plan will be reviewed by the SMT and approved by the Executive Director”.

Policy		allocation within the overall budget of the organization of UNEP's PoW".	
WHO	MS&IPA 0.5 "The Director-General shall appoint a technically qualified head of the Evaluation Office after consultation with the Executive Board. The Director-General shall likewise consult the Executive Board before any termination of the incumbent of that office"	MS 1 "The Executive Board shall approve the biennial Organization-wide evaluation workplan, including its budget "	MS&IPA 0.5 "The workplans shall be submitted to the Executive Board for approval through the Programme, Budget and Administration Committee" BUT: The workplan will be finalised by including mandatory evaluations and those "significant for the organization and the management" (interview).
<i>Source:</i> WHO Evaluation Policy 2018			
UNESCO	MS&IPA 0.5 "The Director of IOS is appointed by the Director-General. The Director-General shall take decisions concerning the appointment, extension, renewal and termination of appointment of the Director of IOS in consultation with the Executive Board (Staff Regulations and Staff Rules, Regulation 4.5.3)".	MS&IPA 0.5 "The Director-General ensures that adequate resources are allocated to implement the quadrennial evaluation plan" AND "Executive Board ensures that adequate resources are allocated to implement the quadrennial corporate evaluation plan".	MS&IPA 0.5 "The Evaluation Office establishes the quadrennial plan in consultation with UNESCO senior management... The Executive Board may also request that specific topics be included in the evaluation plan". "The IOS Director maintains the ultimate authority for approving or modifying the corporate evaluation plan".
<i>Sources:</i> 2015 UNESCO Evaluation Policy Internal Oversight Charter-UNESCO			

Appendix 2: Interviewee list and positions

Interview list

No.	Interview Label	Organization Label	Target Group	Interviewee Position	Date
1	E1	IO 1	Evaluation Unit	Director	June 2018
2	E2	IO 2	Evaluation Unit	Director	March 2019
	E3	IO 3	Evaluation Unit	Director	June 2018
4	E4	IO 4	Evaluation Unit	Director	June 2018
5	E5	IO 5	Evaluation Unit	Director	June 2018
6	E6	IO 6	Evaluation Unit	Director	July 2018
7	IPA1	IO 1	IPA	Senior Programme Officer	October 2018
8	IPA2	IO 1	IPA	Programming Unit Head	April 2019
9	IPA3	IO 2	IPA	Unit Head	February 2019
10	IPA4	IO 2	IPA	Unit Head	February 2019
11	IPA5	IO 2	IPA	Senior Management Officer	March 2019

12	IPA6	IO 2	IPA	Senior Advisor	March 2019
13	IPA7**	IO 3	IPA	Senior Advisor	June 2018
14	IPA8	IO 3	IPA	Division Director	January 2019
15	IPA9	IO 4	IPA	Country Office Head	October 2018
16	IPA10	IO 4	IPA	Senior Manager	November 2018
17	IPA11	IO 4	IPA	Programme Head	October 2018
18	IPA12	IO 5	IPA	Department Director	June 2018
19	IPA13	IO 5	IPA	Department Director	June 2018
20	IPA14	IO 5	IPA	Senior Management Officer	June 2018
21	IPA15	IO 6	IPA	Division Director	July 2018
22	IPA16	IO 6	IPA	Senior Advisor	April 2019
23	MS1	IO 1, 3, 4, 5*	MS	Division Head	June 2018
24	MS2	IO 2	MS	First Secretary	February 2019
25	MS3	IO 2	MS	Deputy Permanent Representative	February 2019
26	MS4	IO 2	MS	First Secretary	February 2019
27	MS5	IO 2	MS	Minister	February 2019
28	MS6	IO 2	MS	Senior Advisor to Ambassador	February 2019
29	MS7	IO 3	MS	First Secretary	January 2019
30	MS8	IO 1, 3, 4, 5*	MS	Ambassador	June 2018
31	MS9	IO 1, 3, 4, 5*	MS	First Secretary	June 2018
32	MS10**	IO 1, 3, 4, 5*	MS	Minister Counselor	June 2018
33	MS11	IO 5	MS	First Secretary	June 2018
34	MS12**	IO 6	MS	Ambassador	July 2018
35	MS13	IO 6	MS	Deputy Permanent Representative	July 2018

* The interviewed country representation covers more than one IO from our sample, e.g. in Geneva.

** Upon request of the interviewee, this interview was not recorded. Instead, we took notes throughout the interview and uploaded our script into MAXQDA.

Equal levels of position seniority covered across IOs

Please note that we conducted interviews with officials from three target groups. First, we spoke with heads of each IO's evaluation unit. Thus, for evaluators, the level of seniority is held constant.

Second, we interviewed member state representatives focusing on high-level officials (ambassadors, deputies, minister counsellors) or medium-level officials (senior advisors, division heads) as they have the best overview of evaluation related dynamics in respective IOs' governing bodies. In some cases, lower-level diplomats were interviewed as well, for instance, when they were specifically responsible for evaluation-related issues in an IO. As the table

below demonstrates, all IOs are covered equally in terms of interviewees' position levels (sorted based on interview list above). The only exception is IO6 where we did not get access to medium-level officials. However, this should not bias the results as officials from high-level positions are covered.

IO	Interviews with member state representatives
IO1	High level (ambassador, minister counsellor) Medium (division head) Lower (first secretary)
IO2	High level (dep. permanent representative, minister) Medium (senior advisor to ambassador) Lower (2x first secretary)
IO3	High level (ambassador, minister counsellor) Medium (division head) Lower (2x first secretary)
IO4	High level (ambassador, minister counsellor) Medium (division head) Lower (first secretary)
IO5	High level (ambassador, minister counsellor) Medium (division head) Lower (2x first secretary)
IO6	High level (ambassador, dep. permanent representative)

Third, we conducted interviews with officials from IO administrations, also focusing on senior management staff – either at the high-level (chef de cabinet, programme heads, directors) or at the medium-level (unit heads, advisors). Again, the table below demonstrates that all IOs were covered equally. In IO5, we did not get access to medium-level management staff, yet we were able to speak with high-level officials.

IO	Interviews with IPA officials (level, position)
IO1	High level (Senior Programme Officer) Medium level (Programming Unit Head)
IO2	High level (Senior Management Office) Medium level (2x Unit Head, Senior Advisor)
IO3	High level (Division Director) Medium level (Senior Advisor)
IO4	High level (Country Office Head, Programme Head) Medium level (Senior Manager)
IO5	High level (2x Department Director, Senior Management Officer)
IO6	High level (Division Director) Medium level (Senior Advisor)

Appendix 3: Semi-structured interview questionnaire

Evaluation System*	<ul style="list-style-type: none"> - Who and how approves the agenda of centralized evaluation system? [evaluators] - Who and how approves the budget of centralized evaluation system? [evaluators] - Who and how approves the staff of centralized evaluation unit? [evaluators]
Evaluation Process	<ul style="list-style-type: none"> - How are the terms of reference formulated for centralized evaluations? [evaluators] - Who conducts centralized evaluations? [evaluators] - How are the reports drafted? [evaluators] - Who has access to draft evaluation results? [evaluators] - How do you present evaluation results to the stakeholders? [evaluators] - What happens after draft evaluation reports are finalized (management response, approval of recommendations, follow-up)? [evaluators] - What are the main challenges in your work? How do you deal with them? [evaluators] - To what extent are you [your unit/department/mission] involved in [IO] evaluation activities (central evaluation function)? - Is there any contestation about the evaluation process or results before they get officially confirmed? - How satisfied are you with the process of centralized evaluations in [IO]?
Evaluation Use	<ul style="list-style-type: none"> - In your view, what purposes does evaluation fulfil in [IO]? - How would you describe the use of evaluation in [IO]? - How would you describe member states/IPA interest and involvement in evaluation activities?

* We double-checked the results of our analysis of structural evaluations factors (agenda setting and budgetary resources) based on primary evaluation documents with the staff of evaluation units.

Appendix 4: Coded statements per IO/interviewee group

Overview of total coded statements on political evaluation use (dV2) across the three interviewee target groups in six international organizations:

	International Organizations					
	IO 1	IO 2	IO 3	IO 4	IO 5	IO 6
Member states	1	14	8	2	12	11
IPA	6	11	6	6	24	10
Evaluation Unit	11	2	6	10	10	5
Total coded statements:	18*	27	20	18*	46	26

* In IO 1 and IO 4, member states made only few statements on political evaluation use and the total number of statements is also lower than in other organizations. In our view, this implies that the political use of evaluation is less pronounced in these organizations. Note that every target group in every IO was interviewed in a balanced way (see Appendix II).

The dependent variable 1 (evaluation unit orientation) was coded based on interviews with evaluation unit staff (heads of units).

Appendix 5: IOs' general characteristics

The three clusters include international organizations which vary in their policy fields, capacities (staff and budgetary resources), the level of operability (staff outside HQ) as well as the dependency on voluntary contributions (see Table below). Given that this variation holds both within the clusters and across them, we hold that such broad IO characteristics cannot account for the observed variation in the dependent variables.

Cluster	IO	Policy Field	Staff		Budget	
			Total	Decentralized (%)	Total revenue (USD)	Non-assessed (%)
1	ILO	Social	3171	70%	691,926,212	44%
	UNDP	Development	7013	82%	5,517,025,115	100%
	UNICEF	Human rights	14474	87%	6,675,758,606	100%
2	IOM	Humanitarian aid	4888	94%	1,862,731,880	97%
	UNHCR	Humanitarian aid	10197	93%	4,338,294,301	99%
	IAEA	Security	2547	3%	667,839,306	38%
3	WHO	Health	8153	71%	2,901,381,847	83%
	UNESCO	Social	2206	48%	683,830,086	51%
	UNEP	Environment	910	18%*	741,749,419	68%

*Data as of 2018, based on IO websites and UN System Chief Executives Board for Coordination data (UN CEB). Policy fields coded manually based on IO websites (IOs' self-description). *Data as of 2010.*

Paper D

Is Control Always Better Than Trust? Determinants of Organizational Agency in Global Public Policy

This is the submitted version of the article

Under review in: *International Organization*

Abstract: The rise of international organizations (IOs) has also led to the rise of international bureaucracies, i.e. international public administrations (IPAs). What role do these bureaucracies play in global governance? Most dominant explanations refer to Principal-Agent (PA) theory which sees IPA agency merely as a by-product of (lacking) member states' control: the higher the control over the agent, the less likely is the agency slack (and vice versa). This study challenges these basic PA assumptions as they ignore agent characteristics exogenous to the delegation structure. We argue that to fully grasp IPA agency in global policy making one should consider both its formal discretion and informal behavioural orientations. Drawing on primary documents and 131 interviews with IPA staff, we separately measure control over the IPAs and their (informal) opportunistic behaviour in eight major IOs. The findings show that these two factors do not always align as PA theory would expect: more control does not necessarily constrain the IPAs' opportunism and less control does not necessarily induce the IPAs to slack. By providing a typology of how these factors interact, we thus argue that trusting rather than controlling the agent might sometimes be a cheaper and equally effective approach for the principal.

Keywords: principal-agent theory, agency slack, control, international organizations, international public administrations

Introduction

There is no doubt that over the last decades, we observe the evolution of an increasingly relevant additional level of international governance that complements and interacts with subnational, national, and supranational governance levels in highly complex and multifaceted ways. This development is reflected in a steep rise in the establishment of International Organizations (IOs). The institutionalization of policy-making structures beyond the nation state, mostly in the form of IOs, has been accompanied by the creation of differentiated bureaucratic substructures within these organizations. The rise of IOs is hence also a rise of international bureaucracies. In many ways, these International Public Administrations (IPAs) fulfil functions comparable to those of national ministerial bureaucracies and agencies. Similar to their national counterparts, IPAs play important roles in the initiation, preparation, and implementation of public policies. Yet, despite a growing interest in IPAs, it is still debated if and to what extent IPAs can be considered an independent source of influence in global policy making (Eckhard and Ege, 2016; Knill and Bauer, 2016).

From the outset, there seems to be a straightforward answer to this question, at least from the perspective of principal-agent (PA) theory, which has evolved as one of the main analytical tools to study IOs and global governance. It has advanced our understanding of why states delegate authority to IO administrations and what issues they face once delegation takes place, providing a well-structured perspective on how actors behave in a hierarchical setting. From a PA perspective, IPA agency is determined by the difference between delegated authority and political control. The more states delegate authority to IPAs and the less they control how IPAs use their granted powers, the higher is the potential for IPA agency in global policy-making: i.e., IPAs might behave in ways that deviate from the preferences of their principals. By contrast, the risk of such agency slack decreases with tighter principal's control. In short, to judge whether an IPA is an important or negligible agent in policy-making beyond the nation state, we simply need to focus on the principals: How much authority did they delegate and how much do they invest in control?

Yet, the nature of straightforward answers is that they are sometimes too simple. We indeed claim in this paper that PA theory provides an incomplete picture of IPA agency (and agent behaviour more generally). More specifically, we argue that the blind spot of PA theory is that it is in essence a theory that concentrates on the principal and simply assumes that the potential of agency slack is exclusively determined by the latter. First, PA assumes that agents will always slack, whenever they can. Second, PA assumes that the reduction of agency slack is

merely a matter of more control; agents will not extend their activities beyond the limits of formal controls if controlled appropriately. Both assumptions ignore that agent behaviour may also be determined by sources that are exogenous to the theorized PA relationship. How agents behave is not only a matter of delegation and control, but also affected by the behavioural dispositions that might vary from agent to agent. Some agents might simply act in line with the preferences of their principals although they have vast opportunities to slack. Other agents, by contrast, might constantly attempt to slack and work around even tightest controls. To capture such behavioural dispositions, we focus on organizational routines guiding IPA behaviour; i.e., informal and collectively internalized practices and processes characterizing the way an IPA engages in the formulation and implementation of public policies. In this regard, IPAs can display a more or less pronounced servant or entrepreneurial orientation. While the former pattern reduces the probability of agency slack even in the absence of tight control, entrepreneurial IPAs display a considerable potential of agency even in the face of strong member state oversight.

The central argument advanced in this paper is hence that the potential for IPA agency is determined by the interplay of two independent sources: (1) endogenous features of the PA relationship defined by the degree of delegation and control, and (2) agent orientations exogenous to this relationship that remain outside the radar of political control. In short, IPA agency can be conceived of as the aggregate of two different sources: formal discretion and administrative routines. To better grasp the risk of agency slack in delegation relationships, we should therefore consider both the endogenous (lack of) control and the exogenous nature of the agent. Adding an exogenous source of agency slack to the model potentially distorts the linearity of the relationship between control and agency slack and casts doubt on the assumed effectiveness of control to constrain agents' opportunism. As control is always costly, trusting the agent might in certain constellations be a preferable option for the principal.

In developing our argument, we contribute to the state of the art in theoretical, conceptual and empirical terms. Theoretically, this paper constitutes a first step towards rethinking how our theoretical view on delegation and agency slack could be modified to account for both principal control and agent characteristics. Building on earlier accounts that agent characteristics exogenous to the delegation structure matter (Hawkins and Jacoby, 2006), such a perspective could moreover help to achieve synergies between approaches linked to the principal-agent framework and those emphasizing informal bureaucratic dynamics. Conceptually, this paper contributes to the literature by presenting a novel and encompassing measure of direct and indirect IPA control as well as a measure for IPAs' opportunism, proxied by the prevalent

organizational routines throughout an IOs' policy cycle. Empirically, we investigate exogenous and endogenous sources of agency slack as well as their interaction focussing on the bureaucracies of eight major IOs. Drawing on primary document analysis and 131 in-depth interviews with IO staff, we provide four case studies (focusing on the IMF, NATO, ILO, and UNHCR bureaucracies) that display pronounced variation in the different ways how control by member states interact with IPAs' organizational routines. Our findings show not only that IPAs display pronounced differences in organizational routines and hence their potential of agency slack, but also that the observed levels of political control are partially ineffective or insufficient in light of agent characteristics.

In the next section, we first turn to detailing the two aforementioned basic PA assumptions (Section 2). We then go on to sketch out how the introduction of exogeneous agent characteristics into this relationship changes the understanding of agency slack (Section 3). In Section 4 and 5, we outline our measurement concepts and operationalizations for IPAs' control and opportunism potential, before putting our theoretical conjectures to an empirical test (Section 6). Section 7 concludes.

Basic assumptions of Principal-Agent theory

Delegation of authority is defined as “a conditional grant of authority from a *principal* to an *agent* that empowers the latter to act on behalf of the former” (Hawkins et al., 2006a: 7). The resulting relationship structure is characterized by power and information asymmetries. While the ultimate power remains with the principal who can rescind agent's authority, the expert knowledge favours the agent. To explain how these actors interact with each other, PA theory makes a set of assumptions about their behaviour.

Essentially, PA theory assumes interest divergence between the principal and the agent, the latter being a self-interested utility maximizing actor. Although Hawkins et al. (2006a: 7) argue that PA theory “does not imply any particular assumptions about the preferences of the actors”, the authors at the same time assume that these preferences diverge, so there is a “natural and perhaps inevitable conflict of interest between the parties” (Hawkins et al., 2006a: 24; see also Kiewiet and McCubbins, 1991: 24). Cortell and Peterson (2006: 256) also conclude that “PA approaches generally take agents' preferences as given and assume they will conflict with those of principals”. Indeed, if there was no preference divergence, PA models would “lose their power” (Tamm and Snidal, 2014: 137; Miller, 2005: 205).

Importantly, the prevailing goal divergence creates incentives for the agent to pursue its own agenda and slack, i.e. engage in an “independent action [...] that is undesired by the principal” (Hawkins et al., 2006a: 8). At the same time, the embedded information asymmetry creates opportunities for the agent to slack (Kiewiet and McCubbins, 1991). Hence, the risk of agency slack is structurally endogenous to all PA relationships. From the perspective of PA theory, merely trusting the agent is hence a hardly recommendable option.

The key concern of the principal, then, is how to ensure the agent’s faithfulness and thereby minimize agency slack potential. In this regard, PA theory assumes that the extent to which agents will slack primarily depends on the principal’s control. Control is defined as mechanisms used by the principal to observe the agent’s behaviour and direct it towards the principal’s interests (Miller, 2005; Kiewiet and McCubbins, 1991). Essentially, control reduces the agent’s autonomy and thus “the range of potential independent action available to an agent” (Hawkins et al., 2006a: 8). The central goal of PA theory has thus been to clarify “how particular institutional forms can be used to increase the likelihood of compliant behaviour by bureaucrats” (Huber and Shipan, 2002: 27), and, conversely, “minimize actual [agency] losses through control mechanisms designed into the P-A contract” (Hawkins and Jacoby, 2006: 226). Therefore, PA approaches have focused on agency slack mostly as an outcome of missing or imperfect control.

Albeit painted with a rather broad brush, we can distil two interrelated assumptions from the above. Firstly, as self-interested opportunistic actors, agents tend to exploit their autonomy and choose to work towards their own preferences over those of their principals. This creates the need for control. Hence, secondly, the tools of control employed by the principal are assumed to constrain the agent’s opportunism and reduce the risk of agency slack. The agent will not extend its activities beyond the limits of formal autonomy if controlled appropriately. As Hawkins et al. (2006a: 31) put it, “the principals can always reduce slack by tightening oversight”. Or in more relative terms, “the more closely the principal monitors the agent’s behavior [...] – the lower the risk of agency slack” (Abbott et al., 2016: 723; Lake and McCubbins, 2006; Kassim and Menon, 2003). The inverse relationship between agency slack and control is illustrated in Figure 1 below.

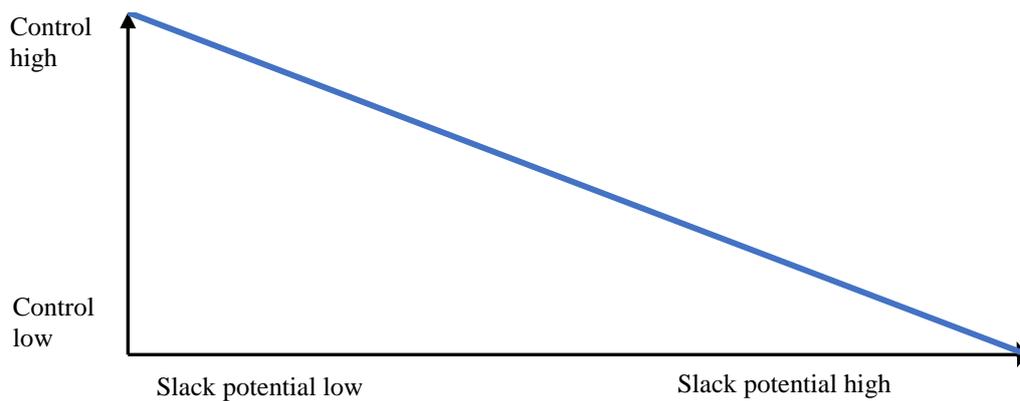


Figure 1: Assumed relationship between the risk of agency slack and control according to P-A theory.

Which position on the control continuum (y-axis) a principal will choose has been extensively theorized. Most prominently, since control is costly, there is a potential cost-benefit trade-off. While perfect control and low risk of agency slack therewith would bring about optimal outcomes in absolute terms, lower control could be preferred if control costs exceed the expected benefits in outcomes (Hawkins et al., 2006a: 31). Furthermore, control might undermine the agent's competences to fulfil its tasks (Abbott et al., 2020). Depending on the concrete configuration of how the trade-off between competence and control is settled politically there might be considerable leeway for independent agency beyond existing control options on the side of the principal. Hence, in determining the (ideal) level of control, principals are generally faced with political and budgetary constraints as well as the danger to undermine the agent's competence.

Endogenous and exogenous sources of agency slack

Whereas we continue to learn more about the variation in principal control, we know very little – in both theoretical and empirical regards – about the variation in the agent's potential to slack (x-axis). In fact, the twin assumptions summarized above view the risk of agency slack as a direct function of principal control. It is therefore hardly surprising that most scholarship on IO delegation still puts its predominant focus on the principals' side of the relationship (Pollack, 2003; Miller, 2005; e.g. Heldt, 2017).

This, however, only tells part of the story. When considering only those factors that are endogenous to the delegation contract, i.e., built into the delegation structure by the principal, such as control arrangements, studying variation in slack potentials per se is rendered of limited use. More recently, attention has been drawn to those characteristics that are exogenous to the delegation structure (Martin, 2006; Hawkins and Jacoby, 2006). Research has demonstrated that IPAs develop a life of their own beyond the IO's delegation contract and controls that bind them. Studies found that IPAs establish informal working routines (Bayerlein et al., 2020b) and vary in their self-interest with some IPAs willingly following member states' commands (Knill et al., 2016) and orientating towards mandate implementation (Knill et al., 2019; Biermann and Siebenhüner, 2009; Yi-chong and Weller, 2008). In practice, it thus appears that agents' opportunism, i.e. propensity to engage in slacking behaviour, is not solely determined by how tightly principals control them.

Taking exogenous factors such as informal behavioural patterns seriously in studying slack potential has profound consequences. First, it challenges the assumption that agents will slack if only given the opportunity to do so. From a solely endogenous perspective on delegation, preference divergence is a theoretical precondition for control to be necessary in the first place. The extent to which a given agent might develop and act upon diverging preferences is a constant that is seldomly explicitly theorized. If opportunistic behaviour varies exogenously, however, a highly autonomous agent does not necessarily need to slack while a tightly controlled agent may not automatically become 'a perfect handmaiden' (Cortell and Peterson, 2006). Second, acknowledging that agents differ in their opportunism independently from principal control challenges the PA assumption of a linear relationship between slack and control. Even when controlled similarly tightly, agents may vary in whether and to what extent they try to work around their principal's rules. In sum, when introducing exogenous variance into the basic P-A model, neither of the two basic assumptions summarized in Figure 1 holds as a constant (cf. Waterman and Meier, 1998).

Against this backdrop, we argue that the risk of agency slack can be conceived of as an aggregate of two different sources: the endogenous (lack of) control and the exogenous patterns of opportunistic behaviour. Taken together, they make up for a more precise picture of agency slack potential in delegation relationships. They might affect – but do not determine – each other. The literature emphasizes rather different determinants for variation in exogenous and endogenous factors. While the latter are primarily explained against the background of principals' preferences, institutional path dependencies and functionalist reasoning (Pierson, 2000; Koremenos et al., 2001; Hawkins et al., 2006), the former have their roots in factors such

as socialization, esprit de corps, and administrative perceptions of external challenges related to organizational fields or political threats (Knill et al., 2019; Bayerlein et al., 2020).

The argument can be summarized in an unusual yet effective metaphor. Imagine the principal as someone who is walking a dog, that is, the agent. Ideally, the dog would just walk beside the person, at the same pace, and in the same direction. Dogs, however, more often than not have their own agenda, which does not necessarily align with the dog walker's. Because the human principal anticipates the dog potentially running away (slipping) or lying down in the middle of the street (shirking), they put it on a leash to prevent or at least impede such behaviour. We can think of this leash as a control instrument. The basic PA twin assumption, which views agency slack as endogenous to the A relationship, focusses almost exclusively on the leash. The shorter it is, the less the animal agent will engage in unwanted behaviour. Conversely, when kept on a long leash, the dog will roam around freely, minding its own business. What is overlooked here is that not all dogs are equal, i.e. that beyond their appearance they have different habits and characters. This 'nature of the beast' can be thought of analogously to agents' opportunistic behaviour beyond their formal constraints. What makes the dog tick is independent from the leash and the principal (exogenous to the delegation structure). The length of the leash does not determine it as such¹ but can only constrain its expression. It is thus well possible that a stubborn dog pulls on its short leash and still does not walk neatly at the side of the principal, the same way it is possible that a dog stays right next to the principal not caring for where it could potentially go given its long leash. The overall outcome of the walk, its pleasantness, thus depends on both the leash and the dog.

If we accept the argument that the risk of agency slack is not determined by control alone, which the mundane example above sought to illustrate, the implication is that the inverse relationship between the level of control and the extent of an agent's opportunistic behaviour should be tested rather than taken as given. The central question that arises is thus how these two sources interact. In the sections to come, we investigate these issues empirically, first testing our conjecture that the twin assumptions do not hold across the board before exploring the interaction of exogenous and endogenous factors as sources of agency slack in delegation relationships.

¹ Except possibly after extended periods of time, but this goes beyond the argument.

‘The nature of the beast’: Conceptualizing and Operationalizing IPAs’ Organizational Routines

How can we grasp the agent’s slack potential – the ‘nature of the beast’ – independent of the delegation contract? Essentially, an opportunistic agent’s behaviour reflects activities which are initiated by the agent itself and follow its self-interest rather than the principal’s orders.² To capture the extent to which an IPA will tend to behave more or less opportunistically, we focus on prevalent behavioural routines. These informal patterns capture what makes public administrations ‘tick’ when it comes to their role in initiating, formulating, and implementing the policies of their respective IOs. Such behavioral routines, or administrative styles (2019: 85–86; see also Knill, 2001; Knill et al., 2017), can be defined as “an institutionalized informal *modus operandi* that materializes as a guiding principle over time and by repetitions, routinization, and subsequent internalization”.

They might entail very distinct orientations that are informative of the kind of influence an IPA may exert. Depending on their characteristics, these behavioural patterns “correspond to an ideal typical characterization of IPAs’ ‘styles’ as entrepreneurial or servant-like” (Knill et al., 2019: 85). IPA behavior can be shaped simply by the goal of fulfilling all their principals’ wishes and needs (prevalence of servant routines) or reflect more entrepreneurial orientations characterized by the desire to enhance the policy effectiveness or institutional consolidation and legitimacy (prevalence of entrepreneurial routines). The extent to which an IPA displays a more servant or entrepreneurial orientations has strong implications for its potential of agency slack. While a servant IPA can generally be expected to routinely align its behavior with the preferences of the principal even in the absence of tight control, an entrepreneurial IPA is more likely to slack and to work around even tight political oversight.

The advantage of this approach is that administrative styles grasp IPAs’ behaviour beyond their structural features and institutional characteristics. This allows us to measure IPAs’ opportunism separately and then compare its levels with the extent of control. By contrast, existing PA approaches follow a deterministic logic where the high risk of agency slack is not measured distinctly but is automatically treated as an outcome of low control. Yet, as Knill et al. (2019: 84) put it, “knowing about what IPAs are able to do in principle (their discretion) does not tell us much about what they actually do”. If the described mainstream PA assumptions

² We make no normative judgement as to whether IPAs’ opportunistic behavior is right or wrong. Sometimes, by strengthening own position, IPAs may stay true to their mandate although deviating from political preferences of IO member states.

were correct, we should observe systematic congruence between the level of applied control and IPAs' opportunistic behaviour.

To measure the level of IPAs' exogenous opportunistic behaviour, we draw on six indicators used to measure IPAs' entrepreneurialism (Knill et al., 2019). These indicators cover relevant IPAs' activities during all stages of policy-making, i.e. policy initiation, drafting, and implementation (Table 1). The more actively an IPA is engaged in each of the activity, the more closely it corresponds to the ideal-type of an entrepreneurial style bureaucracy. If an IPA engages in a certain activity on a regular basis, the indicator value is high and coded 1; if its engagement is occasional or uneven across organizational units, the indicator value is medium and coded 0.5; if an IPA's engagement is rare and passive, the indicator value is low and coded 0. For instance, if an IPA actively interacts with third parties (e.g. NGOs) and uses communicative strategies to mobilize their support for its own proposals, the chance that it will leverage its authority and resources regardless of states' preferences increases (see Abbott et al., 2015). Accordingly, the indicator 'Support Mobilization' would be coded 1. Similarly, if an IPA puts significant efforts in regular identification of new policy problems and then actively proposes own solutions, the risk of agency slack increases as the IPA is better able to steer the IO agenda towards own bureaucratic preferences. In such cases, the indicator 'Issue Emergence' is coded 1. Together, the values of six indicators constitute an additive index of an IO bureaucracy's opportunism or, in other words, a kind of behaviour that is likely to deviate from the principal's wishes.

Indicator for IPA activity	Description
<i>Support Mobilization</i>	Mobilization activities to gain support from political or societal actors in policy initiation phase (e.g. interest groups, NGOs, media)
<i>Mapping of Political Space</i>	Use of political facilitation mechanisms in policy initiation phase to identify member states' interests
<i>Issue Emergence</i>	Identification of policy problems and initiation of own (bureaucratic) policy solutions and developments
<i>Political Anticipation</i>	Strategic anticipation of member states' interests to bypass political rejection by member states in policy drafting phase
<i>Strategic Use of Formal Powers</i>	Manipulative use of bureaucratic powers in policy implementation phase to defend bureaucratic interests
<i>Policy Promotion</i>	Engagement in capacity-building and policy promotion of an IPA's programs and agenda

Table 1: indicators and observable implications for measuring IPA's entrepreneurialism as opportunistic behavior.

'The leash': Conceptualizing and Operationalizing IPA Control

The challenge of conceptualizing and measuring control has been approached from different angles, all of which have advantages yet also drawbacks. Bauer and Ege (2016) provided a useful measurement of IPAs' autonomy, which however largely focuses on bureaucratic capacities rather than explicitly addressing the means of control applied by the principals. Hooghe and Marks (2015) measured delegation and pooling to conceive of IO authority, which is however conditioned by IOs' policy portfolio and membership and hence bears some difficulties when trying to systematically compare the extent of agency control across different IOs. Brown (2010) developed a metric for IO delegation, yet covering only few control mechanisms. Nielson and Tierney (2003) described in detail how member states controlled the World Bank, yet their findings remain case-specific. Further studies examined single types of control such as selection of IO staff (Parížek, 2017) or bureaucratic oversight (Grigorescu, 2010).

To overcome these weaknesses, we propose an additive control index which complements existing studies by addressing some of the blind spots mentioned. Importantly, it allows for comparison of member states' control across IOs. Drawing on existing IO research, we differentiate between three main types of control: *oversight and monitoring*, *screening and selection*, and *sanctioning* (Hawkins et al., 2006b; Conceição-Heldt, 2013; Nielson and Tierney, 2003; Heldt, 2017; Pollack, 1997; Graham, 2014; Kiewiet and McCubbins, 1991). We disaggregate these control types into distinct control mechanisms (indicators), which may be either specific and *direct* or broader and *indirect*.

First, to measure *direct oversight and monitoring*, Grigorescu (2010) outlined six mechanisms: inspection, evaluation, investigation and whistle-blower, conflict of interest, and financial disclosure policies. Today, most of these mechanisms already exist in all major IOs. Measures which exhibit no variance across the units of analysis add no analytical leverage. It is also important to consider whether oversight units report to member states (hierarchical monitoring) or to the IO's administrative head (self-monitoring). Accordingly, we include two indicators which provide for variation across IOs and account for reporting lines. First, direct IPA oversight can be conducted through centralized evaluation units, which conduct assessments of

IO performance. The indicator value is coded 1 if evaluation unit reports directly to member states (as in the IMF), 0.6 in case of a dual reporting as in the FAO, or 0.3 when reporting only to the administration head as in the IOM (see also Eckhard and Jankauskas, 2019).³ Second, direct oversight can be conducted through advisory oversight units. They consist of external experts who oversee an IO's general functioning and advise member states (coded 1), both member states and the IPA head (coded 0.6), or only the IPA head (coded 0.3).

To measure *indirect* oversight, we code the frequency of governing body meetings (see Table 2). Although various issues are discussed in these meetings, the intensity of interaction among member states and with the administration fosters information exchange and thus bureaucratic oversight. For instance, if a governing body meets weekly (like NATO's Council) member states' oversight is significantly higher than if such meetings occur only twice a year (like in the IOM).

Second, to measure *direct screening* efforts in IOs, we code formal member states' involvement in appointing an IPA's staff. If the head of the organization appoints his/her team in consultation with member states, member states have higher control (coded 1) than if the IPA head appoints the team unilaterally (coded 0). To measure *indirect screening* in IOs, we use the ratio of (professional) staff nationals of five main state donors relative to the total (professional) staff. The higher the percentage, the higher the indirect control over IO staff by the (biggest) member states.

Finally, to measure member states' *direct sanctioning* capabilities, we consider formal obstacles for them to agree on sanctioning the IPA. Such control is higher in IOs where the voting system in the governing body is weighted (coded 1), because fewer states can make the decision. In contrast, control is coded 0.5 if unweighted majority voting is needed and 0 if voting is consensus-based. Regarding *indirect sanctioning*, we consider an IPA's financial fragility – the extent to which an IPA is dependent on member states' flexible funding, coded as a ratio of earmarked (unassessed) contributions to the total IO budget.

Table 2 summarizes the control indicators. Indicator values range from 0 to 1 and are unweighted to allow comparison across control dimensions.⁴

³ The value is 0, if no evaluation unit exists.

⁴ We use an average value for the two direct oversight indicators to maintain equal weighting across the dimensions.

Type of control	Direct	Indirect
Oversight and Monitoring	<p><i>Evaluation: existence and reporting line</i></p> <p>1: Head of evaluation reports to member states 0.6: Head of evaluation reports to both member states and the head of IPA 0.3: Head of evaluation reports to the head of IPA 0: no evaluation unit</p>	<p><i>Governing body (Executive Board) meetings: intensity</i></p> <p>1: Board meets at least weekly 0.6: Not weekly, but at least three times a year 0.3: Board meets two times a year 0: Board meets once a year</p>
	<p><i>Advisory oversight: existence and reporting line</i></p> <p>1: Head of advisory oversight unit reports to member states 0.6: Head of advisory oversight unit reports to both member states and the head of IPA 0.3: Head of advisory oversight unit reports to the head of IPA 0: no advisory oversight unit</p>	
Screening and Selection	<p><i>IPA staff appointment: member states' involvement</i></p> <p>1: Head of the IPA appoints his/her team in consultation with member states 0: Head of the IPA appoints his/her team unilaterally</p>	<p><i>IPA staff nationality: member states dominance</i></p> <p>Ratio of (professional) staff nationals of five main donor states relative to total (professional) staff, financed by regular budget and other funds</p>
Sanctioning	<p><i>Formal obstacles for sanctioning: voting system in Executive Boards</i></p> <p>1: weighted voting system 0.5: unweighted majority voting 0: consensus</p>	<p><i>Threat of sanctions: financial fragility</i></p> <p>Ratio of earmarked (unassessed) contributions to total budget</p>

Table 2: Conceptualization and operationalization of control over an IPA.

Empirical Results and Discussion

This section presents descriptive results for principal control and IPAs' opportunism in eight IOs: the Food and Agriculture Organization of the United Nations (FAO), the International Labour Organization (ILO), the International Monetary Fund (IMF), the International

Organization for Migration (IOM), the United Nations Environment Programme (UNEP), the United Nations High Commissioner for Refugees (UNHCR), the North Atlantic Treaty Organization (NATO), and the World Health Organization (WHO). Our case selection was strongly confined by the availability of qualitative data on IPAs' administrative behaviour and we are aware that they differ widely in their global reach, issue area, principal composition, membership in the UN and the like. Given that we make a descriptive, rather than causal argument, however, we do not consider this a problem.

In terms of data collection, for the measurement of IPAs' opportunistic behaviour, we draw on 131 semi-structured expert interviews, conducted with IO staff over the last five years. Each interview lasted about 45 minutes, ranging from minimum 11 to maximum 23 interviews per IO from the sample. Interviews loosely followed the list of indicators presented above, were recorded and transcribed. While only officials in IO headquarters were interviewed, different IO departments and three hierarchical levels were covered.⁵ For more information on interview data see Appendix I. For the measurement of IPA control indicators, we relied on primary IO documents (e.g. annual reports, budget reports or staff rules), data retrieved from IO websites, and official statistics provided by the UN System Chief Executives Board for Coordination (CEB). Most recent data were collected, ranging mostly from 2016 to 2018, depending on document availability. In line with our argument, which for now does not make conjectures about temporal dynamics, the analysis thus captures the last five years and remains static.

Turning to the descriptive results, Figure 2 presents the varying levels of member states' control across eight IO administrations in descending order. If we compare the results across the IOs, the highest level of control is found in the IMF, where the administration faces strictest direct and indirect control mechanisms. The IOM and ILO also score high on the IPA control index, whereas FAO and WHO are positioned in the middle. Member states' control is lowest for NATO, UNEP, and UNHCR.

⁵ As research on administrative behaviour at the national level indicates, bureaucratic staff in the headquarters are better informed about the overall organizational functioning and are still involved in the planning of field operations (Knill (2001))

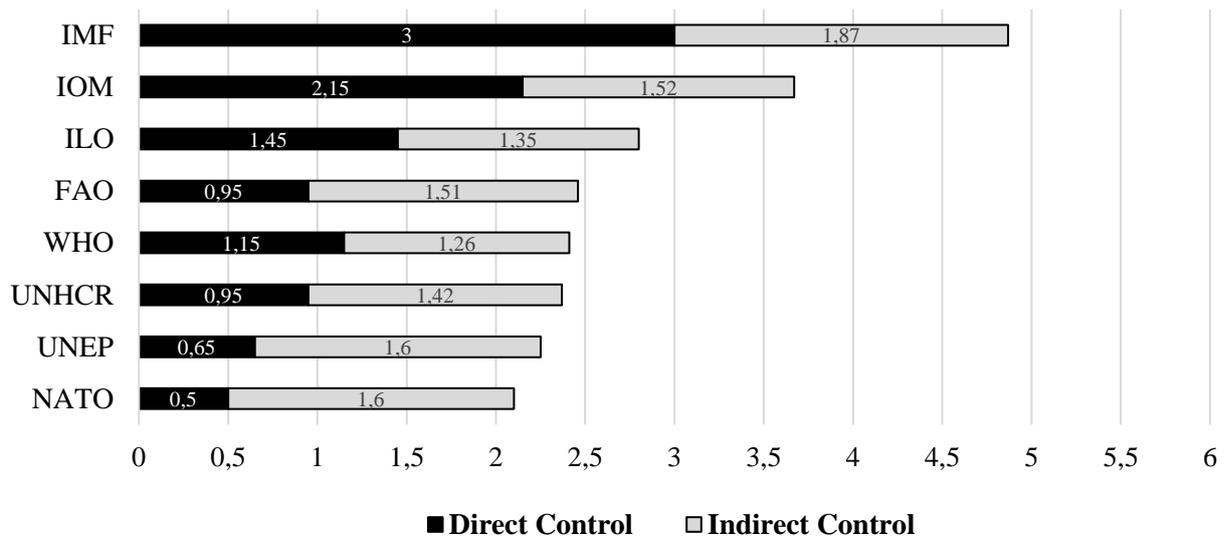


Figure 2: Aggregated values of member states' control over IPAs.

Second, Figure 3 presents the empirical results of IPAs' opportunistic behaviour measurement. Low values correspond with a servant administrative style and thus lower opportunism, whereas higher values indicate a more entrepreneurial style and thus propensity to opportunism. In the sample, UNHCR and the IMF demonstrate highest values. Indeed, both organizations actively engage in entrepreneurial activities, despite their structural differences and issue fields. UNEP, FAO, IOM, and WHO take middle positions, whereas NATO and the ILO score significantly lower. The two organizations also differ in their activities and institutional design but were both found to exercise low opportunism throughout the policy cycle (ibid.).

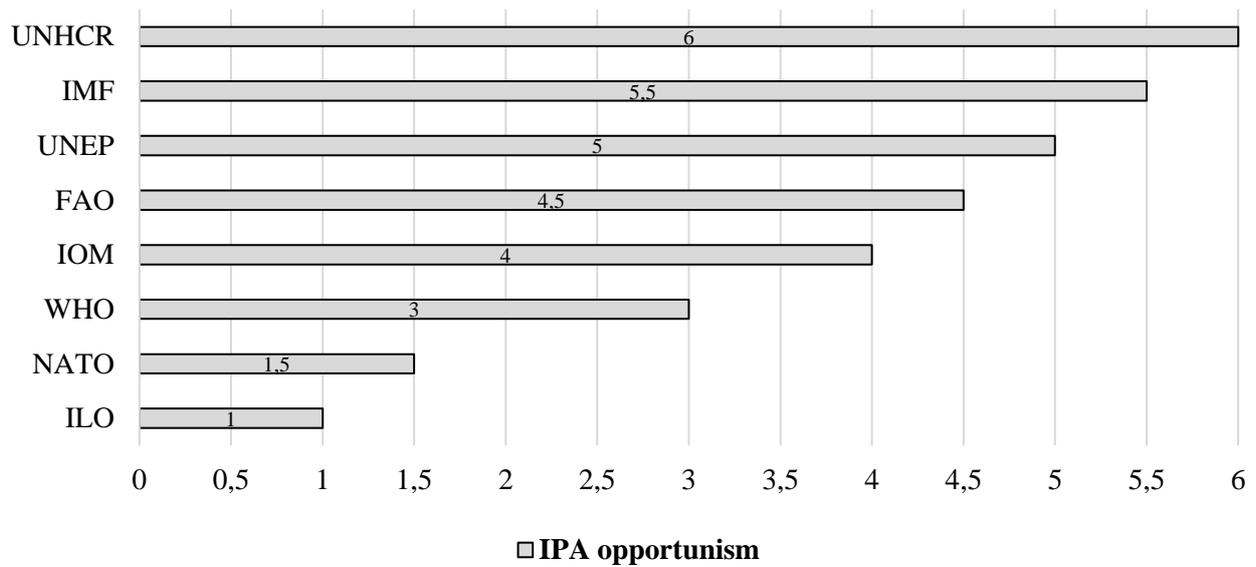


Figure 3: Aggregated values for IPAs' opportunistic behaviour based on their entrepreneurialism.

Having assessed the exogenous and endogenous sources of agency slack separately, to what extent these results support the twin assumptions derived from the basic PA model? A typical case for an IO bureaucracy according to PA theory would be that of high control and therefore constrained opportunism *or* low control and thus enhanced opportunistic behaviour. We do find such typical cases, namely the ILO, which is controlled relatively tightly and exhibits the least values of opportunism among the IOs, as well as UNHCR, whose control is comparatively low while the opportunistic behaviour is the highest among the IOs. It seems that there are indeed cases that do not fall neatly with conventional PA model predictions. However, we also find cases that do not fall neatly with conventional PA model predictions. For instance, NATO secretariat enjoys low control and yet remain servant to member states, whereas the IMF administration is stringently controlled yet still maintains opportunism.

Based on these findings, how the exogenous and endogenous sources of agency slack interact? We have seen variance in the length of the leash, i.e., the strictness of control principals impose on their agents, as well as in how much the self-interested dog will pull the leash, i.e. how opportunistic the agent behaves. As these two factors do not always align, and are theoretically independent, it is possible to cross-tabulate the two dimensions and conceive of four scenarios of how exogenous and endogenous sources of agency slack interact (Table 3). Expected risk of agency slack should depend on the extent to which an IPA agent will pull the leash, i.e. engage in opportunistic behaviour, and on the length of the leash, i.e. the extent to which control

arrangements will mediate opportunistic behaviour. For each of the four scenarios, an IO that exemplifies the respective configuration is discussed in more detail.

		Exogenous sources of agency slack	
		<i>Opportunism low</i>	<i>Opportunism high</i>
Endogenous sources of agency slack	<i>Control low</i>	NATO (Tame dog on a long leash) → little agency slack	UNHCR (Wild dog on a long leash) → substantial agency slack
	<i>Control high</i>	ILO (Tame dog on a short leash) → no agency slack	IMF (Wild dog on a short leash) → occasional agency slack

Table 3: Endogenous and exogenous sources of expected agency slack.

Beginning with a typical case in terms of conventional PA expectations, high control over the IO administration is combined with weak manifestation of opportunistic behaviour (*tame dog on a short leash*). In this case, tight control decreases the possibility space for the IPA to successfully engage in slack even if it wanted to (which it does not, given its low opportunism). The occurrence of agency slack is thus unlikely. Both endogenous and exogenous sources for agency slack mitigate its potential. Consider the ILO – it is a relatively high controlled organization, scoring third in our sample with well-balanced direct and indirect control mechanisms in place. Its centralized evaluation unit reports to the Director-General (ILO, 2017: 42), although its Independent Oversight Advisory Committee reports to both the Director-General and the Governing Body (direct oversight) (ILO, 2019b). The latter meets three times a year (indirect oversight) (ILO, 2019a). Member states have to be consulted for the appointments of Deputy and Assistant Director-Generals and some other high-level IPA positions (direct screening) (ILO, 2018b). Moreover, 33.2 percent of total professional ILO staff are nationals of five main donor states,⁶ thus allowing for indirect screening. Finally, voting procedures in the Governing Body are usually consensus-based (direct sanctioning) (ILO, 2016) and 42.8 percent of the total budget is earmarked voluntary contributions (indirect sanctioning) (ILO, 2018a).

⁶ Own calculation based on UN CEB data.

At the same time, the ILO scores lowest on the extent of opportunistic behaviour patterns and resembles a servant-like administration. First, the IPA rarely builds strategic coalitions with civil society or else to promote own issues or initiatives (support mobilization) (ILO 7, 11, 13).⁷ Second, its efforts to map member states' interests are "rather decentred" (ILO 2) and lack a comprehensive strategy. Third, issue emergence usually occurs outside the bureaucracy, based on the demand by the constituents (ILO 5), so the administration usually "react[s] to things" (ILO 4, 10). Fourth, political anticipation in the ILO is indeed pronounced, but as a requirement in assisting member states rather than as a strategic move by the administration (ILO 4, 6, 13). Fifth, the strategic use of formal bureaucratic powers occurs only rarely. The IPA staff described itself as a tool or a "platform" for the cooperation of others (ILO 3, 12), avoiding any opportunistic interpretation of its formal capabilities. Finally, the administration's efforts to promote itself and its policies remain low: the ILO "play[s] it safe" (ILO 3, 12) and follows the lead given by its constituents. Hence, its overall servant-like administrative behaviour as well as high control by member states make agency slack unlikely.

A similar reinforcement of the two factors can be expected for the second scenario (*wild dog on a long leash*). Here, as expected by P-A theory, low control is accompanied by strong opportunism, which makes the risk of agency slack substantial. Not only do loose control arrangements allow the agent to potentially 'run away', the agent has a strong disposition to actually do so. UNHCR provides an illustrative example. On the one hand, it scores relatively low in terms of control. Its evaluation unit reports to the High Commissioner (UNCHR, 2016), whereas its Independent Audit and Oversight Committee assists both the High Commissioner and the Executive Committee (direct oversight) (UNHCR, 2019a). Its Standing Committee meets three times a year (indirect oversight) (UNHCR, 2019c). In comparison to the ILO, UNHCR member states have much less control over selection of the UNHCR staff: the High Commissioner appoints his/her team unilaterally (UNHCR, 1950) and only 17.3 percent⁸ of total (professional) staff comes from five main donor states. Finally, decision making in the UNHCR governing body is usually based on majority voting (direct sanctioning) (UNHCR, 2016) and 65 percent of its total budget comes from earmarked voluntary contributions (indirect sanctioning) (UNHCR, 2019b).

On the other hand, the UNHCR administration demonstrates high opportunism, scoring first in our sample. Regarding support mobilization, UNHCR is very active in its engagement with

⁷ Interview insights are quoted using labels.

⁸ Own calculation based on UN CEB data.

partners, mostly regional NGOs and private donors (UNHCR 3, 6, 7, 10). As one staff official put it, this allows “to have one voice talking to states” when initiating new topics (UNHCR 5). UNCHR is also very active in mapping the political space. It tends to strategically “keep a low profile” until member states’ support is guaranteed (UNHCR 5, 13); or it frames issues in a way that ensures member states’ support based on IPA’s interests (UNHCR 2). The identification of new issues is strongly driven from within the organization, with innovative strategies being applied by the administration (issue monitoring, field analysis, etc.) (UNHCR 1, 4, 10). Regarding political anticipation, the IPA actively seeks to maintain close contact with important member states while drafting a new policy (UNCHR 13) and sends its staff to international meetings to remain aware of “different sensitivities” among membership (UNHCR 4). Also, UNHCR demonstrates a strategic use of its bureaucratic powers, often resorting to behind-the-scenes diplomacy not refraining from conflicts with member states (UNHCR 1, 10, 11, 13). Finally, the IPA displays an active promotion of its own policies and initiatives: both through its direct communication with member states and public promotion with the larger public (UNHCR 3, 9). In sum, substantial agency slack becomes highly likely considering that the administration demonstrates entrepreneurial behaviour, the opportunism of which not mediated by strict control mechanisms.

The third scenario (*wild dog on a short leash*) resembles a combination of high opportunism and high control, thus deviating from a traditionally assumed inverse linear relationship between the two. Despite strict control by the principal, the agent actively strengthens itself beyond its formal constraints. The delegation relationship is likely contentious, and the agent should only occasionally engage in slack successfully (which does not mean it is not trying). The IMF secretariat provides a good illustration. Although it is strictly controlled by the member states, it *remains* highly opportunistic. The IMF scores first on the IPA control index, being way above the IOM. Its centralized evaluation unit reports directly to the Executive Board (IMF, 2015) and its External Audit Committee, responsible for the oversight of internal control functions, also reports to member states through the Board of Governors (direct oversight) (IMF, 2018). Its in-house Executive Board meets several times a week (indirect oversight) (IMF, 2019b). Regarding staff selection, the Managing Director has to consult with member states prior to appointing or dismissing any official equal or above a division chief (IMF, 2019c). Also, more than a third (35.7 percent) of total IMF professional and managerial staff (A9-B5) are nationals of five main donor states (IMF, 2017). Although sometimes votes are held in the Board, the usual way of decision-making is consensus-based (direct sanctioning)

(IMF, 2019a). Finally, 51.2 percent of the total budget comes from voluntary member states' contributions (IMF, 2019d).

Nevertheless, the IMF administration also exerts high levels of opportunism. The administration has an "active communication strategy" and continuously engages with external actors to mobilize support (IMF 2). Also, IMF managers are regularly involved in mapping member states' interests (IMF 2). Regarding issue emergence, the administration monitors global events and develops *own* policy proposals (IMF 1, 2). In policy drafting stage, the IMF actively anticipates the political environment to avoid deadlocks both by engaging with member states and in country surveillance or program work. As one official told us, the managers must know "what can fly and what cannot" (IMF 2). Furthermore, the bureaucracy is highly strategic in using its formal powers (IMF 6). For instance, officials sometimes deliberately "deliver papers about major policy issues to the Board later than they should", so the Board is forced to follow the management approach in emergencies (IMF 3). Lastly, the administration actively works on its high profile for publications and policy recommendations to enable policy promotion across issue fields and topics (IMF 7, 15). Overall, the observed behavioural patterns driven by administrative self-interest maintain despite high control by the member states.

In the fourth scenario (*tame dog on a long leash*), we observe the exact opposite of the third. Although the agent has formal leeway for independent action and could potentially exploit it, its low opportunism and 'servant' nature leads to expect little agency slack. This pattern, also deviant to P-A expectations, applies to the NATO secretariat. NATO demonstrates the lowest levels of control, with only a few (and mostly indirect) control mechanisms in place. It has no centralized evaluation unit, only its Resource Policy and Planning Board advises member states in the North Atlantic Council on resource management and performance assessment (direct oversight) (NATO, 2016b). The Council meets at least once a week (indirect oversight) (NATO, 2017). As for screening and selection, the Secretary-General appoints his/her team unilaterally (NATO, 2016a) and 48 percent of NATO's international civilian staff comes from five main donor states. The decision-making process in the Council is based on consensus (direct sanctioning) and about 12 percent of NATO funding comes from voluntary contributions (indirect sanctioning).⁹

Considering the above, PA theory would expect the NATO administration to use its discretion and exploit informational asymmetries for own purposes. However, as our data demonstrates, NATO follows a pattern of a servant-type administration. First, there is generally low strategic

⁹ Own estimation for the year 2010 due to data availability.

engagement with external partners, except for Public Diplomacy Division, which mainly works with think tanks and universities (NATO 10). Admittedly, the NATO bureaucracy puts significant efforts in mapping member states' interests in policy initiation phase and when drafting new policy proposals (NATO 1, 3, 6, 7). Nevertheless, new issues and topics rarely emerge within the NATO's bureaucracy. Instead, "it is first and foremost nations that bring forward requests" (NATO 1, 8). Similarly, the administration does not engage in bureaucratic venture employing its formal powers and implementation of member states' decisions are an automatism (NATO 1). Finally, no active policy promotion was observed, which goes in line with characteristics of a servant-like administration despite low control.

Overall, the revealed patterns provide support for our argument that the risk of agency slack is not a mere by-product of (missing) control. Recalling the two questions raised in the introduction, the findings demonstrate that agents which are controlled less do not necessarily show more opportunistic behaviour (NATO case). Also, more control by the principal does not necessarily lead to less opportunism by the agent (the IMF case). Instead, the risk of agency slack includes both the endogenous (lack of) control and the exogenous patterns of opportunistic behaviour, which can both reinforce and mitigate each other.

Conclusion

This paper departed from two basic assumptions of PA theory. Assuming a linear relationship between control and agency slack, agents are expected to slack if they are allowed to and principal control is assumed to reduce agents' opportunism and therefore the risk of agency slack. We have argued that these assumptions only hold if agent characteristics exogenous to the model are excluded. Once we consider agents' potential to engage in slack as a function of their (informal) behaviour patterns instead of treating it as the mere negative of control, we are left with two sources of agency slack, both of which may vary across IOs: (endogenous) principal control and (exogenous) opportunistic behaviour. Our empirical analysis has shown that these two factors do in fact not always align. It seems that agents do not always seek to slack, and control cannot account for the likelihood of agency slack on its own. Hence, depending on the prevalence of opportunistic behavioural orientations of the agent, trusting rather than controlling the agent might be a cheaper and equally effective approach for the principal.

These findings have profound implications on PA theory as well as for research on IO delegation. First, we should abandon the dominant view on IPAs as agents with a natural predisposition towards slacking if principals let them. IPAs vary in their behavioural patterns which are not determined merely by the level of control applied by member states. Guessing a dog's behaviour by looking at its leash does not do the trick. Second and related, determining an ideal level of control is not possible without considering the exogenous nature of the agent. In fact, depending on the prevalence of agent's opportunistic behavioural orientations, trusting rather than controlling the agent might be a cheaper and equally effective approach for the principal. Hence, an IPA's trustworthiness rather than its formal discretion appears as a key concern for member states. Knowing how trustworthy the agent is could help solving classical P-A dilemmas such as cost-benefit and control-competence trade-offs in delegation relationships (Abbott et al., 2020a; Hawkins et al., 2006a). For instance, having a 'tame' IPA on a short 'leash' is not optimal both for member states and the IPA since control is costly and reduces its competence (see Honig, 2019), especially considering that a longer 'leash' on a 'tame' IPA would not hurt the principal. IO scholars should thus refrain from focusing merely on the principal side and put more attention on relationship dynamics between member states and IPAs beyond deterministic structural explanations (see also Lall, 2017).

The paper constitutes a first medium-n attempt to test classical PA assumptions and as such does not come without limitations. For one thing, our analysis remains static as we focused on variation across IOs rather than changes over time. This decision was motivated, on the one hand, by the PA assumption that principal-agent relations tend towards stability and, on the other hand, the fact that administrative styles are known as "relatively stable behavioral orientations" (Knill et al., 2019: 85) and that formal controls in IOs are oftentimes difficult to change. Also, we have for the sake of clarity oversimplified the role of the principals as these were treated as unitary actors, which is something that is rather rare in the world of IOs. Moreover, we did not control for alternative explanations of varying control levels such as principals' inability to use existing control mechanisms, even if this seem rather unlikely. These aspects should thus be considered in future studies which could scrutinize our argument in a more in-depth case study analysis, potentially allowing for variation in formal discretion and administrative routines over long periods of time. Finally, future work could shift from studying agency slack potential to testing our assumptions in terms of outcomes.

References

- Abbott KW, Genschel P, Snidal D, et al. (eds) (2015) *International Organizations as Orchestrators*. Cambridge, UK: Cambridge University Press.
- Abbott KW, Genschel P, Snidal D, et al. (2016) Two Logics of Indirect Governance: Delegation and Orchestration. *British Journal of Political Science* 46(4): 719–729.
- Abbott KW, Genschel P, Snidal D, et al. (2020) Competence versus control: The governor's dilemma. *Regulation & Governance* 14(4): 619–636.
- Bauer MW and Ege J (2016) Bureaucratic autonomy of international organizations' secretariats. *Journal of European Public Policy* 23(7): 1019–1037.
- Bayerlein L, Knill C and Steinebach Y (2020) *A Matter of Style: Organizational Agency in Global Public Policy*. Cambridge University Press.
- Biermann F and Siebenhüner B (eds) (2009) *Managers of global change: The influence of international environmental bureaucracies*. Cambridge MA: MIT Press.
- Brown RL (2010) Measuring Delegation. *The Review of International Organizations* 5(2): 141–175.
- Conceição-Heldt E da (2013) Do Agents “Run Amok”? A Comparison of Agency Slack in the EU and US Trade Policy in the Doha Round. *Journal of Comparative Policy Analysis: Research and Practice* 15(1): 21–36.
- Cortell AP and Peterson S (2006) Dutiful agents, rogue actors, or both? Staffing, voting rules, and slack in the WHO and WTO. In: Hawkins DG, Lake DA, Nielson DL and Tierney MJ (eds) *Delegation and Agency in International Organizations*: Cambridge: Cambridge University Press, pp. 255–280.
- Eckhard S and Jankauskas V (2019) The politics of evaluation in international organizations: A comparative study of stakeholder influence potential. *Evaluation* 25(1): 62–79.
- Graham ER (2014) International Organizations as Collective Agents: Fragmentation and the Limits of Principal Control at the World Health Organization. *European Journal of International Relations* 20(2): 366–390.
- Grigorescu A (2010) The Spread of Bureaucratic Oversight Mechanisms across Intergovernmental Organizations. *International Studies Quarterly* 54(3): 871–886.
- Hawkins D, Lake DA, Nielson DL, et al. (2006a) Delegation under anarchy: states, international organizations, and principal-agent theory. In: Hawkins DG, Lake DA, Nielson DL and Tierney MJ (eds) *Delegation and Agency in International Organizations*: Cambridge: Cambridge University Press, pp. 3–38.
- Hawkins DG and Jacoby W (2006) How agents matter. In: Hawkins DG, Lake DA, Nielson DL and Tierney MJ (eds) *Delegation and Agency in International Organizations*: Cambridge: Cambridge University Press, pp. 199–228.

- Hawkins DG, Lake DA, Nielson DL, et al. (eds) (2006b) *Delegation and Agency in International Organizations*. Cambridge: Cambridge University Press.
- Heldt E (2017) Regaining control of errant agents? Agency slack at the European Commission and the World Health Organization. *Cooperation and Conflict* 52(4): 469–484.
- Honig D (2019) When Reporting Undermines Performance: The Costs of Politically Constrained Organizational Autonomy in Foreign Aid Implementation. *International Organization* 73(1): 171–201.
- Hooghe L and Marks G (2015) Delegation and pooling in international organizations. *The Review of International Organizations* 10(3): 305–328.
- Huber JD and Shipan CR (2002) *Deliberate Discretion? The Institutional Foundations of Bureaucratic Autonomy*. Cambridge University Press.
- ILO (2016) *Compendium of rules applicable to the Governing Body of the International Labour Office*.
- ILO (2017) ILO Evaluation Policy: GB.331/PFA/8.
- ILO (2018a) Core Voluntary Funding (RBSA) for ILO development cooperation. Available at: https://www.ilo.org/wcmsp5/groups/public/---dgreports/---exrel/documents/publication/wcms_317826.pdf (accessed 16 July 2019).
- ILO (2018b) *Staff Regulations*.
- ILO (2019a) Governing Body. Available at: <https://www.ilo.org/gb/lang--en/index.htm> (accessed 16 July 2019).
- ILO (2019b) Independent Oversight Advisory Committee. Available at: <https://www.ilo.org/public/english/edmas/ioac/index.htm> (accessed 16 July 2019).
- IMF (2015) Terms of Reference for the Independent Evaluation Office of the International Monetary Fund.
- IMF (2017) Distribution of professional and managerial staff by nationality. Available at: <https://www.imf.org/external/pubs/ft/ar/2017/eng/assets/ar17-web-table-3-1.pdf> (accessed 16 July 2019).
- IMF (2018) The External Audit Committee (EAC). Available at: <https://www.imf.org/external/np/eac/eng/eac.htm> (accessed 16 July 2019).
- IMF (2019a) How the IMF Makes Decisions. Available at: <https://www.imf.org/en/About/Factsheets/Sheets/2016/07/27/15/24/How-the-IMF-Makes-Decisions> (accessed 16 July 2019).
- IMF (2019b) IMF Executive Directors and Voting Power. Available at: <https://www.imf.org/external/np/sec/memdir/eds.aspx> (accessed 16 July 2019).
- IMF (2019c) Rules and Regulations of the International Monetary Fund: N--Staff Regulations. Available at: <https://www.imf.org/external/pubs/ft/bl/rr14.htm> (accessed 16 July 2019).

- IMF (2019d) Where the IMF Gets Its Money. Available at: <https://www.imf.org/en/About/Factsheets/Where-the-IMF-Gets-Its-Money> (accessed 16 July 2019).
- Kassim H and Menon A (2003) The principal-agent approach and the study of the European Union: Promise unfulfilled? *Journal of European Public Policy* 10(1): 121–139.
- Kiewiet DR and McCubbins MD (1991) *The Logic of Delegation: Congressional Parties and the Appropriations Process*. Chicago: University of Chicago Press.
- Knill C (2001) *The Europeanisation of national administrations: Patterns of institutional change and persistence*. Cambridge: Cambridge University Press.
- Knill C, Bayerlein L, Enkler J, et al. (2019) Bureaucratic influence and administrative styles in international organizations. *Review of International Organizations* 14(1): 83–106.
- Knill C, Eckhard S and Grohs S (2016) Administrative styles in the European Commission and the OSCE Secretariat: Striking similarities despite different organizational settings. *Journal of European Public Policy* 23(7): 1057–1076.
- Knill C, Enkler J, Schmidt S, et al. (2017) Administrative Styles of International Organizations: Can we find them, do they matter? In: Bauer MW, Knill C and Eckhard S (eds) *International Bureaucracy: Challenges and Lessons for Public Administration Research*. London: Palgrave Macmillan UK.
- Lake DA and McCubbins MD (2006) The logic of delegation to international organizations. In: Hawkins DG, Lake DA, Nielson DL and Tierney MJ (eds) *Delegation and Agency in International Organizations*: Cambridge: Cambridge University Press, pp. 341–368.
- Lall R (2017) Beyond Institutional Design: Explaining the Performance of International Organizations. *International Organization* 71(2): 245–280.
- Martin LL (2006) Distribution, information, and delegation to international organizations: The case of IMF conditionality. In: Hawkins DG, Lake DA, Nielson DL and Tierney MJ (eds) *Delegation and Agency in International Organizations*: Cambridge: Cambridge University Press, pp. 140–164.
- Miller GJ (2005) The Political Evolution of Principal-Agent Models. *Annual Review of Political Science* 8(1): 203–225.
- NATO (2016a) NATO Secretary General. Available at: https://www.nato.int/cps/en/natolive/topics_50094.htm (accessed 16 July 2019).
- NATO (2016b) Resource Policy and Planning Board. Available at: https://www.nato.int/cps/en/natohq/topics_67653.htm (accessed 16 July 2019).
- NATO (2017) North Atlantic Council. Available at: https://www.nato.int/cps/ic/natohq/topics_49763.htm (accessed 16 July 2019).
- Nielson DL and Tierney MJ (2003) Delegation to International Organizations: Agency Theory and World Bank Environmental Reform. *International Organization* 57(2): 241–276.

- Parížek M (2017) Control, soft information, and the politics of international organizations staffing. *The Review of International Organizations* 12(4): 559–583.
- Pollack MA (1997) Delegation, agency, and agenda setting in the European Community. *International Organization* 51(1): 99–134.
- Pollack MA (2003) *The Engines of European Integration*. Oxford University Press.
- Tamm H and Snidal D (2014) Rational Choice and Principal-Agent Theory. In: Weiss TG and Wilkinson R (eds) *International organization and global governance*: London: Routledge, pp. 132–143.
- UNCHR (2016) UNHCR Policy on Evaluation: UNHCR/HCP/2016/2.
- UNHCR (1950) Statute of the Office of the United Nations High Commissioner for Refugees: General Assembly Resolution 428 (V) of 14 December 1950.
- UNHCR (2016) Rules of procedure: A/AC.96/187/Rev.8. Available at: <https://www.unhcr.org/4201eb512> (accessed 16 July 2019).
- UNHCR (2019a) Applications for Independent Audit and Oversight Committee (IAOC). Available at: <https://www.unhcr.org/5a7337d44> (accessed 16 July 2019).
- UNHCR (2019b) Global Appeal: 2019 Update.
- UNHCR (2019c) Standing Committee meetings. Available at: <https://www.unhcr.org/standing-committee-meetings.html> (accessed 16 July 2019).
- Waterman RW and Meier KJ (1998) Principal-Agent Models: An Expansion? *Journal of Public Admin Research and Theory* 8(2): 173–202.
- Yi-chong X and Weller P (2008) 'To be, but not to be seen': Exploring the Impact of International Civil Servants. *Public Administration* 86(1): 35–51.

Appendix to Paper D

Appendix I: Interview data

IO	Senior managers	Middle managers	Analysts, technical officers	Total
NATO	2	5	5	12
FAO	3	1	19	23
ILO	1	3	12	16
UNEP	3	4	15	22
IOM	1	3	12	16
UNHCR	0	4	12	16
IMF	1	6	8	15
WHO	4	4	3	11
			Total:	131

Table 1.A: Overview of conducted interviews across IOs and interviewee positions.

Please note that ‘senior managers’ category refers to officials at the level of Secretary- or Director-General, directors, their deputies and similar. ‘Middle managers’ category refers to team leaders, heads of units, and similar. ‘Analysts and technical officers’ category refers to various specialists at the technical level of administrations such as policy specialists or technical officials.

The interviewees were asked about internal operating procedures in their IOs along the introduced six indicators (see Table 1 in the main article). To determine the intensity to which the administration is involved in the six activities, we assessed the consistency of interviewees’ statements. In rare cases of conflicting statements, we triangulated them with secondary scholarly literature to decide upon the value of the indicator.

** Please note that interview data used in Paper D utilizes data compiled by a larger project on IPAs’ administrative styles (see Bayerlein et al. (2020b); Knill et al. (2016); Knill et al. (2019)).*