

Political Decisions, Economic Effects:
How International Organizations Govern the Globalized Economy

Valentin F. Lang

Dissertation
zur Erlangung des wirtschafts- und sozialwissenschaftlichen Doktorgrades
„Doctor rerum politicarum“
der Ruprecht-Karls-Universität Heidelberg

Heidelberg

Mai 2018

BETREUER UND ERSTGUTACHTER DER DISSERTATION

Prof. Dr. Axel Dreher

Professor of Economics

Chair of International and Development Politics

Alfred-Weber-Institute for Economics

Ruprecht-Karls Universität Heidelberg

ZWEITGUTACHTER DER DISSERTATION

Prof. Christopher Kilby, PhD

Professor of Economics

Villanova School of Business

Villanova University

Political Decisions, Economic Effects:

How International Organizations Govern the Globalized Economy

Valentin F. Lang

Für Sabine und Achim

L'économie politique est bien une science, c'est bien un type de savoir, c'est bien un mode de connaissance dont il faudra que ceux qui gouvernent tiennent compte. Mais la science économique ne peut pas être la science du gouvernement et le gouvernement ne peut pas avoir pour principe, loi, règle de conduite ou rationalité interne, l'économie.

Michel Foucault

Table of Contents

List of Tables	IX
List of Figures	XI
Introduction	1
Theoretical Approach	6
Empirical Approach	10
Policy Areas and Policy Implications	13
Structure and Summary of the Dissertation	15
1 Buying Votes and International Organizations: The Dirty-Work Hypothesis	19
1.1 Introduction	20
1.2 The Argument	26
1.2.1 Bilateral and Multilateral Aid	26
1.2.2 The Dirty-Work Hypothesis	27
1.2.3 Contributions to Previous Literature	30
1.3 The IMF and the UNSC	34
1.4 Data and Method	37
1.4.1 A New Dataset on UNSC Voting Behavior	37
1.4.2 Empirical Model and Additional Data	40
1.5 Results	44
1.5.1 UNSC Voting and Aid Allocation	44
1.5.2 Main Results	49
1.5.3 Extensions and Robustness Tests	54
1.6 Conclusion	59
1.7 Appendices to Chapter 1	62
1.7.1 Appendix 1.A: Coding of Resolution-specific UNSC Voting Data	62
1.7.2 Appendix 1.B: Descriptive Statistics	65
1.7.3 Appendix 1.C: Data Sources and Definitions	66
2 Room for Discretion: Biased Decision-Making in International Financial Institutions	69
2.1 Introduction	70
2.2 The Debt Sustainability Framework and the Risk Ratings	76
2.3 Data and Method	78
2.3.1 Main Variables and Descriptive Evidence	78
2.3.2 Econometric Model	82
2.4 Results	86
2.4.1 Macroeconomic Determinants	86

2.4.2	Biases.....	89
2.4.3	Alternative Channels	92
2.4.4	Robustness.....	96
2.4.5	Extension: Testing Manipulation of Projections: Density Discontinuity	102
2.5	Conclusion	104
2.6	Appendices to Chapter 2.....	105
2.6.1	Appendix 2.A: Relevance of the DSF.....	105
2.6.2	Appendix 2.B: Variables.....	106
3	Stigma or Cushion? IMF Programs and Sovereign Creditworthiness	109
3.1	Introduction	110
3.2	Potential Channels and Existing Literature.....	113
3.2.1	Adjustment Effects.....	113
3.2.2	Signaling Effects	114
3.2.3	Issues with the Existing Evidence.....	115
3.3	Data and Identification.....	117
3.3.1	Dependent Variable: Sovereign Credit Ratings	117
3.3.2	Treatment Variable.....	119
3.3.3	Endogenous Selection into IMF Programs	119
3.3.4	Identification	123
3.4	Results.....	127
3.4.1	Baseline	127
3.4.2	Other Rating Agencies and Institutional Investor Assessments	130
3.4.3	Robustness.....	132
3.4.4	An Alternative Identification Strategy	134
3.4.5	Channels: Adjustment and Signaling.....	136
3.4.6	Qualitative Evidence.....	142
3.5	Conclusion	144
3.6	Appendices to Chapter 3.....	146
3.6.1	Appendix 3.A: Construction of the Sovereign Ratings Database.....	146
3.6.2	Appendix 3.B: Translation of Sovereign Ratings into Numerical Values	147
3.6.3	Appendix 3.C: Institutional Investor Data.....	148
3.6.4	Appendix 3.D: Control Variables.....	151
3.6.5	Appendix 3.E: Variables.....	152
3.6.6	Appendix 3.F: Results of Robustness Regressions Described in Section 3.4.3 ...	153
3.6.7	Appendix 3.G: Event-Based Identification	155
3.6.8	Appendix 3.H: Exploratory Analysis of Statements by Rating Agencies	156
3.6.9	Appendix 3.I: Systematic Analysis of Statements by Rating Agencies	162

4	The Economics of the Democratic Deficit: The Effect of IMF Programs on Inequality...	167
4.1	Introduction	168
4.2	The Argument	170
4.2.1	The IMF and the Democratic Deficit	170
4.2.2	IMF Conditionality and Inequality	174
4.2.3	IMF Programs and Domestic Democracy	178
4.2.4	Heterogeneous Effects and Channels.....	180
4.3	Method and Data	181
4.3.1	Endogeneity	181
4.3.2	Identification Strategy	184
4.3.3	Econometric Model and Data	186
4.4	Results.....	188
4.4.1	First-Stage Estimates.....	188
4.4.2	Main Results.....	189
4.4.3	Heterogeneous Effects	191
4.4.4	Robustness.....	195
4.5	Conclusions.....	196
4.6	Appendices to Chapter 4.....	198
4.6.1	Appendix 4.A: Variables	198
4.6.2	Appendix 4.B: Interpreting Differences in Gini Coefficients	200
4.6.3	Appendix 4.C: Full Table 4.1	201
4.6.4	Appendix 4.D: Full Table 4.2 and Discussion of Control Variable Results.....	202
4.6.5	Appendix 4.E: Full Table on Long-Term Effects.....	204
4.6.6	Appendix 4.F: IMF Conditionality.....	204
4.6.7	Appendix 4.G: Robustness.....	207
	References	217

List of Tables

Table 1.1 – UNSC Voting and Aid, OLS, 1960-2015	48
Table 1.2 – UNSC Voting and Aid to Friends and Enemies, OLS, 1960-2015.....	50
Table 1.3 – Extensions and Tests for Robustness.....	56
Table 1.4 – Who Controls the IMF?	58
Table 1.5 – Frequency of Words in UNSC Resolution Titles (100 most frequent)	64
Table 1.6 – Descriptive Statistics.....	65
Table 1.7 – Data Sources and Definitions	66
 Table 2.1 – LIC-DSA Mechanical and Actual Risk Ratings.....	 79
Table 2.2 – Overruling and US Friends.....	81
Table 2.3 – Overruling and Avoiding Downgrades.....	82
Table 2.4 – DSA-specific Variables and Macroeconomic Fundamentals	87
Table 2.5 – The Role of Political Interests and Bureaucratic Incentives	90
Table 2.6 – Testing for Alternative Channels	95
Table 2.7 – Robustness: Additional Control Variables	99
Table 2.8 – Robustness: Different Samples and Timing of Elections	100
Table 2.9 – Robustness: Alternative Definitions of the Political Proximity Variable	101
Table 2.10 – Manipulation Test	102
Table 2.11 – Effect of DSF Risk Ratings on World Bank Lending	105
Table 2.12 – Data Sources and Descriptive Statistics	106
 Table 3.1 – Baseline	 129
Table 3.2 – Other Assessments of Creditworthiness.....	131
Table 3.3 – Channels	138
Table 3.4 – Adjustment vs. Signaling	140
Table 3.5 – Translation of Sovereign Ratings into Numerical Values	147
Table 3.6 – Descriptive Statistics	152
Table 3.7 – Various Robustness Tests.....	153

Table 3.8 – Regression Results of the Event-based Identification	155
Table 4.1 – First Stage Regressions	188
Table 4.2 – Main Results.....	190
Table 4.3 – Conditional Means of the Gini Index	192
Table 4.4 – Sample Split.....	193
Table 4.5 – TNA Access and PRGF Programs.....	194
Table 4.6 – Descriptive Statistics and Data Sources	198
Table 4.7 – Baseline, First Stage.....	201
Table 4.8 – Baseline, Second Stage.....	202
Table 4.9 – Long-term Effects	204
Table 4.10 – Conditionality	206
Table 4.11 – Robustness 1: Challenging the Liquidity Variable	213
Table 4.12 – Robustness 2: Challenging the Probability Variable	213
Table 4.13 – Robustness 3: Selection on Unobservables	214
Table 4.14 – Robustness 4: Alternative IV, Alternative Treatment	215
Table 4.15 – Robustness 5: Alternative Inequality Data	216

List of Figures

Figure 1 – IMF Lending, 1973-2013.....	3
Figure 2 – The Input and Output of International Organizations	5
Figure 3 – The Inner Workings of International Organizations	7
Figure 4 – The Long Delegation Chain	8
 Figure 1.1 – Voting against the United States in the UNSC	39
Figure 1.2 – Effect of UNSC Membership on US aid for Varying UNSC voting	46
Figure 1.3 – Effect of UNSC Membership on IMF loans for Varying UNSC voting	46
Figure 1.4 – Effect of UNSC Membership on US Aid for Varying Political Proximity	52
Figure 1.5 – Effect of UNSC Membership on IMF Loans for Varying Political Proximity	53
Figure 1.6 – Google Hits of UNSC Resolutions	63
 Figure 2.1 – Visualizing the Effect of Political Alignment with the United States	91
Figure 2.2 – Testing Manipulation of Projections	103
 Figure 3.1 – Global Importance of IMF Programs, 1973-2013.....	111
Figure 3.2 – Rating Dynamics Around Starts of IMF Programs.....	120
Figure 3.3 – Illustrating the First-Stage Effect	124
Figure 3.4 – The IMF’s Liquidity Ratio	126
Figure 3.5 – Event-based Identification.....	135
Figure 3.6 – Text Analysis of Rating Statements	143
Figure 3.7 – Spurious Correlations between Ratings and IMF Liquidity?.....	154
 Figure 4.1 – The IMF’s Liquidity Ratio	186
Figure 4.2 – Visualized Effect of the IV	189
Figure 4.3 – Long Term Effects.....	191
Figure 4.4 – Spurious Correlations Between Inequality and IMF Liquidity?.....	212

Introduction

Globalization is one of the most momentous economic shifts of the recent decades. In most of the world's countries, cross-border flows of goods, capital, and information have increased in a way that fundamentally transformed their economies.¹ Nowadays, national economic outcomes are less and less the product of national circumstances. Instead, they are increasingly influenced by developments that occur across national borders and thus *inter*-nationally. According to the vast scholarly literature on the topic, the economic consequences of globalization are manifold and substantial.²

This poses a problem for the nation-state. In the era of globalization, national governments increasingly lose their political control over national economic outcomes. They can influence the developments in other states that affect their domestic economies only to a very limited degree when acting unilaterally. To manage the international economic interdependencies that come with globalization, national governments increasingly cooperate and coordinate their policies in a multilateral fashion. To do so, they have created a large number of international organizations (IOs) that deal with these economic interdependencies.³ In the globalized world, such IOs play an important role as powerful 'global' actors of so-called 'global governance.'

While much of the public and academic attention often centers around national politics, international organizations are now active and influential in virtually all policy areas.⁴ In the economic realm, IOs like the World Bank, the Organization for Economic Cooperation and

¹ The underlying definition of "economic globalization" follows Keohane and Nye (2000, 4). See Dreher (2006a) as well as Dreher, Gaston, and Martens (2008) for a discussion and an empirical operationalization of this definition.

² For literature reviews on the economic effects of globalization, see Goldberg and Pavcnik (2007), Grossman and Helpman (2015), de Haan and Sturm (2016), Harrison, McLaren, and McMillan (2010), Kanbur (2013), Potrafke (2015).

³ See Keohane (1984). Note that this is just one out of several reasons for why IOs are created. Dreher and Lang (2016) provide an overview of some of the arguments that are discussed in the scholarly literature.

⁴ Today, the median country is a member of more than 50 international intergovernmental organizations (calculation based on CIA 2018; KOF 2016).

Development (OECD), or the secretariat of the United Nations Framework Convention on Climate Change (UNFCCC) are key actors for tackling some of the world's most pressing contemporary economic challenges: The World Bank is tasked with taking on global poverty, the OECD works on political responses to international tax evasion, and the UNFCCC deals with the economic consequences of climate change. Given the substantial amount of international coordination and cooperation that these and similar global challenges require, it seems highly unlikely that national governments will solve any of these problems without the help of international organizations.

This general development is the key motivation for this dissertation to focus – to paraphrase its title – on *how international organizations govern the globalized economy*. It addresses this overarching question by examining both the input and the output dimension of IOs: In its first part the dissertation looks at IO decision-making (“input”); the second part focuses on IO effects (“output”). The guiding theme behind this dual focus is the perspective that international organizations – to paraphrase the dissertation's title again – make highly *political decisions* with important *economic effects*. The four chapters, which are written as stand-alone papers, shed new light on specific elements of both the decision-making and the effectiveness of international organizations as well as on the links between *political decisions* and *economic effects*.

The empirical focus of this dissertation is on the International Monetary Fund (IMF), as it is often perceived as the most powerful IO when it comes to international economic policy.⁵ IMF expert Randall Stone (2002, 1) even goes as far as to describe the IMF as the “most powerful international institution in history.” What does this organization do to deserve this characterization? According to its self-description, the IMF is “working to foster global monetary cooperation,

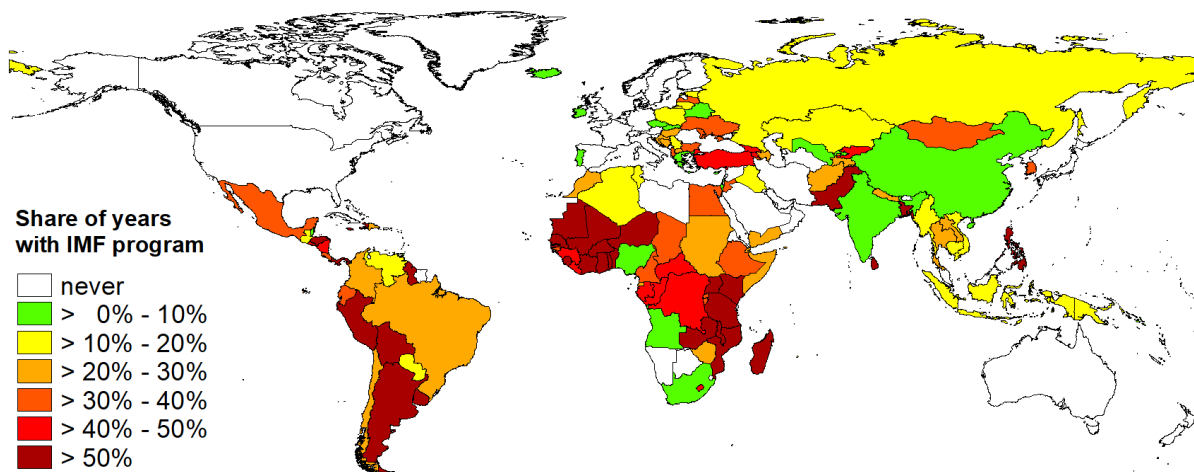
⁵ The only non-regional IO with a similarly large amount of financial resources and similarly potent policy instruments is the World Bank. I focus on the IMF rather than on the World Bank because my research questions center around how international organizations interact with *national* politics and *national* economic outcomes. Whereas the World Bank's power is more relevant at the level of individual development projects, the IMF's power, arguably, refers more directly to the level of the national government. Chapter 2 has a dual focus on both the IMF and the World Bank, as it focuses on one of several policy areas in which the two organizations collaborate.

Many of the theoretical arguments in this dissertation are general enough to relate to international organizations in a general sense and to extend beyond the IMF. Nevertheless, my empirical results, which focus on the IMF, do not allow me to draw the conclusion that they will necessarily hold for other IOs. The fact that no IO is quite like the other requires future research on the extent to which these theoretical arguments apply to other IOs.

secure financial stability, facilitate international trade, promote high employment and sustainable economic growth, and reduce poverty around the world.”⁶ In practice, the IMF attempts to achieve this very broad set of goals primarily by means of its lending and surveillance activities.⁷ Via both of these main activities, the organization can exert substantial power on the national economic developments in its member states.

The IMF is most famous for its lending activities. Originally, IMF loan programs were intended to provide short-term liquidity support to countries with temporary balance-of-payments problems. However, the IMF’s role has evolved. From the 1970s onwards, the IMF began to increasingly lend to countries with more protracted crises or structural economic problems. In fact, IMF programs became so frequent that – with the notable exception of Western Europe and the Anglosphere – there are very few countries that did not have at least one IMF program over the past half-century: When excluding states with less than a million inhabitants, in South Asia the only exception is Bhutan; in Latin America, there are only two exceptions (Cuba and Puerto Rico), and in Sub-Saharan Africa there are only four (Botswana, Eritrea, Namibia, and South Sudan). The world map in Figure 1 illustrates this.

Figure 1 – IMF Lending, 1973-2013



Note: Percentages indicate the share of years with an active IMF program in the 1973-2013 period.

Source: Dreher (2006, updated). Own illustration.

⁶ <http://www.imf.org/en/about> (last accessed: 5 May 2018)

⁷ See Reinhart and Trebesch (2015) for an account of how the IMF oscillated between these two main fields of activity over time. The subsequent paragraph partly builds on this paper.

In these loan programs, the IMF's role is rarely limited to providing short-term liquidity. By making its loans conditional on the implementation of policy reforms, the organization often has a substantial impact on reforms in countries that experience economic crises. For instance, the IMF played key roles in economic adjustments during the Latin American debt crises of the 1980s, in the restructuring of post-Soviet economies after the end of the Cold war, in how countries dealt with the Asian financial crisis in the mid-1990s, and – most recently – in influencing reform agendas during the European debt crisis.⁸ These and many other countries participated in IMF programs in critical crisis periods during which far-reaching reforms seemed necessary. Not least the fact that increasingly specific IMF policy conditions⁹ influenced countries' policy reforms during these critical times made the organization very powerful. For a dissertation that aims to understand *how international organizations govern*, this makes IMF loan programs an important object of investigation. Three of this dissertation's four chapters focus on the IMF's lending activities (chapters 1, 3, and 4).

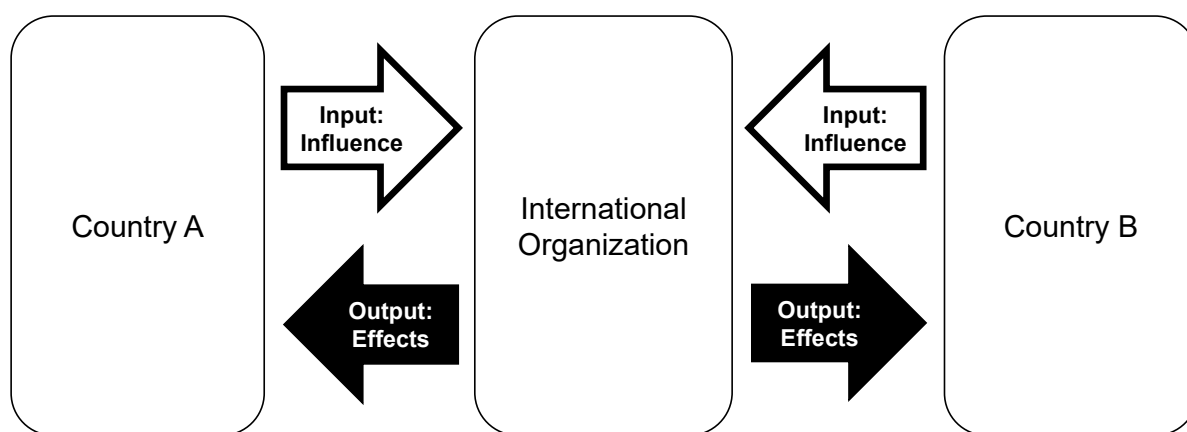
The IMF's second major field of activity is 'surveillance.' The organization gathers statistics, analyzes economic trends and policy reforms, gives policy advice, makes macroeconomic forecasts, and rates countries' economic performance. The output produced by these surveillance activities informs policy debates all around the world – and not only in countries with active IMF programs. What is more, research shows that the output of the IMF's surveillance activities has tangible economic effects. Financial markets and governments react to such outputs as IMF growth forecasts, and public information notices resulting from the IMF's Article IV consultations (Beaudry and Willems 2018; Fratzscher and Reynaud 2011). In light of these findings, it helps this dissertation to paint a more comprehensive picture of the IMF's activities that – in addition to its focus on IMF lending in three chapters – one chapter also examines the organization's surveillance activities (chapter 2).

⁸ See Klein (2008), Stiglitz (2002), Stone (2002), Vreeland (2003) for background information on many of these cases.

⁹ See Polak (1991), Dreher (2009a, 2009b), as well as Kentikelenis et al. (2016) for accounts of how IMF conditionality evolved over time. The general view in this literature is that IMF conditionality over time covered an increasingly large number of policy areas and became much more specific. Other channels of IMF influence discussed in the literature are the disbursed loans themselves and the policy advice the IMF gives to countries participating in its programs (Dreher 2006b).

While addressing different fields of IMF activity, this dissertation consists of two main building blocks: a) the analysis of the IMF’s decision-making, and b) the analysis of the IMF’s effects. Underlying this dual focus is the idea that international organizations are most comprehensively understood when approaching them from both an “input” and from an “output” perspective. As “input” I consider all processes that influence the decisions that IOs make and that lead to the “output” that they eventually produce. As “output” I consider all effects that IOs have on their environment (see Figure 2).

Figure 2 – The Input and Output of International Organizations



In the first half of this dissertation (chapters 1 and 2) the primary focus is on the input dimension; in the second half (chapters 3 and 4) it is on the output dimension. There are two overarching themes in these two parts: First, I provide support for the view that the decision-making of IOs is to a substantial degree *political*. The IMF is not an apolitical, independent, and technocratic institution.¹⁰ Instead, its decision-making (“input”) is deeply embedded in world politics and the IO is frequently used as a political tool to help its most powerful member countries achieve their political goals. The dissertation’s chapters zero in on specific variants of such political influencing. The second overarching theme concerns the output dimension: IOs can have important *economic* effects. They are not weak and unimportant institutions.¹¹ Instead, this dissertation shows that an IO like the IMF can affect economic outcomes like inequality, growth, and capital market access, and thus influence the daily lives of many people very directly and very substantially.

¹⁰ See Keohane, Macedo, and Moravcsik (2009) for a paper that emphasizes this perspective on IOs.

¹¹ See Mearsheimer (1994) for a paper that emphasizes this perspective on IOs.

In short, international organizations make *political decisions* with *economic effects*. Each chapter of this dissertation makes both theoretical and empirical contributions to specific elements of this overarching perspective. I discuss the theoretical and the empirical approach underlying these contributions in turn.

Theoretical Approach

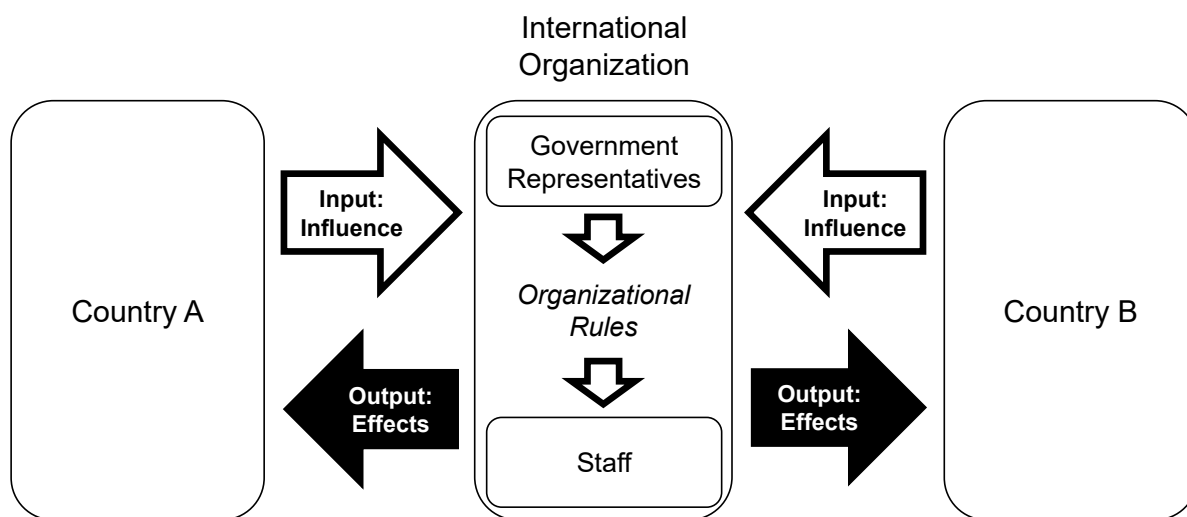
To analyze the input and output of IOs, I apply a particular theoretical framework. IOs are considered from a political economy perspective. As I discuss in a related paper with Axel Dreher (Dreher and Lang 2016), this theoretical approach goes beyond viewing IOs and their member states as monolithic actors. Instead, it focuses on the individual, rational actors that constitute IOs and their member states. The focus thus shifts from “states” to politicians, voters, as well as domestic political interest groups, and looks at how these actors interact within the domestic political system. Inside the IO, this perspective considers actors like government representatives and IO staff, and examines their relationship within the IO’s organizational set of rules. Each of the dissertation’s four chapters zeros in on a particular aspect of this complex network of actors that surround international organizations.

Chapter 1 begins the analysis of the input dimension and examines the influence that the governments of the most powerful member countries have on the IMF. It finds that the United States (US) government is particularly powerful and able to exploit its influence over the IMF in a very specific way. When the exertion of political influence on other countries involves ‘dirty work’ that voters could punish electorally, the US government uses the IMF as political cover. The empirical evidence in this chapter suggests that the United States uses IMF loans to buy the votes of hostile countries in the United Nations Security Council (UNSC). As such deals between governments made via international organizations like the IMF are less visible for the public and less clearly attributable to individual governments, states can use IOs to hide and “launder” unpopular foreign policy decisions.¹² Such “laundering” activities appear to constitute an important ‘input’ element of the political influence exerted on IOs.

¹² This argument goes back to Vaubel (1986). See also Abbott and Snidal (1998).

Chapter 2 retains the focus on the influence that different actors exert on the IMF but zeros in on the inner workings of the organization. This helps the chapter to shed light on the channels that allow such political influence as identified in chapter 1 to materialize. Building on the political economy perspective, it further breaks up the IMF as a monolithic actor and looks at its internal, organizational rules of decision-making (see Figure 3, where blank arrows symbolize influence).

Figure 3 – The Inner Workings of International Organizations



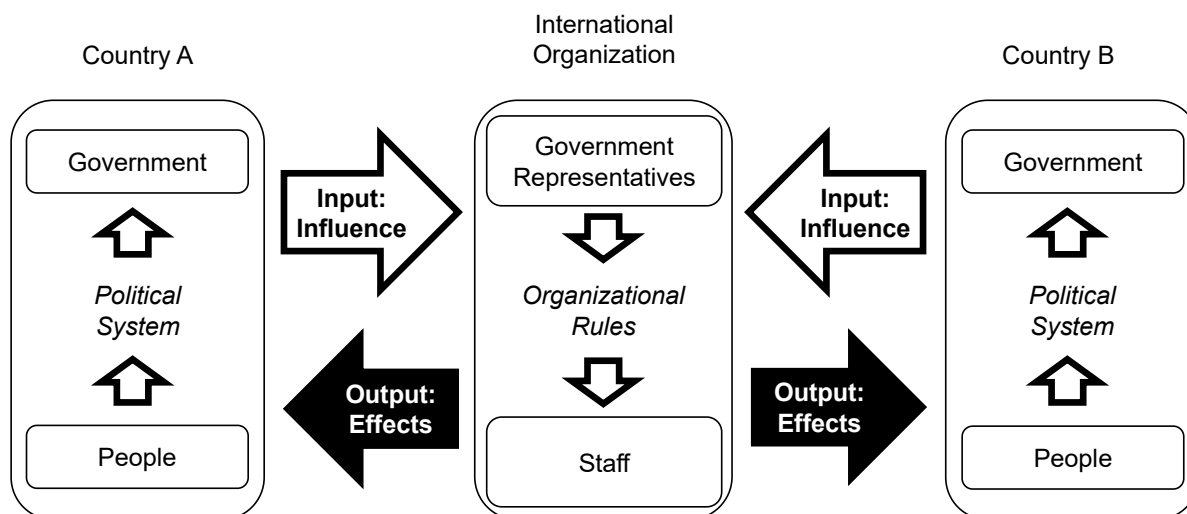
The chapter differentiates between government representatives and staff, and analyzes how these actors interact when decisions are made. More specifically, the chapter provides explicit empirical evidence for the influential “informal governance” model (Stone 2008, 2011, 2013), according to which both formal and informal rules regulate IOs’ decision-making processes. It tests the model in the context of IMF-World Bank Debt Sustainability Analyses and shows that the formal process of debt rating is more often overruled in favor of the rated country when this is in the interest of the United States. The results provide evidence for both sets of rules and suggest that the ‘room for discretion’ that the formal rules leave gives political influence the opportunity to enter via informal rules.

Chapter 3 then shifts the attention from the input dimension to the output dimension. Its focus is on the IMF’s effect with regards to one of its most fundamental goals: helping countries overcome balance-of-payments problems. To gauge the IMF’s effectiveness in this respect, this chapter analyses how assessments of sovereign creditworthiness change when countries enter into IMF

loan programs. The chapter suggests that these effects are complex. IMF programs, on average, have economically contractionary effects but also send positive signals to market participants, which prevent the program country's creditworthiness from deteriorating despite these contractions. While the appropriateness of the contractionary adjustments is debatable,¹³ the chapter suggests that to the extent that IMF programs are intended "to help cushion the impact of adjustment" (IMF 1998) they work as intended.

Chapter 4 continues the analysis of the IMF's output but turns to an *unintended* consequence by connecting the IMF's decision-making with its effects. To do so, it further complexifies the agent structure involved in the decision-making process of IOs (see Figure 4).

Figure 4 – The Long Delegation Chain



On the one hand, the chapter retains the analytical differentiation between staff and government representatives inside the IMF. Additionally, however, it also differentiates between governments and voters inside member states as well as between different relationships between governments and voters of member countries with different political systems. The chapter thus takes into consideration the entire 'chain of delegation' from voters to IO staff. As the blank arrows in Figure 4 illustrate, this chain runs from voters via their domestic political systems

¹³ See, for instance, Blanchard and Leigh (2013) for evidence suggesting that the IMF systematically underestimates the contractionary effects of fiscal austerity. See also Przeworski and Vreeland (2000), Barro and Lee (2005), and Dreher (2006b), for critiques of the IMF on the basis of its programs' negative effects on economic growth. See Bas and Stone (2014) for evidence pointing in the other direction.

(which includes additional actors like parties, interest groups, parliaments) to their governments; from governments to their delegates in the IMF's executive board; and from the IMF's board via a set of formal – and informal (chapter 2) – organizational rules to the IMF's staff. The argument in chapter 4 is based on the observation that this delegation chain from the “ultimate principal” affected by the IMF to those who make the decisions in the IMF is long and fraught with several principal-agent problems (chapter 1, chapter 2, Nielson and Tierney 2003).¹⁴ The chapter examines the argument that this can lead to situations where the IMF's policy output deviates from the preferences of the people that are affected by it. It provides evidence for this argument by showing that IMF programs increase income inequality in democracies.

Chapter 4 thus brings together theoretical arguments on IO input from chapters 1 and 2 with the examination of IO output, which began in chapter 3: Chapter 1 points to the principal-agent problem arising from the fact that some members have disproportionately strong influence. Chapter 2 adds that much of this political influence is informal and establishes that staff preferences that arise out of bureaucratic incentives also influence the decision-making, thereby adding another principal-agent problem. Chapter 4 then links these principal-agent problems to unintended effects of IMF programs. In sum, in a literature that is often divided in studies that examine the determinants of the IMF's activities (input) and in studies that examine their effects (output) this dissertation points to new links between input and output.¹⁵ The specific decision-making structure of the IMF – in which political and bureaucratic interests, informal governance, and principal-agent problems play an important role – influences its output's effectiveness. In short, *it matters for the IMF's economic effects, that its decision-making is political.*

¹⁴ See also Hawkins et al. (2006) and Vaubel (2006).

¹⁵ I do not claim to be the first to make the argument that IMF decision-making and IMF effectiveness are connected. Stone (2002) as well as Chapman et al. (2015) make the related argument that US influence on the organization interferes with the organization's effectiveness. To the extent that more extensive IMF conditionality leads to better policy output, the results in Dreher, Sturm, and Vreeland (2015), which suggest that political influence can reduce the scope of IMF conditions, can also be considered as providing support for this link. Nevertheless, the connection between the principal-agent problems discussed in chapter 4 and unintended economic effects is rarely made explicit in empirical studies on the IMF. Rather, one set of studies looks at the determinants of IMF decisions-making (for the key studies in this literature see chapter 2), and another looks at the IMF's effects (see chapter 4 for the key studies). Note that the argument is also related to the finding that political motivations in aid allocation reduce aid effectiveness. Kilby (2015) provides evidence for this argument by examining World Bank projects. Dreher, Eichenauer, and Gehring (2016) provide evidence in the context of bilateral aid.

Empirical Approach

Having discussed the dissertation's broad theoretical framework, I now turn to its empirical approach. Each chapter puts its theoretical predictions concerning the specific research question it poses to the empirical test. The guiding principle behind the choice of the particular empirical approach of the chapter is to get as close as possible to estimating causal effects without letting the quest for causality hurt the relevance of the research question that is analyzed.

Over the course of the last decade, the academic fields that this dissertation contributes to have gone through what Angrist and Pischke (2010) called the "credibility revolution." Empirical economists now increasingly apply "design-based approach[es] that emphasize [...] the identification of causal effects" (Angrist and Pischke 2010, 1). This dissertation, on the one hand, follows this trend in the discipline because it aims to make causal rather than correlational statements on the effects and mechanisms its hypotheses address. On the other hand, however, the dissertation does not go as far as some recent studies, which – out of the desire for internal validity and causality – limit the scope of their research question to a more marginal level at which the study's external validity and its (policy-)relevance are impaired.¹⁶

In particular, it does not adopt the view that real experiments, which are sometimes referred to as the 'gold standard,'¹⁷ are necessarily the best or the only way to generate credible empirical knowledge. As for this dissertation's research questions conducting real experiments is not feasible – the IMF would, for instance, not agree to allocate its loan programs randomly – applying such a radical 'randomista' perspective would lead the researcher to either dramatically shift the focus and limit the scope of the research question or refrain from studying it altogether. Faced with the choice of not studying these topics or studying them in the best way possible, I decide for the latter.¹⁸

¹⁶ See Deaton (2009) for a general critique on design-based approaches along these lines.

¹⁷ In his critique of the "pervasive view among statisticians [...] that glorifies randomization as the 'gold standard' of causal inference" Heckman (2005, 148), for instance, uses this term.

¹⁸ This choice also reflects my personal conviction that these topics are politically too important to be left unstudied. If scholars decide against studying research questions that are highly relevant for politics and the public, unfounded speculations and simplistic arguments will dominate political and public discourses. This is why in my view, a world in which scholars provide some insights on such questions, even if not in the form of definite causal statements, is preferable to a world in which these scholarly insights are absent.

When ‘real experiments’ are infeasible and gathering anecdotal or correlational evidence will not give satisfactory answers to the research questions, scholars now increasingly make use of ‘natural experiments’ for causal inference. This is the approach I adopt in this dissertation. I draw on a variety of statistical tools that make use of sources of quasi-random variation that occur in the real world for the identification of causal effects. The dissertation as a whole applies a range of identification strategies: It makes use of a *natural experiment* that leads to *quasi-random treatment*, an *instrumental variable* strategy, settings that resemble *difference-in-differences* approaches, another setting that exploits a particular institutional rule, a *regression discontinuity design*, and an *event-based identification* building on the usage of *two-dimensional fixed effects*.

Perhaps the most straightforward identifying assumption features in the *natural experiment* exploited in chapter 1. It is based on the finding that the timing of temporary membership in the United Nations Security Council (UNSC) is conditionally exogenous to relevant outcome variables (Bueno de Mesquita and Smith 2010; Dreher et al. 2014). Primarily due to the adherence to a turn-taking norm (Dreher et al. 2014), in this setting, it suffices to condition on a limited set of control variables and panel fixed effects to make endogeneity bias unlikely.

Instrumental variables are part of the empirical analysis in chapters 3 and 4. To identify the effect of IMF programs they exploit the differential effect of changes in the IMF’s liquidity on loan allocation depending on a country’s history of IMF program participation as a source of exogenous variation. As these chapters will discuss in detail, the identifying assumption underlying this approach, which uses an interaction term as an instrumental variable, resembles the identifying assumption of a *difference-in-differences* approach.

A related strategy is one of the building blocks of chapter 2. It exploits the fact that the timing of national elections introduces quasi-exogenous temporal variation in the setting under study. In both of these settings, the time-varying variables (IMF liquidity and elections) lead to ‘differences’ in the association between the respective outcome and an explanatory variable. Under standard *difference-in-differences* assumptions, this ‘difference’ has a causal interpretation (Bun and Harrison 2018; Nizalova and Murtazashvili 2016).

In chapter 2, this identification strategy is implemented to enhance the plausibility of the assumption underlying its baseline analysis: The chapter’s core empirical design is based on a

particular institutional feature of the IMF's Debt Sustainability Framework (DSF) that makes endogeneity bias in the baseline regressions unlikely. The same chapter, in addition, includes a *regression discontinuity design* that exploits a second particular institutional feature of the DSF: The theoretical considerations suggest that the density function of projections used in the DSF could exhibit a discontinuity at particular values. Such a 'density discontinuity' (see Calonico, Cattaneo, and Titiunik 2014; Cattaneo, Jansson, and Ma 2016; McCrary 2008) at this particular threshold would be highly unlikely to result from an omitted factor and, thus, makes a causal interpretation of the finding plausible.

An empirical test that supplements the main analysis and is based on an alternative identifying assumption is also part of chapter 3. It builds on an *event-time specification* that isolates temporal variation of high-frequency data within individual country-year observations by making use of *two-dimensional fixed effects*. The identifying assumption underlying this approach is that these fixed effects control for all unobserved omitted variables that could bias the coefficients of interest. The fact that this assumption is entirely different to the identifying assumption of the chapter's main analysis, but results in a finding that is consistent with the main analysis should increase confidence in the chapter's empirical results.

Key ingredients of the quantitative data used for these statistical analyses are original. Chapter 1 is based on a newly coded dataset on UNSC voting behavior with resolution-specific information.¹⁹ Chapter 2 is based on the mechanical reconstruction of the ratings of the Debt Sustainability Framework from an internal IMF database consisting of macroeconomic projections. Chapter 3 uses, inter alia, an original set of monthly data on IMF program approvals to exploit monthly variation in data on sovereign credit ratings in a new way.²⁰ For chapter 4, I compiled a new dataset of different measures of the IMF's liquidity for the 1973-2013 period.

The dissertation complements these statistical analyses with supplementary qualitative evidence. All chapters include anecdotal evidence from exemplary cases that illustrate the statistical results.

¹⁹ Vreeland and Dreher (2014) use a preliminary version of the dataset that we introduce in this paper in some regressions.

²⁰ The monthly credit ratings come from Fuchs and Gehring (2017). They deserve the credit for coding these data. However, in their study, the key explanatory variables are at the yearly level. Chapter 3 of this dissertation, which is co-authored with Gehring, involves an empirical setting where the key explanatory variable is at the monthly level and thus allows us to exploit this monthly variation in their data in a new way.

Chapter 3 additionally includes a qualitative text analysis. More generally, several arguments made in this dissertation are inspired by background information that I gathered in interviews at the IMF's headquarters in Washington, DC. In particular, the theoretical argument in chapter 2 benefitted from insights gained in these interviews, which were conducted in November 2016 and November 2017.

Policy Areas and Policy Implications

The four chapters of this dissertation address three different policy areas. In light of the aforementioned perspective that international organizations have gained importance not least because of intensifying global interdependencies, it is no coincidence that these policy areas are all directly related to the economic challenges and opportunities that come with globalization.

Chapters 2 and 3 focus on countries' access to global capital, a core theme in many contemporary political debates in both advanced economies and developing countries. The conditions under which countries can refinance themselves played an important role, for instance, in the context of the recent European debt crisis. During the crisis, countries like Greece, Portugal, and Ireland effectively lost market access, and many placed their hopes on the IMF to help these countries regain creditworthiness. The question of the extent to which the IMF is able to do so, is thus highly topical. However, the issue is also relevant for low-income countries. As Presbitero et al. (2016) show, many of these countries increasingly gain access to global capital markets while integrating in the globalized economy. The findings of this dissertation suggest that the IMF plays an important role in this policy area: both in influencing countries' creditworthiness (chapter 3) and in assessing (and thereby also indirectly influencing) it (chapter 2). For rich and poor countries alike, the IMF's activities in the area of sovereign access to capital markets, have important economic consequences as access to credit markets can promote economic development (Berensmann, Dafe, and Volz 2015; Fink, Haiss, and Hristoforova 2003). By pointing to some of the benefits but also to some of the problems connected to the IMF's activities in this policy area, these two chapters aim to inform the debate on how the IMF can help countries to enhance (or regain) their access to global capital most effectively. A direct policy implication of chapter 2 is that a cutback of the "room for discretion" that formal IMF and World Bank rules leave would

help reduce the opportunities for political biases to take effect. From chapter 3 governments of crisis countries can learn that capital markets are unlikely to react adversely when they decide to ask the IMF for financial support.

The policy area that chapter 1 addresses is related. It also looks at international capital flows but focuses exclusively on official flows. More specifically, it examines explanations for the allocation patterns of bilateral aid flows and the IMF's multilateral loans. In the globalized economy, these official flows play an important role: On the one hand, because they are considered as a means to reach global development goals.²¹ On the other hand, because they are political tools that donors use to reach their strategic goals in other countries in an increasingly interconnected world. The results of chapter 1 place the focus on the latter aspect and suggest that IMF loans are sometimes used to hide unpopular policies when the United States exerts political influence on a developing country. By uncovering this mechanism, this chapter points to an important hidden motive behind the allocation of multilateral loans. As both multilateral and bilateral aid were shown to be less effective for economic development when politically motivated (Dreher et al. 2013; Dreher, Eichenauer, and Gehring 2016; Kilby 2015), a direct policy implication of this result is to find ways to reduce the opportunities of powerful donor governments to use multilateral organizations as tools for laundering their "dirty work." Increasing the transparency of internal IO decision-making processes seems to be an important starting point.

The final chapter tackles a policy area that is often at the center of public debate and, at least since Piketty's (2014) *Capital in the 21st Century*, of the academic debate as well: inequality. Recent reports by the World Bank (2016a) and the World Inequality Lab (2017) confirm rising trends of economic inequality in many countries. While the evidence suggests that globalization is at least one driver of this trend,²² the growing literature on the determinants of inequality increasingly stresses the role of changing national policies and institutions (Dabla-Norris et al. 2015; OECD

²¹ The Sustainable Development Goals (SDGs), for instance, see a key role for bilateral aid (see SDG target 17.2.). On its relevance for the SDGs the IMF states: "The IMF is committed, within the scope of its mandate, to the global partnership for sustainable development. The IMF has launched a number of initiatives to enhance its support for its member countries in crucial ways as they pursue the SDGs" (IMF 2018b).

²² In a paper, which I wrote with Marina Mendes Tavares during my time as a PhD student, but which is not part of this PhD dissertation, we show that economic globalization tends to increase income inequality (Lang and Tavares 2018). See the literature quoted therein for further literature on this topic.

2011). Chapter 4 contributes to this literature by suggesting that the IMF has, during critical periods in many countries, contributed to policy reforms that promoted rising inequality. To link this finding to the current political debate inside the IMF, the institution seems to be increasingly determined to avoid adverse distributional effects: Its Managing Director, Christine Lagarde, recently claimed that “reducing excessive inequality [...] is not just morally and politically correct, but it is good economics” (IMF 2015). The final chapter’s empirical result thus highlights an unintended consequence of the IMF’s lending arrangements and may encourage the Fund to continue revising its policy advice and conditionality with regards to their distributional implications. The chapter also suggests that the IMF’s recent internal reforms that increase more participative decision-making processes and emphasize national ownership of reforms under IMF programs can help reduce this adverse distributional effect.

In sum, I hope that the findings of this dissertation in the policy areas of sovereign market access, official international flows, and income inequality will not only be of interest for scholars but will also be of help for policy practitioners that aim to promote inclusive economic development.

Structure and Summary of the Dissertation

The remainder of this dissertation consists of four chapters and proceeds as follows:

Chapter 1 shows how major shareholders can exploit their power over international organizations to hide their foreign-policy interventions from domestic audiences. The chapter is part of an ongoing project with Axel Dreher, B. Peter Rosendorff, and James Raymond Vreeland. We argue that major powers exert influence bilaterally when domestic audiences view the intervention favorably. When domestic audiences are more skeptical of a target country, favors are granted via international organizations. We test this theory empirically by examining how the United States uses bilateral aid and IMF loans to buy other countries’ votes in the United Nations Security Council (UNSC). To do so, we introduce a new dataset on voting behavior in the UNSC that covers a total of 36,460 individual votes on 2,524 proposed resolutions along with resolution-specific information. Our results show that US “friends” receive more bilateral aid when voting in line with the United States in the UNSC, while concurring votes of US “enemies” are rewarded with loans from the IMF. Temporary UNSC members that vote against the United States do not

receive such perks. These results suggest that the United States chooses bilateral channels to reward friends but relies on obfuscation via international organizations when buying favors from hostile countries. In sum, we argue that powerful countries can use international organizations as political tools to do their “dirty work.”

Chapter 2 then sheds light on a more fine-grained mechanism that can, *inter alia*, help understand the kind of political influence on the IMF for which chapter 1 provides empirical evidence. It is co-authored with the IMF economist Andrea F. Presbitero. His institutional affiliation helped us to get access to the internal data used for the analysis and allowed us to conduct interviews with IMF and World Bank staff to gather background information. The paper was published in the *Journal of Development Economics* before this dissertation was submitted (Lang and Presbitero 2018).

In this chapter, we examine how powerful states exert “informal” influence on the IMF and the World Bank. We exploit the degree of discretion embedded in the World Bank-IMF Debt Sustainability Framework to understand the decision-making process inside the two organizations. Our unique, internal dataset covers all debt sustainability analyses conducted between 2006 and 2015 for low-income countries. These data allow us to identify cases where the risk rating implied by the application of the DSF’s mechanical rules was overridden to assign a different official rating. In contrast to much of the previous literature, we can thus shed light on the *internal* decision-making processes of IOs. We can directly examine the extent to which informal influence *interferes* in technocratic rules instead of only comparing how they relate to differences in outcomes. Our results show that formal rules are adhered to but also that political interests and bureaucratic incentives influence the decision to intervene in the mechanical decision-making process. Countries with political ties to the institutions’ major shareholders are more likely to receive improved ratings; especially in election years and when the mechanical assessment is not clear-cut. Supplementary evidence from a regression discontinuity design additionally suggests that the macroeconomic projections underlying the ratings are also biased. These results suggest that the “room for discretion” IOs have can be a channel for informal governance and a source of biased decision-making.

While in chapter 2 we also provide some correlational evidence suggesting that the DSF ratings matter for sovereign market access, a rigorous empirical test of how the IMF affects the conditions under which countries have access to capital markets follows in chapter 3. This paper is co-authored with Kai Gehring. An earlier version of it was published as a *CIS Working Paper* before this dissertation was submitted (Gehring and Lang 2018).

The chapter starts with the observation that the loan programs of the IMF are often considered to carry a “stigma” that triggers adverse market reactions. Countries are thus sometimes hesitant to engage with the Fund.²³ Our results, however, suggest that the presence of an IMF program sends a *positive* signal to investors that prevents creditworthiness from falling despite substantial economic contractions under IMF programs. Using monthly data on credit ratings from various agencies and professional investors for 100 countries in the 1987-2013 period as measures of creditworthiness, we apply three different methods to circumvent the severe endogeneity problem associated to studying the IMF’s effects. For our main identification strategy, we exploit the differential effect of changes in the IMF’s liquidity on loan allocation as a source of exogenous variation. We find that adjustments under IMF programs reduce GDP growth in the short-run, but do not lower creditworthiness. When examining the underlying channels, we find evidence of a *positive signaling* effect on perceptions of creditworthiness that offsets the *negative adjustment* effect on aggregate output. IMF involvement appears to create positive expectations about the program country’s future policy path and thereby ‘cushions’ the drop in creditworthiness that countries undergoing such contractionary adjustments would usually suffer from. Event-time specifications exploiting monthly variation within country-year observations and a systematic text analysis of rating statements support this interpretation.

In addition to investigating this relationship, chapter 3 uses the high-frequency monthly data on sovereign credit ratings also to illustrate the severe endogeneity problem that studies focusing on the IMF’s effects need to deal with. We document that countries usually experience substantial drops in credit ratings in the months before they enter IMF programs. Our results also suggest that controlling for observables is unlikely to remove the entire bias resulting from this. They

²³ This is a key insight I gained at the IMF headquarters: IMF staff repeatedly told me that they were worried that the IMF loses global relevance because countries fear such a stigma effect. See also Andone and Scheubel (2017).

underline the importance of employing an appropriate identification strategy when examining the IMF's effects. This is also a key issue in the empirical analysis of chapter 4.

Chapter 4 is single-authored. An earlier version of it was published as a *Heidelberg University Discussion Paper* (Lang 2016). Like chapter 3, it examines an economic effect of IMF programs. The reader will notice that the chapter places emphasis on the theoretical argument whereas chapter 3 focuses more on the empirical analysis.²⁴ But like chapter 3, chapter 4 also makes both empirical and theoretical contributions to the literature.

Empirically, the main contribution is introducing the new instrumental variable for IMF programs, which is also one of the identification strategies applied in chapter 3. For this IV, I exploit exogenous time variation in the IMF's liquidity and cross-sectional variation in a country's probability of having a lending arrangement with the IMF. The interaction term of the two variables is excludable to country-specific economic characteristics and allows determining a causal effect via quasi-random assignment relative to country-year-specific economic outcomes. This construction of a plausible counterfactual by means of the IV is important for testing the theoretical predictions made in this chapter. I argue that the IMF, not least because of the specific principal-agent problems in its delegation chain, which are also addressed in chapters 1 and 2, suffers from a relative lack of accountability to the people affected by its policy output. By interfering in domestic politics, it might thus restrict the responsiveness of democratic governments to the distributional preferences of their citizens. I develop the hypothesis that democracies under an IMF program – relative to a plausible counterfactual of the same democracies in the same circumstances not under an IMF program – will experience increases in income inequality. Using panel data for 155 countries over the 1973–2013 period, the results show that IMF programs substantially increase income inequality. The effect is driven by the democracies in the sample, and there is no evidence for such an effect in non-democracies. The evidence also suggests that the IMF's recent enhancement of accountability mechanisms is associated with a mitigation of its inequality-increasing effect.

²⁴ This has to do with the fact that the paper was submitted to a general interest political science journal.

1 Buying Votes and International Organizations: The Dirty-Work Hypothesis

Note: This paper builds on Dreher and Vreeland (2011) and Dreher, Rosendorff, and Vreeland (2013). It is part of an ongoing project with Dreher, Rosendorff, and Vreeland.

1.1 Introduction

After the collapse of the Soviet Union, one of the United States' key geostrategic goals was to reduce the threat Russia posed to the country's security. To do so, the United States attempted to influence political developments in Russia during the 1990s, and in the early years of the decade bilateral aid was one of the main policy instruments: US disbursements of bilateral aid to Russia amounted to one billion US dollars (USD) in 1993 and 2.5 billion USD in 1994.²⁵ In 1994, these aid packages came under increasing popular pressure at home. According to a Congressional Research Service report, "concerns regarding the US budget deficit [and] the unpromising outcome of the December 1993 Russian parliamentary elections," amongst others, led to substantial reductions in US aid. Between 1996 and 1998 annual disbursements of US aid to Russia were reduced to about half a billion USD.

At about the same time, the International Monetary Fund (IMF) became heavily involved in Russia. In 1995, it approved a 6 billion USD loan program, increased it to more than 10 billion the next year and to an extraordinarily large 18 billion USD loan in 1998. The United States strongly supported this. US President Clinton stated: "I believe the loan will go through, and I believe that it should. I do support it strongly."²⁶ And Russian President Yeltsin said that to get the IMF to commit to these loans "[w]e had to involve Clinton, Jacques Chirac, Helmut Kohl, and [John] Major."²⁷ Further anecdotal evidence that the United States put pressure on the IMF abounds (Congressional Research Service 2002; Goldgeier and McFaul 2005; Stone 2002). As Goldgeier and McFaul (2005, 152) put it: "[t]he Clinton administration wanted to use the IMF to support Yeltsin in his time of need; the IMF obliged." And more generally: "[i]n essence, the Clinton administration transferred the responsibility for assisting Russia's economic transformation from the United States to the IMF" (p. 100).

²⁵ Some US politicians justified these large aid disbursements by arguing "that the U.S. defense budget would be \$100 billion greater in the next year if the Soviet Union still existed as a military threat" (Congressional Research Service 2002, referring to a speech on March 4, 1993, by US Senator Patrick Leahy, Chairman of the Senate Foreign Operations Subcommittee, in which he called for a one-billion-dollar aid package).

²⁶ Quoted in Goldgeier and McFaul (2005, 152).

²⁷ <https://www.nytimes.com/1996/02/23/world/russia-and-imf-agree-on-a-loan-for-10.2-billion.html> (last accessed May 20, 2018).

This episode seems to suggest that the United States initially used bilateral aid to pursue a key geopolitical goal. When directly giving its own aid became increasingly difficult to justify domestically, it switched to the IMF and used the organization to support the country with international aid.²⁸

Of course, the example of Russia might be an isolated case rather than representative of a general pattern. The scholarly literature that compares bilateral and multilateral aid typically argues that political interests are less prevalent for multilateral aid, and takes the relative absence of political motives as a reason why multilateral aid is more effective for promoting development (Derek 2008; Milner and Tingley 2013).²⁹ Much of the recent literature therefore concludes that donors use multilateral aid to promote development and other international public goods, while they use bilateral aid to promote their own political agenda (Schneider and Tobin 2016).³⁰

The focus of this recent literature on multilateral aid as a largely apolitical instrument of burden-sharing for promoting development goals stands in contrast to the literature on international organizations. Woods (2003) and McKeown (2009) document that the United States virtually controls major decisions at the IMF and the World Bank. According to Rieffel (2003, 28–29), “[t]he IMF is an instrument of the G-7 countries. There is no example that comes easily to mind of a position taken by the IMF on any systematic issue without the tacit, if not explicit, support of the United States and the other G-7 countries.” Quantitative evidence supports the view that multilateral lending reflects the interests of international organizations’ major shareholders (Dreher, Sturm, and Vreeland 2009a; 2009b; Kilby 2013a; 2013b; Kuziemko and Werker 2006; Vreeland and Dreher 2014).³¹

Even though a large number of papers has investigated the importance of political motives for multilateral and bilateral aid giving, no empirical work exists that can easily reconcile the two different strands of literature. Taken at face value, these two literatures suggest that donor countries use multilateral aid for pursuing their own political agendas while, at the same time,

²⁸ We use the term “aid” for all forms of official support channeled to recipient countries, including Official Development Assistance (ODA), Other Official Flows (OOF), and multilateral loans and credits.

²⁹ For the importance of geo-strategic motives for the effectiveness of aid see Dreher, Eichenauer, and Gehring (2016).

³⁰ Also see Dietrich (2013). According to Dietrich, governments use multilateral (and other non-state) aid when they want it to promote development in recipient countries with low governmental quality.

³¹ For a broad overview of this literature see Dreher and Lang (2016).

bilateral channels seem more politicized than multilateral channels. In our view, this begs the questions how states decide between bilateral and multilateral channels for exerting political influence and why multilateral aid is often perceived as less political than bilateral aid. In this paper we offer an explanation that sheds new light on these questions.

We argue that donor governments use multilateral channels for exerting political influence when the domestic public is hostile towards supporting the recipient. Bilateral channels, on the other hand, will primarily be used for countries that the domestic audiences view more favorably. Our argument is based on the idea that multilateral organizations can be used to do their major shareholders' "dirty work" (Vaubel 1986). Some governments have substantial influence over multilateral organizations, which they can exploit to pursue policies vis-à-vis other states without drawing on bilateral channels. Multilateral organizations can thus help to "launder" governments' political activities that are unpopular with domestic audiences when conducted bilaterally but acceptable when executed via multilateral organizations (Abbott and Snidal 1998; Vaubel 1986). As politics inside multilateral organizations is difficult to observe for the public, governments can implement their preferred policies without having to fear adverse electoral consequences. Conversely, when they aim to give aid to friendly countries they can use the more visible bilateral channel.

For such "laundering," major shareholders exploit multilateral organizations' reputation as a politically neutral donor to hide unpopular policies from their voters. To keep this reputation alive, governments thus rarely interfere in their decision-making, and the organizations grant loans according to need, *on average*. Politics thus seem to be less prevalent in the allocation of multilateral aid compared to bilateral aid. But rather than being free of political motives, multilateral aid is highly political – governments just use it for political purposes in selected salient cases when it would be politically costly to draw on bilateral resources. This is also why political motives in the lending of these organizations are more difficult to detect empirically – for voters and scholars alike. The *average* multilateral loan is more likely to be given for non-political reasons, compared to bilateral aid. And the fact that politically motivated loans are given to 'strange bedfellows' (i.e., recipients that are not among the donors' traditional allies) makes it even harder to detect the underlying political motives.

This is why testing this theory requires a new empirical setting.³² After all, we aim to detect patterns that are deliberately hidden. We test it focusing on vote buying in the United Nations' most powerful organ, the United Nations Security Council (UNSC). Specifically, we examine how voting behavior in the UNSC is linked to the allocation of bilateral aid flows and loans from multilateral financial institutions. Dissenting votes in the UNSC are rare and major powers use incentives and disincentives for other members to avoid them. We expect governments to use increases in aid as reward for loans and reductions in aid as punishment. Countries that vote against powerful governments in the UNSC thus should receive less bilateral and multilateral support. However, patterns of bilateral aid are easily observable by domestic audiences. Those of multilateral aid are not; as we discuss in more detail below, publics tend to perceive international organizations as independent actors and do not easily attribute their lending to the influence of their own government. We thus expect that bilateral aid will be used to buy favors from befriended governments, which domestic audiences can easily agree to provide aid to. In cases where domestic audiences are likely hostile towards the recipient government, we expect powerful shareholders to use their power over international organizations to extend support. They will increase multilateral aid rather than bilateral aid. This is the case that fits the Russian example above.

To test our theory, we compile a new dataset that covers the universe of UNSC votes that were cast by all member states in the seven decades over the 1946-2015 period. We record a total of 36,460 individual votes on 2,524 proposed resolutions. We consider all available UNSC proposals – those that have passed (resolutions) and those that have failed (vetoed resolutions and failed majorities). To our knowledge, this is the first such dataset, which we collected from the United Nations (UN) Library in Geneva, as well as from UN web pages. Along with each member state's decision, we code resolution-specific information, such as the policy area concerned and the amount of media attention the resolution generated.

Armed with the new data on UNSC voting we test our theoretical argument and find considerable support for it. First, the evidence is consistent with the view that votes in the world's

³² The "dirty-work" hypothesis goes back to Vaubel (1986) but has never been tested in a large-n setting. His own work uses "the methodology of example giving" (Vaubel 1986, 45).

most important international institution are for sale. We find that temporary members of the UNSC that vote in line with the United States receive both more bilateral aid from the United States and larger IMF loans than non-members. Countries that vote against the United States in the UNSC do not receive such perks during their time as temporary members. We then turn to testing our argument on the choice of bilateral versus multilateral channels. We find that the United States uses bilateral aid to buy the votes of UNSC members it is politically close to and multilateral loans to buy the votes of members to which it is politically more distant (as measured by voting coincidence in the UN General Assembly prior to entering the UNSC). While our main analyses focus on the trade-off between bilateral US aid and IMF loans, we also show that the results hold for World Bank loans.

Our results add important nuances to the literature on how donors chose between bilateral and multilateral support. Rather than being less political than bilateral aid, donors benefit politically from not influencing multilateral aid *on average*, so that multilateral organizations maintain their reputation as politically neutral. They can then use them to perform their dirty work in strategically important cases. We see this theory as complementing rather than contradicting previous work and, interestingly, results obtained in previous research support our argument. Milner (2006) finds that right-wing governments give more multilateral aid than left ones. She concludes that “it is hard to understand this result” which is “robust and puzzling” (2006, 132). It cannot be explained with Milner’s theory which expects donors to use bilateral aid to pursue political goals. Given that right-wing governments typically pursue more aggressive foreign aid policy than left ones (e.g., Milner and Tingley 2010), it is however exactly what our theory would predict. Strand and Zappile (2015) proxy donor interest in a country with its economic aid, following Fleck and Kilby (2006). They expect countries that receive more economic aid from a member of a multilateral development bank to also receive more aid from the bank itself but find the opposite. This result is contrary to the authors’ expectations but is exactly what we would expect to find when multilateral aid is used in countries where bilateral aid is difficult to give. In addition to adding important insights to the literature on donors’ choice among bilateral and multilateral aid, our results speak to several other literatures: First, our paper links to the literature on associations between aid flows and voting in the UN, which has so far focused on

the United Nations General Assembly (UNGA) (Thacker 1999; Stone 2008; Dreher, Nunnenkamp, and Thiele 2008; Kilby 2013; Carter and Stone 2015). Our results suggest that vote buying extends beyond the UNGA and also relates to the UN's most powerful organ, the UNSC.

Second, we qualify the 'UNSC effect.' Multiple recent studies have shown a relationship between temporary UNSC membership and favorable treatment from aid donors and multilateral organizations (Dreher, Sturm, and Vreeland 2009a; 2009b; 2015; Kilby 2013b; Kuziemko and Werker 2006; Mikulaschek 2017b; Reynolds and Winters 2016; Vreeland and Dreher 2014). For the case of US aid and IMF loans we show that those temporary members of the UNSC that vote in line with the United States rather than membership itself drive this effect. Our results thus add more direct evidence for the conjecture that the larger aid flows to UNSC members are used for vote buying.

Third, and closely related, which countries control the IMF is contested in the literature. While some argue that the United States is by far the most powerful of the IMF's shareholders (e.g., Stone 2008), others hold that "the United States does not unilaterally control IMF lending; rather, the G5 countries as a group exercise substantial influence" (Copelovitch 2010a, 73). We find that temporary members of the UNSC whose voting behavior resembles the voting behavior of the other four permanent five UNSC members do not receive more IMF loans than non-members. This link only emerges as far as voting similarity relative to the United States is concerned. We interpret this as evidence that primarily the United States is responsible for the association between UNSC membership and increased access to IMF resources and, thus, as support for the view that the United States is the dominant shareholder in the IMF.

The remainder of this paper proceeds as follows. We develop our theoretical argument in section 1.2. Section 1.3 provides some background on the IMF and the UNSC. In section 1.4, we present the new dataset on UNSC voting behavior along with the other data used for the empirical analysis as well as our method of estimation. The results of this analysis are presented in section 1.5. Section 1.6 concludes this chapter.

1.2 The Argument

1.2.1 *Bilateral and Multilateral Aid*

A number of recent papers have investigated the conditions under which donors prefer bilateral over multilateral aid. According to the standard view, multilateral aid allows different donors to share the burden of aid-giving, at the cost of losing control over how exactly the aid is spent (Milner and Tingley 2013; Reinsberg, Michaelowa, and Knack 2017).³³ To the extent that donor preferences are aligned, donors prefer multilateral aid over bilateral aid, as it is cheaper and more cost efficient compared to fragmented aid from different donors (Milner and Tingley 2013). As holds true for multilateral cooperation at large, multilateral aid can realize efficiency gains, pool risks, materialize economies of scale, and encourage wide cost sharing (Abbott and Snidal 1998). Over bilateral aid, on the other hand, donors have more direct control and can, thus, use it as a tool to promote their own political interests in other countries. Evidence on political motivations in bilateral aid giving abounds (for a survey of this literature see, e.g., Hoeffler and Outram 2011; Fuchs, Dreher, and Nunnenkamp 2014).

Overall, the literature sees multilateral aid as less politicized than bilateral aid and as more effective (Derek 2008; Milner 2006; Schneider and Tobin 2016). Governments use the two types of aid as substitutes to achieve the same foreign policy goals, on average (McKeown 2009; Milner and Tingley 2013; Schneider and Tobin 2016). The results of this literature, however, stand in stark contrast to the literature that focusses on aid from multilateral organizations.

According to this literature, there is now a good deal of evidence that the United States uses its influence at multilateral organizations like the IMF and the World Bank to favor governments of developing countries it considers strategically important. Again, anecdotal evidence abounds (e.g., Andersen, Harr, and Tarp 2006; McKeown 2009). The first scholar to provide systematic quantitative evidence is Thacker (1999), who shows that IMF programs are more likely to go to governments that move towards the United States in terms of their voting at the United Nations General Assembly. Dreher and Sturm (2012) show that the correlation holds across the G7

³³ These costs can be minimized by delegating to an international organization with an aid portfolio that closely matches the donor's preferences (Schneider and Tobin 2016).

countries, while Copelovitch (2010a) stresses the importance of the G5 as a group. Stone (2002; 2004) shows that governments favored by the United States in terms of bilateral aid, which is well-known to be influenced by international politics, receive lighter punishments for noncompliance with IMF conditionality. According to Stone (2002, 62), “[a]lthough the United States holds a minority of votes, it does indeed call the shots at the IMF, as critics allege.” Fratianni and Pattison (2005) summarize evidence showing that the G7 are in control of the IMF on the most important issues and that staff autonomy is restricted to areas that are of marginal interest to its shareholders. This conclusion is consistent with Stone’s (2008; 2011) ‘informal governance’ model, according to which powerful shareholder use their informal power to intervene in IMF decision-making in cases that are of strategic interest to them and in normal times leave the organization governed by its formal rules (see also chapter 2). In addition to the IMF, there is also substantial evidence that the major shareholders’ political interests are also reflected in World Bank decisions (e.g., Kersting and Kilby 2016b; Kilby 2009; 2013).

The two strands of literature combined provide an interesting puzzle. The literature on donors’ choice of multilateral versus bilateral support sees multilateral aid mainly as a way to share the burden of aid-giving and realize efficiency gains, while the literature on the IMF and the World Bank characterizes the organizations as political tools of their major shareholders, and in particular of the United States. How can multilateral aid be perceived as non-political and highly politicized at the same time? If bilateral aid is used as a political tool, why is there also evidence on political interests in multilateral aid? If both bilateral and multilateral channels are used to shape political developments in other countries, how do governments decide between them?

1.2.2 The Dirty-Work Hypothesis

We expect governments to prefer multilateral aid over bilateral aid when the benefits of doing so exceed the costs. As politicians are interested in winning elections and gaining popularity, a key benefit we see in supporting recipients via multilateral organizations is the ability to hide one’s support for an unpopular recipient (Vaubel 1986).³⁴ Governments might even publicly shame an

³⁴ As a further benefit that is not central to our argument, multilateral aid reduces the scope for citizens to compare the success of aid across different donors, so that donor governments cannot be held accountable for failures (“yardstick

international organization for supporting regimes that are unpopular with its domestic audience, while secretly pressing international organizations to provide support. As Voigt and Salzberger (2002, 295) point out, “an ability to shift blame will bring a legislator to prefer delegation.” In line with Vaubel (1986) as well as Voigt and Salzberger (2002), we thus expect aid to be channeled via an international organization when it is unpopular at home. The main costs of channeling the aid to multilateral organizations is the damage this imposes on the reputation of the organization as politically independent and neutral actor. Only when the issue at stake is sufficiently salient to the donor, and bilateral action would be sufficiently costly, will the benefits of using the organization exceed the costs. We thus expect multilateral aid to prevail when donor governments want to channel resources to countries that its own public would not want it to support. We expect governments to use international organizations to shield them from the views of their domestic audiences.

Our argument rests on two main pillars which we briefly discuss in turn. First, we argue that domestic audiences in donor countries have sufficiently strong preferences against supporting certain types of regimes with aid for their governments to take note. Second, we argue that domestic audiences know little about the decision-making processes of multilateral organizations. Not least because these processes are often non-transparent for the public, they perceive these organizations as largely independent organizations so that the role of their own government in granting aid to a specific country largely goes unnoticed.

We expect governments to be sensitive to the foreign policy preferences of their domestic audience (Moravcsik 1997). Recent evidence suggests that the public has an aversion to providing bilateral aid to hostile countries (Heinrich and Kobayashi 2018). As they point out, “people believe that aiding such regimes is morally unacceptable as it signifies complicity in promoting harmful policies” (2018, 2). In short, “voters abhor giving aid to such regimes” (Heinrich and Kobayashi 2018, 3). What is more, publics prefer humanitarian aid over political aid (Milner 2006). Among the examples given by Milner (2006, 118), Lumsdaine (1993, 43) points out that “[p]ublics when asked consistently said aid should go to needy countries that would use it well rather than

competition,” Vaubel 2006). What is more, of course, multilateral aid is “cheaper” from an individual donor’s perspective.

being used to promote narrow national interest.” She also cites McDonnell et al. (2003, 20) pointing out that “[i]n most cases, the overwhelming [public] support for foreign aid is based upon the perception that it will be spent on remedying humanitarian crises.” According to Milner and Tingley (2015, 3), “[e]fforts at using positive inducements, such as economic aid and trade, have often found bitter opposition within the United States.” They point out that the US President is unsuccessful in obtaining congressional approval on a crucial foreign policy in around one third of the time. What is more “Congress, interest groups, and the public constrain the president in foreign policy, but, crucially, the extent of this constraining varies across policy instruments” (Milner and Tingley 2015, 13). Overall, domestic audiences care about the type and recipients of aid, so that it becomes difficult to channel political aid to hostile recipients.

Domestic audiences in donor countries know little about the IMF. As one example, consider IMF Managing Director Christine Lagarde’s threat to pull out of Greece ahead of a 2016 meeting of Eurozone finance ministers (discussing an extension of a Greek bail-out package). Her threat was taken at face value in newspapers discussing the bail-out.³⁵ The fact that Christine Lagarde could hardly take such decision against the will of the major IMF shareholders has largely gone unnoticed. According to Vaubel (1986) voters are to some extent rationally ignorant, so that governments can use international organizations to increase voters’ information costs. As Vaubel (2006, 134) points out, “[m]ost parliaments and voters know very little about their executive directors.” He cites Gerster (1993: 101), according to whom “the manner in which executive directors and their domestic authorities regularly report to parliament and the public on their participation in the Fund and the Bank is [...] only poorly developed.” Gerster (1993, 107) concludes that “there is an institutionalized bias against public accountability of executive directors.” In addition, Grigorescu (2013) finds that a certain ‘culture of secrecy’ is visible in many international organizations. He argues that this is, amongst others, the case as “[g]overnments want to maintain their monopoly over information to control IOs [and] [i]nformation on bargaining is too sensitive to release to domestic audiences” (Grigorescu 2007, 630). Similarly,

³⁵ One representative example is a May 6, 2016 article in The Guardian, see <https://www.theguardian.com/world/2016/may/06/imf-threatens-greece-eurozone-christine-lagarde> (accessed May 10, 2018).

Stasavage (2004) suggests that such secrecy allows member states to blame the international organization for unpopular decisions.

We expect governments to make use of voters' lack of knowledge about international organizations and use them to hide unpopular policies. Governments collude with pressure groups at the expense of their voters. They hide the costs of concessions to interest groups (such as domestic banks) and shirk domestic responsibilities for unpopular policies (such as a bail-out).³⁶ International organizations raise the costs of information for voters, but not for well-organized interest groups. As a consequence, "[t]o the extent that foreign aid is unpopular in the donor countries, the multilateral aid institutions help the national politicians to collude against their voters and to avoid responsibility for specific grants and the inevitable scandals" (Vaubel 1986, 50). The longer chain of control along the principal-agent relationship from donor populations to recipient populations weakens citizens' ability to achieve their will (Nielson and Tierney 2003; Vaubel 2006; see also chapter 4).

1.2.3 Contributions to Previous Literature

Our argument is closely related to recent literature on the allocation of aid. As Heinrich and Kobayashi (2018) point out, "by simply giving less aid, the donor can *distance* itself from the nasty policies of the recipient." According to our argument, while the donor will indeed give less *bilateral* aid to unpopular regimes, we expect the donor will simply use *multilateral* aid instead. Governments give "aid to nasty regimes because they tend to be the optimal target to bribe for concessions" (Heinrich and Kobayashi 2018, 3; also see Bueno de Mesquita and Smith 2010). Given this, it would be surprising if donor governments would not try to find alternative channels of influence in cases in which bilateral aid is unpopular to use. Indeed, Heinrich and Kobayashi (2018, 5) posit that "donor governments could attempt to *divert* the public's attention from the recipients' nasty policies and thus not have to give up the policy concessions" that it can buy with its aid. We argue that multilateral aid achieves exactly this.

³⁶ For instance, Germany's insistence on involving the IMF in the highly unpopular bail-out for Greece during the European debt crisis can be considered from this perspective.

The importance of international organizations' "laundering function" has also been discussed before. In addition to Vaubel (1986) and Voigt and Salzberger (2002), Abbot and Snidal (1998) come to mind. However, there are subtle but important differences to what we argue here. According to Abbott and Snidal (1998), states structure international organizations so that they further their powerful members' interests but also incentivize weaker states to participate. They designed the IMF so that, on average, neutral economics guide its policies, but in a way that they can use it for their own geostrategic interests in cases that are important to them. In line with what we argue here, Abbott and Snidal (1998, 19) observe that "[p]owerful states face a tension between the immediate advantages of dirty laundering versus the long-run costs of jeopardizing IO independence." According to Abbott and Snidal (1998, 18), however, this function mainly serves as a tool to implement policies in recipient countries without being blamed by recipient audiences. They expect multilateral action to reduce the impact of domestic lobby groups, and leading to *less* politicized actions. This is the opposite of what we expect in this paper. The fact that unpopular policies can be hidden from *domestic* audiences should strengthen the role of domestic lobbies and lead to *more* politicized actions.

Our characterization of the donor-recipient relationship in a principal-agent framework is likewise not new. According to the seminal study by Milner (2006), multilateral aid is used to signal to its own voters the donor governments' commitment that the aid is used for non-political or non-commercial goals. She points out that "[g]iving (more) aid to a multilateral forum ties the leader's hands relative to that aid but also makes the voters more likely to approve of greater aid overall" (Milner 2006, 119). Our argument is the opposite. Rather than tying donors' hands to humanitarian goals, multilateral aid allows donors to exert political influence in cases this would be too costly to do bilaterally.

Hicks, Parks, Roberts, and Tierney (2010) focus on the principal-agent relationship as well. They argue that donors use multilateral aid to tie their hands *ex ante* to be able to provide public goods. They argue that the threat to withhold bilateral aid from strategic allies in terms of non-compliance with developmental policies would not be credible, so that donors use multilateral

aid in case they aim to link their aid to developmental goals.³⁷ Again, this is contrary to our own expectation, where we expect multilateral aid to be as politicized as bilateral aid.

We would like to emphasize that our theory does not contradict these previous contributions. Rather, we would expect the different theories to be all true, but at different points in time. In order to exploit an international organization's reputation as an independent actor, major shareholders must invest in such reputation and refrain from interfering with its policies too frequently. According to Milner (2006), governments channel resources through multilateral organizations to assure voters the aid is beneficial, *on average*. She shows that donors give aid via non-state actors when recipient country government quality is low, so as to maximize the impact of aid. We argue that these same governments can use multilateral organizations for their geostrategic purposes *in specific cases* of importance to them. They thus create multilateral organizations that give their publics the impression of impartiality and benevolence but use these organizations for their "dirty work" when needed. The importance of political donor motives in the allocation of average multilateral aid will be less easy to detect compared to bilateral aid. We need to focus on specific cases to do so. These cases are situations where donors have a strong interest to give aid to countries which they would not want to be seen giving to.

Finally, it is important to compare our theory to the insights of the burgeoning literature on informal governance (Stone 2008; 2011; 2013). As we do, this literature sees an important role for international organizations in being useful in particular cases, while being sufficiently technocratic to incentivize minor powers to participate in them. As Stone (2008, 590) explains, "[i]nformal influence must be exercised with discretion, however, in order to avoid undermining the legitimacy of the organization." He argues that "[t]he other leading states tolerate these practices, so long as they are not exploited too frequently" (p. 590). According to Stone (2013, 125-126), "the design features of international organizations are equilibrium outcomes of a game that balances the power and interests of the leading state, or group of states, and the rest. Informal meddling by great powers is made costly enough, usually in terms of bad publicity, to keep manipulation from becoming routine." Stone's expectations sound familiar but are different from

³⁷ The Samaritan's Dilemma is also analyzed in Hagen (2006). In Hagen's model the donor is highly interested in providing a collective good, so that it is not credible to withdraw aid in case of non-compliance.

what we argue here. According to Stone, powerful shareholders cannot intervene in international organizations too often, as otherwise the other member states would no longer “tolerate these practices” (Stone 2008, 590). According to the theory we test in this paper, intervention in international organizations’ policies is costly not because it affects other countries’ willingness to engage in these organizations, but because such intervention taints the organizations’ legitimacy in the eyes of *domestic* audiences and consequently makes them less useful tools of foreign intervention.

We would like to conclude this section in pointing out that the argument we discuss in this paper can explain a number of puzzling results in the recent aid allocation literature. We introduced two of them above: First, according to Milner (2006, 132), “it is hard to understand” why right-wing governments give more multilateral aid than left-wing ones. However, given that right-wing governments typically pursue more aggressive foreign aid policy than left ones (e.g., Milner and Tingley 2010), this is exactly what our theory would predict.

Second, the same holds for Strand and Zappile’s (2015) analysis that shows that countries receiving bilateral economic aid – their proxy in a donor’s interest in a specific country – obtain less aid from multilateral development banks. While their result is contrary to the authors’ own expectations, it is exactly what we would expect to find: multilateral aid is used in situations where bilateral aid is difficult to give. Rather than being substitutes for each other, bilateral and multilateral aid are used as complements.

A third puzzle our paper speaks to has recently been raised by Schneider and Tobin (2016). They find that donors chose among a number of different multilateral organizations, so that the preferences of the organization about how to allocate aid match those of the donor. Schneider and Tobin (2016, 658) conclude with a puzzle: “If [...] governments pursue goals with bilateral and multilateral aid that are largely similar, why do they use both venues instead of either going fully bilateral or fully multilateral? The similarity of bilateral and multilateral aid portfolios provides an important puzzle that needs to be addressed in future research.” In this paper, we provide a simple answer. A large number of international organizations hide the costs of unpopular policies. While governments can obtain the same allocation of their aid via multilateral and bilateral aid alike, their support is hidden when using the former, but highly visibly when

using the latter. Governments use bilateral aid to signal their support of a recipient to their domestic audiences and use multilateral aid to hide such support. Even when the preferred allocation of aid is exactly the same, both types of aid continue to serve their purpose.

1.3 The IMF and the UNSC

In order to test our theory, we look at two international organizations: the IMF and the UNSC. We focus on UNSC voting, because donor governments care a great deal about these votes. UNSC voting is thus sufficiently salient for us to expect our theory to have some traction. We focus on the IMF rather than all multilateral aid because the United States has substantial influence over IMF loans and IMF loans are sufficiently large to be considered as substitutes for US bilateral support – both from the donor and the recipient perspective.³⁸ We discuss them in turn and begin with how the United States can go through the IMF to exert political influence.

First, power on the IMF Executive Board is explicitly linked to the financial contributions that they provide to the organization. With nearly 17 percent of the total votes, the United States has veto power over certain decisions that require an 85 percent majority. Beyond this formal power, the United States also has a degree of informal influence over the institution (Stone 2008; 2011; chapter 2). The IMF Executive Board typically operates according to a consensus rule, which gives the management agenda-setting power. The management, in turn, is subject to pressure from the United States, both because proposals are shaped to avoid US opposition and because – as the IMF headquarters are located in Washington – representatives of the US Federal Government are actively involved in important IMF meetings. A further channel of US influence is through the US Congress, which must periodically approve increases in US contributions to the IMF (Broz 2008; 2011; Broz and Hawes 2006). As the United States is the largest contributor and influences other contributors on whether to approve increases, IMF management and staff pay due attention to the preferences of US policy-makers.

³⁸ We test robustness focusing on the World Bank rather than the IMF. In the empirical section we also turn our attention to organizations where the United States is less powerful and use them as placebo tests. See also the introduction and chapter 3 for a discussion of the global relevance of IMF loans.

While the IMF can be used for exerting influence in many regards, we focus on buying favors in one of the world's most powerful international institutions, the United Nations Security Council. The UNSC is the primary organ of the United Nations with responsibility for the maintenance of international peace and security. The Security Council is the only UN body with the power to make binding resolutions. It may adopt legally binding measures in order to maintain or restore international peace – including the investigation of international disputes, the imposition of economic sanctions, and the use of armed forces.

Historically, when the United States acts in concert with the UNSC, it bears a smaller share of the burden of international campaigns (Hartley and Sandler 1999). So it stands to reason that the United States should care about UNSC resolutions. Yet, the elected members of the UNSC have a limited impact on passing them. Veto power on the Security Council belongs to each of the five permanent members (the victors of World War II: China, France, Russia, the United Kingdom, and the United States). The ten elected members, which represent various regions of the world, are rarely pivotal (O'Neill 1996). Still, nine total votes are required for a resolution to pass, and since permanent members sometimes abstain, upwards of four out of the ten elected members must vote in favor.

A likely reason to care about the votes of elected UNSC members, beyond their formal voting power, is legitimacy (Hurd 2007; Voeten 2005; Vreeland and Dreher 2014). As Hurd (2007) explains, the elected members serve the purpose of giving voice to the “rest of the world” on the Security Council. And the legitimizing effect of the Security Council extends beyond the international level and into domestic politics: Chapman and Reiter (2004) find that US Presidents enjoy higher levels of public support for actions endorsed by the UNSC, an effect not found for any other international organization they test. In the absence of UNSC legitimacy, domestic public support might be more difficult to achieve and US Congress might be more recalcitrant (Hurd 2007; Hurd and Cronin 2008; Voeten 2001). Voeten (2001) provides examples. He cites the memoirs of James Baker (Baker 1995, 278), emphasizing domestic support to be the main reason for the US government to seek a multilateral solution to the Gulf War. He also cites Malone (1998, ix), arguing that it was easier for the Clinton administration to secure the support of the UNSC as compared to that of the US Congress. Mikulaschek (2017b) shows that the signal incorporated in

UNSC resolutions is most valuable in terms of popular support when it is unanimous, as it signals consensus among foreign elites. There is thus a premium for getting unanimous votes, and every single vote matters.³⁹

Although no one has systematically studied UNSC voting behavior to see if it is related to aid, Kuziemko and Werker (2006) show that temporary members on the UNSC receive substantial increases in US aid. As their argument goes, the United States desires influence on the UNSC. The governments of some developing countries may care more about the aid than they care about the global security issues considered important by the US government. If major donors like the United States value the voting behavior of developing countries more than their aid, votes-for-aid trades are possible. Like all subsequent studies on the benefits that come with temporary UNSC membership, Kuziemko and Werker test their vote-buying argument without data on actual voting behavior.⁴⁰ Among these studies, the ones that are most closely related to our empirical analysis are Dreher, Sturm, and Vreeland (2009b; 2015), who show that elected members of the UNSC are more likely to participate in IMF programs and the conditions attached are fewer in number and narrower in scope than for other countries. They take this as evidence of IMF favoritism for UNSC members.

In addition, there is a substantial body of circumstantial evidence that the United States regularly engages in vote-buying at the Security Council. Eldar (2008) provides examples. For one, the United States promised to support a World Bank loan for China in return for support on the Security Council for the first Gulf War in 1991. As another example, the United States helped China obtain World Bank loans (and provided security guarantees regarding Taiwan) in return for allowing a UNSC resolution to restore democracy in Haiti in 1994. More generally, Eldar (2008, 17) argues that in order to get UNSC support for the Gulf War, the United States made “a promise of financial help to Columbia, Côte d’Ivoire, Ethiopia and Zaire; a promise to the USSR

³⁹ The legitimacy may derive from the idea that UNSC members have been elected to represent their respective regions and also from the idea that UNSC votes represent informed decisions. Members of the UNSC have access to sensitive documents and private discussions regarding the importance of taking international action. For more on these informational theories, see Fang (2008), Chapman (2007), and Thompson (2006a).

⁴⁰ These studies include Dreher, Sturm, and Vreeland (2009a; 2009b; 2015); Kilby (2013); Mikulaschek (2017b); and Reynolds and Winters (2016). Vreeland and Dreher (2014) use a preliminary version of the dataset that we introduce in this paper in some regressions.

to keep Estonia, Latvia and Lithuania out of the November 1990 Paris Summit conference and to persuade Kuwait and Saudi Arabia to provide it with hard currency.” He further argues that before the second Gulf war, the United States again attempted to buy votes of temporary UNSC members. Another example was published in the memoirs of US Secretary of State James Baker. Baker points out that the United States cut all foreign aid to Yemen when their government failed to support the UNSC resolution that authorized the use of force in Iraq in 1990 (Baker 1995, 278). Baker was quoted saying “[t]hat is the most expensive vote you have ever cast” and the United States subsequently cut all of its USD 70 million in aid (Bandow 1992).

The most recent ‘smoking gun’ is from late 2017: On December 18, the United States vetoed a Security Council resolution that called for the withdrawal of US President Donald Trump’s recognition of Jerusalem as the capital of Israel. The resolution was supported by all remaining 14 UNSC members. Two days later, Donald Trump threatened to cut foreign aid to countries that vote against the United States at the United Nations. He stated: “these nations that take our money and then they vote against us at the Security Council [...]. We’re watching those votes. Let them vote against us, we’ll save a lot.”⁴¹

In this study, we shed light on the general patterns behind such remarks by means of the following data and method.

1.4 Data and Method

1.4.1 A New Dataset on UNSC Voting Behavior

The previous literature on vote buying in the UNSC primarily built on a binary variable indicating UNSC membership for a given country i in a year t (Kuziemko and Werker 2006; Dreher, Sturm, and Vreeland 2009a, b). Testing our theory however requires data on how countries *voted* during their time as temporary members. One of this paper’s contributions to the literature is to introduce new data that allow such tests.

⁴¹ The full passage of the statement reads: “For all of these nations that take our money and then they vote against us at the Security Council or they vote against us potentially at the [General] Assembly. They take hundreds of millions of dollars and even billions of dollars and then they vote against us. Well, we’re watching those votes. Let them vote against us, we’ll save a lot. We don’t care.” <https://www.theguardian.com/us-news/2017/dec/20/donald-trump-threat-cut-aid-un-jerusalem-vote> (last accessed: 28 April 2018).

We have collected data on voting behavior in the United Nations Security Council from various sources. Voting behavior on successful resolutions is available from the United Nations Bibliographic Information System (UNBISNET).⁴² We added information on vetoed resolutions from the official United Nations veto list (UN document A/58/47, Annex III, for the 1946-2004 period), from archival research in the UN Library in Geneva, and from the online archive of the Dag Hammarskjöld Library.⁴³ Most difficult to obtain are data on failed majorities. We include voting behavior on these failed majorities obtained from our archival research in the UN library and from searching for keywords in UNSC meeting minutes.⁴⁴

Overall, we obtained data on the votes of all UNSC members in 2,524 decisions (2,259 resolutions, 230 vetoes, and 35 failed majorities) over the seven decades of the 1946-2015 period.⁴⁵ This translates into 36,460 individual votes. We also record the title of the proposed resolution, its number (if it passed), and the date of the decision. In addition, we collected and coded additional resolution-specific information to categorize the proposed resolution's policy area and to proxy its political importance. We describe these data in more detail in Appendix 1.A.

First, we use these data to calculate a member -year specific count of how often member countries voted against the United States in the UNSC in a given year. Figure 1.1 shows a histogram of the distribution of this count variable. As in the UNSC (in contrast to the UNGA) the vast majority of resolutions are unanimously adopted, this variable is positively skewed and often equals zero. In light of this distribution, we code two variables *UNSCall_{it}* and *UNSCnotall_{it}* that indicate whether or not a specific recipient country that served on the UNSC has voted in line with the United States on all votes in a year.⁴⁶ Given the large number of unanimous decisions, one disagreement per year indicates a notable deviation in articulated preferences over foreign policy. Furthermore, Mikulaschek (2017a) shows that domestic audiences value unanimity in the

⁴² See <http://unbisnet.un.org/> (last accessed May 3, 2018).

⁴³ The archive of the Dag Hammarskjöld library is available online: <http://research.un.org/en/docs/sc/quick/> (last accessed May 3, 2018). We also identified one veto that was cast in a secret vote via searching for keywords in UNSC meeting minutes.

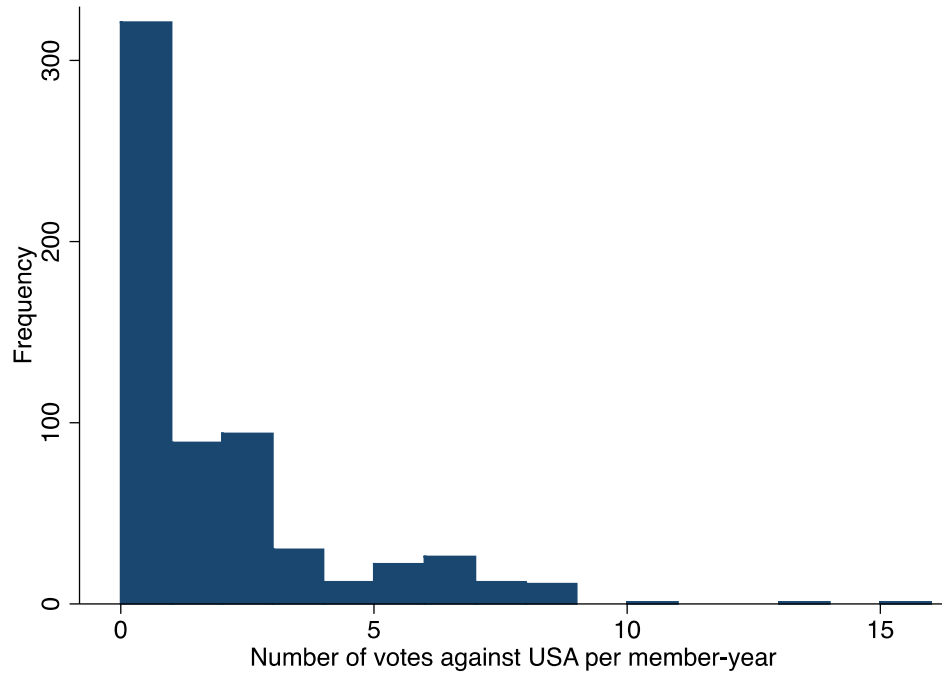
⁴⁴ Unfortunately, we cannot guarantee that the data on failed majorities are complete.

⁴⁵ In our dataset the indicator for temporary UNSC membership is coded one for 620 observations. This reflects the fact that the UNSC had six temporary members between 1946 and 1965 and ten such members between 1966 and 2015.

⁴⁶ For temporary members, the mean of *UNSCall_{it}* is 1.42 (the standard deviation is 2.19). Of 620 member-year observations this variable equals one in 321 cases.

UNSC,⁴⁷ so that the United States is likely to have an interest in temporary members *always* agreeing. Thus, we expect this binary variable of voting alignment to capture much of the variation in voting behavior that we are interested in.

Figure 1.1 – Voting against the United States in the UNSC



Notes: The figure shows the histogram of the number of votes per temporary-UNSC-member-year where a country's votes differ from those of the United States.

To exploit more information compared to what is contained in this binary indicator, in alternative regressions we additionally code continuous measures of country-specific UNSC voting alignment variables, following the literature on voting behavior in the UNGA. We calculate the number of votes in which a member disagrees with the United States relative to the total number of votes that were cast in a given year (*ShareAgainst_{it}*). In the construction of this variable we follow the approach proposed by Kegley and Hook (1991) for measuring voting alignment in the

⁴⁷ Mikulaschek (2017: 25) finds that “the unanimous endorsement of a U.S. military intervention by the UN Security Council increases popular support for the use of force by six to ten percentage points, in comparison to the Council’s approval of the same action despite dissent.”

UN General Assembly and discard abstentions or absences.⁴⁸ We exclude unanimous votes when we construct the share of votes against the United States. As the UNSC often decides unanimously, this ensures that we exclude decisions on relatively uncontroversial matters and thereby significantly reduce the noise in this measure of voting alignment. When running regressions with this variable, we include a binary variable indicating UNSC membership ($UNSC_{it}$) and its interaction with the share of votes against the United States ($UNSC_{it} * ShareAgainst_{it}$).⁴⁹

1.4.2 Empirical Model and Additional Data

Armed with these key explanatory variables, we turn to the first set of regressions we estimate. They are at the recipient-year-level and take the following form:

$$y_{it} = \beta_1 UNSC_{it} + \beta_2 UNSC_{it} * ShareAgainst_{it} + \beta_3 GDPpc_{it-1} + \beta_4 Population_{it} + \beta_5 War_{it} (+\beta_6 pastIMF_{it}) + \gamma_i + \tau_t + \varepsilon_{it}, \quad (1)$$

$$y_{it} = \beta_1 UNSC_{it} + \beta_2 UNSC_{it} * ShareAgainst_{it} + \beta_3 GDPpc_{it-1} + \beta_4 Population_{it} + \beta_5 War_{it} (+\beta_6 pastIMF_{it}) + \gamma_i + \tau_t + \varepsilon_{it}, \quad (2)$$

In these regressions, we consider two different outcome variables, y_{it} , that are both aid amounts to recipient country i in year t : US bilateral aid on the one hand, and multilateral IMF loans, on the other.⁵⁰

We build our regressions on those in Vreeland and Dreher (2014). IMF loans are therefore logged commitments in millions of current SDR (Special Drawing Rights, the IMF's unit of account).⁵¹

⁴⁸ Our results are robust to employing the approach proposed by Wittkopf (1973), who includes abstentions and absences and codes agreements for both countries abstaining and both being absent.

⁴⁹ In our full dataset the indicator for temporary UNSC membership is coded 1 for 620 observations. This reflects the fact that the UNSC had six temporary members between 1946 and 1965 and ten such members between 1966 and 2015.

⁵⁰ To make the sample of the two sets of regressions with the two different outcome variables comparable we restrict the sample to countries that according to the OECD are eligible to receive Official Development Assistance (ODA) in year t . As the OECD does not provide the list of ODA eligible countries for the early years of our sample, we follow the OECD definition and denote a country i in year t as ODA eligible if it has not “exceeded the high-income threshold for three consecutive years” according to the World Bank’s definition and is neither a member of the European Union nor of the G8 (OECD 2018a).

⁵¹ We add one before we take the natural logarithm to avoid losing zero observations. Note that our regressions include fixed effects for years, which capture changes in the overall level of prices (inflation). We therefore prefer to not deflate the original IMF data or convert them to USD.

IMF loan commitments are better suited to test the influence of major donors on IMF loans compared to disbursements, as disbursements are typically made in equal tranches and mainly depend on borrowers' compliance with IMF conditions. While US influence could also be important to receive loans in spite of non-compliance, compliance is likely endogenous and can depend on the borrowers' standing with major powers, their economic development, as well as on their political willingness to implement IMF-mandated policy reforms. The loan size the IMF commits to, however, is determined before the program starts. It is here that we expect US influence to be most visible.⁵² In our largest sample, the data cover the years 1960 to 2015. During this period, 143 different countries participated in IMF programs. In these countries, a total of 2,536 out of 7,352 possible country-year observations – and thus roughly a third of the years in these countries – are under an IMF program. For observations with an active loan program, the mean IMF loan size in our sample is 422 million SDR (roughly 600 million USD in 2015).

When turning to regressions of US bilateral aid, we again follow Vreeland and Dreher (2014), and measure US aid as logged disbursements (in constant 2015 million USD) rather than commitments.⁵³ Unlike loans from the IMF, disbursements of US aid follow no clear pattern relative to commitments, do typically not depend on compliance with specific ex post policy conditions, and are often substantially delayed, so that we assume favoritism to shorten these delays and thus to materialize at the disbursement rather than the commitment level. As Carter and Stone (2015) show, the US executive branch makes use of its discretion to deviate from previously committed aid levels to use aid for political purposes. Net US aid disbursement data come from the OECD and cover the 1960-2015 period. In this period, a total of 150 countries have received ODA from the United States. Of these countries, the average country has received a total of 4.6 billion USD (in constant 2015 dollars) over the entire period.

We test robustness to using various alternative measures, including binary indicators for IMF programs, IMF purchases, and US aid commitments.

⁵² The IMF usually does not disburse more than what was originally agreed upon, so political pressure is likely to be exerted when the loan size is decided. Additional regressions show that our results hold when we substitute the IMF loan variable with a binary variable indicating the start of an IMF program. This supports the expectation that political interests are exerted at the design stage of a program.

⁵³ Again, we add one before taking the natural logarithm to avoid losing zero observations.

We include a number of important control variables. Previous research has argued that the timing of being elected to the UNSC is “not random [but] largely unrelated to aid and political and economic development” (Bueno de Mesquita and Smith 2010, 72). In their analysis of the determinants of election to the UNSC, Dreher et al. (2014) find that “turn-taking is likely an exogenous source of variation” while noting that for such settings their results also “suggest the importance of controlling for population and income” (p. 80). We follow this advice and add the natural logarithm of $Population_{it}$ size and per capita GDP ($GDPpc_{it-1}$) as control variables to all regressions.⁵⁴

Dreher et al. (2014) and Vreeland and Dreher (2014) also find the involvement in warfare to reduce the likelihood of being elected to the UNSC. We therefore also add a country-year specific War_{it} indicator.⁵⁵ Furthermore, as previous participation in IMF programs is one of the strongest predictors of receiving IMF loans (Sturm, Berger, and de Haan 2005; Dreher, Sturm, and Vreeland 2009b; Moser and Sturm 2011) and increases the precision of the estimation without reducing the size of our sample, we add a variable indicating previous IMF participation in the regressions focusing on IMF loans ($pastIMF_{it}$). We include country fixed effects γ_i and year fixed effects τ_t in all regressions to rule out that time-invariant country characteristics and global trends that affect all countries equally drive the results. Estimation is by ordinary least squares (OLS); ε_{it} represents the error term.⁵⁶

Both sets of models arguably allow us to make the identifying assumption that temporary UNSC membership is conditionally exogenous. The coefficients on the membership indicator $UNSC_{it}$ will thus not be biased by endogeneity. As regards the possibility to interpret our results as causal there are nevertheless two important caveats.

First, while membership itself can be considered exogenous, UNSC voting behavior cannot. It is likely to be correlated with potential determinants of receiving aid (like a country’s general political orientation, its economic conditions, etc.). Therefore, our estimates do not allow to infer

⁵⁴ We lag $GDPpc$ by one year to avoid that any economic effects resulting from UNSC membership introduce endogeneity bias (Bueno de Mesquita and Smith 2010; Dreher, Eichenauer, and Gehring 2016).

⁵⁵ The variable is set to one for country-years with more than 1000 battle-related deaths. Removing the variable does not affect the results.

⁵⁶ Appendix 1.B reports descriptive statistics of all variables. Appendix 1.C contains sources and definitions.

whether the links between voting behavior and aid allocation are also causal. What we can test, however, is whether or not any causal effect of UNSC membership on aid allocation is driven by countries that exhibit a certain kind of voting behavior. Some countries' votes might be easier to buy, some might vote with the United States in any case for reasons we do not capture in our models. Essentially, the UNSC voting variable is an interaction between membership and voting because it is not observed for non-members. It thus indicates whether the causal effect of UNSC membership differs for countries with different kinds of voting behavior (and potentially unobserved variables correlated with it).

The second caveat concerns the order of events. We do not observe the exact order of votes and commitments or disbursements of aid, and thus cannot test whether decisions at the level of the UNSC precede decisions at the level of the IMF and the donor government. Even if we find that IMF loans or aid disbursements precede a change in UNSC membership and voting behavior we could not know whether the loan is paid as a reward or rather as a bribe. Even if the loan precedes the vote, it could well be paid in anticipation of a positive vote rather than a bribe. For testing our argument, we are interested in whether bilateral and multilateral aid allocation is influenced by geopolitical considerations. Whether aid is used to change the voting behavior of countries in the UNSC or countries are rewarded for their voting behavior is of secondary importance.

To test our core hypothesis, we modify the above model:

$$y_{it} = \beta_1 UNSC_{it} + \beta_2 UNSC_{it} * Proximity_{it} + \beta_3 UNSCnotall_{it} + \beta_4 UNSCnotall_{it} * Proximity_{it} + \beta_5 Proximity_{it} + \beta_j CONTROL_{it} + \gamma_i + \tau_t + \varepsilon_{it}. \quad (3)$$

This model differs from our baseline model (1) in that we introduce a proxy for each recipient country's political proximity to the United States – $Proximity_{it}$ – that we interact with our indicators $UNSC_{it}$ and $UNSCnotall_{it}$. We code $Proximity_{it}$ as a moving average of the share of votes that a country casts in line with the United States in the UNGA over the period from $t-5$ to $t-2$. We do not include the years of UNSC membership (t and, potentially, $t-1$), so that potential

changes in UNGA voting behavior that may result from UNSC membership do not bias the estimates.⁵⁷

We use this measure because voting positions in the UNGA have clear relevance for whether or not a country is perceived as an ally of the United States. According to the US Department of State (1985), examining UN votes makes it possible “to make judgments about whose values and views are harmonious with our own, whose policies are consistently opposed to ours, and whose practices fall in between.” A report from the same department in 2000 states that “a country’s behavior at the United Nations is always relevant to its bilateral relationship with the United States, a point the Secretary of State regularly makes in letters of instruction to new U.S. ambassadors” (US Department of State 2000).

In the regressions of US bilateral aid our theory predicts a positive coefficient for the interaction of $UNSCall_{it}$ and $Proximity_{it}$ – countries that are close to the United States should be rewarded with more bilateral aid when they vote in line with the United States in the UNSC. Conversely, we expect a negative coefficient for the same interaction in the regressions of IMF loans. This reflects our expectation that the US will buy or reward the Security Council votes of countries that are politically distant to the United States by means of IMF loans. Finally, we do not expect temporary members that vote against the United States in the UNSC to receive more aid or loans than non-members.

1.5 Results

1.5.1 UNSC Voting and Aid Allocation

Table 1.1 sets the stage. Columns 1 to 6 investigate determinants of US aid, columns 7 to 12 report the results of the analogous regressions of IMF loans. Across all regressions, richer countries receive less aid and smaller loans, at the one percent level of significance. At the ten percent level, larger countries receive more aid from the United States (while population size is not associated

⁵⁷ We prefer voting coincidence – which measures actual voting behavior on the specific topics up for voting in each year – over countries’ ideal point distance, which takes account of differences among topics over time (Bailey, Strezhnev, and Voeten 2017). Our results are however robust to using either of them.

with the size of IMF loans). The coefficient of *War* is insignificant; and countries that had IMF programs in the past on average tend to receive significantly larger IMF loans in the present.

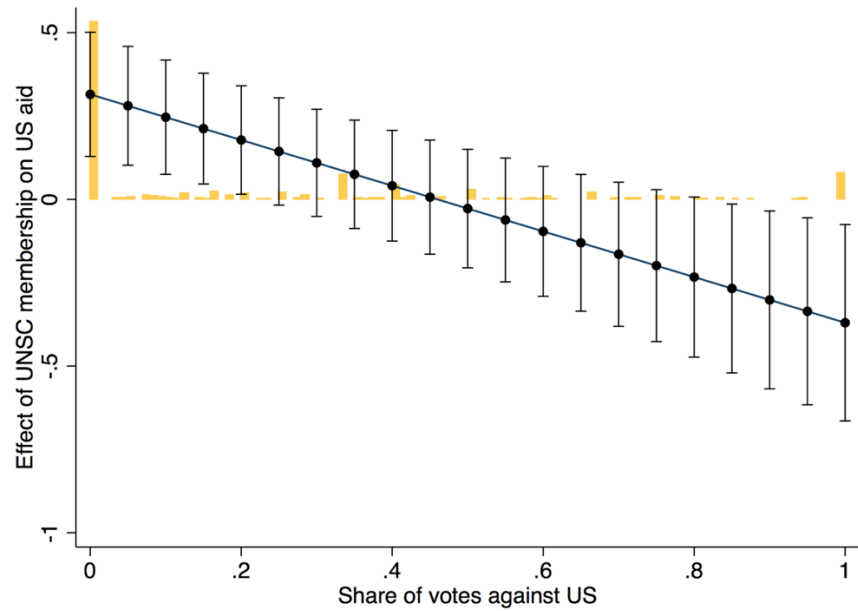
Turning to our variables of interest, we start with including a binary indicator for temporary membership in the UNSC ($UNSC_{it}$), along with our control variables, country-, and year fixed effects. While US aid (column 1) increases with UNSC membership (significant at the 10 percent level), IMF loans (column 7) do not.⁵⁸

Column 2 shows results for equation 2 above. The regression includes both the UNSC membership indicator and our first measure of UNSC voting similarity ($ShareAgainst_{it}$). As voting behavior is only observed for members, voting similarity is implicitly an interaction with the UNSC variable. Accordingly, the two variables must be interpreted jointly: The coefficient on $UNSC_{it}$ provides the estimate for the effect of UNSC membership on aid when $ShareAgainst_{it}$ equals zero. The coefficient on $ShareAgainst_{it}$ then estimates the extent to which voting against the US changes the size of the UNSC effect. The results show that both UNSC membership and its interaction with the share of votes a country casts against the United States are statistically significant at the one percent level. Figure 1.2 illustrates the marginal effect of UNSC membership on US aid along the range of $ShareAgainst_{it}$. As can be seen, the positive effect of UNSC membership on aid is positive for members that regularly vote in line with the United States and turns insignificant (at the 10 percent level) for members that vote against them in more than 20 percent of controversial UNSC decisions. The marginal effect is negative for countries that vote against the United States in at least forty percent of the votes. A significantly negative effect is visible for the very small set of observations for which we record a share of voting against the United States in controversial decisions that is larger than 80 percent.

Column 8 reports the analogous regression for IMF loans. Neither the coefficient of UNSC membership nor its interaction with the vote share are statistically significant. Figure 1.3 shows that the marginal effect of UNSC membership on the size of IMF loans decreases with the share

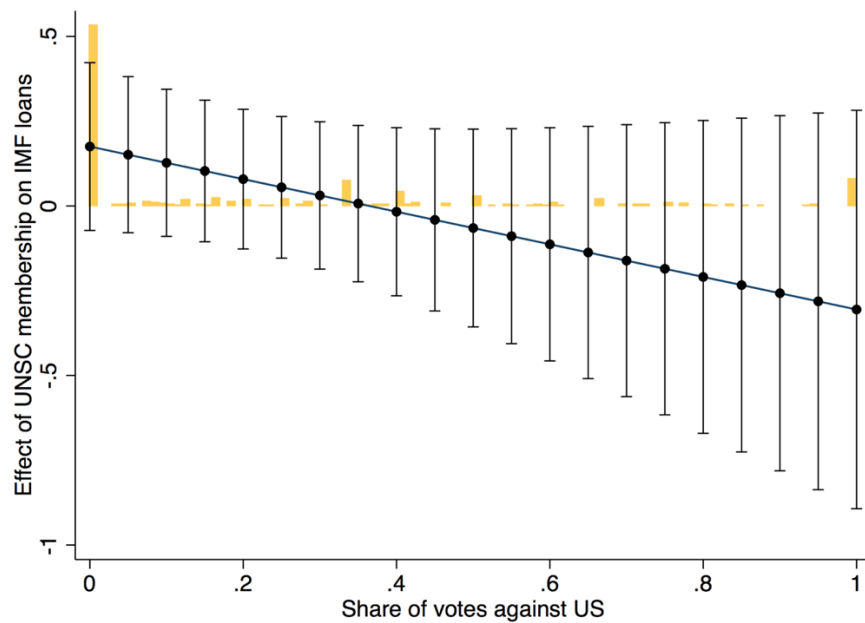
⁵⁸ Note that the latter result does not contradict previous research. Dreher et al. (2009b) find that temporary membership in the UNSC affects the probability to be under an IMF program, but not the amount of loan commitments. We turn to IMF programs below.

Figure 1.2 – Effect of UNSC Membership on US aid for Varying UNSC voting



Notes: The figure shows the marginal effect of UNSC membership on US aid for different levels of political proximity, based on the regression in Table 1.2, column 1, in concert with the 90 percent confidence interval. The histogram shows the distribution of political proximity to the United States.

Figure 1.3 – Effect of UNSC Membership on IMF loans for Varying UNSC voting



Notes: The figure shows the marginal effect of UNSC membership on IMF loans for different levels of political proximity, based on the regression in Table 1.2, column 5, in concert with the 90 percent confidence interval. The histogram shows the distribution of political proximity to the United States.

a country votes against the United States in the UNSC; it is however not significant at conventional levels. Columns 3-6 and 9-12 turn to our binary measures for voting with the United States – $UNSCall_{it}$ and $UNSCnotall_{it}$ (equation 1 above). Columns 3 and 9 focus on all votes, while the remaining columns report regressions for which we differentiate between votes according to their importance. We define importance in three different ways. First, we code the number of *Google hits* that appear when searching for “United Nations Security Council Resolution [number]” via the *Google* search engine.⁵⁹ We consider a resolution to be important if its number of *Google hits* is above the median of all resolutions of a given year. In addition, all votes that did not produce a resolution because of a veto or a failure to reach the required majority are also coded as important.

Our second definition of importance includes votes on topics related to Israel exclusively.⁶⁰ Resolutions related to Israel stand out as the single most important topic in the UNSC. 140 out of the 2524 resolutions included in our sample refer to this key US ally. Resolutions against Israel are particularly vigorously debated, typically with large majorities voting against the United States (as in the aforementioned example regarding Donald Trump’s recognition of Jerusalem as the country’s capital). The US government and public clearly care about these votes (Becker et al. 2014; Hillman and Potrafke 2015).

Our third definition of importance follows Kuziemko and Werker (2006), who argue that UNSC membership is more valuable in years in which the institution is of major geopolitical importance. They proxy importance with the number of New York Times (NYT) articles that include the words “United Nations” and “Security Council” and separate the years into different categories of importance. We do the same for our sample period based on the NYT online archive.⁶¹

⁵⁹ We do this for all resolutions from 1 to 2259 and enter the search term in quotes, thereby ensuring that the words appear in this exact order on the webpages that *Google* lists. For this we use the *Google Custom Search Engine* and run it via a program written in *Python*. See Appendix 1.A for details.

⁶⁰ To determine which resolutions concern Israel, we code the title of each resolution and search for the keywords “Israel,” “Palestine,” “Jerusalem,” and “Golan.” See Appendix 1.A for details. For future research, our data also include variables indicating resolutions that concern Lebanon, Cyprus, humanitarian issues, tribunals, sanctions, the admission of new members, and those that extend an existing resolution. This set of variables could easily be expanded.

⁶¹ Contrary to Kuziemko and Werker (2006), who do not differentiate between members with different kinds of voting behavior, we only use two instead of three categories of importance to reduce the number of categories when the voting variables are added to the regressions, but the results are qualitatively similar when three categories are used. Our cutoff value that defines the two categories is the median.

Table 1.1 – UNSC Voting and Aid, OLS, 1960-2015

	USA (1)	USA (2)	USA (3)	USA (4)	USA (5)	USA (6)	IMF (7)	IMF (8)	IMF (9)	IMF (10)	IMF (11)	IMF (12)
UNSC member	0.169* [0.097]	0.315*** [0.113]					0.063 [0.123]	0.175 [0.150]				
UNSC member * Share of votes against US		-0.685*** [0.207]						-0.481 [0.418]				
UNSC, voted all with US			0.350*** [0.115]	0.324*** [0.113]	0.607*** [0.190]				0.403** [0.169]	0.382** [0.171]	0.099 [0.178]	
UNSC, voted not all with US			0.009 [0.137]	0.030 [0.135]	0.008 [0.131]				-0.229 [0.171]	-0.214 [0.169]	-0.168 [0.228]	
UNSC, voted all with US, important years (NYT)						0.480*** [0.128]						0.576** [0.242]
UNSC, voted all with US, unimportant years (NYT)						0.143 [0.205]						0.096 [0.317]
UNSC, voted not all with US, important years (NYT)						0.253 [0.206]						-0.281 [0.263]
UNSC, voted not all with US, unimportant years (NYT)						-0.125 [0.161]						-0.202 [0.201]
GDP/capita (ln, t-1)	-0.95*** [0.280]	-0.96*** [0.278]	-0.95*** [0.280]	-0.95*** [0.280]	-1.11*** [0.309]	-0.952*** [0.279]	-0.34*** [0.127]	-0.35*** [0.128]	-0.34*** [0.127]	-0.34*** [0.127]	-0.42*** [0.147]	-0.34*** [0.127]
Population (ln, t-1)	1.264* [0.671]	1.217* [0.661]	1.254* [0.670]	1.256* [0.670]	1.287* [0.717]	1.253* [0.670]	-0.002 [0.394]	-0.014 [0.395]	-0.018 [0.395]	-0.016 [0.395]	0.010 [0.453]	-0.023 [0.394]
War	0.022 [0.249]	0.039 [0.243]	0.021 [0.249]	0.021 [0.249]	-0.136 [0.266]	0.020 [0.249]	-0.311 [0.206]	-0.311 [0.206]	-0.313 [0.206]	-0.313 [0.206]	-0.403* [0.209]	-0.316 [0.205]
Past IMF program							1.525*** [0.159]	1.532*** [0.160]	1.516*** [0.159]	1.517*** [0.159]	1.507*** [0.167]	1.517*** [0.159]
p-value (all with vs. not all with)			0.036	0.059	0.008	0.316			0.008	0.012	0.331	0.008
R-squared	0.136	0.137	0.137	0.137	0.124	0.138	0.124	0.123	0.125	0.125	0.144	0.126
Observations	6142	6066	6142	6142	4222	6142	5826	5757	5826	5826	4051	5826

Notes: OLS regressions with country- and year fixed effects. Standard errors clustered at the country-level in brackets. Significance levels * p < 0.1; ** p < 0.05; *** p < 0.01

The results of Table 1.1 paint a clear picture. Countries voting exclusively in line with the United States in the UNSC receive more aid and larger IMF loans than non-members. Specifically, US aid increases by approximately 42 percent ($e^{0.350} - 1 \approx 0.42$) for members that voted with the United States on all votes (at the one percent level of significance), but not for members that did defect at least once (column 3). Investigating the difference between the two coefficients shows that members that always vote in line with the United States receive more aid than members that do not, at the five percent level of statistical significance.

The coefficient of voting exclusively with the United States for the definition of importance based on Google hits (column 4) is similar in magnitude, with the coefficient indicating that voting exclusively with the United States increases aid by 38 percent. As expected, the effect on US aid is starkest when it comes to votes on Israel (column 5). Voting exclusively in line with the United States increases aid by more than 83 percent, at the one percent level of significance. The New York Times-based definition of importance shows that voting exclusively in line with the United States increases aid by 62 percent, while there is no significant increase in unimportant years or for countries that do not always vote in line with the United States (column 6).

Results for IMF loans are similar, both in terms of statistical significance and magnitude. Countries voting always in line with the United States on all votes receive an increase in IMF loans by 50 percent, at the five percent level of significance (column 9). The corresponding increases are 46 percent for voting on important votes according to the Google-based definition (column 10), and almost 78 percent according to the definition based on the New York Times (column 12). Only the coefficient for resolutions on Israel fails to be significant at conventional levels (column 11). Overall, our results clearly show that membership on the UNSC is associated with more aid from the United States and larger loans from the IMF – but only for countries that permanently vote with the United States.

1.5.2 Main Results

Table 1.2 turns to our core regressions (equation 3 above). Columns 1–4 investigate US bilateral aid; column 5–8 focus on IMF loans. Before introducing the measures of UNSC voting behavior, we interact *Proximity* with the simple UNSC membership indicator.

Table 1.2 – UNSC Voting and Aid to Friends and Enemies, OLS, 1960-2015

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	USA	USA	USA	USA	IMF	IMF	IMF	IMF
UNSC	-0.404**				0.314			
	[0.196]				[0.247]			
UNSC * Political proximity to US	2.010***				-1.107			
	[0.710]				[0.733]			
UNSC, voted all with US		-0.373	-0.398	-0.511		0.897**	0.872**	0.911*
		[0.256]	[0.258]	[0.359]		[0.364]	[0.361]	[0.469]
UNSC, voted all with US * Political proximity to US		1.959**	1.953**	2.426**		-1.541*	-1.543*	-2.361**
		[0.928]	[0.931]	[1.039]		[0.879]	[0.881]	[1.131]
UNSC, voted not all with US		-0.436	-0.441	-0.321		0.089	0.085	-0.574
		[0.323]	[0.321]	[0.314]		[0.300]	[0.299]	[0.472]
UNSC, voted not all with US * Political proximity to US		2.098	2.221*	2.213*		-1.740	-1.653	1.953
		[1.300]	[1.266]	[1.276]		[1.225]	[1.220]	[1.826]
Political proximity to US	3.172***	3.172***	3.176***	2.918***	0.027	-0.017	-0.016	0.221
	[1.024]	[1.029]	[1.029]	[1.091]	[0.526]	[0.531]	[0.531]	[0.570]
Votes	all	all	important	Israel	all	all	important	Israel
Observations	5113	5113	5113	3344	4982	4982	4982	3341
R-squared	0.176	0.176	0.176	0.157	0.113	0.116	0.116	0.132

Notes: OLS regressions with country- and year fixed effects. Includes GDP per capita, Population, and War. IMF regressions also include Past IMF program. Standard errors clustered at the country-level in brackets. Significance levels * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

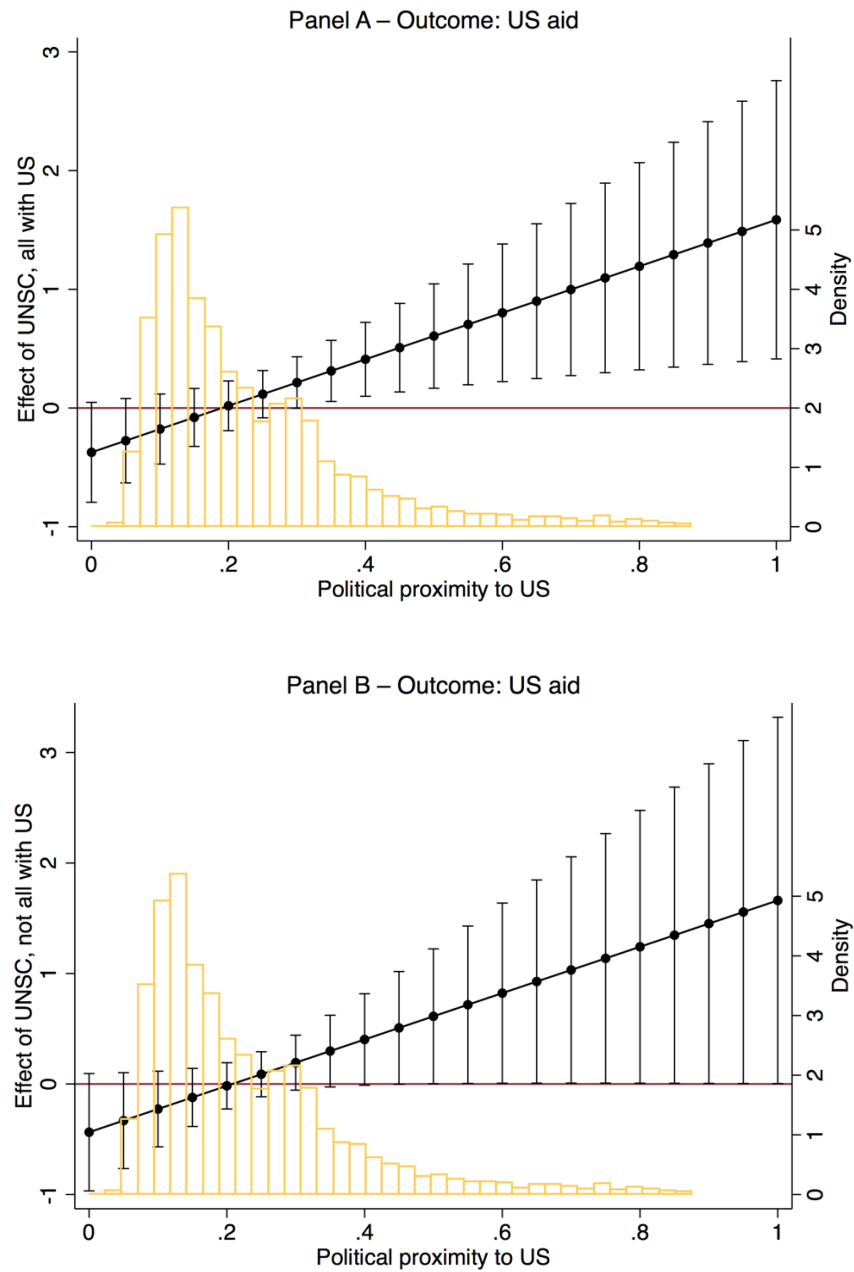
For US aid we find a positive and statistically significant coefficient on the interaction (column 1). Jointly interpreted with its constituent terms it suggests that only UNSC members that are politically close to the United States benefit from more US aid. The marginal effect of UNSC membership on US aid is positive only for countries that vote with the United States in the UNGA in more than 20 percent of the votes. When it comes to IMF loans (column 5), the coefficient is negative, as expected, but fails to be significant at the ten percent level (p -value= 0.133).

The remaining columns of Table 1.2 again separate UNSC members that exclusively voted with the United States from those that did not and test out core hypotheses. The results paint a clear picture that is in line with these hypotheses. Column 2 shows that countries that are politically close to the United States and vote exclusively in line with it in the UNSC receive more aid. This result holds when we focus on important votes in column 3 (Google definition) and column 4 (Israel definition).⁶² The results of these regressions are best illustrated graphically. Panels A and B of Figure 1.4 visualize the result for the specification including all votes, in concert with the 90 percent confidence interval (column 2). The plots show that UNSC members that always vote in line with the United States receive more US aid when political proximity to the United States is high. Countries that are politically more distant to the United States do not receive more US aid when they serve on the UNSC and always vote in line. Panel B shows a similar picture for UNSC members that do not always vote in line with the United States. While the confidence interval is wider, it seems that sufficiently close friends of the United States can benefit from US aid also when they vote against the United States in the UNSC at least once.

Columns 5-8 replicate the analysis for IMF loans. In line with our theory, we find the opposite pattern as compared to bilateral aid. The effect of receiving larger IMF loans when serving on the UNSC and consistently voting with the United States *increases* with political distance to the United States. Panel A of Figure 1.5 visualizes these results for all votes (column 6). Only countries that are politically distant to the United States receive larger IMF loans when they serve on the UNSC and – in spite of their political distance – consistently vote with the United States. Countries that do not always vote with the United States do not receive larger IMF loans. On the contrary,

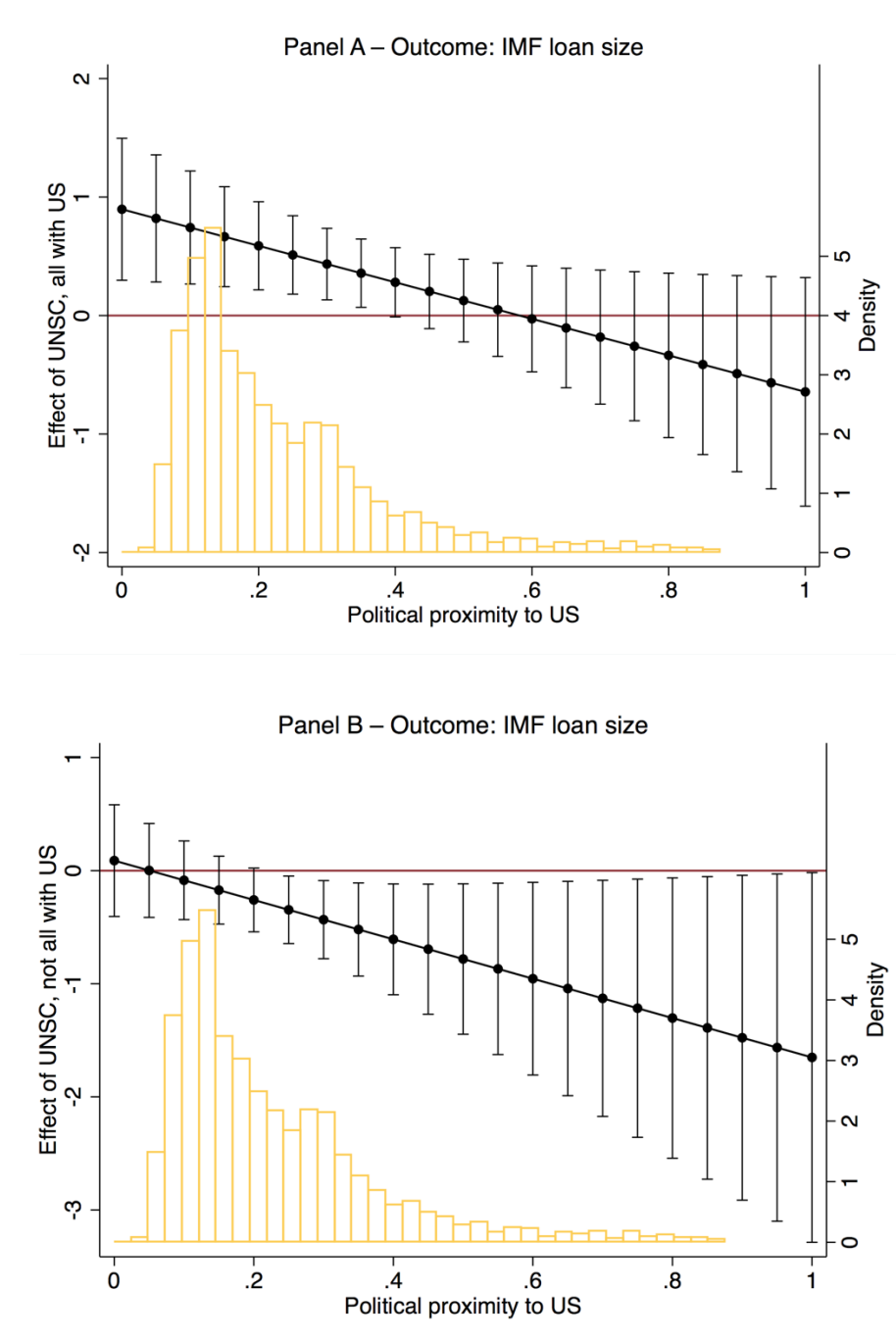
⁶² We do not include the New York Times-based definition of importance which would result in a triple interaction with eight interaction coefficients to estimate and would thus be difficult to interpret.

Figure 1.4 – Effect of UNSC Membership on US Aid for Varying Political Proximity



Notes: The figure shows the marginal effect of UNSC membership on US aid for different levels of political proximity, based on the regression in Table 1.2, column 2, in concert with the 90 percent confidence interval. Panel A focuses on the marginal effect of UNSC membership for countries that always voted with the United States in a year; panel B shows those for countries that voted against the United States at least once. The histogram shows the distribution of political proximity to the United States.

Figure 1.5 – Effect of UNSC Membership on IMF Loans for Varying Political Proximity



Notes: The figure shows the marginal effect of UNSC membership on IMF loans for different levels of political proximity, based on the regression in Table 1.2, column 6, in concert with the 90 percent confidence interval. Panel A focuses on the marginal effect of UNSC membership for countries that always voted with the United States in a year; panel B shows those for countries that voted against the United States at least once. The histogram shows the distribution of political proximity to the United States.

for close allies of the United States that vote against them, the ‘UNSC effect’ turns negative (Panel B of Figure 1.5).

We consider these results as strong evidence for the hypothesis that the channel used for buying UNSC votes depends on the donor’s political proximity to the ‘trading partner.’ In short, the United States uses bilateral aid to buy or reward the votes of its friends and multilateral aid when it comes to its enemies. Friends can be paid off openly, as reputational costs for giving aid to allied countries are low. For enemies, however, reputational costs will be high. For these countries the IMF is used for obfuscation and laundering ‘dirty work.’

1.5.3 Extensions and Robustness Tests

We extend the analysis in a number of ways. First, previous results have shown that temporary membership in the UNSC increases the probability to be under an IMF program, but not the size of IMF loan commitments. Our regressions offer an explanation for this puzzle. Given that some temporary members of the UNSC vote against the United States, average commitments for members do not necessarily increase. The frequently cited example of Yemen introduced above comes to mind. Yemen was a temporary member in 1990 and failed to support the UNSC resolution that authorized the use of force in Iraq in 1990 (Baker 1995). Though being a member of the UNSC, Yemen received less rather than more aid from the United States and the IMF. With our new data on *voting* in the UNSC, we find results for commitments that previous work was unable to detect (Dreher, Sturm, and Vreeland 2009b). Still, we think it is interesting to replicate the analysis focusing on IMF *programs* rather than loan size. These additional regressions also allow interesting insights as to whether countries with existing IMF programs receive larger loans when voting with the United States in the UNSC (intensive margin) or whether countries receive additional programs (extensive margin).

More importantly, we investigate commitments and disbursements in more detail. Remember that the above regressions focus on US aid disbursements and on IMF commitments, in line with previous work (Carter and Stone 2015; Vreeland and Dreher 2014). In this section we test whether and to what extent our theory holds for commitments of US aid and IMF “purchases” (i.e., the

amount of the loan that the program countries draws on). For completeness, we also investigate whether the results discussed for US aid above are driven by the intensive or extensive margin. Third, we investigate the allocation of World Bank aid. While this paper has focused on the IMF, the United States has substantial power over the World Bank as well (e.g., Kilby 2013b), so that our theory should hold for the Bank. The dependent variable for this regression is the World Bank's commitment of ODA to country i in year t .

Fourth, we show regressions that focus on international organizations where the United States cannot plausibly be expected to exert dominant influence on loan allocation. We investigate the effect of voting in line with the United States on aid from the Asian Development Bank, the European Bank for Reconstruction and Development, and the Islamic Development Bank. While political influences in these organizations are certainly important (Ben-Artzi 2005; Hernandez and Vadlamannati 2017; Lim and Vreeland 2013), and the United States has some influence in the Asian Development Bank (Lim and Vreeland 2013), and usually nominates the vice-president of the European Bank for Reconstruction and Development (Babb 2009), this influence is arguably not large enough to influence the allocation of their loans in line with our theory. These regressions thus offer an important placebo test. Given that the United States is unlikely to have insufficient influence over the lending patterns of these organizations, significant interactions with voting in line with the United States would cast doubt on our interpretation of results.

Finally, we extend the analysis beyond the United States. In principle, our theory applies to all donors, to the extent that they are able to sufficiently influence IMF lending. Of the other permanent UNSC members, the United Kingdom and France are often considered to be influential IMF shareholders (Copelovitch 2010a).⁶³ As previous research has focused on UNSC *membership* rather than UNSC *voting* it was thus unable to investigate which country is responsible for the link between temporary UNSC membership and increasing access to Fund's lending. Our data allow such test. We therefore show results for IMF loans, when we focus on temporary members' voting behavior relative to the other four permanent UNSC members. As far as voting relative to Russia and China is concerned we consider these regressions as placebo

⁶³ While in principle we could also analyze voting behavior relative to *temporary* UNSC members that have some influence on the IMF (e.g., Germany and Japan), this would substantially reduce the sample.

Table 1.3 – Extensions and Tests for Robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
UNSC, voted all with US	1.032*** [0.372]	0.557 [0.456]	-0.366 [0.372]	-0.721 [0.598]	-0.160 [0.170]	-0.048 [0.167]	0.574** [0.278]	-0.029 [0.202]	0.051 [0.177]	0.001 [0.163]
UNSC, voted all with US * Political proximity to US	-2.739** [1.179]	-0.313 [1.169]	1.653 [1.635]	2.296 [2.039]	0.923 [0.600]	-0.145 [0.733]	-2.225** [0.984]	0.185 [0.539]	-0.092 [0.470]	-0.021 [0.335]
UNSC, voted not all with US	0.383 [0.485]	-0.073 [0.593]	0.015 [0.313]	-0.683 [0.603]	-0.359 [0.234]	-0.405 [0.248]	-0.077 [0.265]	0.050 [0.134]	0.026 [0.034]	0.109 [0.144]
UNSC, voted not all with US * Political proximity to US	-0.143 [2.053]	-2.948 [2.402]	-0.719 [1.299]	3.209 [2.707]	1.373 [0.937]	1.482 [1.050]	0.768 [1.196]	-0.388 [0.625]	0.145 [0.363]	-0.472 [0.571]
Political proximity to US	1.510* [0.896]	-0.412 [0.859]	0.821 [0.605]	5.727*** [0.942]	1.496** [0.684]	3.137*** [1.157]	2.178** [0.839]	0.779 [0.665]	1.185** [0.560]	0.811** [0.380]
Country FE, Year FE, Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4295	1993	4795	3493	3993	4926	4926	5113	4926	5113
R-squared		0.135	0.088		0.230	0.163	0.074	0.082	0.049	0.194
Dependent Variable	IMF program	IMF loan size	IMF purchases	US aid indicator	US aid disburs.	US aid commit.	World Bank aid	AsDB aid	EBRD loans	IsDB aid
Sample	full	IMF prog. active	full	full	US aid recipient	full	full	full	full	full

Notes: OLS regressions with country- and year fixed effects. Conditional logistic regressions (conditioned on country fixed effects) if the outcome variable is binary (columns 1 and 4). Includes GDP per capita, Population, and War. IMF regressions also include Past IMF program. Standard errors clustered at the country-level in brackets. Significance levels * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

tests, as Russia and China do not have sufficient influence in the IMF to influence its lending patterns. We should thus not see the same pattern as for the United States.

Table 1.3 shows the results of the first set of additional regressions. As can be seen, our results hold for the presence of an IMF program, but neither for IMF purchases nor IMF commitments for programs that already exist.⁶⁴ UNSC voting thus seems to affect the extensive but not the intensive margin. As we argued above, we expect US influence to be more visible in IMF commitments compared to disbursements, given that disbursements of IMF loans are typically made in equal tranches and mainly depend on borrowers' compliance with IMF conditions. What is more, unlike for bilateral aid, loan commitments typically determine the maximal size of the loan, which only in exceptional cases exceeds initial commitments. For the average loan, where recipients comply with conditions and agreed upon tranches are disbursed absent any US influence, there might just not be sufficient leeway in IMF decisions for US influence to be measurable.

Table 1.3 also shows that none of the interactions is significant at conventional levels when we focus on a binary indicator for US aid recipients, additional commitments for previous US aid recipients nor overall US aid commitments. The interaction of voting exclusively with the United States and political proximity comes closest to statistical significance for loans to preexisting recipient countries (column 5, p -value = 0.126). In concert with our main results regarding aid disbursements, UNSC voting seems to affect the intensive margin but not the extensive margin. As argued above disbursements of US aid, unlike IMF loans, follow no clear pattern relative to commitments, do typically not depend on compliance with specific ex post policy conditions, are often substantially delayed, and can easily exceed initial commitments of aid. This is in line with arguments in Carter and Stone (2015, 15) who explicitly design their variables so that they "can be interpreted as discretionary deviations by the executive branch from appropriated aid levels."

⁶⁴ In additional regressions, we find that the results for IMF programs hold if we only consider the start of such programs. (This alternative indicator variable is set to one only for program starts rather than for the whole period in which the IMF program is active.) This supports the view that political influence is particularly important when IMF programs are prepared and decided upon. See also Kilby (2013b) for related evidence linking shorter preparation periods of World Bank projects to US political interests.

The further results summarized in Table 1.3 show that our results for the IMF hold for the World Bank. Countries that vote always with the United States in the UNSC receive more World Bank aid if they voted less than 25 percent of the times with the United States in the UNGA before entering the UNSC.

The results also show that UNSC voting behavior is not associated with loans from the ADB, EBRD, and IsDB, regardless of a recipient country's proximity to the United States. Given that the United States does neither have sufficient influence nor interest in these international organizations to shape their allocation of loans, this result is in line with expectations.

Table 1.4 investigates the influence of the other permanent UNSC members on IMF loan allocation. We replicate the baseline specification of equation 1 (*voted all* vs. *voted not all*, Table 1.1 column 9) for all permanent UNSC members. Column 1 initially reports the regression results for the United States for comparison. The results in the subsequent columns show that voting in line with the United Kingdom, France, China, and Russia in the UNSC is not associated with larger IMF loans. This suggests that the vote buying activities of the United States are behind the link between temporary UNSC membership and increased access to IMF loans.

Table 1.4 – Who Controls the IMF?

	(1)	(2)	(3)	(4)	(5)
	<i>j</i> = US	<i>j</i> = UK	<i>j</i> = France	<i>j</i> = China	<i>j</i> = Russia
UNSC, voted all with <i>j</i>	0.403**	0.206	0.146	0.054	0.182
	[0.169]	[0.158]	[0.137]	[0.115]	[0.162]
UNSC, voted not all with <i>j</i>	-0.229	-0.260	-0.308	0.112	-0.134
	[0.171]	[0.179]	[0.192]	[0.386]	[0.193]
Country FE, Year FE, Controls	Yes	Yes	Yes	Yes	Yes
Observations	5825	5825	5825	5825	5825
R-squared	0.125	0.124	0.124	0.124	0.124

Notes: Dependent variable: IMF loan size (ln). OLS regressions. Controls include GDP per capita, Population, War and Past IMF program. Standard errors clustered at the country-level in brackets. Significance levels * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

1.6 Conclusion

This paper investigates how major shareholders can exploit international organizations to hide their policies from domestic audiences (“dirty work”). The argument that international organizations can be used in this way goes back to Vaubel (1986). The theory explains how multilateral organizations can be used to hide governments’ costs of concessions to interest groups when such concessions are unpopular with domestic audiences. However, the theory has never been confronted with data. When “national politicians [...] try to get rid of their ‘unpleasant’ activities, their ‘dirty work’” Vaubel (1986, 48), then the allocation of multilateral aid should be in line with the political interests of their major shareholders. Previous empirical analyses confirm that IMF and World Bank lending indeed follows the interests of their major shareholders (Dreher et al. 2009a; 2009b; Vreeland and Dreher 2014).

The recent literature investigating the allocation of bilateral and multilateral aid, however, comes to the opposite conclusion. It shows that multilateral aid is less political and more effective compared to bilateral aid (Milner 2006; Schneider and Tobin 2016). The results of these literatures stand in some contrast to each other and thus offer an interesting puzzle.

Our theory addresses this puzzle and reconciles the two strands of literature. We argue that major powers exert influence bilaterally when domestic audiences view the intervention favorably. When domestic audiences are more skeptical of a recipient, favors are granted via international organizations. They will use their power over international organizations selectively, so that the average loan is not affected by donors’ political considerations in an obvious way. The previous literature indeed investigated the *overall* allocation of multilateral aid versus bilateral aid. It is thus unsurprising that politics turned out as less important in the allocation of multilateral aid. We are instead not interested in overall aid portfolios, but in whether international organizations can be used in particular cases that are of importance to the donor to pursue their geostrategic interests, even though they are designed not to, on average.

We test our theory focusing on US aid and IMF loans. Using new data on UNSC voting over the 1960-2015 period, our results show that US “friends” receive larger bilateral aid when voting in line with the United States in the UNSC, while positive votes of “enemies” are rewarded with

loans from the IMF. Multilateral aid is thus highly political in important cases where the preferences of politicians differ from those of their domestic audiences.

Our results have important implications for the nature of multilateral interventions. According to Milner (2006, 110), “[d]onor governments desire to use foreign aid for political and economic purposes that are related to donor interests. Publics, however are more interested in addressing the needs of the recipient countries.” According to her results, publics are more confident that multilateral aid is developmental compared to bilateral aid, so that governments can give more aid when making use of multilaterals in the presence of skeptic publics. She finds that public opinion is indeed an important determinant of the choice of how to allocate aid. When the public dislikes aid, more aid is given multilaterally rather than bilaterally. According to Milner (2006, 111), “[m]ultilateral aid thus helps solve a domestic principal-agent problem. Domestic politics may be a reason that governments chose to use multilateral international institutions.” We certainly agree to the latter, but strongly disagree with the former. According to our argument, multilateral aid makes domestic principal-agent problems worse (see also chapter 4). Multilateral aid is given via international organizations when publics dislike aid, not to make it more developmental, but rather because it is easier to hide from the donor’s domestic audience.

Our results also add to the literature on who controls the IMF. According to our results, voting in line with the United Kingdom, France, China, and Russia in the UNSC is not rewarded with IMF loans. Of the five permanent UNSC members, only the United States seems to have sufficient power over, and interest in, exploiting IMF loans to further its political goals.

Furthermore, our results can explain why governments have an interest in founding new international organizations and make them seem legitimate (see also Rocabert et al. 2017). Schneider and Tobin (2016) argue that governments prefer large numbers of international organizations so that they can delegate to those organizations with an aid portfolio that most closely matches the government’s preferences. Our results show that multilateral aid allows donors to obfuscate payments to a country that the donors’ voters do not want to support.

Our results provoke us to be bold and make predictions about the future development of the international aid architecture. A May 2018 poll by the institute Infratest Dimap shows that 59 percent of the (German) respondents are in favor of reducing foreign aid to countries that do not

cooperate sufficiently in taking back refugees – a position that German Minister of Development Gerd Müller is clearly opposed to.⁶⁵ Similarly, substantial shares of the populations in major countries of the European Union are opposed to a Greek bail-out, while leading academics and politicians see such support as a necessary condition to maintain the Euro and potentially the European Union.⁶⁶ We expect this difference in views to make multilateral aid attractive from a politician's perspective. This can explain the insistence of German politicians to keep on involving the IMF in the Greek bail-out, which large parts of the German electorate are not in favor of (see a 2010 Poll cited in Schneider and Slantchev 2018, 21).⁶⁷

In the same vein, we thus predict that major European donors will react to the recent refugee crisis by channeling larger shares of foreign aid through the budget of the European Union. The degrees of freedom that politicians gain from the existence of an international organization also explain political support for the creation of new organizations, and their resistance towards abolishing existing ones.⁶⁸ The recent creation of the Asian Infrastructure Investment Bank and the New Development Bank are cases in point. We do not expect these organizations to make Chinese support less political and more developmental, on average, but to allow China to better obfuscate its political influence. We also expect a 'European Monetary Fund' to be called in existence in due course, and additional European organizations in charge of foreign aid and loans to follow later. The potential benefits of international organizations in pursuing policies that domestic audiences strongly dislike seem too strong for national governments to resist.

⁶⁵ See <https://www.welt.de/politik/deutschland/article176217850/Migration-Mehrheit-will-unkooperativen-Staaten-Entwicklungshilfe-kuerzen.html> (in German, accessed May 10, 2018).

⁶⁶ See for example a 2015 YouGov survey "Greece: Germans and Finns back a hard line, but support for Grexit wanes," <https://yougov.co.uk/news/2015/07/10/germans-and-finns-public-prefer-hard-line-support-/> and a May 29, 2017 article in the Journal of International Affairs "Germany's Domestic Politics Complicate the Greek Debt Crisis", <https://jia.sipa.columbia.edu/online-articles/germanys-domestic-politics-complicate-greek-debt-crisis> (last accessed May 23, 2018).

⁶⁷ See also again the May 29, 2017 article in the Journal of International Affairs, <https://jia.sipa.columbia.edu/online-articles/germanys-domestic-politics-complicate-greek-debt-crisis> (last accessed May 23, 2018).

⁶⁸ According to Haberler (1974, 156) "international institutions may change their names or lose their function but they never die" (cited in Vaubel 2006, 127). Also see Gray (2018).

1.7 Appendices to Chapter 1

1.7.1 Appendix 1.A: Coding of Resolution-specific UNSC Voting Data

As described in the main text, we initially measure the importance of a vote following Kuziemko and Werker (2006), who argue that UNSC membership is more valuable in years in which the institution is of major geopolitical importance. They proxy “importance” with the number of New York Times articles that include the words “United Nations” and “Security Council” and separate the years into different categories of importance. We code the same variable for our sample updating it until 2015 based on the New York Times online archive. Unlike Kuziemko and Werker (2006) we focus on two rather than three categories of importance, to reduce the number of categories when the voting variables are added to the regressions.⁶⁹ Our threshold for important years is the median number of New York Times articles.

In addition to that, we propose other ways of identifying relevant votes for measuring voting alignment in the UNSC. The fact that we use data on the resolution-level allows us to additionally exploit resolution-specific rather than only year-specific information.

First, we exploit information contained in the resolution’s title. To this end, we identified key words that frequently appear in resolution titles, using word counting software. This allows coding variables that indicate the policy area the resolutions address. Table 1.5 shows the 100 most frequent keywords. For this study we only show regressions that restrict the sample of resolutions to those that concern Israel.⁷⁰ A relatively large number of UNSC decisions focus on this key US ally (140 out of 2524), and our expectation is that the United States will consider these decisions as particularly important.

Second, for all resolutions we code the number of *Google hits* that appear when searching for “United Nations Security Council Resolution [number]”⁷¹ via the *Google* search engine. Figure 1.6 illustrates these data and shows that there is no visible time trend in this variable. We then

⁶⁹ Our results are qualitatively similar when we use three categories of importance.

⁷⁰ To determine which resolutions concern Israel, we code the title of each resolution and search for the keywords “Israel,” “Palestine,” “Jerusalem,” and “Golan.”

⁷¹ We do this for all resolutions from 1 to 2259 and enter the search term in quotes, thereby ensuring that the words appear in this exact order on the webpages that *Google* lists. For this we use the *Google Custom Search Engine* and run it via a program written in *Python*.

consider a resolution as important if its number of *Google hits* is above the median of a given year. In addition, all votes that did not produce a resolution because of a veto or a failure to reach the required majority are also coded as important. When using this information for the analysis on the country-year level we then only consider the “important” votes when aggregating.

Figure 1.6 – Google Hits of UNSC Resolutions

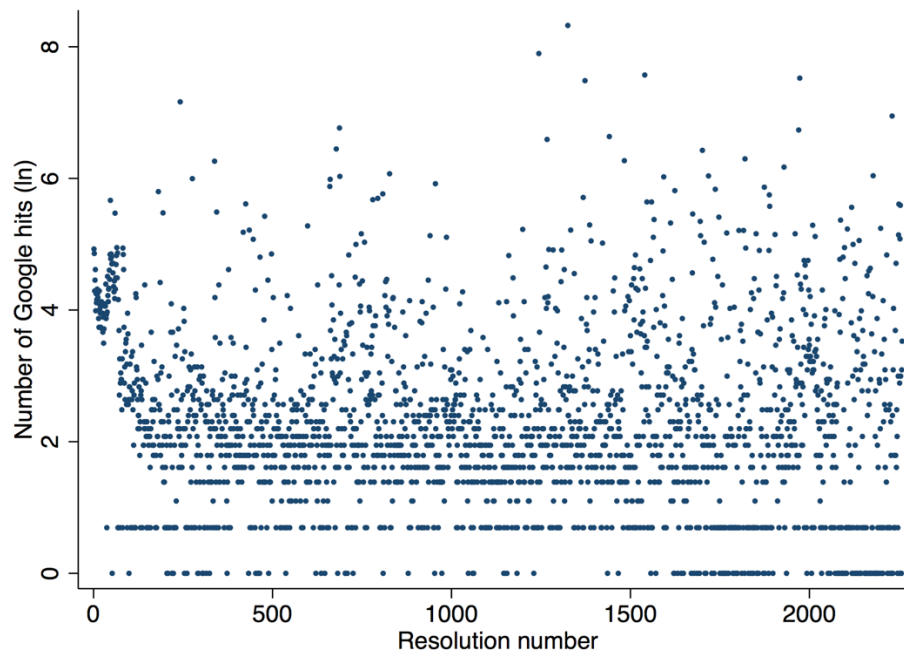


Table 1.5 – Frequency of Words in UNSC Resolution Titles (100 most frequent)

744 (6%): un	78 (1%): western	42 (0%): afghanistan
651 (6%): mandate	77 (1%): somalia	42 (0%): settlement
475 (4%): extension	76 (1%): measures	41 (0%): court
415 (4%): mission	75 (1%): disengagement	41 (0%): fire
342 (3%): situation	75 (1%): liberia	41 (0%): observers
342 (3%): force	74 (1%): application	41 (0%): stationing
190 (2%): membership	74 (1%): sudan	41 (0%): deployment
187 (2%): peace	71 (1%): sahara	40 (0%): complaint
177 (2%): against	68 (1%): assistance	40 (0%): all
171 (1%): observer	67 (1%): middle	40 (0%): armed
167 (1%): republic	67 (1%): rwanda	39 (0%): commission
150 (1%): Cyprus	65 (1%): bosnia	39 (0%): concerning
137 (1%): security	64 (1%): africa	39 (0%): forces
129 (1%): admission	64 (1%): operation	38 (0%): calling
127 (1%): establishment	64 (1%): herzegovina	38 (0%): minurso
125 (1%): extends	63 (1%): former	37 (0%): states
121 (1%): lebanon	62 (1%): secretary	37 (0%): territories
111 (1%): south	61 (1%): between	37 (0%): southern
108 (1%): question	60 (1%): referendum	36 (0%): central
105 (1%): resolution	59 (1%): humanitarian	36 (0%): israeli
103 (1%): military	56 (0%): arms	36 (0%): imposed
95 (1%): general	50 (0%): d'ivoire	35 (0%): group
91 (1%): east	50 (0%): côte	35 (0%): rhodesia
89 (1%): congo	48 (0%): agreement	35 (0%): authorization
88 (1%): council	48 (0%): cease	34 (0%): justice
86 (1%): keeping	48 (0%): african	34 (0%): peacekeeping
86 (1%): tribunal	48 (0%): monitoring	33 (0%): under
84 (1%): angola	48 (0%): haiti	33 (0%): palestinian
84 (1%): renewal	46 (0%): embargo	33 (0%): process
83 (1%): democratic	45 (0%): conflict	32 (0%): office
83 (1%): sanctions	44 (0%): israel	
81 (1%): implementation	44 (0%): leone	
80 (1%): iraq	44 (0%): protection	
79 (1%): yugoslavia	44 (0%): criminal	
79 (1%): interim	43 (0%): sierra	

1.7.2 Appendix 1.B: Descriptive Statistics

Table 1.6 – Descriptive Statistics

Variable	Obs.	Mean	S.D.	Min	Max
US aid disbursements (million USD, ln)	6142	2.54	2.07	0.00	9.51
IMF loan size (million SDR, ln)	6142	1.09	1.95	0.00	10.36
UNSC member	6142	0.06	0.23	0.00	1.00
Share of votes against US	6066	0.01	0.09	0.00	1.00
UNSC, voted all with US	6142	0.03	0.16	0.00	1.00
UNSC, voted not all with US	6142	0.03	0.17	0.00	1.00
UNSC, voted all with US (important Google)	6142	0.03	0.16	0.00	1.00
UNSC, voted not all with US (important Google)	6142	0.03	0.17	0.00	1.00
UNSC, voted all with US (important Israel)	4222	0.02	0.15	0.00	1.00
UNSC, voted not all with US (important Israel)	4222	0.02	0.14	0.00	1.00
UNSC, voted all with US (important year NYT)	6142	0.02	0.13	0.00	1.00
UNSC, voted all with US (unimportant year NYT)	6142	0.01	0.10	0.00	1.00
UNSC, voted not all with US (important year NYT)	6142	0.01	0.10	0.00	1.00
UNSC, voted not all with US (unimportant year NYT)	6142	0.02	0.14	0.00	1.00
Political proximity to US	5114	0.22	0.14	0.00	0.88
IMF program	5826	0.38	0.49	0.00	1.00
IMF purchases (million SDR, ln)	5494	1.12	1.88	0.00	9.78
US aid indicator	6142	0.75	0.43	0.00	1.00
US aid commitments (million USD, ln)	5798	2.70	2.07	0.00	9.92
World Bank aid commitments (million USD, ln)	5798	1.56	2.22	0.00	8.36
IsDB aid commitments (million USD, ln)	6142	0.20	0.67	0.00	6.46
AsDB aid commitments (million USD, ln)	6142	0.48	1.34	0.00	7.57
EBRD aid commitments (million USD, ln)	6142	0.13	0.72	0.00	7.29
GDP per capita (ln)	6138	7.57	1.11	4.75	10.04
Population (ln)	6141	15.39	2.02	9.11	20.99
War	6142	0.06	0.23	0.00	1.00
Past IMF program	5826	0.72	0.45	0.00	1.00

Note: The sample used for calculating these statistics is the sample of column 1 of Table 1.1.

1.7.3 Appendix 1.C: Data Sources and Definitions

Table 1.7 – Data Sources and Definitions

Variable	Source	Description
US aid disbursements (million USD, ln)	OECD (2018)	Bilateral US net disbursements of Official Development Assistance.
IMF loan size (million SDR, ln)	Dreher et al. (2009a), own update with data from IMF (IMF 2018a)	Total amount agreed of IMF loan. IMF (2018) provides the total amount of the agreed upon loan. We divide this number by the years of subsequent program duration, assuming equal phasing of the loan over the program period.
UNSC member	Dreher et al. (2009b), own update	Binary, indicating observations in which country i was a temporary UNSC member in year t .
Share of votes against US	multiple sources (own coding, see main text)	The number of UNSC votes country i cast in line with the United States in year t divided by the number of UNSC votes in year t . Unanimous votes are excluded.
UNSC, voted all with US	multiple sources (own coding, see main text)	Binary, indicating observations in which country i was a UNSC member in year t , and voted in line with the United States in all votes of year t .
UNSC, voted not all with US	multiple sources (own coding, see main text)	Binary, indicating observations in which country i was a UNSC member in year t , and voted against the United States in at least one vote of year t .
UNSC, voted all with US (important Google)	multiple sources (own coding, see main text)	As above, but only considering UNSC votes on resolutions whose number of hits on the Google search engine surpasses the yearly median and UNSC votes that did not produce a resolution (see Appendix A for details).
UNSC, voted not all with US (important Google)	multiple sources (own coding, see main text)	As above, but only considering UNSC votes on resolutions whose number of hits on the Google search engine surpasses the yearly median and UNSC votes that did not produce a resolution (see Appendix A for details).
UNSC, voted all with US (important Israel)	multiple sources (own coding, see main text)	As above, but only considering UNSC votes on resolutions whose title is related to Israel (see Appendix A for details).
UNSC, voted not all with US (important Israel)	multiple sources (own coding, see main text)	As above, but only considering UNSC votes on resolutions whose title is related to Israel (see Appendix A for details).

UNSC, voted all with US (important year NYT)	multiple sources (own coding, see main text)	As above, but the indicator is set to zero if the year's number of New York Times articles that include the words "United Nations" and "Security Council" is below the median of the observation period (see Appendix A for details).
UNSC, voted all with US (unimportant year NYT)	multiple sources (own coding, see main text)	As above, but the indicator is set to zero if the year's number of New York Times articles that include the words "United Nations" and "Security Council" is above the median of the observation period (see Appendix A for details).
UNSC, voted not all with US (important year NYT)	multiple sources (own coding, see main text)	As above, but the indicator is set to zero if the year's number of New York Times articles that include the words "United Nations" and "Security Council" is below the median of the observation period (see Appendix A for details).
UNSC, voted not all with US (unimportant year NYT)	multiple sources (own coding, see main text)	As above, but the indicator is set to zero if the year's number of New York Times articles that include the words "United Nations" and "Security Council" is above the median of the observation period (see Appendix A for details).
Political proximity to US	Bailey, Strezhnev, and Voeten (2017)	A country's share of votes in line cast with the United States in the United Nations General Assembly, moving average from $t-5$ to $t-2$.
IMF program	Dreher et al. (2009a), updated with data from Kentikelenis et al. (2016)	IMF program active at any point in year t .
IMF purchases (million SDR, ln)	World Bank (2018)	Amount of the IMF loan "purchased" by the IMF program country.
US aid indicator	OECD (2018b)	Binary, indicating country-years with positive US aid disbursements.
US aid commitments (million USD, ln)	OECD (2018b)	Bilateral US commitments of Official Development Assistance.
World Bank aid commitments (million USD, ln)	OECD (2018b)	World Bank commitments of Official Development Assistance.
IsDB aid commitments (million USD, ln)	OECD (2018b)	Islamic Development Bank commitments of Official Development Assistance.
AsDB aid commitments (million USD, ln)	OECD (2018b)	Asian Development Bank commitments of Official Development Assistance.

EBRD aid commitments (million USD, ln)	OECD (2018b)	European Bank for Reconstruction and Development commitments of Official Development Assistance.
GDP per capita (ln)	World Bank (2018)	Gross Domestic Product per capita, constant 2010 USD.
Population (ln)	World Bank (2018)	Population size.
War	Uppsala Conflict Data Program (2015)	Binary, indicating years with more than 1000 battle-related deaths in year t in country i .
Past IMF program	Dreher et al. (2009a), updated with data from Kentikelenis et al. (2016)	Binary, indicating countries that had an IMF program in any of the years prior to year t .

2 Room for Discretion: Biased Decision-Making in International Financial Institutions

Note: This paper is co-authored with Andrea F. Presbitero. It was published in the *Journal of Development Economics* (2018), Issue 130, Pages 1-16.

2.1 Introduction

A considerable body of literature suggests that the major international financial institutions (IFIs) are not the independent, technocratic organizations many expect them to be. Instead of making decisions exclusively based on objective and economic criteria, their behavior also reflects the particular interests of individual actors, which appear to bias the organizations' decision-making in their favor.⁷² Focusing on the two most powerful IFIs, the International Monetary Fund (IMF or Fund) and the World Bank, a first strand of literature shows that the political interests of major shareholders play a significant role in their lending decisions. Countries that are politically aligned with or important for the United States and other "G5" governments⁷³ have privileged access to financial assistance from the Fund and the Bank (chapter 1; Barro and Lee 2005; Dreher and Jensen 2007; Dreher, Sturm, and Vreeland 2009a; 2009b; 2015; Kilby 2009; 2013a; 2013b; Reynaud and Vauday 2009; Stone 2008; Thacker 1999). A second strand of literature identifies bureaucratic incentives as a specific source of bias in the IMF and World Bank decision-making processes. According to this view, which follows standard public choice models and approaches underlining the importance of organizational culture, staff in international organizations aim to increase their budget, power, prestige and independence. As a result, the decisions IFIs make are not always economically optimal and technocratic, but can also reflect the bureaucracy's particular interests and beliefs (Barnett and Finnemore 2004; Chwioroth 2013; Copelovitch 2010a; Nelson 2014; Stone 2008; Vaubel 1986; 1996; 2006).

In this paper we exploit a specific feature of the design of the Debt Sustainability Framework (DSF) for low-income countries (LICs), developed jointly by the World Bank and the IMF to assess debt sustainability, as an ideal set-up to understand the decision-making of IFIs. We use a unique dataset on the application of the DSF that allows us to reconstruct internal decision-making processes and to identify decisions that deviated from the mechanical application of formal rules. In particular, the DSF assigns a rating (low risk, moderate risk, high risk) for the risk of debt distress to a given country according to the projected evolution of its debt levels over a 20-year period, with respect to policy-dependent debt thresholds. While the assessment of the risk of debt distress is based on a mechanical, model-based rule, the final rating may

⁷² For a recent survey of this literature see Dreher and Lang (2016).

⁷³ In addition to the United States, this group of powerful shareholders of the Bank and the Fund comprises the United Kingdom, France, Germany and Japan.

involve the use of judgment by staff. The DSF explicitly allows staff to override the mechanical rating if they consider that country-specific circumstances justify this choice. In practice, in a number of cases this *room for discretion* translates into deviations of the final risk ratings from the ones obtained by mechanically applying the rule. Our data allow us to identify the cases in which such overrides of the mechanical rating took place. In the empirical analysis we examine the determinants of these overrides.

Specifically, we test for the presence of political and bureaucratic biases by looking at whether the country's political alignment with the organizations' major shareholders and the staff's desire to maintain its existing risk rating are related to overrides of the mechanical risk rating. In addition, we also investigate the role that other macroeconomic variables play in determining the actual risk rating, with the aim of shedding light on the design of the DSF, whose simplicity—a result of the need to make it accessible to a wider set of stakeholders—could result in a limited capacity to take into account important macroeconomic developments in the mechanical risk rating.

In the context of the literature on the political economy of international organizations (IOs) our empirical approach allows us to explicitly test the influential model of “informal governance” (Stone 2008; 2011; 2013). This model, which has been initially developed for the IMF, posits that the influence of powerful states on international organizations primarily runs through informal channels.⁷⁴ While formal rules and a relatively autonomous bureaucracy regulate the IOs' day-to-day operations, powerful states retain influence through informal practices that allow them to intervene in formal processes when urgent strategic interests are at stake.⁷⁵ The model furthermore suggests that such informal influence must only be selectively exercised to avoid undermining the IO's legitimacy, which in turn is an important reason for powerful states to act through IOs in the first place (see also chapter 1). Transferred to our setting, the *room for discretion* embedded in the DSF is a potential channel for such informal political influence. Furthermore, it gives bureaucrats the opportunity to influence decision-making according to the bureaucracy's preferences. It would be consistent with this

⁷⁴ See Stone's (2011) book and the 2013 special issue of the *Review of International Organizations* (volume 8, issue 2) on “Informal Governance in International Organizations.” See also Copelovitch (2010a) for a related model, which posits “common agency” by the largest shareholders and the bureaucracy.

⁷⁵ In the words of Stone (2013, 124) “all international organizations operate to some degree at variance with their formal rules. The formal rules—standard operating procedures, voting rules, organizational chains of command, written policies—provide stable and predictable policy outputs. Derogations from these standard procedures are made to safeguard the interests of powerful states.”

model's predictions if political interests and bureaucratic incentives were reflected in the use of judgment and if, in addition, the political bias were more significant a) when overruling is less clearly opposed to the formal rules, and b) when political interests are particularly strong. We test the former hypothesis by allowing for heterogeneity of the political interests effect depending on how clear-cut the mechanical rating is, and the latter by examining whether the political bias is stronger in election years of the rated country.

In addition to the explicit test of the informal governance model in a particularly suitable setting, the main contribution of our approach is threefold. First, we can directly examine the extent to which political economy variables explain interference in technocratic rules instead of only comparing how they relate to differences in outcomes. The deviations from the mechanical rule can be interpreted as a direct measure of the Bank and Fund's discretion in assessing the borrowing capacity of a country. In contrast to most of the previous literature, we can thus shed light on the internal decision-making processes that lead to the outcomes that IFIs produce.

Second, in contrast to many studies that analyze IMF and World Bank lending, the focus on risk ratings minimizes selection bias. Different countries might have different levels of demand for Fund and Bank resources. The fact, for instance, that countries that disagree with major shareholders on issues of foreign policy are also less likely to have an IMF program could be due to the lack of willingness to engage with the Fund, rather than to unfavorable treatment from the Fund and Bank. Since the DSF is a standard toolkit of IMF surveillance that is regularly applied in *all* low-income countries and all countries have the same interest in obtaining the most favorable risk rating under the DSF (as it would assure better borrowing terms, see below), there are no differences on the demand-side in this setting.

Third, while the analysis of the DSF is particularly well suited to study the decision-making process of international financial institutions, examining potential biases in DSF risk ratings is also a relevant policy question *per se*. The debt sustainability analyses—and the final risk ratings—provide a unique and relevant source of information for a variety of stakeholders on the sustainability of the fiscal stance in LICs, which are typically not covered by the major

sovereign credit agencies.⁷⁶ The risk ratings have financial and macroeconomic consequences, as they can directly influence the size and the terms of borrowing of low-income countries.⁷⁷ Since most LICs do not have regular access to international capital markets, development finance is a major source of their external financing. The DSF risk ratings determine the terms under which countries receive financial assistance from multilateral institutions. For instance, the World Bank's International Development Agency (IDA) reduces its allocation to countries with weak ratings by up to a fifth and makes the loans-grants mix conditional on the risk rating.⁷⁸ In addition, the World Bank uses the DSF risk ratings to design non-concessional borrowing limits for LICs. To test whether the World Bank adheres to these rules in practice, we regress a country's volume of lending received from the World Bank (as a fraction of GDP) on its DSF risk rating and a basic set of control variables. We find that "medium risk" and "high risk" countries indeed receive significantly less lending; on average and holding other factors constant, "high risk" countries receive between 17 and 35 percent less financing than "low risk" countries (see Table 2.11 in Appendix 2.A).⁷⁹ In the same spirit, regional development banks and bilateral aid agencies base their grant and lending decisions on DSF ratings. For instance, according to the so-called "Lagarde rule" (from when the current IMF Managing Director was Minister of Finance in France), the Agence Française de Développement only lends to countries at low risk of debt distress and continues to lend for a

⁷⁶ Our study thus also contributes to the literature on the determinants of credit ratings more broadly. Recent contributions have examined bias in sovereign credit ratings assigned by private rating agencies (Bartels and Weder di Mauro 2013; Fuchs and Gehring 2017).

⁷⁷ We also contribute to the literature that goes beyond the analysis of IMF/World Bank lending activities, where the existing evidence on biasedness is much scarcer. Dreher et al. (2008) find that IMF inflation forecasts are systematically biased and favor countries that are politically close to the United States. Fratzscher and Reynaud (2011) find that countries with more political influence in the IMF and in the UN receive more favorable Public Information Notices of Article IV consultations from the Fund. Other than that, there is little systematic evidence as to whether political and bureaucratic interests bias Fund and Bank decision-making in areas that are – at least *a priori* – less political and more technocratic in nature.

⁷⁸ Being classified as high risk is associated with 100 percent grants, medium risk with 50 percent grants and 50 percent loans, while low risk is associated with 100 percent loans and zero grants. Grants come with a 20 percent reduction in available resources to minimize moral hazard. See: <http://ida.worldbank.org/financing/debt-sustainability-grants> (last accessed November 16, 2016). The countries' preference for higher volumes of World Bank lending appears to dominate the relative allocation between grants and loans. One reason could be related to a substantial time discount rate that politicians attach to such financing due to political cycles and political incentives. Loans are highly concessional and come with very long maturities and grace periods. As of April 2017, the regular IDA loan has a 38-year maturity and a 6-year grace period. In this case, a politician could prefer borrowing 100 percent rather than receiving 80 percent in grants, and ignore the repayment problem, as it will be effective only after 6 years.

⁷⁹ These regressions also show that the negative effect of the official DSF rating on World Bank lending holds when controlling for the mechanical rating, suggesting that potential biases in the deviations from the mechanical rating would directly affect lending.

year in case a country is downgraded to moderate risk, while it does not lend at all to high risk countries. Finally, risk ratings affect debt conditionality under IMF-supported programs. Countries at moderate or high risk of debt distress have different kinds of debt limits, while program conditionality in countries at low risk of debt distress normally does not include limits on public external borrowing (IMF 2015).

In recent years, an increasing number of LICs have started gaining market access, mostly through syndicated bank loans and Eurobond issuances (Presbitero et al. 2016). For these economies, the DSF risk ratings are also a critical source of information for market participants and can affect the availability and the terms of commercial lending. A descriptive look at our data illustrates this. Out of 14 low-income countries that have issued sovereign bonds since 2014, ten were classified at a low risk of debt distress, only four had a moderate risk (Cameroon, Cote d'Ivoire, Ghana and Honduras), and no country with a high risk of debt distress has been able to issue. The average spread at issuance has been about 140 basis points higher for countries at moderate risk than for countries at low risk of debt distress, even taking into account the size and maturity of the bonds. Moreover, comparing countries with similar *Institutional Investor* country risk ratings but different DSF risk ratings shows that a worse DSF risk rating is indeed reflected in higher sovereign bond spreads. For instance, in 2014 and 2015, Ghana had a similar country risk rating as Mongolia, Senegal and Zambia, but it was the only country classified at moderate risk of debt distress and faced an average premium ranging between 110 and 228 basis points compared to the other three countries at low risk of debt distress.⁸⁰

A similar picture emerges when considering cross-border bank lending. Over the 2007-2015 period the financial markets platform *Dealogic* records 803 syndicated loans to countries with an outstanding DSF risk rating, 59.4 percent of which went to countries with a low risk, 34.1 percent to countries with a moderate risk and only 6.5 percent to countries with a high risk of debt distress. This distribution is even more skewed towards low risk countries when

⁸⁰ More precisely, the country risk ratings by *Institutional Investor* range between 0 and 100 with 100 indicating the least likelihood of default, see <http://www.institutionalinvestor.com/Research-and-Rankings.html> (last accessed November 16, 2016). We compare Ghana—at moderate risk—with Mongolia, Senegal and Zambia—at low risk—over the period 2014-2015. In that period, the four countries had, on average, a very similar *Institutional Investor* country credit rating, but faced different sovereign bond spreads. In particular, the credit rating was 33.6 in Ghana, 33.2 in Mongolia, 33.3 in Senegal and 35.3 in Zambia, while the average spreads were 670 bps in Ghana, 556 in Mongolia, 441 in Senegal and 559 in Zambia. See also chapter 3 for additional information on *Institutional Investor* data.

considering loans to the private sector, three quarters of which have had a company in a low risk country as borrower. If we look at dynamics over time and focus on countries that have seen a worsening of their debt rating, we can see that countries that have been regular borrowers (Mali, Mongolia, Mozambique, Niger, and Sri Lanka, with at least one loan before and after a debt sustainability analysis) have on average experienced a reduction in the number of syndicated loans when considering a window of one year before and one year after the publication of the risk rating.⁸¹

Our results indicate that there is evidence for the presence of systematic biases. First, the debt ratings for countries that are politically aligned with the Bank's and the Fund's major shareholders are more often overridden (and improved) as compared to the result of the strict application of the DSF's mechanical model. The effect is stronger when the mechanical assessment is less clear-cut, suggesting that biasedness is more likely when there is more *room for discretion*. Also, the effect is mainly driven by the years in which an election took place in the rated country, suggesting that allies of the major shareholders are particularly likely to be treated favorably when their governments particularly care about positive assessments.⁸² Second, bureaucratic incentives also explain the use of judgment in the DSF. In a large number of cases staff make use of the *room for discretion* in order to avoid deteriorations of the debt sustainability rating relative to the previous assessment. This is consistent with the expectation that the bureaucracy has an incentive to "keep" the ratings assigned in previous analyses as this is more likely to be in line with the established in-house view and avoids potential confrontations with senior staff. Finally, we find that some macroeconomic variables are associated with the decision to use judgment, suggesting that the DSF is "too simple" and Bank and Fund staff have to rely on additional information that is not part of the DSF to assess debt sustainability. We show that these results are unlikely to be driven by endogeneity, unobserved country-specific heterogeneity, and we rule out many other channels and alternative factors that could explain the override of the risk rating. We also corroborate our statistical results with the help of anecdotal evidence that we gathered in interviews with involved Bank and Fund staff.

⁸¹ The relatively small size of cross-border lending to LICs and the fact that we cannot disentangle demand from supply effects would call for caution when interpreting these numbers.

⁸² This result is consistent with recent findings by Kersting and Kilby (2016), who show that World Bank lending accelerates for allies of the United States when domestic elections approach, and interpret this as evidence for the World Bank's engagement in "global electioneering that serves U.S. foreign policy interests."

While we believe that our results document the presence of systemic biases, they are silent about the trade-off between rules versus discretion. We point to some of the potential costs of the discretionary bias, but we do not assess its benefits. For instance, overridden ratings could reduce the rate of false alarms and better reflect *ex-post* debt sustainability. Because endogeneity prevents us from such an assessment—as future debt dynamics critically depend on actual risk ratings—we have to leave the analysis of the potential benefits of the *room for discretion* for future research.

2.2 The Debt Sustainability Framework and the Risk Ratings

The DSF was introduced in 2005 to guide the borrowing decisions of LICs in a way that matches their financing needs with their current and prospective repayment ability, taking into account each country's specific circumstances. Under the framework, debt sustainability analyses (DSAs) are conducted regularly with the aim of assessing the risk of external debt distress and providing policy recommendations that limit the risk of debt distress through prudent borrowing and lending strategies.⁸³ All DSAs must be prepared jointly by the World Bank and the IMF and submitted to their respective executive boards. Staff of the two institutions coordinate closely in preparing the DSAs from the early stages to the final approval. They are supposed to agree on the key macroeconomic projections and assumptions on new borrowing, with World Bank staff generally taking the lead on growth prospects and IMF staff focusing on medium-term macroeconomic projections. As a general rule, one DSA should be produced at least once every year (not necessarily at a given point in time for all countries), as it is the basis to determine the IDA credit-grant allocation. Exceptions to this time schedule may happen under specific circumstances, such as a request for IMF financing that involves exceptional access (IMF 2013). To empirically check whether this regular annual schedule is adhered to in practice, we look at the number of months between two consecutive DSAs for a given country. In only 19 cases two DSAs are separated by less than six months—possibly because the DSA has been triggered by a specific event. As DSAs appear to generally

⁸³ Since we recognize that the terminology may appear confusing, we would like to point out that “although the terms ‘DSF’ and ‘DSA’ are sometimes used interchangeably, they are in fact distinct: the DSF is the framework within which a DSA is produced for a particular country” (IMF 2013, 6).

follow a regular, pre-determined schedule we can reliably rule out selection bias resulting from selection into receiving a rating.⁸⁴

The DSA assigns a risk of debt distress (low risk, medium risk, high risk, or in debt distress) to a country depending on the evolution of five debt indicators compared to some policy-dependent thresholds, assuming that breaches of these thresholds would signal the presence of debt vulnerabilities. Specifically, the mechanical rule for assigning debt risk ratings is based on the projections of five key macroeconomic indicators for the subsequent 20 years: the present value (PV) of external debt as a percentage of GDP, exports, and revenue as well as the external debt service as a percentage of exports and revenue. For each of these five variables there is one baseline scenario and eight stress test projections with alternative macroeconomic assumptions for the 20-year period. The projected values are compared to the thresholds, that depend on the country's quality of policies and institutions, as measured by a three-step version of the Country Policy and Institutional Assessment (CPIA)—an index of institutional quality and macroeconomic stability produced by the World Bank.⁸⁵ According to the mechanical rule, the risk rating 'moderate risk' is assigned if one of the stress test projections exceeds the corresponding threshold, while the rating 'high risk' is assigned if this is the case for one of the baseline scenarios (IMF 2013). In sum, these data allow us to reconstruct the rating suggested by the mechanical rule and to compare it to the rating actually assigned.⁸⁶

The actual rating may differ from the mechanical one because of the use of judgment by IMF and World Bank staff. One of the distinctive features of the DSF is to explicitly allowing for judgment as a source of flexibility to ensure that ratings are not excessively affected by short term macroeconomic fluctuations and that they also take into account country-specific

⁸⁴ To minimize concerns about selection bias, we show that our results are robust to limiting the sample to the DSAs that were conducted between 6 and 18 months after the previous one and thus strictly followed the pre-determined schedule (see Section 4.4).

⁸⁵ These policy-dependent thresholds are based on the estimation of a set of simple probit models on a large sample of low and middle-income countries over the 1970-2007 period—one for each debt indicator—where the probability of debt distress is a function of the debt indicator, the CPIA indicator as a measure of policies and institutional quality, and GDP growth as a proxy for economic shocks. See IMF and World Bank (2012) for additional details.

⁸⁶ For an overview of the LIC DSF, see <http://www.imf.org/external/np/exr/facts/jdsf.htm> (last accessed November 23, 2016). For a more detailed discussion on the underlying method and its application, see IMF (2013) and Berg et al. (2014). After its introduction, the DSF has been reviewed three times (2006, 2009 and 2012), but its main structure has remained the same. The two most important changes introduced by the 2012 review have been: 1) the revision of two of the five debt thresholds (debt over exports and debt service over exports), to take into account remittances in the denominator of the ratio for countries highly dependent on remittances, and 2) the introduction of the "probability approach," which uses country-specific information to complement the standard debt sustainability assessment in borderline cases.

characteristics that are not reflected in the macroeconomic projections (IMF 2009). In particular, the guidance notes of the LIC DSF state that “[a]lthough the indicative thresholds play a fundamental role in the determination of the risk rating, they should not be interpreted mechanistically. The assessment of risk needs to strike a balance between paying due attention to debt levels rising toward or above thresholds and using judgment. Thus, a marginal or temporary breach of a threshold may not necessarily imply a significant vulnerability. Conversely, a near breach should not be dismissed without careful consideration” (IMF 2013). More precisely, the assessment done by the country team can deviate from the rule in the presence of minor and temporary breaches of the debt thresholds, while protracted (and large) breaches are more worrisome. Also, staff should pay attention to the pace of debt accumulation and give the right weight to some of the stress scenarios, which—being standardized—in some circumstances may not be very realistic. Moreover, careful consideration should be given to the country’s ability to repay external debt that is not captured in the framework. For instance, the availability of a large pool of international reserves could counteract some negative indication coming from debt service ratios.

2.3 Data and Method

2.3.1 Main Variables and Descriptive Evidence

Our analysis is based on a unique dataset covering *all* 377 debt sustainability analyses undertaken in low-income countries between December 2006 and January 2015. The dataset includes, for each DSA, the five debt ratios in the year of the DSA and their projections for the 20-year period ahead, as well as all projections for the key macroeconomic variables that are used to determine the evolution of the country’s level of debt flows and stocks. These projections are the ones used by IMF and World Bank staff in the macroeconomic framework produced by the country team at the time when the DSA was produced. Some information on the projections over the medium term (but not the yearly data for the entire 20-year period) for the key variables are publicly available in the country report and in the associated DSA write-up.

We are interested in the determinants of the decision to make use of the discretion embedded in the DSF. In Table 2.1 we take a first, descriptive look at our data on the DSAs by comparing mechanical and actual ratings. First, the table shows that in the majority of cases the actual

rating is equivalent to the mechanical rating. Of the 367 DSAs for which we have all necessary data,⁸⁷ 272 end up with the rating that the mechanical rule suggests (~74 percent). Making use of the *room for discretion* in the DSF is the exception rather than the rule. Second, it becomes evident that the vast majority of deviations are improvements of the risk rating (i.e., a lower risk) with respect to the ones mechanically determined. Almost all of these deviations are one-notch improvements (i.e., from high risk to medium risk or from medium risk to low risk); only one country received a low risk rating when the projections suggested assigning high risk. The actual rating is worse than the mechanical rating in only nine cases.⁸⁸ This first piece of evidence would already suggest the possibility that the almost unidirectional use of judgment is consistent with the presence of staff incentives to provide a good rating of the country they work on, as well as with a positive bias that could be explained by political pressures to obtain an improvement in the risk rating.

Table 2.1 – LIC-DSA Mechanical and Actual Risk Ratings

		Mechanical risk rating							
		Low		Moderate		High		Total	
Actual risk rating	Low	81	92.0%	47	30.7%	1	0.8%	129	35.1%
	Moderate	7	8.0%	104	68.0%	38	30.2%	149	40.6%
	High	0	0.0%	2	1.3%	87	69.0%	89	24.3%
	Total	88	100.0%	153	100.0%	126	100.0%	367	100.0%

Note: The table shows the absolute frequencies of mechanical and actual DSF risk ratings. The columns distinguish between the three categories of the mechanically produced ratings while the rows distinguish between the actual ratings countries have received. To the right of the absolute frequencies are relative frequencies that show the distribution of actual ratings given the mechanical rating.

Against this backdrop, for the remainder of our analysis we code a binary dependent variable that indicates whether the final DSF rating was more favorable than what the mechanical rating suggested and examine the determinants of such rating improvements. While we consider multiple potential determinants, we are primarily interested in two hypotheses that emerge from the political economy literature on international organizations.

⁸⁷ The last DSF revision introduced remittances-augmented debt thresholds and projections to capture enhanced repayment capacity of countries receiving large remittances. Because of incomplete data for the ten DSAs done in 2014 and 2015 that use these augmented values, we exclude them from the analysis.

⁸⁸ We looked at the justification for the ratings of these nine DSAs in detail and found that the analysts in all cases referred to country-specific vulnerabilities. Examples include the 2014 DSA for Afghanistan, which concluded that risk was high because of “significant vulnerabilities” including the country’s heavy reliance on donor grants, and the 2012 DSA for Burkina Faso, which set the risk rating from low to moderate because of vulnerabilities related to the country’s gold exports. Because of the small sample size, we do not test for potential biases in these nine decisions.

First, we test the hypothesis that the use of judgment in debt sustainability ratings is biased in favor of countries that are politically affiliated with the Bank's and the Fund's major shareholders. For such a bias to be present, shareholders do not have to interfere directly with the decision-making processes at the staff level where DSAs are produced. The mechanisms may be more subtle. Preemptive obedience to the views of the shareholders' delegates (Executive Directors) may play a role as all DSAs have to be approved by the Board of Directors. In our background research at the institutions' headquarters in Washington, staff suggested that both the Fund and the Bank have certain "narratives" for most countries and treat some countries as role model for others. Internally, DSAs are expected to fit these preexisting "stories" for individual countries. Such narratives and role models can originate at the political level and trickle down via senior management to the staff level.⁸⁹ While the exact mechanisms remain speculative, it is clear that such favorable treatment of the major shareholders' political allies would constitute a bias that is not reconcilable with a correct application of the DSF that is solely based on economic criteria.

To measure political proximity we use data on voting behavior in the United Nations General Assembly (UNGA). In our baseline, we rely on estimates of the distances between the "ideal point" of rated countries and the United States (Bailey, Strezhnev, and Voeten 2017).⁹⁰ Similar to Vreeland and Dreher (2014) we code a country as a "US-friend" if it is in the lowest quintile of the distribution of the ideal point distance (meaning that it is among the closest friends of the United States), as we expect variation at the upper tail of the political alignment distribution to be more likely to make a difference than variation around the mean. As our study focuses on LICs, most of which score low on political alignment with the United States, we believe that minor differences in political alignment between the United States and the rated country are uninformative. By contrast, we expect benefits to materialize only for "close friends," which are the ones that are significantly different from the average LIC. We test the robustness of this assumption using alternative cutoff values as well as the mean distance to ideal points of the United States, the United Kingdom, Germany, France and Japan, the so-called G5. In Table 2.2 we descriptively compare US-friends to other countries and find that

⁸⁹ Source: interviews with World Bank and IMF officials (November 2016).

⁹⁰ In contrast to other voting affinity scores such as S-scores these ideal points are better at removing noise by using information on changes in the UNGA's agenda (Bailey, Strezhnev, and Voeten 2017). The variables are defined as ideal point distances, such that a smaller value indicates greater voting affinity.

the former are indeed more likely to benefit from the use of judgment in DSAs (42 vs. 29 percent)—a first indication of bias but nothing close to rigorous econometric evidence.

Table 2.2 – Overruling and US Friends

		US friend					
		No		Yes		Total	
Rating improved through overruling	No	158	71.2%	30	57.7%	188	68.6%
	Yes	64	28.8%	22	42.3%	86	31.4%
	<i>Total</i>	222	100.0%	52	100.0%	274	100.0%

Note: The table shows both absolute and relative frequencies of improved ratings given whether the country is classified as a US friend or not. Only DSAs with “moderate risk” and “high risk” mechanical ratings are considered as only these can be improved. Note that we lose five observations because of missing data on UNGA voting.

Second, we consider the role of bureaucratic incentives. Existing literature and our background research at the World Bank and the IMF suggests that there often is a strong disincentive for staff to increase a country’s risk of debt distress relative to its previous rating. First, as a worse rating implies that lending from the World Bank must be reduced, the “pressure to lend” (Kilby 2000, 2009) incentivizes the Bank bureaucracy to “show that its customers are solvent.”⁹¹ Second, an increase in the DSF risk rating constitutes a break with past analyses and suggests that previous projections had to be adjusted. Staff involved in producing new DSAs confirmed that they have an incentive to agree with the conclusions drawn in previous analyses, as any change requires a stronger justification vis-à-vis management than maintaining the *status quo*. Moreover, some countries function as important role models for the institutions and staff may fear that increasing the debt risk ratings for such countries may undermine the consistency and credibility of the institutions’ policy analysis and advice. To statistically test whether staff use the *room for discretion* in the DSF to avoid increasing countries’ debt risk ratings, we code a binary variable (*Risk Increased*) that indicates that the mechanical rating is worse than the actual rating assigned in the previous DSA (i.e., increasing from low risk to moderate/high risk, or from moderate risk to high risk). We look at these cases in Table 2.3. It shows that out of the 80 DSAs in which assigning the mechanical rating would have meant increasing the country’s risk rating relative to its previous DSA, in 64 DSAs (80 percent) the mechanical rating was overridden to keep the rating the same as in

⁹¹ Source: Interview with World Bank official (November 2016)

the previous DSA.⁹² The descriptive evidence thus strongly supports the hypothesis that the IMF and the World Bank make use of judgment in order to make DSF risk ratings more persistent at lower risk levels than the strict application of the formal framework would suggest.⁹³

Table 2.3 – Overruling and Avoiding Downgrades

		Mechanical rating: risk increased					
		No		Yes		Total	
Mechanical rating	No	177	88.9%	16	20.0%	193	69.2%
adjusted downward	Yes	22	11.1%	64	80.0%	86	30.8%
	Total	199	100.0%	80	100.0%	279	100.0%

Note: The table shows both absolute and relative frequencies of improved ratings given whether the mechanical rating was worse than the country's previous DSF risk rating. Only "moderate risk" and "high risk" mechanical ratings are considered as only these can be improved.

While this descriptive evidence suggests the presence of both a bureaucratic and a political bias, we can neither rule out that this bias is driven by confounding factors, nor can we illuminate any underlying channels. In order to do so, we turn to a more rigorous econometric analysis.

2.3.2 Econometric Model

We run binary response regressions of the variable indicating an improvement of the DSF risk rating employing a conditional logit estimator that controls for unobserved time-invariant heterogeneity across regions or countries via fixed effects. We also include year fixed effects to absorb the impact of the global business cycle and other global factors that can affect all LICs equally. We naturally exclude all observations whose ratings cannot be further improved and also control for the mechanical rating as the conditional propensity to override might be

⁹² When contrasting this with the mirror case, we find that of the 24 cases in which the mechanical rating suggested a lower risk than before, eight (33 percent) were overridden to leave the rating unchanged. As this fraction is positive yet substantially lower than in the opposite case, this evidence suggests that there is both a conservative bias to keep the previous rating and a bias of avoiding to increase risk ratings.

⁹³ We know the subsequent rating for 50 of the 64 DSAs that were overruled when the mechanical rating suggested a deterioration. 21 of these 50 reverted mechanically back to a better rating in the subsequent DSA while for the remaining 29 the mechanical rating again suggested a higher risk. Note also that in the data there is no regular pattern of how many times a worse mechanical rating is tolerated before a worse official rating is assigned.

different for adjustments from high to medium as compared to adjustments from medium risk to low risk. Formally we estimate:

$$\Pr(A_{itn} < M_{itn} | \alpha_{r/i}, \delta_t, X_{itn}, M_{itn})$$

where subscripts i , t , r , and n denote countries, years, regions and DSAs respectively. M and A denote the mechanical and the actual risk rating, α and δ are full sets of region (or country) and year fixed effects, and X_{itn} is a vector of explanatory observable variables that are either country-year-specific or DSA-specific.⁹⁴ The explanatory variables of interest in X_{itn} are the country-year specific measure of political proximity (*US friend_{it}*) and the DSA-specific variable indicating that the mechanical rule suggests assigning a worse rating relative to the previous DSA (*Mechanical Rating: Risk Increased_{itn}*). The effect of the latter variable can be estimated with country fixed effects, since it naturally varies significantly over time. However, given that political alignment with the United States is highly persistent over time, we choose region fixed effects over country fixed effects in the baseline regression with the *US friend* variable. Moreover, since regional spillovers are a key element of the spread of financial (and sovereign debt) crises, controlling for regional effects serves to capture these region-wide unobserved effects. We expect the effect of political proximity to the United States in a ten-year panel to be driven primarily by cross-country variation, which would be absorbed by country fixed effects.⁹⁵

To exploit within-country variation over time and to address endogeneity concerns in these regressions, we include interaction effects with time-varying variables. First, as theoretical considerations suggest that the political bias is stronger when the mechanical rating is less clear-cut, we allow the effect to be heterogeneous depending on the DSA-specific, time varying number of threshold breaches. Second, we interact the measure of political alignment with major shareholders with a variable (*Election_{it}*) indicating years in which an election took place in the rated country, as we expect governments to be particularly interested in favorable

⁹⁴ Note that we use the subscript *itn* because in some, rare cases there are multiple DSAs for the same country within one year. Countries i are nested in regions r .

⁹⁵ Another advantage of using region fixed effects is that with country fixed effects we lose almost half the sample because several countries never receive an improved rating. However, we include country fixed effects in additional specifications where we interact the *US friend* variable with a dummy for the presence of elections.

assessments in election years.⁹⁶ In addition to exploiting within-country variation over time—allowing us to employ country fixed effects and thus significantly mitigating the concern that there could be unobserved country-specific, time-invariant factors affecting the likelihood of obtaining better ratings—this approach has a second methodological advantage. Under the assumption that the timing of elections is exogenously predetermined, the interaction can be interpreted in a causal way. Even if political proximity were endogenous to the decision to improve the risk rating after conditioning on all covariates, the main effect of political proximity would be biased, but the interaction would still have a causal interpretation (Nizalova and Murtazashvili 2016). We consider endogenous election timing in our setting to be highly unlikely, as we are not aware of any instances in which an election was rescheduled because of an upcoming DSF rating.⁹⁷ In addition, even if there were unobserved variables correlated with both elections and overrides of mechanical DSF ratings—due to anticipation effects—such endogeneity would only bias the interaction effect if it was conditional on political proximity to the United States.⁹⁸ In sum, we believe that the identifying assumption necessary for a causal interpretation of the interaction is plausible.

In addition to the explanatory variables of interest, we augment X_{itm} with the following covariates.⁹⁹ We first add control variables that derive naturally from the DSF’s method. The mechanical rating is a function only of threshold breaches and the CPIA. We test whether these two variables also affect the decision to override the mechanical rating. According to the DSF guidance note (IMF 2013), staff are explicitly supposed to take the number of threshold breaches into account: “[l]arge, protracted breaches are more worrisome than small, temporary ones. Breaches of multiple thresholds suggest greater vulnerabilities than a single breach.” We thus expect more overrides in DSAs with fewer threshold breaches. Specifically, we use the number of threshold breaches in the baseline and stress test projections, scale these

⁹⁶ In an alternative specification we use a variable indicating that an election took place in the six months following the publication of a DSA. Arguably, political interests in good ratings are not only generally strong in election years but especially strong *in the run-up* to an election.

⁹⁷ In similar settings, Faye and Niehaus (2012) as well as Kersting and Kilby (2016) do not find evidence for endogenous election timing with regard to aid inflows. This suggests that endogenous election timing with regard to upcoming DSF ratings is also unlikely.

⁹⁸ Instead of assuming $E(\text{election} \cdot \varepsilon) = 0$ one would then assume $E(\text{election} \cdot \varepsilon) = E(\text{election} \cdot \varepsilon \mid \text{US friend})$. For econometric details on this point, see Bun and Harrison (2018) and Appendix S4 in Dreher, Eichenauer, and Gehring (2016).

⁹⁹ See Appendix 2.B for sources, descriptions and descriptive statistics of all variables.

two figures by the number of possible breaches of the baseline test ($5 \times 20 = 100$) and the stress tests ($8 \times 5 \times 20 = 800$) and add the two fractions up to generate the variable *Breaches*.¹⁰⁰

Second, we control for the continuous CPIA index, as the DSF's method emphasizes the importance of policies and institutions by making the thresholds contingent on the CPIA, and by emphasizing that the "quality of policies and institutions has a large influence on the macroeconomic return of public investment" (IMF 2013, 24). We expect that the continuous version of the CPIA provides additional information that is not captured by the discrete, three-step version of the index used for the determination of the debt thresholds. Moreover, Fund and World Bank staff may consider the CPIA as a proxy for "other country-specific considerations" that should be taken into account when gauging whether a few breaches warrant downgrades in risk ratings.

We add other macroeconomic fundamentals to test if they also influence the use of judgment and to rule out that they confound our findings. We focus on variables that the literature identifies as typical determinants of sovereign risk ratings (Hill, Brooks, and Faff 2010) and control for the logarithm of *GDP per capita*, *GDP growth*, and the logarithm of *population*, as larger countries are less sensitive to economic shocks than small ones. In additional regressions we also control for the level of total public debt, scaled by gross national income, and for the current account balance over GDP. Then, as the guidance note advises to take into account whether the country has large foreign exchange reserves when applying judgment (IMF 2013), we test if the level of international *reserves* can contribute to explaining the overrides. In the same spirit, we control for *natural resource rents* to account for the idea that debt sustainability could hinge on countries' future revenues from natural resources, information which cannot easily be incorporated in the DSF and which therefore could require the use of judgment.¹⁰¹ In a robustness test, we also look at the effect of *IMF growth projections*, as optimistic growth scenarios could not only lower debt ratios but could also motivate an overrule of the mechanical risk rating.¹⁰²

¹⁰⁰ Fund staff suggested that the stress test based on the historical scenario (which fixes some key macroeconomic variables at their historical averages to compute the 20-years ahead projections) is, in practice, often considered to be implausible given that many countries are expected to be on a structurally different growth path compared to the mid-1990s and/or early-2000s. To address this concern, we also ran regressions in which we ignore this stress test. Our inferences are not affected.

¹⁰¹ Note that we lag all macroeconomic explanatory variables by one year to make sure that they were observed at the time the DSA was produced.

¹⁰² All of these variables are taken from the World Bank's World Development Indicators (WDI), with the exceptions of growth projections which are taken from the IMF World Economic Outlook (WEO) database (different vintages).

As the DSF's method explicitly leaves room for the application of judgment, it would be in accordance with an unbiased application of the DSF's method if any of these control variables were related to the decision to override the mechanical result of the DSF. In general, however, we expect the omitted variable bias that is driven by macroeconomic fundamentals to be low, given that the projections used to produce the DSAs already incorporate most of this information. Any correlations between the macroeconomic fundamentals and the decision to manually adjust the risk rating would show that staff rely on information that is not incorporated in the mechanical projection models. This would suggest that the DSF's method is "too simple" and would support a refinement of the econometric underpinnings of the DSF (Berg et al. 2014).

2.4 Results

2.4.1 *Macroeconomic Determinants*

Before turning to our main variables of interest, in Table 2.4 we examine the effect of macroeconomic fundamentals and variables that are internal to the DSAs. Some of these variables could be correlated with overrides if the DSF is applied as intended, given that the number of threshold breaches and "country-specific considerations" are supposed to be taken into account. These regressions also help us to identify a suitable set of control variables.

As column 1 shows, the share of breached thresholds is a statistically significant predictor of the decision to override the risk rating. As one would expect, the higher this share, the less likely it is that the rating is improved. The continuous version of the CPIA also appears to influence this decision. Given that the three-step version of the CPIA plays an important role in setting the relevant thresholds in the DSF, but forces the framework to ignore some relevant information contained in the continuous measure of the CPIA, it is not surprising that Bank and Fund staff also rely on the additional information conveyed by the continuous version when deciding on whether to improve the rating. We also find that DSAs in which the mechanical rule suggests that a "high risk" rating be assigned are more likely to be improved, when the other variables are kept fixed. All three DSA-specific variables are statistically significant at the one percent level and enter with consistent signs across various specifications in Table 2.4 (and Table 2.5).

Table 2.4 – DSA-specific Variables and Macroeconomic Fundamentals

	(1)	(2)	(3)	(4)
Breaches	-0.116*** (0.029)	-0.102*** (0.027)	-0.105*** (0.028)	-0.105*** (0.030)
CPIA	1.519*** (0.394)	1.232*** (0.421)	1.337*** (0.450)	1.647*** (0.509)
Mechanical Rating: High	2.116*** (0.486)	2.222*** (0.500)	2.412*** (0.524)	2.488*** (0.539)
GDP per capita (ln, t-1)		1.060*** (0.297)	1.178*** (0.322)	0.993*** (0.360)
GDP growth (t-1)		-0.017 (0.033)	-0.032 (0.033)	-0.050 (0.035)
Population (ln, t-1)		0.407*** (0.136)	0.461*** (0.152)	0.380** (0.170)
Debt/GNI (t-1)			0.004 (0.008)	0.003 (0.010)
Current Account Balance (% GDP, t-1)			-1.044 (1.900)	-1.355 (1.961)
Reserves/Debt (t-1)				0.004 (0.005)
Natural Resources (% GDP, t-1)				0.023 (0.015)
Observations	263	263	255	247
Pseudo R-squared	0.189	0.242	0.258	0.272

Note: dependent variable indicates an improved debt rating through overruling; conditional logistic regressions that control for region fixed effects; reported are the coefficient estimates of the regressions; standard errors, robust to clustering at the region level, in parentheses. All regressions include year fixed effects. Significance levels: * $p < .10$, ** $p < .05$, *** $p < .01$

In column 2 we add GDP per capita, GDP growth and population as explanatory variables. As one would expect, GDP per capita appears as a strong, robust and statistically significant predictor of the outcome variable and enters with the expected sign in Table 2.4 (and Table 2.5). The statistical evidence for the effect of the growth rate is less robust (but in Table 2.5 has the expected, positive sign in the regressions in which it appears as a statistically significant predictor of overrides). Country size as measured by logged population is positively correlated with rating improvement.¹⁰³ While this could be interpreted as indicating a bias in favor of geopolitically important countries, large countries are more likely to be economically

¹⁰³ In auxiliary regressions, which are available upon request, we find that the latter effect is driven by the largest countries in the sample.

diversified and thus less vulnerable to external shocks. Moreover, larger countries could rely relatively more on local currency borrowing (Berensmann, Dafe, and Volz 2015), therefore being less exposed to the risk of external default. In our view, this finding may thus be economically justified and is not necessarily an indication of bias. In column 3, we include two additional macroeconomic variables that are typical determinants of sovereign credit risk. The current account balance (as a share of GDP) and the external-debt-to-GNI ratio neither enter with statistically significant coefficients, nor do they affect the coefficients on the other explanatory variables. In column 4 we add the country's amount of reserves (scaled by country debt) and its rents from natural resources (as a share of GDP). Countries with larger amounts of reserves and natural resources have the option to use these assets in order to mitigate the risk of debt distress. In particular, reserves are highly liquid and can be used to service external debt, making some breaches of the debt service indicators less relevant for the DSF risk rating. Although the signs on the two variables are, as could be expected, positive, there is no statistically significant empirical evidence that suggests that this plays a role for the use of judgment in DSAs.

We conclude that in addition to the variables that are internal to each DSA (number of breached thresholds, CPIA, level of the mechanical rating), GDP per capita, and to a lesser extent growth and country size, help predicting improvements of the risk rating. Other macroeconomic fundamentals do not seem to matter. As the projections and the threshold breaches on which the ratings are based include a large amount of country-specific economic information, the finding that other macroeconomic fundamentals are not correlated with overrides of the mechanical rating is not surprising. Instead, it shows that the method underlying the DSF is successful in incorporating a large amount of country-specific information that is typically considered to be informative for assessments of debt sustainability. The fact that at least the CPIA and GDP per capita robustly predict rating overrides indicate that there is room for further improving the DSF. Nevertheless, the results presented so far suggest that, to a significant extent, the DSF as well as the *room of discretion* embedded in it are employed as intended.

2.4.2 Biases

In Table 2.5 we augment our baseline model with a set of variables that, if statistically associated with overrides of the mechanical DSF ratings, indicate the presence of bias. In these regressions our baseline specification builds on the covariates included in column 2 of Table 2.4, given that these variables allow us to keep the entire sample, while the additional controls are not robustly related to the dependent variable.¹⁰⁴

Initially, we examine the role played by politics. In column 1 we add the *US friend* indicator¹⁰⁵ and find that countries that are politically affiliated with the United States are more likely to benefit from the use of judgment in DSAs. In column 2 we allow this effect to be heterogeneous depending on the number of threshold breaches and find that the effect is much stronger when the projections breach only relatively few thresholds. Figure 2.1 illustrates this result. While for US friends the conditional likelihood for receiving an improved rating is more than 80 percent when less than 2 percent of all projections breach thresholds, the likelihood for other countries is around 40 percent. The more breaches that are projected the smaller the difference between these two groups of countries. The difference vanishes at around six percent of breached thresholds.¹⁰⁶ These findings suggest that risk ratings are more likely to reflect geopolitical preferences the less clear-cut the mechanical rating is.

While this distinct pattern gives us confidence that our *US friend* indicator does not pick up an unobserved confounder, we aim to get closer to a causal interpretation of the effect in column 3. We interact *US friend* with a variable that indicates whether the DSA was published in an election year in the rated country. The results of specification 3 show that the US friends' benefit from the use of judgment is driven by the election years in the sample. Elections in other countries are not related to rating improvements. As discussed above, even if we allow

¹⁰⁴ All other variables we consider suffer from missing data and thus result in list-wise deletion of the respective observations. Given that our baseline regressions include 263 observations, we argue that the bias introduced by potentially non-random list-wise deletion of dozens of observations is likely to be more severe than omitting covariates that are found to not be robustly correlated with the outcome. In a robustness tests (Table 2.7), we show that the results do not depend on this choice.

¹⁰⁵ Note that these variables are lagged by two years instead of one as this avoids losing 37 observations and because most votes in the UNGA take place in November and December. As the correlation of the variables indicating ideal point distances from the G5 for $t-1$ and $t-2$ is $r = 0.94$ this is unproblematic and in the robustness section we show that our results do not depend on this choice.

¹⁰⁶ As the underlying density function of the percentage of breached thresholds shows, there are very few observations for this percentage being above ten. The results for these values are thus of no economic significance. The reason why the graph for US friends approaches zero for values close to and greater than ten is that there are no observations of US friends that receive an upgrading with this number of breaches. The model therefore correctly predicts a zero probability of overriding.

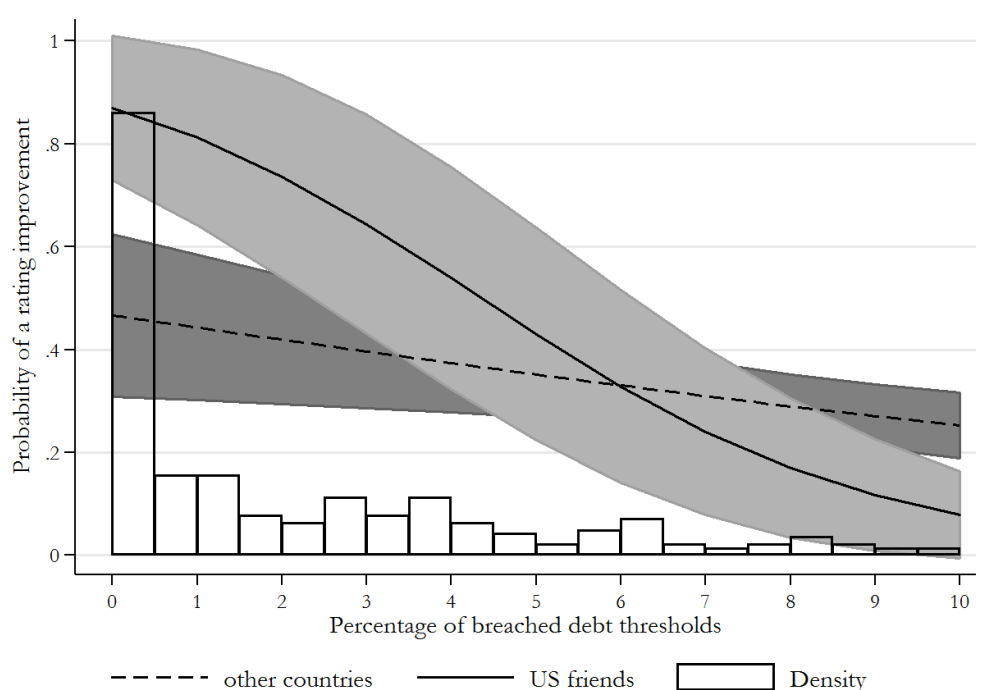
Table 2.5 – The Role of Political Interests and Bureaucratic Incentives

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
US friend (t-2)	0.421*** (0.122)	2.166*** (0.386)	-0.110 (0.164)		0.280 (0.377)	1.660*** (0.415)	-0.005 (0.443)	-2.905 (1.867)		-4.011 (2.763)
US friend (t-2) X Breaches		-0.394*** (0.114)				-0.426** (0.175)				
US friend (t-2) X Election			2.333*** (0.565)				1.054*** (0.355)	5.222*** (1.593)		6.665*** (2.004)
Election			-0.406 (0.254)				-0.367 (0.501)	-1.287 (0.788)		-0.980 (0.711)
Mechanical Rating: Risk Increased				4.466*** (0.343)	4.456*** (0.341)	4.409*** (0.301)	4.396*** (0.321)		3.630*** (1.174)	3.387*** (1.008)
Breaches	-0.100 (0.061)	-0.088** (0.043)	-0.109 (0.068)	-0.094*** (0.022)	-0.093*** (0.022)	-0.092*** (0.019)	-0.096*** (0.022)	-0.295*** (0.059)	-0.250** (0.102)	-0.344*** (0.122)
CPIA	1.252* (0.723)	1.231* (0.674)	1.350* (0.727)	0.007 (0.746)	0.020 (0.760)	0.096 (0.683)	0.092 (0.743)	17.041*** (5.928)	14.941 (9.384)	15.314* (7.825)
Mechanical Rating: High	2.201*** (0.428)	2.496*** (0.409)	2.387*** (0.466)	1.495*** (0.081)	1.502*** (0.092)	1.862*** (0.314)	1.558*** (0.054)	6.571*** (1.521)	4.753** (1.857)	5.712*** (2.030)
GDP per capita (ln, t-1)	1.017*** (0.109)	0.924*** (0.153)	1.085*** (0.138)	0.551*** (0.207)	0.521*** (0.199)	0.541** (0.221)	0.553*** (0.198)	19.810 (14.736)	19.959 (21.567)	20.603 (20.867)
GDP growth (t-1)	-0.012 (0.026)	-0.016 (0.017)	-0.016 (0.019)	0.041*** (0.010)	0.043*** (0.009)	0.039*** (0.006)	0.041*** (0.007)	-0.121 (0.170)	-0.082 (0.209)	-0.095 (0.220)
Population (ln, t-1)	0.406*** (0.053)	0.338*** (0.039)	0.436*** (0.056)	0.180* (0.102)	0.176* (0.094)	0.158* (0.090)	0.190** (0.092)	-10.478 (35.515)	-10.971 (18.994)	-12.593 (19.460)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Logit conditioned on	Region	Region	Region	Region	Region	Region	Region	Country	Country	Country
Observations	259	259	259	263	259	259	259	132	134	132
Pseudo R-squared	0.236	0.288	0.256	0.547	0.542	0.565	0.545	0.660	0.742	0.760

Note: dependent variable indicates an improved debt rating through overruling; conditional logistic regressions that control for region or year fixed effects; reported are the coefficient estimates of the regressions; standard errors, robust to clustering at the region/country level, in parentheses. Significance levels: * p<.10, ** p<.05, *** p<.01

the assumption of conditional exogeneity of political proximity to be violated—and thus cannot infer that US friends are on average favorably treated—we can still conclude that elections cause differences in the favorable use of judgment only for the group of countries classified as US friends.

Figure 2.1 – Visualizing the Effect of Political Alignment with the United States



Notes: The line chart shows the estimated probability (and the 90 percent confidence intervals) of a risk rating improvement for countries classified as “US friends” and for other countries, depending on the percentage of debt thresholds breached in the DSA. The line chart is based on the results of a logistic regression with region dummies, rather than the ones of the conditional logistic regression, as reported in Table 2.5 column 2. The results of these two regressions are almost identical except that the standard errors are smaller when the conditional logit estimator is used. This is why the graph understates the precision of the estimation. All control variables are set at their sample means. The bar chart shows the density function of the breached debt thresholds, at different 0.5 percent bins.

Next, we consider the effect of bureaucratic incentives (column 4). As the descriptive evidence already suggested, the probability that the mechanical rating is overruled increases significantly whenever the projections indicate a deterioration of debt sustainability. The *room for discretion* is used to avoid assigning higher risk ratings relative to the previous assessment.

The effect is economically substantial; if debt risk increases, the probability that the rating is overruled increases by 45 percentage points.¹⁰⁷

In column 5-7 we jointly test our two main hypotheses by replicating specifications 1-3 and adding the *Risk Increased* indicator. While in column 5 the coefficient on the *US friend* indicator is imprecisely estimated, all the other results, which exploit time variation, are robust to this modification. We conclude that – consistent with the “informal governance” model – bureaucratic and political biases are two distinct patterns.

Finally, in column 8-10 we replicate our main results (columns 3, 4, and 7) exploiting the within country variation and saturating the model with country fixed effects. While this specification is quite demanding, especially in a short panel, and substantially reduces the sample size, our results are still valid, so that we can confidently rule out that our main findings are driven by unobserved time-invariant country-specific characteristics.¹⁰⁸

2.4.3 Alternative Channels

In Table 2.6 we investigate and rule out further alternative channels that may explain the likelihood of obtaining an improved risk rating. According to a strand of the literature on the political economy of international organizations, major shareholders influence IOs not only for geopolitical reasons but also to pursue their economic interests. The literature highlights the particular role that commercial banks play in this context.¹⁰⁹ Gould (2003, 2006) argues that private financial institutions are able to influence conditionality in favor of “bank-friendly conditions” because they often provide necessary supplementary financing for IMF loans to borrowing countries. In a similar vein, Oatley and Yackee (2004) find that countries that are indebted to US banks receive larger IMF loans. As these loans are often used to repay debts to commercial banks (Bird 1996; Broz 2005), this suggests that IMF loans, in part, serve these actors’ financial interests. Broz and Hawes (2006) build on these findings to show that private creditors influence IMF decisions through the US government. Copelovitch (2010a) further qualifies these empirical results by demonstrating that both the size of IMF loans and the

¹⁰⁷ Effect size estimated via a linear probability model with country fixed effects based on column 9.

¹⁰⁸ As expected, the controls that do not vary much over time lose their statistical significance in this specification that absorbs all time-invariant cross-country variation (e.g., GDP per capita, growth, population, and the US friend indicator).

¹⁰⁹ See also Dreher and Richert (2017) as well as Malik and Stone (2017) for the influence that private companies have on the World Bank.

extent of conditionality are related to the financial exposure that banks from G5 countries have in IMF program countries. We follow Copelovitch (2010a) and proxy these economic interests with the financial exposure that banks from major shareholder countries have in the rated countries. The data we use come from the Bank for International Settlements (BIS) and include all cross-border consolidated claims of BIS reporting banks from the United States.¹¹⁰ Our results do not show any statistically significant evidence that bank exposure is related to the use of judgment to improve DSF risk ratings (column 1), even though the sign of the coefficients on the bank exposure variable and its interaction with the number of breaches are consistent with the presence of US economic and financial interests (column 2).

In the subsequent regressions we examine whether the political bias is in fact a cultural bias. In the literature on the political economy of sovereign credit ratings, Fuchs and Gehring (2017) have recently shown that the cultural distance between the home country of ratings agencies and the rated country affects the assessment of credit risk. Rating agencies perceive countries that are more distant in cultural and linguistic terms to be at higher risk than other comparable countries. Given that the World Bank's and the IMF's headquarters are both based in the United States, their staff are familiar with the United States' culture and language. If familiarity in cultural and linguistic terms indeed introduces familiarity bias that leads to a more favorable perception of rated countries and if culturally close countries are also politically closer to the United States we might misattribute a cultural familiarity bias to a political bias. In columns 3 and 4 we thus add a variable measuring a country's distance-adjusted ethno-linguistic fractionalization (DELFI) relative to the United States to the regressions (Kolo 2012).¹¹¹ Again, we find no statistically significant evidence for such a bias and rule out that the political bias we identify is in fact driven by cultural and ethno-linguistic familiarity.

In the literature on the political economy of the IMF it is well-established that countries with political ties to the United States are more likely to receive loans from the IMF. The finding that "US-friends" are more likely to benefit from the use of judgment in DSAs might thus be an artefact of the fact that the same countries are also more likely to be under an IMF arrangement.

¹¹⁰ We run the same regressions with data on bank exposure for banks from the European Union and find the same results.

¹¹¹ While this variable is the best measure we have for the cultural distance between the rated country and the IMF and World Bank, we admit that it is a rough proxy as their staff are very international. These results should thus be taken with a grain of salt.

There are several reasons that explain why countries that receive IMF resources might be treated differently in DSAs. In line with the reasoning in Dreher, Marchesi, and Vreeland (2008) and Reynaud and Fratzscher (2011), IMF staff might have an interest in ‘defensive rating,’ causing them to cast countries that receive IMF resources in a good light. At the same time, however, they might want to show that the country is at risk of debt distress to justify the existence (and a possible extension) of the program. When we include a variable indicating the presence of an IMF program in the year of the DSA, we initially find that such countries are less likely to receive improved ratings. As countries under IMF programs, however, see their mechanical DSF ratings deteriorate less frequently than countries not receiving Fund resources (39 vs. 61 percent), we control for *Risk Increased* and then find this effect to disappear. These findings are consistent with the following interpretation: The Fund adjusts the debt sustainability outlook for program countries less frequently than for non-program countries, since this would mean contradicting the projections on which the design of their programs was based. This is why overrides, on average, are also less frequent. However, if the projections are still adjusted under IMF programs, then judgment is used to avoid a downgrade: In fact, out of the 32 DSAs which suggested a deterioration of debt sustainability while the country was under an IMF program, 28 (87.5 percent) were overruled in order to keep the old rating.

In sum, we interpret this result as further evidence for the bureaucratic bias and—given that the coefficient on the US-friend variable is not affected—we can rule out the possibility that the political bias is due to the higher propensity to participate in IMF programs.

Table 2.6 – Testing for Alternative Channels

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
US friend (t-2)	0.427 (0.312)	1.865*** (0.609)	0.409** (0.190)	2.069*** (0.496)	0.323*** (0.092)	2.110*** (0.383)	0.263 (0.391)	1.652*** (0.312)
US friend (t-2) x Breaches		-0.356*** (0.099)		-0.378*** (0.128)		-0.393*** (0.111)		-0.433*** (0.166)
US banks (%GDP, t-1)	-0.260 (12.335)	10.916 (9.708)						
US banks x Breaches		-1.699 (4.336)						
DELF to US			-0.096 (1.059)	-0.627 (1.653)				
DELF to US x Breaches				0.087 (0.131)				
IMF program					-1.041*** (0.348)	-0.682** (0.301)	-0.209 (0.538)	0.295 (0.271)
IMF program x Breaches						-0.065 (0.057)		-0.066 (0.046)
Mechanical Rating: Risk Increased							4.405*** (0.426)	4.413*** (0.414)
Controls	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
Observations	259	259	259	259	259	259	259	259
Pseudo R-squared	0.236	0.294	0.236	0.289	0.258	0.316	0.543	0.574

Note: dependent variable indicates an improved debt rating through overruling; conditional logistic regressions that control for region fixed effects; reported are the coefficient estimates of the regressions; standard errors, robust to clustering at the region level, in parentheses. Baseline controls include: Breaches, CPIA, Mechanical Rating: High, GDP per capita (ln, t-1), GDP growth (t-1), Population (ln, t-1). All regressions include year fixed effects. Significance levels: * p<.10, ** p<.05, *** p<.01

2.4.4 Robustness

In Tables 2.7-2.9 we present a set of robustness exercises. We start by considering a number of additional factors that could motivate the use of judgment and possibly bias our estimates of the importance of political interests and bureaucratic incentives (Table 2.7).

First, in columns 1-3 we show that adding the larger vector of covariates (from Table 2.4, column 4) does not affect the results, notwithstanding a significant reduction in sample size. In addition to the current account balance, the debt-to-GNI ratio, the country's level of international reserves and its rents from natural resources, in these regressions we also consider the inflows of official development assistance from the US to rule out that the favorable treatment of US allies is due to the larger amount of US aid that such countries typically receive. We also add the interaction *CPIA x Breaches* to control for a potential non-linearity of the CPIA's effect depending on the number of threshold breaches.

Second, we include information on whether the rated country participated in the Heavily Indebted Poor Countries (HIPC) initiative and on whether it issued sovereign bonds at the time of the assessment. The former could signal a propensity to override the mechanical debt rating for reasons that have nothing to do with a political bias but are instead related to a benign view of countries which received debt relief, for which the IMF and the World Bank may have incentive to show that post-HIPC debt levels are sustainable. With regards to the latter, the fact that countries that issue sovereign bonds have easier access to external financing could also influence the assessment of debt sustainability and, in particular, the use of judgment. In fact, the specific structure of sovereign bonds—with bullet repayment at maturity—translates into sharp but temporary spikes in debt service, which could result in breaches of the relevant debt thresholds. Nevertheless, staff do not generally consider those breaches as relevant for debt sustainability and judgment is often used to overrule the mechanical rating. Those two variables, however, are not statistically significant and, more importantly, the coefficients on our variables of interest are not affected (columns 4-6).¹¹²

An alternative source of bias could arise from the fact that country teams can easily improve the risk rating by modifying the macroeconomic framework to incorporate overly optimistic

¹¹² However, in the first specification, which almost exclusively exploits cross-country variation, the standard error increases, resulting in the coefficient on *US friend* marginally missing statistical significance at the ten percent level ($p = .105$).

growth projections. Not controlling for IMF growth projections may bias the estimate of our key variables, especially because IMF growth projections could themselves be politically biased (Dreher, Marchesi, and Vreeland 2008). To rule out that a bias in growth forecasts drives our findings, we add the difference between IMF growth projections and the actual growth rate to the baseline regressions.¹¹³ Again, our results are robust to this modification (columns 7-9). One may also argue that ratings could improve because of a political bias in the CPIA. However, this is unlikely to be an issue in this context, given that the CPIA scores are computed by the World Bank for its aid allocation and — differently from growth projections — they enter the DSA as an exogenous input, over which the country team producing the DSA does not have any control. In any case, if there were a political bias in the IMF growth projections and in the CPIA, these biases would work against our results, so that our estimates can be safely considered a lower bound.

Finally, in columns 10-12 we remove all control variables except the DSA-specific variables and region and year fixed effects to compare how much the covariates affect the coefficient of interest. Following the method proposed by Altonji, Elder, and Taber (2005), we can then calculate “selection ratios” that indicate how much of the effect is explained by selection on observables. Comparing the “full” specifications (columns 1-3) to the specifications that only include the core covariates (columns 10-12, same sample) we find the selection ratios to be 5.4 for the simple effect of the US friend indicator, 4.3 for the level effect (and 30.2 for the interaction effect) when allowing for non-linearity in threshold breaches, and 10.1 for the effect in election years. This suggests that selection on unobservables in the first two specifications would have to be more than four times larger than selection on observables for the effects to be zero. Given that we already control for a large number of potential confounders, we think that it is unlikely that selection on unobservables drives this entire effect. The coefficient on the variable indicating US-friends with domestic elections is only marginally affected by the inclusion of covariates; the finding that selection on unobservables would have to be more than ten times as large as selection on observables to wash away this effect is consistent with our assumption that this interaction term is conditionally exogenous.

¹¹³ We can compute this difference only over a one-year horizon, because of data availability (extending the horizon would implicate a dramatic reduction in sample size, as we would have to focus only on “old” DSAs). However, we argue that if there is a strategic manipulation of growth projections, this should be already reflected in the first year.

In sum, these findings are in line with our expectation that the endogeneity bias potentially induced by omitted variables is not substantial in this setting. After all, the DSF projections already include most of the relevant country-specific information, explaining why many macroeconomic fundamentals we observe do not substantially affect the decision to overrule. Furthermore, both the fact that the non-linearity of the effect with respect to the number of breaches behaves as theory predicts, and the evidence that the effect is driven by election years—whose timing is exogenously predetermined—make us confident that our results are not due to endogeneity.

As discussed in the introduction, the timing of DSAs is also generally exogenous and follows an annual cycle. However, under specific circumstances, DSAs can be triggered by requests for exceptional IMF financing. While these cases are relatively rare in our sample, in Table 2.8 we test the robustness of our findings by excluding the DSAs not undertaken between six and 18 months after the previous one and thus potentially outside the regular, pre-determined schedule. Results are again confirmed (columns 1-3). In column 4, we recode the election dummy to consider exclusively those elections that took place within the six months following the publication of a DSA. We hypothesize that political interests in good ratings should be especially strong in the run-up to an election. It supports this hypothesis that with this more restrictive definition of the election dummy, we find an even stronger positive association between the *US friend* indicator and the likelihood of obtaining an improved risk rating in election periods.¹¹⁴

Finally, Table 2.9 reports the results of additional robustness exercises that use alternative definitions of the variable indicating political proximity to the major shareholders of the Bank and the Fund. First, we set alternative cutoff values to be classified as a “US friend” (top 25 percent and top 15 percent instead of top 20 percent). Second, we take a one-year lag instead of a two-year lag, at the cost of reducing the sample size. Third, we consider the mean of ideal point distances to all G5 countries instead of only the United States. Our main findings are robust to all these modifications.

¹¹⁴ In this specification, countries not classified as US friends appear to have a slightly lower probability to benefit from improved risk ratings in the run-up to an election. This could be considered weak evidence for attempts to politically weaken incumbent governments that are not politically aligned with the United States (Kersting and Kilby 2016).

Table 2.7 – Robustness: Additional Control Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
US friend (t-2)	0.559*** (0.097)	1.885*** (0.258)	0.117 (0.232)	0.416 (0.257)	2.194*** (0.593)	-0.152 (0.292)	0.302* (0.162)	1.951*** (0.560)	-0.325 (0.235)	0.686*** (0.150)	2.463*** (0.199)	0.244 (0.173)
US friend (t-2) x breaches		-0.380*** (0.144)			-0.386*** (0.120)			-0.378*** (0.135)			-0.393*** (0.110)	
US friend (t-2) x election			2.320*** (0.730)			2.400*** (0.651)			2.597*** (0.854)			2.091*** (0.744)
Election			-0.190 (0.458)			-0.393 (0.241)			-0.377 (0.236)			-0.199 (0.339)
Current Account Balance (% GDP, t-1)	-1.683 (1.383)	-0.829 (1.809)	-1.530 (1.367)									
Debt/GNI (t-1)	0.006 (0.004)	0.006* (0.004)	0.006* (0.003)									
Reserves/Debt (t-1)	0.006 (0.006)	0.005 (0.007)	0.006 (0.005)									
Natural Resources (% GDP, t-1)	0.015 (0.016)	0.008 (0.013)	0.012 (0.017)									
US aid (% GDP, t-1)	-3.931 (3.287)	-3.384 (3.247)	-3.934 (3.160)									
CPIA X Breaches	0.132** (0.066)	0.130 (0.089)	0.135** (0.068)									
HIPC (DP)				-0.526 (0.746)	-0.655 (0.794)	-0.492 (0.702)						
Sovereign Bonds				-0.931 (0.872)	-0.524 (1.045)	-1.121 (0.813)						
Error in IMF Growth Forecasts							0.007 (0.035)	0.012 (0.042)	0.014 (0.037)			
Additional Controls	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	DSA- specific	DSA- specific	DSA- specific
Number of Observations	244	244	244	259	259	259	251	251	251	244	244	244
Pseudo R-squared	0.301	0.334	0.317	0.244	0.293	0.265	0.242	0.287	0.267	0.198	0.253	0.216

Note: dependent variable indicates an improved debt rating through overruling; conditional logistic regressions that control for region fixed effects; reported are the coefficient estimates of the regressions; standard errors, robust to clustering at the region level, in parentheses. Baseline controls include: Breaches, CPIA, Mechanical Rating: High, GDP per capita (ln, t-1), GDP growth (t-1), Population (ln, t-1). DSA-specific controls include: Breaches, CPIA, Mechanical Rating: High. All regressions include year fixed effects. Significance levels: * p<.10, ** p<.05, *** p<.01

Table 2.8 – Robustness: Different Samples and Timing of Elections

	only regularly timed DSAs (between 6 and 18 months after previous)			only elections in 6 months following DSA
	(1)	(2)	(3)	(4)
US friend (t-2)	0.718*** (0.128)	2.297*** (0.363)	0.169 (0.150)	-0.049 (0.153)
US friend (t-2) X breaches		-0.357*** (0.065)		
US friend (t-2) X election			2.710*** (0.621)	
Election			-0.141 (0.322)	
US friend (t-2) x Election (in 6 months following DSA)				3.298*** (0.899)
				-0.559* (0.321)
Election (in 6 months following DSA)				
Controls	Baseline	Baseline	Baseline	Baseline
Number of Observations	200	200	200	259
Pseudo R-squared	0.275	0.315	0.299	0.264

Note: dependent variable indicates an improved debt rating through overruling; conditional logistic regressions that control for region fixed effects; reported are the coefficient estimates of the regressions; standard errors, robust to clustering at the region levels, in parentheses. Baseline controls include: Breaches, CPIA, Mechanical Rating: High, GDP per capita (ln, t-1), GDP growth (t-1), Population (ln, t-1). All regressions include year fixed effects. Significance levels: * p<.10, ** p<.05, *** p<.01

Table 2.9 – Robustness: Alternative Definitions of the Political Proximity Variable

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
US friend (top25%)	0.488*	1.829***	0.105									
	(0.257)	(0.582)	(0.266)									
US friend (top25%) x Breaches		-0.318**										
		(0.144)										
US friend (top25%) x Election			1.878**									
			(0.909)									
US friend (top15%)				0.522***	2.379**	-0.218						
				(0.154)	(0.970)	(0.307)						
US friend (top15%) x Breaches					-0.370							
					(0.229)							
US friend (top15%) x Election						2.560***						
						(0.972)						
US friend (t-1)							0.763***	2.139***	0.125			
							(0.293)	(0.588)	(0.371)			
US friend (t-1) x Breaches								-0.302*				
								(0.168)				
US friend (t-1) x Election									2.091**			
									(0.880)			
G5 friend (t-2)										0.698***	2.521***	0.266
										(0.115)	(0.425)	(0.163)
G5 friend (t-2) x Breaches											-0.403***	
											(0.156)	
G5 friend (t-2) x Election												2.057***
												(0.680)
Controls	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
Observations	259	259	259	259	259	259	229	229	229	259	259	259
Pseudo R-squared	0.237	0.288	0.252	0.236	0.274	0.256	0.309	0.344	0.328	0.240	0.297	0.256

Note: dependent variable indicates an improved debt rating through overruling; conditional logistic regressions that control for region fixed effects; reported are the coefficient estimates of the regressions; standard errors, robust to clustering at the region level, in parentheses. Baseline controls include: Breaches, CPIA, Mechanical Rating: High, GDP per capita (ln, t-1), GDP growth (t-1), Population (ln, t-1). All regressions include year fixed effects. Significance levels: * p<.10, ** p<.05, *** p<.01

2.4.5 Extension: Testing Manipulation of Projections: Density Discontinuity

The projections of debt levels underlying the DSF are based on a standardized econometric framework. Nevertheless, IMF and World Bank economists have several degrees of freedom as to how this framework is applied: minor changes in assumptions concerning, e.g., the expected growth rate, can substantially affect the number of threshold breaches. During our background research, staff thus suggested that biases could not only be present in decisions to override the mechanical rating, but also in the projections underlying the mechanical rating. We argue that if such manipulation took place it would be reasonable to expect bunching of projections just below the relevant threshold. In other words, the density of a variable

$$\Delta = p - t,$$

measuring the difference between projections p and relevant thresholds t would be low for values just above zero and high for values just below zero.

To test this hypothesis we rely on manipulation testing in a regression discontinuity design – an idea introduced by McCrary (2008). We employ a nonparametric test for a discontinuity in the density of Δ at the threshold (in our case zero). Specifically, we use the manipulation test developed by Cattaneo et al. (2016a, 2016b), which is based on a local polynomial density estimator that does not require pre-binning the data (see also Calonico et al., 2014). This choice allows us to take a purely data-driven approach to estimating the density near the cutoff and to test hypotheses regarding the density's discontinuity.

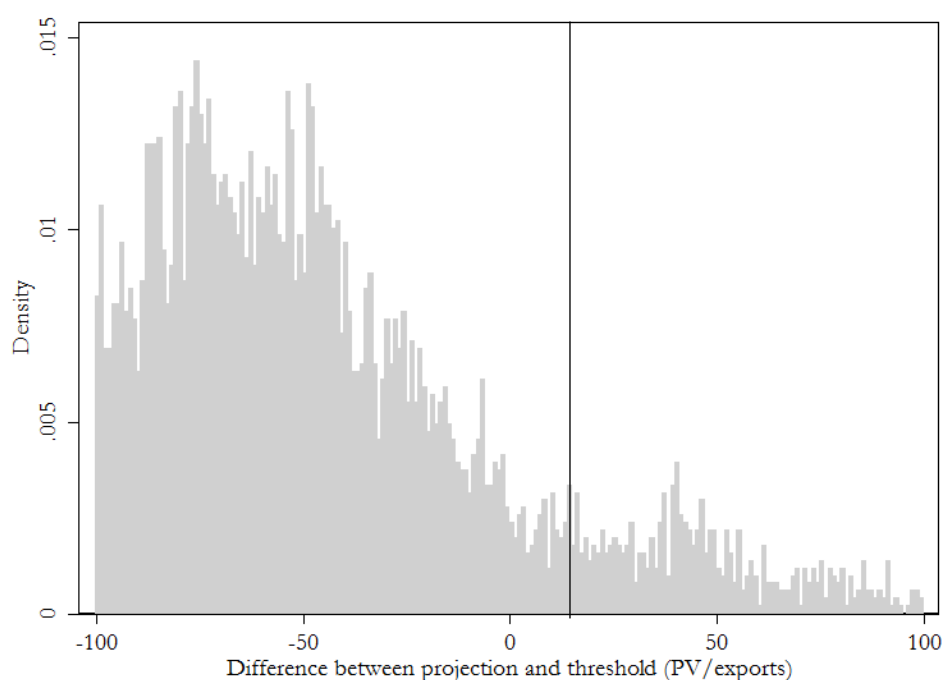
Table 2.10 – Manipulation Test

Projection	Observations (left and right of cutoff)	Bandwidth values (left and right of cutoff)	p-value
PV of debt / GDP	910; 130	6.642; 8.337	.364
PV of debt / Exports	872; 168	30.894; 33.529	.088*
PV of debt / Revenue	1027; 13 †	30.413; 35.868	.056*
Debt Service / Exports	993; 47	3.302; 3.708	.410
Debt Service / Revenue	1030; 10 †	3.354; 3.897	.018**

Notes: regression discontinuity manipulation test using local polynomial density estimation; p-value corresponds to a test of the null hypothesis that the density is not discontinuous; † test reports a warning that the bandwidth may be too low because of too few observations ($N_{\text{right}} < 20$); significance levels: * $p < .10$, ** $p < .05$

We run this test for each of the five baseline projections (Table 2.10). In the sample of US-friends the hypothesis that the density is continuous at the cutoff can be rejected for two of the five projection sets at the 10 percent level and for one set at the 5 percent level. For two of these three sets of projections with potential discontinuities (*PV of debt over revenue* and *debt service over revenue*), however, the density estimator for values exceeding the threshold is based on very few observations, leading to potentially unreliable results. For the projections of *PV of debt over exports*, however, the test suggests a discontinuity of the variable's density at the relevant threshold. Interestingly, this indicator is the one that exceeds the threshold most often and seems to have the strongest informative content among the five, since it signaled the risk of debt distress in about 80 percent of all high-risk cases.

Figure 2.2 – Testing Manipulation of Projections



Note: The graph is a histogram of a variable that measures the difference between the projected value of the debt indicator “present value of debt over exports” and the respectively relevant threshold. It visualizes the discontinuity of the density around 0 that the manipulation test (Table 2.10) reports.

Figure 2.2 shows the histogram of Δ for *PV of debt over exports*, whose particular shape around zero is consistent with the hypothesis of bunching just below the relevant thresholds. Nevertheless, we find the evidence too weak to reliably exclude the possibility that this distribution of projections around the thresholds came about by chance. In sum, as we can only reject the null hypothesis at the 10 percent level for one of the five tests, we do not

interpret this as definitive evidence for manipulation of projections. With reference to our main research question, however, we note that the size of the biases we find is, if anything, a lower bound as there is both anecdotal and weak statistical evidence that these biases also affect the estimation of projections and not only the decision to override the mechanical rating.

2.5 Conclusion

Political interests and bureaucratic incentives influence the decision-making of international financial institutions. While our findings confirm previous research that has documented such biases, our approach and data allow us to reveal evidence on the specifics of how these biases operate. Our results suggest that the *room for discretion* embedded in the technocratic rules that aim to ensure objectivity enables political interests and bureaucratic incentives to influence (and bias) the decision-making of IFIs. We find that the influence of political interests is stronger when formal rules are less clear-cut. This evidence supports scholars claiming that “informal governance” and “unwritten rules” in international organizations enable powerful stakeholders to intervene in their decision-making (Kilby 2013a; Koremenos 2013; Stone 2011). In the particular empirical setting we consider, we find that the assessments under the World Bank’s and the IMF’s Debt Sustainability Framework for low-income countries—the only tool that provides these countries with such debt risk ratings—are not free of bias. Both the geopolitical interests of the organizations’ major shareholders and bureaucratic incentives appear to be reflected in the ratings. These countries, however, have a strong and legitimate interest in unbiased and objective assessments of their debt sustainability.

These biased ratings might misguide the lending decisions of various—public and private—creditors and could thus entail adverse economic effects. The use of discretion may also, however, reduce the number of false alarms. Assessing the actual presence and extent of positive and negative economic effects is complicated by a number of identification challenges. First and foremost, the fact that the rating *per se* affects future economic and financing conditions makes future debt sustainability endogenous. Absent such an assessment and a discussion of the potential benefits of the use of discretion, our results point to some potential costs of discretion, but do not provide a normative statement of the trade-off between rules and discretion. We leave this exercise for future research.

2.6 Appendices to Chapter 2

2.6.1 Appendix 2.A: Relevance of the DSF

Table 2.11 – Effect of DSF Risk Ratings on World Bank Lending

	(1)	(2)	(3)
Moderate Risk	-0.010*	-0.015***	-0.020***
	(0.005)	(0.005)	(0.007)
High Risk	-0.011	-0.023***	-0.023**
	(0.007)	(0.008)	(0.011)
GDP/capita (ln)		-0.007	-0.005
		(0.004)	(0.004)
GDP growth		-0.002**	-0.002**
		(0.001)	(0.001)
Population (ln)		-0.005***	-0.005**
		(0.002)	(0.002)
Mechanical Rating: Moderate Risk			0.012
			(0.007)
Mechanical Rating: High Risk			0.004
			(0.010)
Year FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Observations	347	344	341
R-squared	0.155	0.226	0.238

Notes: OLS fixed effects regressions, dependent variable is World Bank lending (%GDP) in year t+1; robust standard errors in parentheses; significance levels: * p<.10, ** p<.05, *** p<.01

2.6.2 Appendix 2.B: Variables

Table 2.12 – Data Sources and Descriptive Statistics

Variable	Obs.	Mean	S.D.	Min	Max	Source and Description
Breaches	367	7.79	12.83	0	85.56	IMF DSF Dataset; percentage of breached debt thresholds
CPIA	364	3.37	0.43	2.43	4.36	IMF DSF Dataset; Country Policy and Institutional Assessment
Current Account Balance (%GDP, t-1)	370	-0.09	0.12	-0.55	0.43	IMF (2016d); balance on current account (%GDP)
Debt/GNI (t-1)	353	41.91	34.23	4.29	457.8	World Bank (World Bank 2016b); external debt stocks (% of GNI)
DELFI to US	370	0.8	0.15	0.52	1	Kolo (2012); distance adjusted ethno-linguistic fractionalization index (DELFI)
DSF rating improved	367	0.23	0.42	0	1	IMF DSF Dataset; indicator = 1 if mechanical risk rating is higher than official rating
Election	370	0.23	0.42	0	1	Beck et al. (2001), updated with information on election dates from Wikipedia; indicator = 1 if legislative or executive election took place in year t
Election (6 months later)	370	0.15	0.36	0	1	Beck et al. (2001), updated with information on election dates from Wikipedia; indicator = 1 if legislative or executive election took place in the six months following the publication of the DSA
Error in IMF Growth Forecasts	358	0.39	3.93	-9.17	34.33	IMF (2016d) different vintages; Growth rate projection for year t done in year t-1, minus actual GDP growth in year t
GDP growth (t-1)	368	4.84	4.85	-37.01	21.02	World Bank (World Bank 2016b); growth of GDP
GDP per capita (ln, t-1)	366	7.02	0.82	5.37	8.99	World Bank (World Bank 2016b); natural logarithm of GDP per capita in constant 2005 US\$:

Variable	Obs.	Mean	S.D.	Min	Max	Source and description
HIPC (DP)	370	0.54	0.5	0	1	IMF DSF Dataset; indicator = 1 if country participated in the heavily indebted poor countries initiative and reached at least the “decision point”
IMF program	370	0.53	0.5	0	1	IMF (2016b); indicator = 1 if IMF program in place in year t for at least 1 month
Mechanical Rating: High	367	0.34	0.48	0	1	IMF DSF Dataset; indicator = 1 if mechanical risk rating is “high”
Mechanical Rating: Risk Increased	369	0.22	0.41	0	1	IMF DSF Dataset; indicator = 1 if mechanical risk rating higher than mechanical risk rating in country’s previous DSA
Natural Resources (%GDP, t-1)	359	13.41	14.16	0.03	67.52	World Bank (World Bank 2016b); total rents from natural resources (% of GDP)
Population (ln, t-1)	370	15.52	1.92	9.2	18.99	World Bank (World Bank 2016b); natural logarithm of total population
Reserves/Debt (t-1)	352	60.27	50.28	3.39	431.49	World Bank (World Bank 2016b); total reserves in % of total external debt
Sovereign Bonds	370	0.06	0.25	0	1	Bloomberg; indicator = 1 if country issued sovereign bonds
US aid (%GDP, t-1)	369	0.02	0.09	0	0.8	USAID (2016); total US foreign (economic and military) assistance (%GDP)
US friend (t-2)	274	0.19	0.39	0	1	Bailey et al. (2017); indicator = 1 if UNGA voting ideal point distance to the United States is in the bottom quintile of the distribution in the sample
US banks (%GDP, t-1)	366	0.02	0.08	0	0.82	BIS (2016); foreign claims by BIS reporting US banks (%GDP), immediate borrower basis

3 Stigma or Cushion? IMF Programs and Sovereign Creditworthiness

Note: This paper is co-authored with Kai Gehring. An earlier version of it is available as *ETH and University of Zurich CIS Working Paper No. 98* (2018).

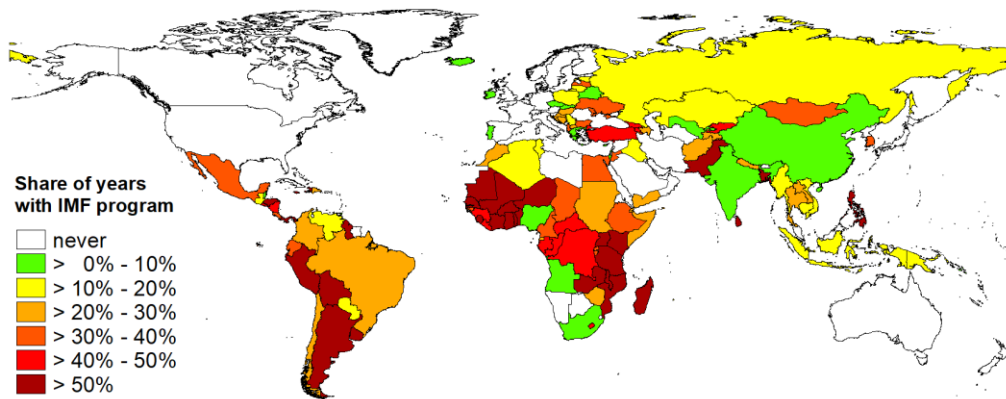
3.1 Introduction

In the early 2000s, the International Monetary Fund (IMF) was widely considered to be in terminal decline (Dieter 2006). The demand for its loan programs at a record low, the IMF reduced the size of its staff and focused on its “surveillance” activities (Reinhart and Trebesch 2015). The 2008 global financial crisis and the ensuing sovereign debt crises, however, re-established the crucial role that the IMF plays for the global economy. With the IMF’s financial commitments reaching new all-time highs in 2010, pressing questions about the role and effectiveness of the “most powerful international institution in history” (Stone 2002, 1) re-emerge.

We take this resurgence of the IMF’s lending activities as a motivation to re-evaluate how successful the IMF is in achieving one of its core mandates, namely to help countries overcome balance-of-payments problems. As these problems usually manifest themselves in both the government and private companies facing severe limitations in access to foreign capital, we focus on restoring market and investor confidence as a key outcome to evaluate the IMF’s success. We consider this an urgent task for economists not only because of the IMF’s widespread engagement (see Figure 3.1) but also because the IMF’s effectiveness in this regard has recently been questioned in policy circles. Out of fear of a “stigma” associated with the use of Fund resources triggering adverse market reactions, countries are often hesitant to enter IMF programs and question their benefits (Andone and Scheubel 2017; IMF 2017). Probably not least due to the alleged decline of the IMF – but also because of the empirical challenges associated with assessing its effectiveness – economists so far have no clear answer to this.

We begin our analysis of this question by illustrating the problem of endogenous selection into IMF programs. To measure market confidence in a country’s creditworthiness, we use a large monthly panel data set of sovereign credit ratings from different US and Non-US agencies as well as assessments from professional investors. Combined with start dates of IMF programs, these data unambiguously indicate that countries typically sign IMF agreements while their creditworthiness is in severe decline. The fact that countries tend to experience economic crises and negative trends in their main economic fundamentals when IMF programs begin, results in a substantial negative selection effect that biases downwards any estimates of the IMF’s effect on creditworthiness (and related measures) that do not adequately account for this.

Figure 3.1 – IMF Lending, 1973-2013



Note: Number in parentheses indicates share of years with an active IMF program in the 1973-2013 period.
Source: Dreher (2006, updated)

We apply multiple empirical approaches to circumvent this endogeneity problem. Our main identification strategy is based on an instrumental variable (IV) that combines temporal variation in the IMF's liquidity with cross-sectional variation in a country's prior probability of participating in an IMF program (see chapter 4). The IMF's liquidity varies primarily because of an institutional rule that requires the IMF to review the financial contributions of its members ("quotas") every five years and as a consequence of large individual loan repayments. For identification, we exploit the finding that the IMF tends to expand its regular clientele in years in which its liquidity is higher, so that countries with an initially lower participation probability are more likely to receive a program in these years.

Using annualized panel data for a maximum of 100 countries over the 1987-2013 period, we find that the simple correlation of IMF programs with sovereign ratings is strongly negative. As one would expect in the presence of a downward bias, the OLS coefficient, while remaining negative, moves increasingly close to zero when conditioning step-by-step on country and year fixed effects as well as lagged macroeconomic and political controls. We then show that the remaining small, negative relationship turns statistically insignificant (and positive) when switching to the IV approach. This pattern emerges irrespective of whether we focus on ratings issued by Standard & Poor's, Moody's or Fitch, from non-US rating agencies based in Europe and Asia or when employing assessments by *Institutional Investors*, which are based on surveys among professional investors and analysts at banks, money management and securities companies.

When turning to the mechanisms, we find that the absence of a significant aggregate effect masks important underlying dynamics. Our evidence suggests that the economic adjustments under IMF programs substantially reduce economic growth in the short run. Given that both official rating agency methods and empirical evidence show that ratings are directly (positively) influenced by GDP and growth, these contractionary effects would usually result in lower credit ratings. The fact that they are not affected suggests that IMF programs have a *positive signaling effect*. This signal (“seal of approval”) creates positive expectations about the country’s future policy path and “cushions” the drop in creditworthiness that countries undergoing such contractionary adjustments would usually suffer from.

We corroborate this result in three ways: First, specifications that ignore “bad control” problems reveal that when controlling for GDP dynamics, the conditional effect of IMF programs becomes substantially positive and statistically significant; IMF-induced GDP contractions, on the other hand, do not seem to significantly affect credit ratings. Second, we use the data on credit ratings at a monthly frequency and information on the exact date of IMF agreements, and isolate variation within individual country-year observations with the help of country-times-year fixed effects. Event-based specifications then show that ratings deteriorate before IMF agreements, but begin to improve in the month after the programs start. We argue that these immediate improvements cannot plausibly be attributed to the success of economic adjustments and political reforms but only to a positive *signaling* effect. Third, we conduct a systematic text analysis of statements about the IMF’s influence on sovereign credit ratings available on the news database *Dow Jones Factiva*. Out of 117 statements from rating agencies that mention the IMF, 84 indicate a positive influence of an active IMF program on their assessment while only one mentions a negative influence. A majority of these statements refer to the anticipated positive effects on investor confidence of policy reforms implemented as part of the programs.

In the remainder of this paper, we first examine theoretical expectations regarding potential mechanisms based on the existing literature in section 3.2. Section 3.3 presents our identification strategies and data. We report and discuss the empirical results in section 3.4. Section 3.5 concludes this chapter.

3.2 Potential Channels and Existing Literature

To increase creditworthiness, IMF programs need to increase investors' confidence in the "ability and willingness of an issuer [...] to meet its financial obligations in full and on time" (Standard and Poor's 2016; see also Panizza et al. 2009; Tomz and Wright 2007). We argue that it is helpful to differentiate between two main channels. First, we define *adjustment effects* as effects on ratings that are consequences of short-term changes in the country's economic and political fundamentals under IMF programs. Second, we consider *signaling effects* as changes in ratings caused by the signals about the country's expected future policy path that the presence of an IMF program sends to credit rating agencies and investors. As we build on this conceptual separation when empirically analyzing the channels driving the effects we find, we discuss theoretical considerations and existing evidence on both channels in the following.

3.2.1 Adjustment Effects

Consider *adjustment* effects first. A sovereign's creditworthiness as measured by credit ratings is most strongly influenced by the country's economic and political fundamentals. In the empirical literature on the determinants of ratings, gross domestic product (GDP) per capita, GDP growth, inflation, and external debt are found to be robust predictors (Afonso 2003; Cantor and Packer 1996; Fuchs and Gehring 2017; Hill, Brooks, and Faff 2010). Several political indicators like the political regime type, partisanship, and the rule of law have also been found to correlate with rating outcomes (Archer, Biglaiser, and DeRouen 2007). Together, these variables explain a large share of the variance in sovereign ratings. These results in the scholarly literature are in line with official rating manuals, which agencies publish to comply with regulatory standards.

The previous literature on the IMF examines several of these key determinants of creditworthiness as the outcomes of IMF programs (for reviews of this literature see Dreher and Lang 2016; Steinwand and Stone 2008). To the extent that IMF programs affect such outcomes, they influence creditworthiness via the *adjustment* channel. In terms of economic fundamentals, the focus of many such studies has been on economic growth. While some studies find a positive (Bas and Stone 2014) or insignificant (Atoyan and Conway 2006) relationship between IMF programs and growth, the majority of empirical studies suggest negative immediate effects (Barro and Lee 2005; Dreher 2006b; Easterly 2005; Marchesi and Sirtori 2011; Przeworski and Vreeland 2000). Beyond growth, monetary stability, debt

management and the containment of external arrears are key goals of IMF programs (Kentikelenis, Stubbs, and King 2016). IMF programs are associated to reduced inflation and monetary growth, less risk of currency crises and banking crises, and improved market performance of banks (Dreher and Walter 2010; Papi, Presbitero, and Zazzaro 2015; Steinwand and Stone 2008).

In addition to these economic effects, IMF programs also appear to affect political outcomes. Several scholars link IMF programs to political instability and suggest that they increase the risk of civil war onset (Hartzell et al. 2010), coup d'états (Casper 2017), and government crises (Dreher and Gassebener 2012). One explanation for these politically destabilizing effects of IMF programs is that the burdens of economic adjustments under IMF programs are often distributed unequally (chapter 4; Vreeland 2002).

In sum, the existing evidence suggests some positive adjustment effects regarding financial and monetary indicators, but mostly negative adjustments regarding reduced growth and political instability. It is thus an open empirical question as to whether the immediate implementation of adjustment policies resulting from IMF interventions leads to improvements or deteriorations in creditworthiness. Before we turn to testing this empirically, we distinguish these *adjustment* effects from *signaling* effects.

3.2.2 Signaling Effects

Sovereign credit ratings, as assessments of a future default probability, are based not only on information about a country's current economic and political performance, but also on expectations of the country's future development (Fuchs and Gehring 2017). As economic indicators like GDP and inflation are imperfect and noisy measures, it is rational for investors and rating agencies to use other signals to infer information and adapt their assessment. Any signal that gives an indication about the country's future policy path will influence this expectation. IMF programs can plausibly serve as such a signal.

On the one hand, they could function as a "seal of approval" (Polak 1991). The Fund itself claims that "IMF resources provide a *cushion* that eases the adjustment policies and reforms that a country must make to correct its balance of payments problem" (IMF 2016a, emphasis added). Accordingly, the provided liquidity is intended to enable a period of IMF-approved adjustments. In addition, the Fund can "lend credibility" (Stone 2002) to the announced reforms, and function as a commitment device to overcome time consistency problems (Dreher

2009a).¹¹⁵ To the extent that the IMF agrees to arrangements only if it approves of the country's policy agenda, its engagement also indicates reform quality (Dreher 2009a; Marchesi and Thomas 1999). For these reasons, the Fund's engagement can positively affect expectations about the reforms' effects on sustainability and macroeconomic performance (Edwards 2006; Mody and Saravia 2006; Corsetti et al. 2006; Morris and Shin 2006).¹¹⁶

On the other hand, IMF programs may convey negative information (Andone and Scheubel 2017; Bas and Stone 2014; Ito 2012). If IMF programs are perceived as indicating that the country's financial problems are more severe than official indicators suggest, they can act as a negative signal. The IMF (2014a) itself, for instance, is worried that countries under its loan programs carry a "stigma" (see also Reinhart and Trebesch 2015). Our background research and interviews at the IMF's headquarters with IMF staff revealed that in the recent past several countries indeed hesitated to sign Fund agreements out of fear of such a stigma.¹¹⁷ In a recent statement on lending reforms the IMF (2017) states: "[a] key objective of the lending reform is to reduce the perceived stigma of borrowing from the IMF."

3.2.3 *Issues with the Existing Evidence*

Existing empirical studies linking IMF programs with creditworthiness have produced inconsistent results. We argue that issues with the proxies used as outcome variables and with accounting for selection bias are likely to be behind this inconsistency.¹¹⁸

A first set of studies examines the IMF's effect on inflow of different kinds of capital, mostly foreign direct investment (FDI). In an early review, Bird and Rowlands (2002) conclude that the empirical literature suggests that IMF programs *reduce* countries' access to capital markets. In the following, some studies found a negative effect (Bird and Rowlands 2009; Edwards 2006; Jensen 2004), insignificant results (Rowlands 2001) or evidence for a (conditionally) positive

¹¹⁵ This conjecture is in line with the literature on the effects of membership in international organizations more broadly (Dreher and Lang 2016). Membership in international organizations can improve borrowing conditions and increase inflows of foreign capital (Dreher, Mikosch, and Voigt 2015; Dreher and Voigt 2011; Gray 2009; 2013).

¹¹⁶ An additional signaling effect of IMF programs discussed in the literature is the creditor moral hazard problem. The IMF could lead creditors to increase investments in government bonds of program countries because they anticipate IMF bailouts. Dreher's (2004, 20) literature survey concludes that there is "considerable evidence in favor of the hypothesis that the safety net provided by the IMF creates significant moral hazard with investors." We are not separately examining this aspect, as we are only interested in whether the IMF helps countries to restore creditworthiness, one way or the other.

¹¹⁷ Multiple conversations with several IMF employees in the period between November 2016 and November 2017.

¹¹⁸ Steinwand and Stone (2008) and Bauer et al. (2012) reach the same conclusion in their literature reviews. We refer the reader to these studies for a more detailed overview of this literature.

effect on FDI inflows (Bauer, Cruz, and Graham 2012; Biglaiser and DeRouen 2010; Woo 2013). Jorra (2012) uses an indicator for sovereign default and finds an increased probability of default as a consequence of IMF lending. Another set of studies examines the IMF's effect on government bond spreads. Among these, Mody and Saravia (2006) and Eichengreen, Kletzer, and Mody (2006) find lower bond spreads in IMF program countries.¹¹⁹ Chapman et al. (2015) report that implementing an IMF program is associated with higher bond spreads, but find that the size of the IMF loan, the extent of conditionality, and the political proximity of the program country to the United States all lead to important heterogeneities.

We argue that all of these measures come with important problems that can be avoided when using sovereign credit ratings as a proxy for creditworthiness. Compared to FDI, ratings are a more direct and precise proxy for creditworthiness. FDI flows are an indirect consequence of creditworthiness but are influenced by many other factors like economic openness and, in addition, capture only a fraction of total capital inflows. Compared to using defaults, which are very rare events, as a proxy, ratings provide a more fine-grained assessment and go beyond capturing only the extreme end of the wide spectrum of balance of payment problems. Compared to bond spreads, credit ratings are available for a larger set of countries and remain a reliable measure in times of crisis. Ratings provide a continuous measure of creditworthiness that, unlike bond spreads, is not directly influenced by changes in general market conditions such as shifts in demand for different asset classes (e.g., fixed income vs. equity) and risk categories (e.g., flight into quality), and bond supply effects. If governments under IMF programs adjust the supply of government bonds or when central banks acquire them, bond spreads convey a biased and inaccurate picture of how investors perceive the creditworthiness of a country. In addition, the liquidity of trading, which is crucial for the informational value of the market price, is often low for countries in crises. In sum, bond spreads are the least informative at the time when we are most interested in the information they convey. Ratings, on the other hand, are at all times easily comparable across countries and over time as they proxy for the same latent variable in each case. To the best of our knowledge, in this literature only Cho (2014) uses a measure that is related to credit ratings, and finds that assessments by

¹¹⁹ A different but related series of studies has looked at how government bond spreads react to IMF signals regarding the likelihood of future bailouts (see footnote 116). Some studies find evidence for such "creditor moral hazard" caused by the IMF (Dell'Ariccia, Schnabel, and Zettelmeyer 2006; Lee and Shin 2008). Other studies provide evidence against the argument (e.g., Noy 2008).

Institutional Investor correlate positively with IMF programs in countries with left-wing governments.

The second shortcoming in the literature summarized above is that most studies do not establish causality in a convincing way. As selection into IMF programs is clearly not random, endogeneity severely biases the estimates of analyses that do not account for this. Mody and Saravia (2006, 852), the most cited study in this field, state that due to the difficulty of modeling selection into IMF programs and finding a suitable instrument, “explicit consideration of the selection bias problem is not undertaken.” While Jorra (2012) uses an instrument, its underlying assumption that IMF programs are the only plausible channel that link a country’s political proximity to the United States and default events is unlikely to hold: A country’s economic condition is plausibly related to the political preferences of the country’s government via more direct channels.¹²⁰ Chapman et al. (2015) provide instruments for the extent of IMF conditionality and IMF loan size, but do not instrument the presence of an IMF program.

Most other studies in this literature address endogeneity by controlling for a range of observable factors. As we explain in more detail in the next section, this is unlikely to remove the entire bias and, in addition, often creates a bad control problem. In sum, the astonishing differences in empirical results are potentially also attributable to the lack of plausible identification strategies. The empirical approach we present in the following aims to augment the literature in these respects.

3.3 Data and Identification

3.3.1 Dependent Variable: Sovereign Credit Ratings

Our main proxy to measure the creditworthiness of a country is its sovereign’s long-term foreign-currency rating. In addition to their aforementioned advantages over other measures used in the previous literature, sovereign ratings possess several additional features that make them good proxies for sovereign creditworthiness: First, Reinhart (2002) shows that ratings predict defaults. This makes them an informative measure of creditworthiness for countries

¹²⁰ See chapter 4 for a detailed evaluation of different empirical strategies that have been used in the literature on the IMF’s effects and the need for a new instrument. When we use political proximity to the United States as an IV in our sample we do not find the IV to be sufficiently relevant in the first stage.

with severe payment problems, an important feature for our research question. Second, previous studies have related ratings to changes in government bond spreads (Afonso, Furceri, and Gomes 2012). They thus indicate the terms at which a country can access international capital markets. Third, many investors, in particular pension funds but also insurances and to some degree banks, are bound by internal regulations that restrict investments to investment-grade bonds. In addition to the information effect that bond assessments convey to investors, this “hard-wiring” is another reason why rating changes directly affect refinancing costs of governments. Fourth, ratings serve as a de-facto ceiling for the credit rating of private companies from the respective country (Borensztein, Cowan, and Valenzuela 2013), and hence also capture the private sector’s ease of access to foreign capital. The main criteria applied to assess sovereign bonds are to a large degree comparable across agencies, but there are some differences (Fuchs and Gehring 2017). While our main estimations rely on ratings from Standard and Poor’s (S&P), which offers the broadest country coverage over the longest period, we also use ratings from Moody’s and Fitch to show that the existing differences across agencies do not drive the results.¹²¹ Since the three major rating agencies are based in the US, and cultural distance between an agency and the sovereign it assesses can influence ratings (Fuchs and Gehring 2017), we also consider ratings from agencies based elsewhere: The variable *Ratings(Non-US)* captures the average of all major agencies outside the United States.¹²² This included the Japanese agencies Japan Credit Rating (JCR) and Rating and Investment Information (R&I), as well as the German agency Feri, the Canadian Dominion Bond Rating Services (DBRS) and Capital Intelligence (CI) from Cyprus.

We use hand-collected information on sovereign ratings by most agencies from Bloomberg (see Appendix 3.A, as well as Fuchs and Gehring, 2017 for details). Ratings published by Feri and Fitch are directly from the agencies. To analyze the dynamics around IMF program starts we use data at a monthly frequency. Our panel regressions at the yearly level use ratings at the end of the year. All ratings are translated to a 21-point scale, assigning the highest value for a “AAA” rating, while “C” and below translates into a value of one (see Appendix 3.B).

¹²¹ S&P covers most high and middle-income countries. The IMF itself – jointly with the World Bank – rates the risk of debt distress under the so-called Debt Sustainability Framework (see chapter 2).

¹²² Fitch Rating is dual-headquartered in London, UK and New York, USA.

3.3.2 Treatment Variable

The explanatory variable of interest (or “treatment” variable), *IMFprogram*, is an indicator that takes the value of one if country *i* was under an IMF program for at least five months in year *t* (as in Dreher 2006). Following the previous literature, our definition encompasses all IMF programs under any of the following facilities: Stand-By-Arrangements (SBA), the Extended Fund Facility (EFF), the Structural Adjustment Facility (SAF), or the Poverty Reduction and Growth Facility (PRGF). In alternative specifications, we also use the variable *IMFapproval*, which indicates only the year in which an IMF program was initially approved. To corroborate our arguments concerning biasing factors and channels, we also use an alternative monthly dataset in which we use information on the exact date an IMF program was approved. The latter we coded based on the IMF’s Monitoring of Fund Arrangements (MONA) database (IMF 2016c).

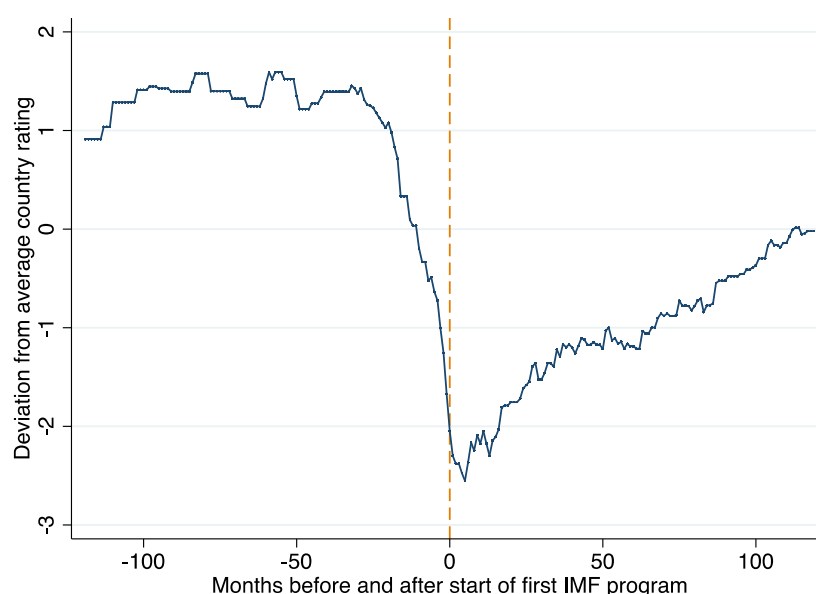
3.3.3 Endogenous Selection into IMF Programs

We want to know whether the presence of an IMF program in year *t* in country *i* affects the country’s credit rating at the end of year *t*. The fundamental methodological issue with this question is that selection into IMF programs is obviously not random. On the contrary, “treated” countries typically experience an economic crisis when entering into a program, and the more severe the crisis is, the more likely it is that a country is under an IMF program. As a consequence, simple comparisons between treated and non-treated country-year observations will not yield causal effects but capture the negative bias resulting from omitted variables and reverse causality. The deteriorating economic conditions that make a country more likely to enter an IMF program negatively affect a country’s creditworthiness, and a country with lower creditworthiness is thus more likely to turn to the IMF. In the following, we show why controlling for selection-on-observables is insufficient and propose an alternative strategy.

To begin with, we use our monthly data on sovereign credit ratings and on the exact date that countries enter into an IMF arrangement to illustrate the problem graphically. Figure 3.2 plots the average behavior of credit ratings around *IMFapprovals*. Specifically, on the y-axis the figure depicts the unweighted average of the month-specific deviations from each country’s mean credit rating in the 1990–2013 period over all countries that received an IMF program at

least once in this period.¹²³ For all countries, “month” on the x-axis is set to zero for the month in which the country’s first IMF program in the observation period started.

Figure 3.2 – Rating Dynamics Around Starts of IMF Programs



Note: The figure plots the unweighted mean across countries of the month-specific deviation from each country’s average S&P credit rating in the 1990-2013 period on the y-axis. The number of months around the start of the country’s first IMF program of this period is on the x-axis. Sample restricted to countries with at least one IMF program.

Several important observations are evident. First, credit ratings appear to capture balance-of-payment crises well. As one would expect, countries enter into IMF agreements several months after economic crises hit and creditworthiness collapses. On average, countries’ credit ratings deteriorate by about three notches in the approximately one and a half years preceding the IMF program’s beginning. Second, IMF programs start at a low point, but creditworthiness continues to fall for several months thereafter. After about a year, ratings seem to begin recovering. Third, this recovery process is on average rather slow: It takes several years until creditworthiness is restored to pre-crisis levels. The figure also illustrates the problem of endogenous selection into the treatment. Credit ratings are at a low level and in an ongoing process of decline during the early months of IMF programs for reasons at least partly unrelated to the IMF program itself. Given that the average IMF program in our sample lasts for about four years (with a large variance), any simple regression of credit ratings on a

¹²³ Examining the deviation from the country mean is equivalent to using country fixed effects in a panel regression.

variable indicating the start or the presence of an IMF program is biased by the fact that IMF programs typically start when ratings are already low and in further decline.

A basic model designed to estimate the effect of *IMFprogram* on *Rating* based on controlling for selection-on-observables looks like the following:

$$Rating_{i,t*} = \beta IMFprogram_{i,t} + \mathbf{X}'_{i,t-1}\gamma + \delta_i + \theta_t + \varepsilon_{i,t} \quad (1)$$

In a regression equation of this type $\mathbf{X}'_{i,t}$ is a vector of country-year specific observable control variables and δ_i and θ_t stand for country fixed effects and year fixed effects, which control for unobserved time-invariant country characteristics and for global shocks that affect all countries equally. $\varepsilon_{i,t}$ is the i.i.d. error term and t^* indicates the value at the end of year t .

Formally, our expectations regarding endogeneity in this setting can be written as:

$$E(IMFprogram_{i,t} \varepsilon_{i,t}) < E(IMFprogram_{i,t} \varepsilon_{i,t} | \delta_i, \theta_t) < E(IMFprogram_{i,t} \varepsilon_{i,t} | \mathbf{X}', \delta_i, \theta_t) < 0 \quad (2)$$

It is natural to expect that fixed effects reduce the negative bias in this estimation: Global business cycles could affect both creditworthiness and the demand for IMF programs and, more importantly, typical IMF program countries tend to have time-invariant characteristics that make them economically weaker, and thus less creditworthy, than the countries that do not tend to receive such programs. Furthermore, it is plausible that country-year specific control variables further reduce this bias because they make treatment and control groups comparable in terms of observables. Nevertheless, such an empirical strategy is insufficient and problematic for at least three reasons.

First, the available cross-country panel data on macroeconomic and political fundamentals are unlikely to capture all information that ratings agencies, national policy-makers, and IMF staff had available at the specific time decisions about creditworthiness and IMF participation were made. This includes information on country-specific economic vulnerabilities or political events that rating agencies consider when assessing creditworthiness and that decision-makers in the IMF and in national governments take into account when deciding on starting or continuing IMF programs. Because many of these vulnerabilities and events are context-specific not all of them are comprehensively reflected in the macroeconomic and political cross-country data that are available for a large panel of countries.

Second, even if all relevant economic and political fundamentals could be observed and measured at the country-year level, this would not necessarily solve the problem. Most of these indicators are available only at the yearly level and ignore the crucial dynamics within a year. Economic and political fundamentals in countries that enter IMF programs are likely to

deteriorate quicker during the year. Figure 3.2, which is based on monthly data, illustrates this. A focus on country-year specific means of observable controls would not be able to control for this unobserved heterogeneity between treatment and control groups.

Third, many of the control variables that in this particular setting are needed to increase the comparability of treatment and control groups are “bad controls” because they are themselves plausible outcomes of the treatment. As discussed above, the literature suggests that ratings are a function of many of the same economic and political fundamentals that the IMF directly affects. If, for instance, IMF programs increased growth and rating agencies improved their credit assessment because of this, holding growth constant would prevent the regression from attributing the positive effect of the IMF on creditworthiness via the growth channel.¹²⁴ Usually, it mitigates this problem to lag these variables by one or two periods, but IMF programs last for multiple years and there is also persistence in both ratings and fundamentals. Lagging the values can thus mitigate but not entirely solve the problem in this setting.¹²⁵

In sum, estimation strategies that rely on controlling for selection on observables cannot adequately address the question at hand. Ideally, we would want a mechanism that randomly assigns countries that are on comparable trajectories to an IMF program. We approach such an ideal assignment mechanism by employing an instrumental variable (IV) that changes the likelihood that a particular country receives a program based on factors that are exogenous to the trajectory of this particular country. This leads to an assignment into treatment that is as-good-as-random.

¹²⁴ To see this, make the uncontroversial assumption that $\gamma > 0$ in equation (1) if \mathbf{X}' is a “positive” economic fundamental like GDP growth. If, for instance, IMF programs indeed increase growth and, thus, $cov(IMFprogram, X) > 0$, the OLS estimator becomes $\hat{\beta} = \frac{cov(Rating, IMFprogram)}{var(IMFprogram)} - \hat{\gamma} \frac{cov(IMFprogram, X)}{var(IMFprogram)}$, such that the second expression is negative and reduces $\hat{\beta}$. Hence, if we want $\hat{\beta}$ to capture the effect of $IMFprogram$ on *Rating* that includes the effect that operates via changes in economic fundamentals we must not control for \mathbf{X}' .

¹²⁵ Additionally, holding \mathbf{X}' constant if IMF programs have an effect on the intermediate outcome \mathbf{X}' creates the problem that treatment and control groups will differ in potential outcomes. The coefficient of the treatment will compare units that are identical in \mathbf{X}' but differ in $IMFprogram$. If, however, the treatment affects \mathbf{X}' , then potential outcomes of these units will automatically differ and this coefficient will not estimate a causal effect. For more details on this point see chapter 9 in Gelman and Hill (2007) or chapter 3.2.3 in Angrist and Pischke (2008).

3.3.4 Identification

Our instrumental variable strategy combines spatial and temporal variation via an interaction term. The IV is based on Lang (chapter 4) and is defined as:

$$IV_{it} = IMFprobability_{it} \times \ln(IMFliquidity_t) \quad (3)$$

IMFprobability is a country's probability of having participated in an IMF program in the past, defined as the share of past years that a country was under an IMF program.¹²⁶ *IMFliquidity* denotes the IMF's time-varying liquidity ratio, defined as the organization's liquid resources divided by its liquid liabilities. It is a measure of the amount of liquid resources the Fund has available for its loan programs in a given year. The data are taken from individual IMF Annual Reports (1973-2013) and from the IMF International Financial Statistics. Chapter 4 provides further details.

To see how the instrument works, note first that variables indicating past IMF participation, like *IMFprobability*, are strong predictors of present participation in IMF programs (Berger, de Haan, and Sturm 2005; Bird, Hussain, and Joyce 2004). Based on this observation, the IV exploits the fact that the influence of prior *IMFprobability* on *IMFprogram* participation differs conditional on the IMF's liquidity ratio. More specifically, a country's IMF participation history is a weaker predictor of *IMFprogram* in years in which the IMF's liquidity is high. Figure 3.3 illustrates this relationship by plotting the marginal effects of *IMFprobability* on *IMFprogram* conditional on the level of *IMFliquidity*.

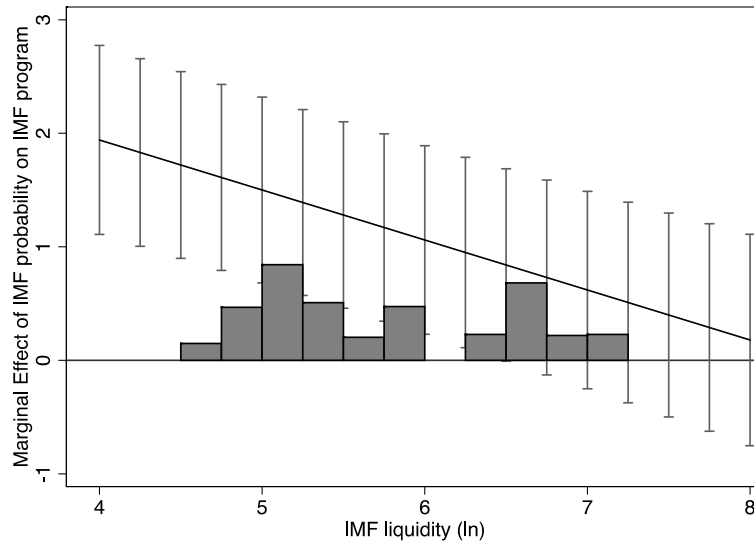
Our interpretation of this finding is as follows. In years with relatively low levels of IMF liquidity, IMF programs tend to go to countries that routinely receive these programs. During these times, *IMFprobability* is thus a strong predictor of *IMFprogram*. In years in which the IMF's liquidity is high, however, the Fund can be more generous and has an incentive to look for additional program countries beyond its more regular clientele. In line with the public choice literature on international organizations (see Dreher and Lang 2016 for an overview), Fund staff face bureaucratic incentives to expand its field of activities. The incentives and the financial opportunities to do so are particularly strong when liquidity is high.¹²⁷ This is likely

¹²⁶ We start the count of years of past IMF participation in 1973 and thus 15 years before our observation period starts. This ensures that the variable does not fluctuate strongly from one year to the next in the early years of the sample. As we calculate the probability based on values from before year t , the variable is predetermined at time t .

¹²⁷ This literature argues that international bureaucrats face incentives to increase their budgets, remits, staff, relevance, and political influence and thereby contribute to the expansion of IOs in size, power and responsibilities in an increasing number of policy areas and countries. These arguments in the scholarly literature are in line with anecdotal evidence we gathered in personal conversations with IMF staff: Inside the IMF, there is a certain concern

to be the explanation for why in these years, the IMF is more likely to give loans to countries that are otherwise less frequently under an IMF program. As a consequence, the relationship between past participation and present participation becomes weaker in years with a higher IMF liquidity ratio, as we see in Figure 3.3.

Figure 3.3 – Illustrating the First-Stage Effect



Note: The figure plots the marginal effects of *IMFprobability* on *IMFprogram* for varying levels of *IMFliquidity*. It corresponds to the first-stage regression of our baseline IV regression, as reported in column 6 of Table 3.1. For more details see the main text (as well as chapter 4).

Our identifying assumption is that this mechanism affects loan allocation at the level of the IMF, but that the isolated interaction effect (conditional on its constituent terms) is orthogonal to other economic developments at the country-year-level. To see why we consider this assumption to be plausible, note first that this identification strategy follows a difference-in-differences logic as in Werker, Ahmed, and Cohen (2009) and Nunn and Qian (2014). As the constituent terms forming the interaction term are controlled for, we need to assume that the interaction term is exogenous conditional on the level of the constituent terms:

$$E(\varepsilon_{i,t} \text{ IMFprobability}_{it} \text{ IMFliquidity}_t | \text{IMFprobability}_{it}, \text{IMFliquidity}_t) = 0 \quad (4)$$

to lose relevance when there are only few active IMF programs. Several staff stated that the frequent re-configuration and re-labelling of Fund facilities, especially in more recent years, is, amongst others, an attempt to make them more attractive and to increase demand for IMF programs (conversations in Washington, D.C., between November 2016 and November 2017).

For the exclusion restriction to be violated, unobserved variables would have to be correlated with the IMF's liquidity and affect creditworthiness differently in countries with different IMF participation histories. This complex relationship is highly unlikely because, for one thing, the main sources of variation in the IMF's liquidity ratio are very distant to events in individual country-years. For another thing, even if unobserved country-specific determinants of creditworthiness were correlated with the overall IMF liquidity ratio, then their effect on country-year specific creditworthiness would, in addition, have to vary depending on the country-year specific level of *IMFprobability*.

Skeptical readers may worry that the IMF's liquidity correlates with global economic trends that, in turn, could affect creditworthiness differently depending on a country's history of participation in IMF programs. We offer both theoretical arguments and empirical evidence to counter such worries. Theoretically, a systematic correlation between the IMF's liquidity and such global economic trends is unlikely because the main source of variation in the IMF's liquidity is an institutional rule in the IMF's Articles of Agreement that requires the Fund to review the quota subscriptions of its members every five years. Following quota reviews that propose increases, members commit more resources leading to an increase in the Fund's liquid resources. The timing of the subsequent spikes in liquidity, which emerge several years after the review, is thus plausibly exogenous to creditworthiness dynamics in individual countries. The second source of variation in the liquidity ratio is repayments of extraordinarily large, individual loans. While the vast majority of individual repayments are too small to significantly affect the IMF's overall liquidity, some repayments of large loans of economically large countries can make a noticeable difference. However, the timing of these transactions is agreed upon years in advance. We consider it implausible that the predetermined repayment schedule the Fund develops with a small number of economically large program countries is associated with individual future creditworthiness dynamics in other countries. Yet, even if such a relationship existed and the IMF's liquidity ratio was, thus, indeed correlated with global economic developments that affect country-year-specific assessments of creditworthiness, then the identifying assumption would only be violated if these global developments affected creditworthiness assessments differently in countries with different IMF participation histories. In our view, such a heterogeneity is highly unlikely.¹²⁸

¹²⁸ For the sake of completeness, note that there are two additional minor sources of variation in the IMF's liquidity ratio. Liquid resources can also vary if the Fund changes the basket of currencies that it considers "usable." The

Figure 3.4 – The IMF’s Liquidity Ratio



Empirically, we investigate correlations between the IMF’s liquidity (plotted in Figure 3.4) and global economic trends. As expected there is no substantial correlation between the liquidity ratio and global GDP growth ($r = -.14$) or the number of systemic banking crises ($r = -.03$). To rule out that these global trends do not interact with a country’s IMF participation history in a way that threatens the exclusion restriction along the lines discussed above, we interact them with *IMFprobability* and add these interactions as control variables in robustness tests. In addition, we examine whether there are particular credit rating trends in sets of countries with different levels of *IMFprobability* that might be correlated with time trends in *IMFliquidity*. Christian and Barrett (2017) show that the findings by Nunn and Qian (2014) could be driven by a spurious correlation between the time-varying constituent term of their interacted IV and a simple time trend in their outcome variable for a set of countries with a specific level of their probability measure. This is why we plot credit rating averages for countries with different levels of *IMFprobability* over time in Appendix 3.F (Figure 3.7). We find no evidence for any non-parallel trends that could threaten our exclusion restriction.¹²⁹ In additional robustness

liquid liabilities’ second source of variation is the Fund’s borrowing from its members. While total borrowing by the Fund is zero in many years, its average share of the liquid liabilities is approximately 15 percent.

¹²⁹ As Christian and Barrett (2017) show, a problem in Nunn and Qian (2014) arises from the fact that the time series of the time-varying constituent term of their interacted IV is remarkably similar to a simple inverse-U shaped trend and does not vary strongly from one period to the next. As *IMFliquidity* exhibits no obvious similarity to any such simple trend and is subject to several idiosyncratic shocks, it is less likely to be correlated with a similar trend in the outcome variable. In addition, to increase the confidence that our first stage does not pick up an artefact, we run placebo regressions in which we randomize the time-varying constituent term of our instrument, *IMFliquidity*. We run 1000 iterations of such regressions based on a randomized order of the actual values of *IMFliquidity*, and find that the resulting first-stage IV coefficients are normally distributed around zero, and that the coefficient’s *t*-statistics are all smaller than in the first-stage regression based on the actual *IMFliquidity*. This increases our confidence in the mechanism driving the first-stage and suggests that it is unlikely that the IV picks up an artefact.

tests, which are reported below, we also use alternative definitions of all key variables and exclude observations that could threaten the exclusion restriction. Our finding that none of these tests changes the econometric results increases our confidence in the identification strategy.

Equipped with this IV we run 2SLS panel regressions over an unbalanced sample of 100 countries in the 1988–2013 period. This gives us the first-stage equation:

$$IMFprogram_{i,t} = \alpha_1 IV_{it} + \alpha_2 IMFprobability_{i,t} + \delta_i + \tau_t + u_{i,t} \quad (5)$$

along with the corresponding second-stage equation:

$$Rating_{i,t*} = \beta_1 IMF\widehat{program}_{i,t} + \beta_2 IMFprobability_{i,t} + \delta_i + \tau_t + \varepsilon_{i,t} \quad (6)$$

Note again that as we assume the excludability of the interaction term conditional on its two levels and, thus, control for *IMFprobability* in both stages while year fixed effects absorb the level effect of *IMFliquidity*. This assumption yields

$$E(IMF\widehat{program}_{i,t} * \varepsilon_{i,t} | IMFprobability_{i,t}, \tau_t) = 0 \quad (7).$$

Under the assumption that the exclusion restriction holds, the exogenous variation in *IMF\widehat{program}* induced by the instrument ensures that selection on unobserved variables ($\varepsilon_{i,t}$) does not bias the coefficient of interest.

3.4 Results

3.4.1 Baseline

We begin by looking at the simple correlation between the treatment variable, *IMFprogram*, and the S&P rating as the outcome. Column 1 in Table 3.1 shows that the correlation is negative with a large coefficient of -6.256. The subsequent specifications support the conjecture of a large downward bias in the coefficient when not accounting for endogenous selection into IMF programs. Conditioning on country fixed effects in column 2, plausibly eliminating an important part of this bias, drastically decreases the point estimate in absolute terms to -1.422. This shows that the unconditional correlation in column 1 largely picks up time-invariant differences between countries. Unobserved global time trends that affect both credit ratings and countries' likelihood of receiving an IMF loan, in contrast, play no major role for endogeneity bias in this setting. We net these out by additionally including year fixed effects in column 3 and find that the coefficient of interest changes only marginally.

Before adding control variables, we restrain the sample to those countries for which all control variables are available in column 4. This barely affects the coefficient, showing that sample selection depending on the availability of control variables is not a concern. The fifth specification then adds a comprehensive set of country-year specific economic and political controls, all lagged by one year (following Fuchs and Gehring, see Appendix 3.D for details). The aim is to condition on the initial state in which countries enter into a year under an IMF program and further reduce the (negative) selection bias. Consistent with this expectation, the coefficient of interest decreases in absolute terms, but remains negative at -0.347 and statistically significant at the five percent level. Thus, when approaching the selection problem via conditioning on observables, we would still conclude that IMF programs have an economically small, yet statistically significant, negative effect on creditworthiness. Nevertheless, as we argue above, this coefficient is likely biased because of this approach's inability to address selection on unobservables like dynamics during the year and because of the potential bad control problem it creates.

The next step implements our instrumental variable approach. The first stage, reported in the bottom panel of the table, shows that the interaction term is negative and statistically significant at the one percent level. This indicates that in high liquidity years, the potential program country's participation history in IMF programs is indeed a less important predictor of receiving a program. The IV passes the underidentification test with a p-value of less than 0.001, and the Kleibergen-Paap (K-P) F-statistic testing for weak identification is about 35 and thus well above the rule of thumb of 10, as well as above the more conservative threshold of 16.66 proposed by Stock and Yogo (2005).

The second stage of this regression shows that the coefficient of interest now turns positive with values of 0.404 and 0.197, depending on whether we choose the whole sample or the one for which all controls are available. Both coefficients are not significant at conventional levels with p-values of 0.648 and 0.819. In sum, the point estimates moved exactly as one would expect in the presence of negative selection bias, which was only partly captured by fixed effects and conditioning on observables. Overall, we thus find no evidence for a negative IMF program effect on a country's creditworthiness; if anything, the point estimates indicate a small but positive relationship. To make sure that our strategy is valid and to shed light on the underlying channels, we corroborate this result with alternative specifications and methods in the subsequent sections.

Table 3.1 – Baseline

Estimation Method	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)	IV (6)	IV (7)
IMF program	-6.256 [0.525] {0.000}	-1.422 [0.282] {0.000}	-1.311 [0.300] {0.000}	-1.235 [0.311] {0.000}	-0.347 [0.168] {0.039}	0.404 [0.885] {0.648}	0.197 [0.860] {0.819}
Country FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	No	Yes	No	No
Observations	2047	2047	2047	1347	1347	2045	1347
Adjusted R-squared	0.206	0.082	0.116	0.137	0.456		

First Stage Results

IMF liquidity × IMF probability	-0.458 [0.076] {0.000}	-0.505 [0.068] {0.000}
IMF probability	3.721 [0.581] {0.000}	4.446 [0.509] {0.000}
Kleibergen-Paap (K-P) underidentification LM-statistic	16.091	20.447
K-P underidentification p-value	0.000	0.000
K-P weak identification F-statistic	35.923	54.720

Notes: The dependent variable is the country's long-term foreign-currency rating by Standard and Poor's. Standard errors clustered at the country level are displayed in brackets, p-values in curly brackets. Appendix D provides a comprehensive list of all economic and political controls added in column 5.

3.4.2 Other Rating Agencies and Institutional Investor Assessments

As a first step, we use alternative outcome variables. Analysts at S&P might have a particular view on the effect of IMF programs that is not generally shared by other analysts and investors. This is why in this section we substitute the S&P ratings with ratings from other agencies and with assessments from professional investors. First, we take the ratings by Moody's and Fitch, the other two major agencies of the "Big Three." Although the credit rating of these three agencies are highly correlated, there are some differences (especially in times of crisis) and we want to be sure that these are not driving our results. Second, as cultural proximity of analysts to rated countries has been shown to affect country ratings (Fuchs and Gehring 2017), we also look at non-US rating agencies as their analysts come from a different cultural background. Analysts at the US-based "Big Three" and at the IMF often have similar educational or professional backgrounds (or have worked for the respective other institution), and might thus share a common ideological mindset that need not represent general investor sentiment. Third, even though we argue that sovereign credit ratings are the most useful measure of a country's creditworthiness for our research question, it would be reassuring if the results hold for alternative measures. While ratings are hard-wired into investment decisions, regulations and company charters, they are officially marketed as mere opinions. Investors with "skin in the game" could deviate from them in cases where they are not bound by regulation. As credit agencies have been blamed for being either too harsh or too reluctant to change ratings in crisis periods, investors might come to different assessments. This is why we digitize and use assessments collected by *Institutional Investor* as an alternative measure of creditworthiness. These are based on surveys among investors and analysts at banks, money management and securities companies, and should also not be affected by bond supply side shocks. We managed to collect data from 1987 onwards covering up to 181 countries (see Appendix 3.C). Table 3.4 presents the regressions that use these four different measures as outcome variables. Panel A of the table replicates the OLS specification with controls (as in column 5 of Table 3.3) and Panel B replicates the baseline IV regression (as in column 6 of Table 3.3). In all four OLS specifications, the coefficients of interest are negative and statistically significant. The point estimates for Fitch is comparable to S&P, the ones for Moody's and the non-US agencies are a bit more negative. For interpreting the regression results of Institutional Investor assessments

note that these ratings range from 0 to 100 rather than from 0 to 21. When adjusting for these different scales, the magnitude of the coefficient in this regression (-2.57) is thus similar to the coefficients for rating agencies and lies between the results for S&P and Moody's. As before, the conditional correlation between IMF programs and measures of sovereign creditworthiness is significantly negative. When turning to the IV approach, the estimates again turn positive and statistically insignificant. This result emerges in all four specifications. Hence, irrespective of how we measure sovereign creditworthiness, we observe the same pattern as before: The negative association between IMF programs and creditworthiness disappears when applying an approach that is able to control for selection on unobservables. Contrary to widespread views in parts of the existing correlational literature and especially in policy circles, we find no evidence for a negative effect on creditworthiness or a "stigma" associated with IMF programs.

Table 3.2 – Other Assessments of Creditworthiness

Dependent variable is rating/assessment by:	Moody's	Fitch	Non-US Agencies	Institutional Investor
Panel A: OLS regressions with controls				
	(1)	(2)	(3)	(4)
IMF program	-0.849 [0.227] {0.000}	-0.348 [0.217] {0.109}	-0.810 [0.302] {0.007}	-2.570 [0.599] {0.000}
Panel B: IV regressions without controls				
	(5)	(6)	(7)	(8)
IMF program	1.539 [1.357] {0.257}	0.494 [1.256] {0.694}	0.320 [0.908] {0.724}	0.063 [4.182] {0.988}
<i>First stage diagnostics:</i>				
K-P underid. LM	12.588	14.771	14.032	24.923
K-P underid. p	0.000	0.000	0.000	0.000
K-P weak id. F	31.067	26.761	38.026	40.899
Observations	1210	1127	855	1912

Notes: Standard errors clustered at the country level are displayed in brackets, p-values in curly brackets. Appendix 3.D lists a comprehensive list of all economic and political controls.

3.4.3 Robustness

Before we turn to alternative empirical methods and to examining channels, we further examine the robustness of this baseline result. First, we want to address potential concerns regarding any of the two constituent terms forming our interaction instrument and aim to enhance the plausibility of the exclusion restriction. Second, we test whether our findings are driven by certain time periods or particular countries. Third, we apply an alternative definition of our treatment variable. In addition to these robustness exercises focusing on the baseline analysis, we also apply an alternative identification strategy that exploits variation at the monthly level. We report the results of these analyses in Appendix 3.F and describe them below.

With regard to the first component of the instrumental variable, *IMFprobability*, we take as an alternative a time-invariant, country-specific measure instead of the cumulative, time-variant probability. This makes *IMFprobability* multicollinear with the year fixed effects. Taking all observations in the sample period into account considers observations from periods $t+1$, $t+2$, ... to compute the probability in t , and thus uses information from the future to explain the present. Although we regard this as conceptually problematic, column 1 in Table 3.4 shows that the estimates are virtually unchanged by this modification. The interaction term in the first stage is of almost the exact same size, showing that the relationship we exploit for identification does not depend on how a country's probability of participating in IMF programs is defined. The significance of the IV, the K-P F-statistic, and the second stage point estimate are also very similar, as compared to the baseline.

Regarding the second component of the instrument, some readers might, as discussed above, question the exogeneity of the IMF's liquidity ratio. Even though individual countries in general are unable to significantly influence the IMF's liquidity, a few countries in the sample repaid extraordinarily large tranches of extraordinarily large IMF loans in some years. While the repayment schedule of Fund resources is usually developed years in advance, we still want to exclude the possibility that such events could lead to a correlation between the liquidity and country-year specific economic fundamentals unrelated to the presence of an IMF program. While this would only threaten the exclusion restriction if this relationship depended on the country's level of *IMFprobability* we still want to be cautious and exclude the country-year observations that could significantly influence the IMF's liquidity. Column 2 excludes the top

five percent of country-year observations with the largest “repurchases” of IMF loans and column 3 excludes all observations from countries in which such relatively large repurchases have taken place. Neither of these regressions yields substantially different results, indicating that individual repurchases in general – including the extraordinarily large ones – do not threaten the exclusion restriction.¹³⁰

Even if we accept the IMF’s liquidity as being plausibly exogenous, the exclusion restriction would be violated if other global trends correlate with it and also affect countries’ creditworthiness with different past probabilities of receiving a program in a heterogeneous way. We consider such a relationship unlikely in particular because we find no time trends in credit ratings across countries with different levels of IMF probabilities that are correlated with the IMF’s liquidity.¹³¹ There is also no evidence of a correlation between relevant global trends such as global growth rates or the number of crises and the IMF liquidity ratio. To nevertheless examine this potential threat further, we interact global GDP growth and the number of banking crises with the country-specific probability and include these terms as control variables in column 4. The fact that neither the relevance of the IV in the first stage nor the F-statistics are affected, provides support for our approach. The result that the point estimate in the second stage barely changes further indicates that violations of the exclusion restriction are unlikely.

In regards to the second concern that certain countries or periods could drive the results, column 5 omits the years following the global financial crisis (GFC) and column 6 excludes all countries that were members of the Eurozone in year t . Arguably, the IMF programs that were implemented in Eurozone countries in the aftermath of the GFC were atypical. First, the IMF designed them jointly with European Union (EU) institutions. Second, default risks in Eurozone countries are potentially assessed differently than in other countries because signals from EU institutions and other EU member states will be taken into account. As columns 5 and 6 show, these restrictions to our sample do not significantly affect our results. The coefficient on *IMFprogram* is again positive and statistically insignificant.

Next, we redefine our treatment variable and use the binary variable *IMFapproval* as an alternative in column 7. This variable indicates only the year in which an agreement with the

¹³⁰ Using only the amount of liquid resources (and thus only the numerator of the liquidity ratio) as the second component of the IV instrument also yields a very similar result.

¹³¹ See: the above discussion in section 3.3.4; Figure 3.7 in Appendix 3.F; Christian and Barrett (2017).

IMF was reached and is set to zero for all other years, including the years during which an IMF program was still in place. Again, we observe a significantly negative OLS coefficient (not shown), which turns positive and insignificant when accounting for endogeneity via our IV strategy (column 7). This is important in two ways. First, it is reassuring that our instrumental variable approach also works for the approval of programs. Second, this allows us to compare the IV-based country-year level results more directly with the following results. These are based on an alternative dataset and an alternative identification strategy.

3.4.4 An Alternative Identification Strategy

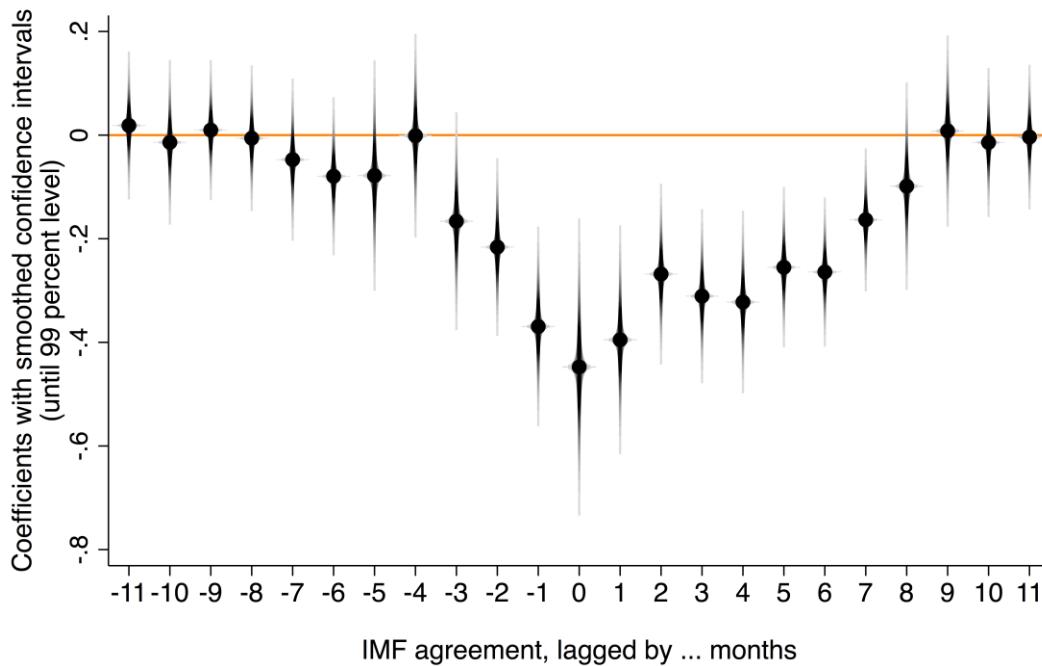
So far, we have used data at the yearly level, because information on the IMF's liquidity ratio, which was needed for the construction of the IV, is only available on an annual basis. For both credit ratings and IMF programs, however, we were able to collect data at a finer-grained level. The variation of these higher-frequency data allows us to apply an alternative method, which – like our IV approach – is able to control for selection on unobservables. The identifying assumption it is based on, however, is completely different.

In these regressions, which are based on data at the monthly level, our dependent variable is the S&P rating at the end of month m . The treatment variable is the country-month specific variable $IMFagreement_{i,m}$, indicating the month in which an IMF program was agreed upon. We employ an event-time specification and add $IMFagreement_{i,m}$, as well as 11 lags (indicated by l) and leads (indicated by $-l$) of it. The monthly frequency of these data allows us to include month fixed effects (μ_m) and, crucially, country-times-year fixed effects (denoted as $\theta_{i,t}$). This means that in the most conservative specification based on these data we estimate:

$$Rating_{i,m} = \sum_{r=-11}^{11} \beta_r IMFagreement_{i,m-r} + \theta_{i,t} + \mu_m + \varepsilon_{i,m} \quad (8)$$

Conceptually, the demeaning implemented by this fixed effect estimator is equivalent to controlling for all potentially biasing factors at the country-year level. The only variation that remains are the dynamics within country-years and we thus compare ratings in different months within the same country-year observation. Accordingly, the coefficients β_r estimate the extent to which the rating in the months around the start of an IMF program deviated from the mean rating of country i in year t .

Figure 3.5 – Event-based Identification



Note: The figure plots the coefficients of different lags and leads of *IMF agreement* in a regression of monthly S&P ratings. See regression equation 8.

Figure 3.5 illustrates the results of this regression by plotting all coefficients β_r and smoothed confidence intervals.¹³² We discuss these coefficients in ‘chronological’ order. First, in the period between eleven months and four months before the start of an IMF program no significant deviation from the country-year mean rating is visible. Then, starting three months before the agreement, a negative pre-trend is visible. It is interesting that this negative pre-trend is statistically significant even though all variation between treated and non-treated country-years is eliminated: Countries whose creditworthiness collapses during the year are thus more likely to receive an IMF program within the next three months. This is consistent with our conjecture that, even if all country-year specific characteristics are held constant, the mean trajectory of treated countries still exhibits certain dynamics around the treatment start that make the treatment more likely but are not caused by the treatment itself.

¹³² The regression output is reported in Table 3.8 in Appendix 3.G. In this table, before we turn to the results of the regression specified in equation 8, we first run the regression with alternative, less conservative sets of fixed effects. These results show that the point estimates become more positive, the more biasing variation we reduce by adding more fixed effects. Note that the idea behind plotting smoothed confidence intervals is to visualize statistical uncertainty without setting arbitrary thresholds for p -values. (This is analogous to our choice in the main tables to not report statistical significance by means of ‘significance stars’ for certain p -value thresholds, but by directly reporting p -values.)

Nevertheless, the regression yields an interesting pattern at the time of the treatment start. The most negative point estimate for β_l appears in the month in which the IMF agreement is approved (i.e., for $l = 0$). As soon as the program starts, the point estimates increase (for $l > 0$) and increasingly move to zero. The trend of decreasing credit ratings goes into reverse exactly when the IMF program starts.¹³³ Eight months after program approval ($l = 8$) the negative deviation from the mean rating of the year is no longer statistically significant.

What does this pattern tell us? First, the negative pre-trend shows that the timing of IMF agreements within a given country-year observation is not random; IMF programs typically start at a low point. Second, while this negative pre-trend negatively biases the absolute values of the coefficient estimates for the variables indicating the (post-)treatment period, they start increasing in relative terms as soon as the IMF program starts. Even though IMF programs are thus endogenously timed in spite of all country-year-specific variation being netted out, it would be an improbable coincidence to observe that the trend reversal happens exactly after the treatment starts if the treatment itself did not have any impact. As in the IV regressions, the evidence is thus not consistent with a “stigma” effect. On the contrary, these specifications suggest that IMF programs succeed in “cushioning” against further deteriorations in sovereign creditworthiness. As changes in economic and political fundamentals (*adjustment* effects) do not take effect within a month, the trend reversal can be best thought of as a *signaling* effect attributed to the IMF intervention. To investigate this further, we look at the differentiation between signals and adjustments in more detail in the subsequent section.

3.4.5 Channels: Adjustment and Signaling

Our theoretical considerations distinguished two ways of how IMF interventions can influence creditworthiness. First, as IMF programs often lead to far-reaching economic reforms they can influence a country’s creditworthiness via the implementation of immediate *adjustments*. In the previous literature discussed above, IMF programs were found to improve certain economic fundamentals but were also repeatedly associated with political instability and lower growth rates. Second, an IMF program is also a *signal* that affects expectations. Independent of its actual economic effects, the mere presence of the IMF conveys information about the country’s future policy path to those assessing its creditworthiness.

¹³³ Note that our background research suggests that agencies take on average one month to update their ratings.

To shed light on these different channels, we investigate the short-term adjustment effects of an IMF program on the most important economic factors determining creditworthiness in our sample.¹³⁴ We focus on GDP per capita, the growth rate of GDP, inflation, the change in government debt and the current account balance, as in the rating literature these are cited as the most important predictors of sovereign credit ratings (Archer, Biglaiser, and DeRouen 2007; Cantor and Packer 1996; Hill, Brooks, and Faff 2010). We can replicate the explanatory power of these variables in our sample. We find significant associations with S&P ratings for all variables except the change in government debt. In a simple OLS regression of S&P ratings the five variables explain 75 percent of the variance. In an analogous fixed-effects regression the within- R^2 equals .31 while the overall- R^2 equals .74. Interestingly, most of the variation is explained by the variables indicating level and growth rate of GDP: These two variables alone explain 71 percent of the variation in an OLS rating regression and 27 percent of the within-country variation in a fixed-effects regression. Having established the important role these variables play for assessments of creditworthiness, we turn to analyzing how IMF programs affect them.

Table 3.6 shows the results based on the baseline IV specification when the outcome variable is substituted with these variables. For inflation, changes in government spending, and the current account balance, the estimates are statistically insignificant. There is, however, a negative effect on growth rates of GDP in the short-run. In this sample IMF programs induce growth rates that are about four percentage points lower compared to the counterfactual.¹³⁵ In the average IMF program country (where growth rates fluctuate more than in countries that never receive IMF programs) this is equivalent to about one standard deviation. It is thus a large, albeit not unrealistically large effect considering previous results in the literature on this relationship (Barro and Lee 2005, Dreher 2006) and considering also the substantial extent of budget cuts, tax increases and other measures with potentially short-run contractionary consequences that IMF programs typically entail. Many program countries feature a large

¹³⁴ Thoroughly studying the long-term effects of IMF programs on other outcomes is beyond the scope of this paper and deserves more attention in future research. Arguably more complex models might be needed to fully understand these dynamics, but we want to highlight the causal direction in the short-run and some important channels. Also note as another caveat that these outcomes are not independent of each other.

¹³⁵ Note that IMF programs last for multiple years, and thus most of the country-year observations with an active program are years in which IMF programs were already active in the year(s) before. The estimates, thus, also includes lagged effects of previous program years.

public sector, whose size IMF conditions often reduce (Rickard and Caraway 2018). IMF staff recently argued that the IMF underestimated the size of the fiscal multiplier in past crises and thus projected smaller negative effects of fiscal austerity on GDP than those that eventually materialized (Blanchard and Leigh 2013). Many program countries also rely on debt-financed growth in the years before they start IMF programs, and cannot maintain such growth under a program as the IMF often sets limits on new debt (Kentikelenis, Stubbs, and King 2016).

Table 3.3 – Channels

Dependent Variable:	GDP Growth	Inflation	Change in Public Debt	Current Account Balance
	(1)	(2)	(3)	(4)
IMF program	-4.187	0.067	1.776	4.187
	[1.292]	[0.044]	[2.265]	[3.432]
	{0.001}	{0.129}	{0.433}	{0.223}
Observations	2032	1796	1840	1808
K-P underid. LM	16.098	15.983	16.397	16.637
K-P underid. p	0.000	0.000	0.000	0.000
K-P weak id. F	35.898	35.383	37.598	38.032

Note: Results are based on the baseline IV regression (Table 3.3, column 6), but with other dependent variables. Standard errors clustered at the country level in brackets, p-values in curly brackets.

To sum up, the economic adjustments that crisis countries under IMF programs typically implement lead to lower growth and no significant improvements in the other major predictors of creditworthiness. As credit ratings are a direct function of changes in GDP (as is stated in the official manuals of all agencies), rating agencies would normally respond to such sharp growth reductions by lowering their assessments of creditworthiness.¹³⁶ Against this backdrop, it is remarkable that the coefficient on IMF programs in a regression of creditworthiness is not negative. This suggests that IMF programs not only cause negative economic *adjustments* that would usually lead to declining creditworthiness, but also convey a positive *signaling* effect that prevents this decline leading to a net neutral effect. The IMF's

¹³⁶ According to the manual published by Standard & Poor's a credit rating can be best understood as a scoring model. There is an economic and a political dimension, which are each composed of different factors. For each factor the country gets assigned a grade, and the factors are summed up to a grade for the given dimension.

presence in a program country when its growth declines appears to have a reassuring effect. In line with the view that IMF programs signal the quality of a country's policy reforms to investors, our results suggest that this "seal of approval" helps program countries to maintain their level of creditworthiness even if economic adjustments cause substantial growth reductions.

It is econometrically difficult to estimate the exact size and significance of this positive signaling effect. Ideally, we would want to estimate the effect of IMF programs while holding GDP growth (and other economic fundamentals) constant. So far, we have not controlled for growth because of its role as a "bad control." The results in Table 3.3 support this choice: IMF programs reduce GDP growth and growth itself is a strong predictor of ratings. GDP growth is thus an intermediate outcome and adding it is likely to bias the estimated coefficient for the variable indicating the presence of an IMF program. Such a regression amounts to a comparison between treated and untreated observations for which growth is fixed. As the treatment itself changes growth, the compared observations will necessarily differ in terms of potential outcomes and we can thus not be sure that this regression estimates a causal effect of IMF programs.¹³⁷

With this caveat in mind, we turn to Table 3.4. In column 1 we report the results of the baseline IV regression when *GDP per capita* and *GDP contraction*¹³⁸ are added as control variables. In column 2 we add the full set of control variables described above. In both specifications, the coefficients on *IMFprogram* indeed turn more positive with point estimates of 1.675 and 1.987 and become statistically significant with p-values of 0.052 and 0.009 respectively.¹³⁹ While for the reasons outlined we cannot claim that these findings indicate causality, they are consistent with interpreting the previous results as an indication for a positive signaling effect of IMF programs: When GDP levels and growth rates are held constant, the instrumented IMF program indicator is positively and significantly associated with creditworthiness.

¹³⁷ See the identification section, Gelman and Hill (2007) as well as Angrist and Pischke (2008) for discussions of the bad control problem.

¹³⁸ We use the variable *GDP contraction*, defined as the additive inverse of *GDP growth*, to simplify the interpretation of results in this table.

¹³⁹ It is interesting that the large set of other control variables, which could also be channels, affect the coefficient only marginally once the GDP variables are included. This supports the view that the changes in GDP are the decisive underlying channel.

Table 3.4 – Adjustment vs. Signaling

	(1)	(2)	(3)	(4)	(5)	(6)
IMF program	1.675 [0.862] {0.052}	1.987 [0.764] {0.009}				
GDP contraction	-0.115 [0.028] {0.000}	-0.118 [0.026] {0.000}	-0.090 [0.020] {0.000}	-0.095 [0.021] {0.000}		
GDP contraction, IMF induced					0.070 [0.211] {0.740}	0.012 [0.148] {0.934}
GDP contraction, residual of IMF induced						-0.121 [0.022] {0.000}
GDP per capita (ln)	6.790 [1.225] {0.000}	4.154 [0.886] {0.000}	6.274 [1.001] {0.000}	4.386 [0.788] {0.000}		
Country and Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Additional Controls	No	Yes	No	Yes	No	No
Observations	2016	1362	2132	1365	2032	2034
Adjusted R-squared			0.264	0.495		
K-P underid. LM	13.314	13.199			7.775	7.775
K-P underid. p	0.000	0.000			0.005	0.005
K-P weak id. F	26.698	24.368			11.888	11.888
Estimation by	IV	IV	OLS	OLS	IV	IV

Notes: The dependent variable is the country's long-term foreign-currency rating by S&P. In columns 1-2, IMF program is instrumented by *IV*; in column 5-6, GDP contraction is instrumented by *IV*. Standard errors clustered at the country level are displayed in brackets, p-values in curly brackets. Appendix 3.D lists all economic and political controls used for columns 2 and 4.

To increase the plausibility of this interpretation we offer another alternative test in columns 3-6. In column 3 we first replicate that GDP levels and growth rates are significantly and positively associated with ratings. This does not change if additional controls are added in column 4. For the next step, we exploit our above findings. As in the IV regressions both the first stage effect and the effect of IMF programs on growth are highly significant, we use the *IV* as an instrument for GDP contractions. We thus have a source of variation in GDP growth that is due to the presence of an IMF program and can test how such IMF-induced variation in growth affects ratings. Finding no effect would support our interpretation. We report the result of this regression in column 5 and find that the first stage diagnostics surpass critical

values and that there is, indeed, no statistically significant effect on ratings in the second stage. In column 6 we additionally include the first stage residuals; they capture the variation in *GDP contractions* that is not explained by the IV and thus *not* induced by the exogenous variation in IMF programs we exploit for identification. We find that these residuals are associated to significantly lower ratings; the part of the variation in GDP contraction that is induced by variation in the IV is not. To be sure, this also does not prove the existence of a positive signaling effect, but it shows the following: While GDP dynamics directly map onto ratings in normal times, the GDP contractions induced by IMF programs have no such effect on ratings. In our view, the most plausible explanation for this finding is that IMF programs come with a positive signaling effect.

In sum, several pieces of econometric evidence suggest that the overall null effect of IMF programs on creditworthiness is a combination of a negative adjustment effect and a positive signaling effect. First, IMF programs reduce contemporary growth rates, which are usually strong predictors of ratings. As the net effect of IMF programs on ratings is slightly positive and statistically indistinguishable from zero, IMF programs must have an additional, positive effect on the creditworthiness assessment of investors and rating agencies. Second, the finding that IMF-induced changes in growth rates leave ratings unaffected is consistent with this interpretation. As long as growth reductions occur under an IMF program, rating agencies do not consider them as a reason for concern. Third, the positive signaling effect becomes visible in a specification in which GDP is held constant. While a “bad control” bias cannot be ruled out in this specification, the positive and statistically significant coefficient it yields is in line with a positive signaling effect. Finally, in the event-time specification based on monthly data and country-times-year fixed effects a clear trend reversal of credit ratings is visible at the exact time of the program start. This provides further support for the idea that assessments of creditworthiness respond positively to signals sent when countries enter IMF programs. These signals ‘cushion’ the negative adjustment effects.

Even though all of these pieces of evidence provide empirical support for a positive signaling effect, none of these estimates is a perfect test of this hypothesis. Because we acknowledge the limits of econometric estimates in this context, we test whether qualitative evidence can shed some more light on the signaling mechanism.

3.4.6 Qualitative Evidence

The econometric results of the previous section suggest that IMF programs send a positive signal to the agencies assessing a country's creditworthiness. To see whether this is reflected in the agencies' reasoning, we evaluate rating statements that are issued when a rating or its outlook are changed. We use the *Dow Jones Factiva* database for this qualitative assessment.

Initially, we study these statements in an exploratory way. (See Appendix 3.H for details and a list of exemplary statements.) It becomes evident that rating agencies indeed often link the IMF's presence to *positive expectations*. We find many statements in which they state that they expect that the IMF program will help the country in the near future. Examples include statements like: "[w]e think the new IMF program [...] will help in addressing fiscal and external imbalances" (S&P on Ghana in 2015), or "the International Monetary Fund program will serve as a policy anchor for fiscal consolidation" (S&P on Albania in 2014).

Furthermore, we noticed that several of these statements emphasize the IMF's role in helping countries to overcome short-term liquidity problems; others emphasize the increased likelihood of successful reform implementation. For example, with regard to Sri Lanka, Moody's stated in 2016: "the IMF program will alleviate Sri Lanka's external liquidity pressures." For Egypt the same agency in the same year expected the IMF program to "support the implementation of fiscal and economic reforms." We find many more such examples in which the liquidity aspect and/or the reform aspect of IMF programs is emphasized.

Based on this initial inspection, we then conduct a more systematic analysis. We extract all available articles on *Factiva* using all combinations of the search terms "IMF/International Monetary Fund," "rating," "program," "reform," in English or German and in the industry category "Rating Agency." We then use a *Python* script to extract the paragraphs before and after statements mention the IMF. This approach yields 117 statements. Two research assistants then coded these statements following a pre-defined codebook (see Appendix 3.I for details). The aim of this coding was, first, to distinguish negative, neutral/mixed and positive associations between IMF programs and changes in credit ratings. Second, it also aimed to differentiate between texts mentioning the liquidity aspect of IMF programs, the reform aspect, or both. The codebook was designed to be conservative in the sense of biasing against findings in support of our priors resulting from the econometric analysis: in case of doubt about the association between IMF programs and the rating, the statement was categorized as

“no clear association with rating.” If it was not obvious whether the statement relates to liquidity or reform aspects of IMF programs, it was put in a residual category.

Figure 3.6 – Text Analysis of Rating Statements

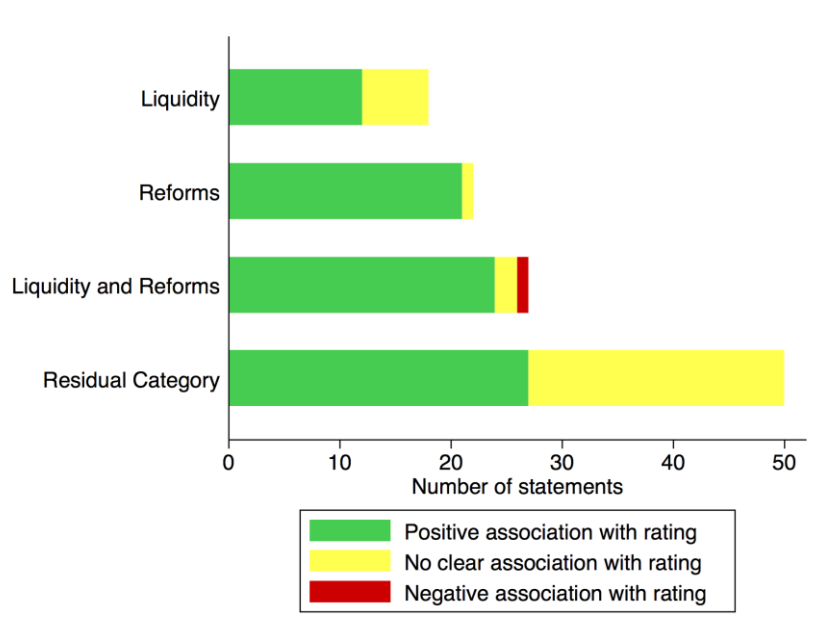


Figure 3.6 graphically illustrates the results of this exercise. The first and most noticeable finding is that the vast majority of statements attribute a positive effect to IMF programs. Of 117 statements, only one statement notes that the IMF’s presence negatively influences the rating. For 32 statements it was not possible to conclusively determine whether the agency considered the IMF program to have a positive or a negative effect on the rating. The second finding is that rating agencies often link the reforms expected to materialize under IMF programs to positive trends in credit ratings. There thus seems to be more to IMF programs for credit agencies than just the temporary increase in liquidity they come with. It is also apparent that the statements about “reforms” or “reforms and liquidity” are with very few exceptions positive. Statements concerning liquidity are slightly more mixed, and the residual category, quite naturally, captures a number of neutral statements in which no clear association could be noted. The expectation of successful reforms thus appears to be a significant part of the IMF’s positive signaling effect on creditworthiness assessments.

Overall, the text analysis is in line with the results of the econometric analysis and exemplary statements like the following illustrate the effect we find: “We view the risk of another default in the next two to three years as diminished due to the Ukrainian authorities’ commitment to the reforms set out in the International Monetary Fund (IMF) program.” S&P made this

statement in October 2015 during a period of substantial GDP contraction under multiple consecutive IMF programs in Ukraine: The country's growth rate stood at -6.6 percent in 2014 and at -9.8 percent in 2015. Nevertheless, S&P improved its credit rating because it expected reforms under IMF programs to enhance sovereign creditworthiness.

Our results in their entirety suggest that this piece of anecdotal evidence is accurately representative of a general pattern: IMF programs, rather than coming with a *stigma*, arouse expectations of successful reform implementation and thereby send a positive signal that, despite substantial economic contractions under IMF programs, *cushions* assessments of sovereign creditworthiness.

3.5 Conclusion

As the international lender of last resort, the IMF's main objective is to help countries resolve their balance-of-payments problems. Its loan programs need to restore the creditworthiness of countries with severely limited access to external financial resources. In light of the IMF's resurgence as the most important multilateral actor in the global financial system, this study re-investigates the IMF's effectiveness in achieving this key goal with new data and new identification strategies. Our results highlight the econometric difficulties associated to examining the IMF's effectiveness, and suggest that these difficulties are likely reasons for mixed results in much of the previous literature. Not accounting for the omitted variable bias due to endogenous selection into IMF programs of countries with already deteriorating economic and political conditions overstates negative consequences. This has contributed to a situation where potential program countries fear a *stigma* associated with participation in IMF programs.

Our results provide evidence against this view and paint an alternative, more nuanced picture. The presence of the IMF does not negatively affect the creditworthiness of the program country. While adjustments under IMF programs are often contractionary, the programs' positive signal prevents creditworthiness assessments from deteriorating. IMF programs, thus, provide a *cushion* that allows program countries a transition period in which they can implement potentially contractionary reforms without having to fear further rating downgrades.

While the results of our various empirical analyses are clear-cut as far as this specific relationship is concerned, they remain silent about the long-term benefits of reforms under IMF programs. First, the successful implementation of reforms that provide a sustainable solution to the country's underlying problems comes with many obstacles along the way. The fact that the IMF's engagement sends a positive signal to financial market participants can provide countries with important time and maneuvering room to overcome crises. This, however, is only a precondition, not a guarantee for success.

Second, our results cannot determine the extent to which the adjustments under IMF programs we observe are necessary and adequate for such long-term success. Not all IMF programs are identical in terms of the policy conditions they come with (Stone 2008), and some adjustment policies might be more effective, and less painful, than others. We consider it an important task for future research to shed more light on this by differentiating harmful conditions from beneficial ones. While the endogeneity of conditions would present yet another methodological challenge, such work could help improve the design of future IMF programs.

3.6 Appendices to Chapter 3

3.6.1 Appendix 3.A: Construction of the Sovereign Ratings Database

The description of how the sovereign rating database was constructed is in most parts identical to the part in the online appendix of Fuchs and Gehring (2017), but reprinted here for the reader's convenience. Fuchs and Gehring also provide more details about the ratings and the individual agencies.

Data on sovereign ratings assigned by CI, Dagong, DBRS, JCR, Moody's, R&I, and S&P have been obtained from Bloomberg. Hence, everybody with access to Bloomberg can replicate the dataset easily. We downloaded the data in late September 2012 in the Princeton University Library and updated all information on June 28, 2013.¹⁴⁰

The approach was the following:

- (1) To access the data, we logged on to a Bloomberg terminal and typed "CSDR." The variables selected are *Foreign Long Term* for CI, *FC LT Sovereign Ratings* for Dagong, *Foreign Currency LT Debt* for DBRS, JCR, Moody's and S&P, and *Foreign Curr Issuer Rtg* for R&I. We followed Bloomberg and collected Moody's foreign currency issuer rating if Moody's had not assigned a foreign-currency debt rating to a country. We took screenshots for each page displaying sovereign ratings.
- (2) Using these screenshots, two student assistants entered the letter-ratings into a database. The double-coding was used to identify and correct typing errors.
- (3) The three-letter ratings were translated to numerical values according to the 21-point scale presented in Appendix 3.B.
- (4) We checked the data for potential errors, for example by examining rating changes by more than two steps. Two obvious mistakes in the R&I data from Bloomberg have been corrected after e-mail correspondence with the agency's chief analyst: (i) India received a "BBB+" rating on 15 June 1998, and a "BBB" rating on 18 November 1998, 20 December 1999, and 30 January 2001, (ii) Ukraine received an "BB-" rating on 18 July 1998, a "B" rating on 28 August 1998, and a "B-" rating on 28 September 1999.

¹⁴⁰ The ratings from Feri and Fitch have been obtained from the companies directly.

3.6.2 Appendix 3.B: Translation of Sovereign Ratings into Numerical Values

Table 3.5 – Translation of Sovereign Ratings into Numerical Values

CI CYP (KWT)	Dagong CHN	DBRS CAN	Feri DEU	Fitch USA (FRA)	Moody's USA	JCR JPN	R&I JPN	S&P USA	Numerical scale
AAA	AAA	AAA	AAA	AAA	Aaa	AAA	AAA	AAA	21
AA+	AA+	AAH	AA+	AA+	Aa1	AA+	AA+	AA+	20
AA	AA	AA	AA	AA	Aa2	AA	AA	AA	19
AA-	AA-	AAL	AA-	AA-	Aa3	AA-	AA-	AA-	18
A+	A+	AH	A+	A+	A1	A+	A+	A+	17
A	A	A	A	A	A2	A	A	A	16
A-	A-	AL	A-	A-	A3	A-	A-	A-	15
BBB+	BBB+	BBBH	BBB+	BBB+	Baa1	BBB+	BBB+	BBB+	14
BBB	BBB	BBB	BBB	BBB	Baa2	BBB	BBB	BBB	13
BBB-	BBB-	BBBL	BBB-	BBB-	Baa3	BBB-	BBB-	BBB-	12
BB+	BB+	BBH	BB+	BB+	Ba1	BB+	BB+	BB+	11
BB	BB	BB	BB	BB	Ba2	BB	BB	BB	10
BB-	BB-	BBL	BB-	BB-	Ba3	BB-	BB-	BB-	9
B+	B+	BH	B+	B+	B1	B+	B+	B+	8
B	B	B	B	B	B2	B	B	B	7
B-	B-	BL	B-	B-	B3	B-	B-	B-	6
CCC+	CCC+	CCCH	CCC+	CCC+	Caa1	CCC+	CCC+	CCC+	5
CCC	CCC	CCC	CCC	CCC	Caa2	CCC	CCC	CCC	4
CCC-	CCC-	CCCL	CCC-	CCC-	Caa3	CCC-	CCC-	CCC-	3
CC	CC	CC	CC	CC	Ca	CC	CC	CC	2
C	C	C		C	C	C		C	1
DDD				DDD		DDD		SD	1
DD				DD		DD			1
D	D	D	D	D		D	D	D	1
				RD		RD			1

Sources: Rating scales from company webpages, except DBRS and Feri. DBRS and Feri scales were obtained from the agencies via personal e-mail communication.

3.6.3 Appendix 3.C: Institutional Investor Data

We use data from Institutional Investor as an alternative assessment of a country's creditworthiness. Regarding the methodology, the company states that "Institutional Investor's Country Credit ratings are based on information provided by senior economists and sovereign-risk analysts at leading global banks and money management and securities firms. The respondents have graded each country on a scale of zero to 100, with 100 representing the least likelihood of default. We weighted participants' responses according to their institutions' global exposure. Names of respondents are kept strictly confidential."

The access to the individual reports is easy for subscribers, or to those with access to a data provider like "EBSCOhost." To access the data, a reader interested in replication or extending this study can go to <http://www.institutionalinvestor.com>, select "Research + Rankings" and then "Country Credit". For most years there exist two reports, one for March and one for September.

Institutional Investor

Investors • Asset Management • Banking + Capital Markets • Research + Rankings • BI

[in](#) [t](#) [f](#) [g+](#) [p](#) [r](#)

Premium Research + Rankings data is accessible by subscribers only. [Click](#)

ALL RANKINGS: GLOBAL RANKINGS

Select one of the following rankings to view its current leaders:

Country Credit - March

Country Credit - September

Global Research Leaders: Year to Date

Top Global Research Firm of the Year

The World's Best Hotels

The World's Largest Custodians

The World's Biggest Sovereign Wealth Funds

Jump to a ranking:

For older ratings, the reports look as follows:

The beat goes on

Shapiro, Harvey D
Institutional Investor; Mar 1996; 30, 3; ABI/INFORM Collection
pg. 135

COUNTRY CREDIT

The beat goes on

Amid broadly positive reviews from international bankers everywhere, a new leader has emerged in *Institutional Investor's* exclusive semiannual rating of sovereign creditworthiness. One banker surveyed suggests that "nostalgia for the deutsche mark, a currency that may soon disappear," helped push Germany's rating up slightly just as Switzerland slipped a bit, easing the Federal Republic into a tie for the No. 1 position. Germany goes to the head of the

Germany moves up as global optimism continues to nudge sovereign creditworthiness higher.

By Harvey D. Shapiro

every financier's favorite part of the world, actually had a small dip in its rating. In fact, there is a good deal of divergence in the views of respondents regarding individual countries, but the consensus is, decidedly, "Onward and upward." The reason is clear, says one British banker: "The economic outlook for the world is reasonably favorable, but the biggest changes are political. A lot of potential theaters of war are being changed by peace processes. This includes southern Africa and the Middle East."

INSTITUTIONAL INVESTOR'S 1996 COUNTRY CREDIT RATINGS

Rank Sept. 1995	Rank March 1996	Country	Institutional Investor credit rating	Six-month change	One-year change	Rank Sept. 1995	Rank March 1996	Country	Institutional Investor credit rating	Six-month change	One-year change
3	1*	Germany	91.5	0.6	0.9	72	70	Vietnam	30.3	0.8	2.7
1		Switzerland	91.5	-0.7	-1.0	65	71	Venezuela	30.1	-1.3	-3.0
2	3	Japan	91.0	-0.6	-0.9	73	72	Swaziland	30.0	0.8	1.5
4	4	United States	90.9	0.2	0.5	70	73	Libya	29.9	-0.1	-0.6
5	5	Netherlands	89.2	-0.1	0.2	68	74	Pakistan	29.5	-1.2	-0.6
6	6	France	88.4	-0.7	-0.6	74	75	Ghana	29.2	0.1	0.0
7	7	United Kingdom	88.2	0.4	0.5	79	76	Estonia	28.9	2.6	3.5
9	8	Luxembourg	85.9	0.4	0.5	76	77	Panama	27.9	1.5	2.2
8	9	Austria	85.7	-0.5	-0.5	78	78	Jamaica	27.7	1.4	2.5
10	10	Singapore	82.8	-1.2	-0.2	80	79	Peru	27.2	1.4	3.5
11	11	Norway	82.0	0.4	1.7	77	80	Kenya	26.9	0.5	2.0
13	12	Denmark	80.3	0.4	1.0	81	81**	Bangladesh	26.5	0.9	1.7
12	13	Canada	79.9	-0.4	-0.3	83	82**	Lebanon	26.5	1.2	1.6
15	14	Belgium	79.5	0.3	0.9	88	83	Seychelles	25.8	1.5	2.1
14	15	Taiwan	78.9	-1.0	-0.8	84	84	Ecuador	25.7	0.6	0.7
18	16	Ireland	74.4	1.0	2.1	82	85	Gabon	25.1	-0.2	-0.7
16	17	Sweden	74.3	0.2	-0.1	89	86	Latvia	24.7	1.3	2.1
17	18	Spain	73.2	-0.5	-0.5	87	87	Syria	24.6	0.0	-0.3
21	19	Finland	72.2	0.8	2.2	84	88	Nepal	23.9	-1.2	-0.5
19	20**	Italy	72.0	-0.3	-0.4	90	89	Lithuania	23.8	0.9	2.1
20	21**	South Korea	72.0	-0.2	0.6	93	90	Bolivia	23.7	1.3	1.2
22	22	Australia	71.0	-0.2	0.1	86	91	Iran	23.6	-1.2	-1.9
23	23	New Zealand	70.3	0.9	2.1	94	92	Bulgaria	23.1	0.9	1.2
25	24	Portugal	68.8	0.4	1.5	92	93	Dominican Republic	23.0	0.4	-0.1
24	25	Malaysia	68.4	-0.7	-0.2						
26	26	Hong Kong	65.4	-1.6	-1.6	95	94	Guatemala	22.4	0.3	0.5
27	27	Thailand	63.4	-0.4	-0.1	91	95**	Algeria	21.5	-1.3	-2.0
28	28	Malta	62.3	0.5	1.3	96	96**	Senegal	21.5	-0.1	-0.1
29	29	United Arab Emirates	60.8	0.0	0.3	102	97	Croatia	21.4	2.9	4.3
30	30	Czech Republic	60.1	1.7	4.3	97	98	El Salvador	20.6	0.5	1.9
32	31	Chile	59.2	1.8	3.6	98	99	Russia	19.9	0.5	0.4
31	32	Iceland	58.3	0.7	1.4	100	100	Malawi	19.8	0.7	1.0
33	33	China	56.4	-0.6	-1.2	99	101	Kazakhstan	19.2	-0.1	0.5
35	34	Cyprus	56.0	1.7	3.4	105	102	Myanmar	18.9	1.6	2.4
34	35	Saudi Arabia	55.1	-0.2	-0.7	101	103	Cameroon	18.5	-0.2	-0.7
37	36	Kuwait	54.1	0.7	1.4	108	104	Tanzania	17.7	1.0	2.2
36	37	Oman	53.8	0.2	0.1	109	105	Honduras	17.3	1.4	1.8

For the newer years, the accessible files look like the following example.

(**) Sept. 2004 to March 2005.
INSTITUTIONAL INVESTOR'S MARCH 2005 COUNTRY CREDIT RATINGS
 Legend for Chart:

A - Rank Sept. 2004
 B - Rank March 2005
 C - Country
 D - Institutional Investor credit rating
 E - Six-month change
 F - One-year change

A	B	C	D	E	F
1	1	Switzerland	94.5	-0.7	-0.7
4	2	Norway	93.7	0.0	-0.3
2	3	Luxembourg	92.8	-1.1	-2.0
11	4(*)	Finland	92.7	0.4	0.4
5	5(*)	U.K.	92.7	-0.9	-1.3
3	6	U.S.	92.4	-1.3	-1.6
8	7	France	92.2	-0.5	-1.1
6	8	Denmark	92.1	-0.9	-0.3
7	9	Netherlands	92.0	-0.9	-1.9

In each year, we use the country assessments as of September. Only in three years we had to revert to using the assessment as of March as the September value was not available. We then import the values into STATA, merge them with country codes and add them to the rest of our data. The ratings range is between 0 and 100, with 100 expressing the highest confidence on behalf of the experts.

3.6.4 Appendix 3.D: Control Variables

As discussed in the main text, we add an extensive set of control variables to some – but not the main – regressions. For this control vector, we build on and combine the sets of explanatory variables employed in Cantor and Packer (1996), Archer et al. (2007) and Hill et al. (2010) to control for the country-specific economic and political factors that should capture countries' ability and willingness to repay their debts.

We therefore add the following variables: the natural logarithm of GDP per capita, the annual GDP growth rate and its square, the inflation rate, the rents from natural resources (over GDP), the log of population, the debt to GDP ratio, the annual change in government debt (over GDP), trade (over GDP), the current account balance (over GDP), external debt (over GDP), the two variables indicating whether the country defaulted ever or within the previous five years, the quality of the rule of law, the degree of democracy (Polity IV), whether an election took place, the number of the government's years in office, the ruling party's political ideology, whether the country was affected by an internal or an external conflict, whether the military played an active role in politics, and an indicator for membership in the Eurozone (see also Fuchs and Gehring 2017).

We also include variables that the literature identified as correlates of IMF programs. Some of them are part of the above list. The variables we include in addition are the occurrence of a systemic banking crisis, the exposure of foreign banks to the country, investment (over GDP), and the similarity of voting with the United States in the United Nations General Assembly (Copelovitch 2010a; Moser and Sturm 2011; Sturm, Berger, and de Haan 2005). These variables are taken from the World Bank's World Development Indicators (WDI), the IMF (Laeven and Valencia 2012), the Database of Political Institutions (Beck et al. 2001), the Polity IV Project (Marshall, Jaggers, and Gurr 2011), and the International Country Risk Guide (ICRG), the Bank for International Settlement (BIS), and Bailey et al. (2017).

Descriptive statistics for all these control variables can be found below, in Appendix 3.E.

3.6.5 Appendix 3.E: Variables

Table 3.6 – Descriptive Statistics

Variable	N	Mean	SD	Min	Max
Rating S&P	1350	13.58	4.99	1.00	21.00
Rating Moody's	1142	14.13	4.98	1.00	21.00
Rating Fitch	1077	14.15	4.98	1.00	21.00
Rating Non-US	847	15.49	4.31	4.00	21.00
Institutional Investor	1335	59.09	21.55	10.50	96.40
IMF program	1350	0.21	0.40	0.00	1.00
IMF agreement	1350	0.08	0.28	0.00	1.00
IMF probability	1350	0.21	0.23	0.00	0.89
GDP/capita (ln)	1349	8.80	1.37	5.69	11.38
GDP growth	1350	3.89	3.75	-17.95	17.51
Inflation	1349	0.06	0.07	-0.05	0.95
Natural resource rents (% GDP)	1350	7.27	12.03	0.00	64.80
Population (ln)	1350	16.62	1.61	12.96	21.02
Debt (% GDP)	1349	48.45	30.41	0.00	238.03
Change in Government Debt (% GDP)	1349	3.16	10.74	-115.42	102.29
Default history (indicator)	1350	0.34	0.48	0.00	1.00
Default in last 5 years (indicator)	1350	0.07	0.25	0.00	1.00
Trade openness	1339	88.00	57.78	14.93	562.06
Current Account Balance (% GDP)	1345	-0.26	8.25	-44.21	44.62
External Debt (% GDP)	1349	21.90	28.31	0.00	189.48
Law and Order	1350	4.14	1.29	1.00	6.00
Democracy (Polity IV)	1348	6.16	5.57	-10.00	10.00
Election	1350	0.22	0.41	0.00	1.00
Honeymoon	1349	5.78	6.80	1.00	46.00
Left government	1350	0.31	0.46	0.00	1.00
Absence of Internal Conflict (ICRG)	1350	9.72	1.62	3.42	12.00
Absence of External Conflict (ICRG)	1350	10.36	1.22	5.17	12.00
Absence of military in politics	1350	4.45	1.44	0.00	6.00
Euro area (indicator)	1350	0.13	0.34	0.00	1.00
Investment (% GDP)	1347	24.01	6.37	8.27	58.15
Systemic Banking Crisis	1261	0.13	0.34	0.00	1.00
Foreign bank exposure	1350	217.92	625.79	0.008	6491.18
UNGA voting	1350	0.20	0.91	-1.66	2.89
Global GDP growth*	1350	0.03	0.02	-0.02	0.08

*Interacted with IMF Probability in the regressions.

Note: Based on the sample used for specification 7 in Table 3.3.

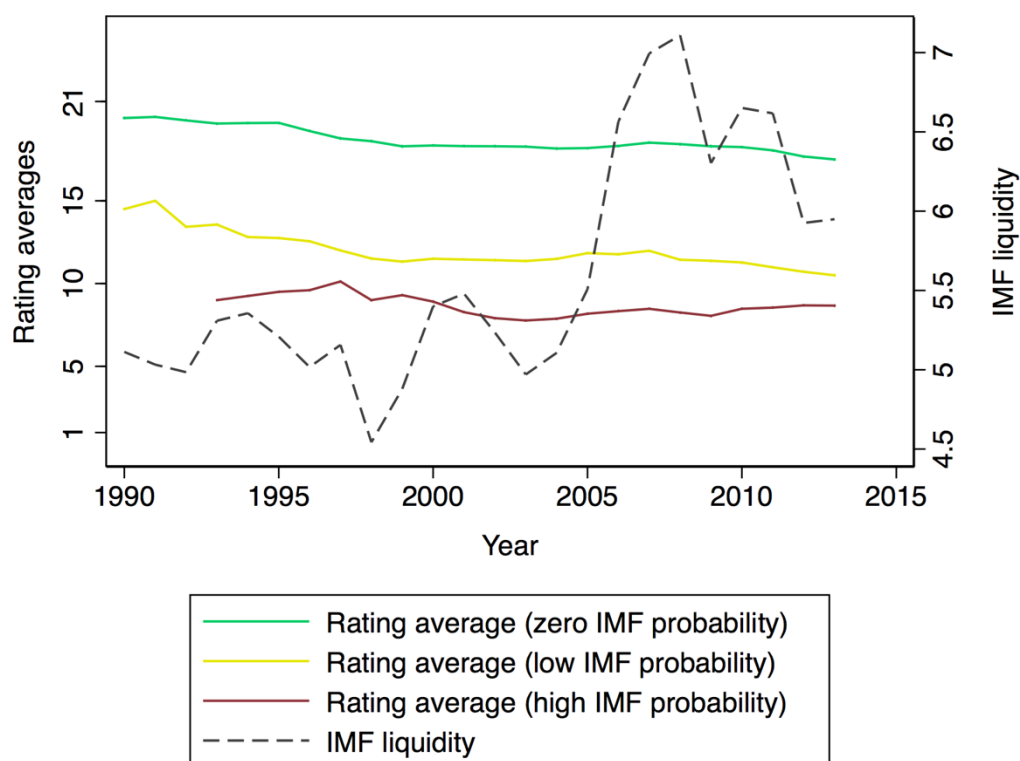
3.6.6 Appendix 3.F: Results of Robustness Regressions Described in Section 3.4.3

Table 3.7 – Various Robustness Tests

	constant probability	excluding large repurchases	excluding countries with largest repurchases	controlling for global trends interacted with IMF probability	excluding GFC	excluding Eurozone	IMF agreements
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
IMF program/ IMF agreement	0.228 [0.766] {0.766}	0.338 [0.844] {0.689}	0.404 [0.774] {0.602}	0.648 [0.834] {0.437}	0.368 [0.871] {0.673}	1.085 [0.907] {0.232}	2.205 [1.838] {0.230}
Country and Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2045	2004	1812	1767	1326	1840	1840
K-P underid. (LM)	17.412	15.569	12.077	16.114	17.573	15.866	16.453
K-P underid. (p)	0.000	0.000	0.000	0.000	0.000	0.000	0.000
K-P weak id. (F)	36.810	36.301	32.441	40.727	34.260	36.002	36.379
First stage							
IMF probability		3.604 [0.567] {0.000}	3.933 [0.681] {0.000}	3.358 [0.616] {0.000}	3.947 [0.645] {0.000}	3.415 [0.571] {0.000}	0.534 [0.295] {0.070}
IMF probability x IMF liquidity	-0.513 [0.085] {0.000}	-0.472 [0.078] {0.000}	-0.514 [0.090] {0.000}	-0.480 [0.075] {0.000}	-0.466 [0.080] {0.000}	-0.455 [0.076] {0.000}	-0.224 [0.037] {0.000}

Notes: The dependent variable is a country rating from S&P measured on a 21-point scale. Standard errors clustered at the country level are displayed in brackets, p-values in curly brackets. The sample contains up to 100 countries and covers the 1992 to 2013 period. Estimations were conducted using the `xtivreg2` command in Stata. Crisis refers to the years 2009-2013. All regressions include country and year fixed effects, as well as the economic and political control variables in t-1.

Figure 3.7 – Spurious Correlations between Ratings and IMF Liquidity?



Note: The figure plots the year-specific cross-country averages of S&P ratings for three sets of countries with different levels of *IMFprobability*. For the top line all countries that never received an IMF program are included for calculating these averages. For the middle line all countries whose *IMFprobability* is in the bottom half of this variable's distribution are included. For the bottom line all countries whose *IMFprobability* is in the upper half of this variable's distribution are included. The dashed line plots the *IMFliquidity* time series for comparison.

3.6.7 Appendix 3.G: Event-Based Identification

Table 3.8 – Regression Results of the Event-based Identification

	(1)	(2)	(3)	(4)
IMF start (t+11)	-4.289***	-0.324*	0.021	0.019
IMF start (t+10)	-4.372***	-0.433**	-0.014	-0.014
IMF start (t+9)	-4.335***	-0.457**	0.004	0.010
IMF start (t+8)	-4.338***	-0.449**	-0.007	-0.006
IMF start (t+7)	-4.459***	-0.516**	-0.055	-0.047
IMF start (t+6)	-4.516***	-0.528**	-0.088	-0.079
IMF start (t+5)	-4.565***	-0.576***	-0.092	-0.078
IMF start (t+4)	-4.560***	-0.593***	-0.013	-0.001
IMF start (t+3)	-4.738***	-0.713***	-0.172*	-0.166*
IMF start (t+2)	-4.839***	-0.745***	-0.236***	-0.216**
IMF start (t+1)	-5.003***	-0.883***	-0.392***	-0.369***
IMF start	-5.023***	-0.933***	-0.470***	-0.447***
IMF start (t-1)	-5.067***	-0.970***	-0.415***	-0.395***
IMF start (t-2)	-5.102***	-1.023***	-0.275***	-0.268***
IMF start (t-3)	-5.070***	-1.017***	-0.313***	-0.311***
IMF start (t-4)	-4.927***	-0.987***	-0.322***	-0.322***
IMF start (t-5)	-4.945***	-1.007***	-0.260***	-0.255***
IMF start (t-6)	-4.916***	-1.000***	-0.277***	-0.264***
IMF start (t-7)	-4.789***	-0.910***	-0.181***	-0.163**
IMF start (t-8)	-4.747***	-0.925***	-0.119	-0.099
IMF start (t-9)	-4.675***	-0.870***	-0.018	0.008
IMF start (t-10)	-4.686***	-0.915***	-0.042	-0.014
IMF start (t-11)	-4.664***	-0.892***	-0.027	-0.004
Constant	14.084***			
Country FE	No	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes
Country x Year FE	No	No	Yes	Yes
Month FE	No	No	No	Yes
Observations	25625	25625	25574	25574
Adjusted R-squared	0.077	0.917	0.995	0.995

Notes: OLS-FE regressions. The dependent variable is the S&P rating at the end of month t ; standard errors not shown. Figure 3.5 is based on the regression in column 4. Significance levels * $p < .10$, ** $p < .05$, *** $p < 0.01$

3.6.8 Appendix 3.H: Exploratory Analysis of Statements by Rating Agencies

In a first step, we conduct an exploratory analysis about the availability of statements on *Factiva*, a commercial database for press articles as well as corporate and business information owned by Dow Jones & Company, and the LexisNexis search engine. We searched for articles containing statements of rating agencies concerning the up- or downgrading of sovereigns based on the (potential) interference of the IMF, using the following search terms independently or in combination with each other: *IMF*, *Sovereign*, *Rating Agency*, *Rating*, *Development*. The statements listed below contain decisions of the three major rating agencies *Standard & Poor's*, *Moody's*, and *Fitch*. Overall, the exploratory search process yielded statements for 14 different countries (in Africa, Asia and Eastern Europe) in the years between 1999 and 2016. In the following, we list statements starting with decisions from *Standard & Poor's*, the agency of primary interest, followed by the ones from *Moody's* and those from *Fitch*. Countries are ordered alphabetically and ascending in years.

Based on this exploratory analysis, which makes no claim of being exhaustive, we designed our systematic text analysis described in more detail after the following statements.

Standard and Poor's:

Albania, 2014

"We revised the outlook to stable because we think that the recently concluded *International Monetary Fund* programme will serve as a policy anchor for fiscal consolidation and support the sustainability of Albania's high government debt," S&P's said."

Source: Balkan Insights, <http://www.balkaninsight.com/en/article/standard-and-poor-s-upgrade-albania-s-rating>

Angola, 2011

"Standard & Poor's (S&P) has raised Angola's sovereign risk rating to BB-, citing [...] the *IMF-recommended fiscal and monetary reforms*, which are expected to help mitigate the downside risks to over-dependence on the hydrocarbon sector."

Source: IHS Global Insight Daily Analysis, accessed via Factiva, 08.06.2017

Bosnia and Herzegovina, 2016

"The *IMF arrangement* will also provide the fiscal space for needed reforms and infrastructure investments. [...] it will anchor fiscal discipline for the authorities and aim to improve revenue collection and the efficiency of government spending."

Source: S&P, http://www.standardandpoors.com/en_US/web/guest/article/-/view/type/HTML/id/1707896

Ghana, 2015

"We think the new *International Monetary Fund (IMF) program* [...] will help in addressing fiscal and external imbalances [...]"

Source: S&P according to African Markets, <https://www.african-markets.com/en/news/west-africa/ghana/s-p-ghana-b-b-ratings-affirmed-on-new-imf-program-outlook-remains-stable>

Sri Lanka, 2009

"[...] (S&P) revised the outlook on its "B" long-term foreign currency rating for Sri Lanka to positive yesterday. The move reflects the country's improved external liquidity position owing to the new International Monetary Fund (IMF) standby loan agreement of US\$2.6 billion. [...]. The *stringent macro-economic consolidation conditions attached to the programme should force the government to reduce its fiscal deficit [...]. The central bank's commitment under the programme to a strict monetary policy including a reduction of advances to the government and a flexible exchange rate should also have a positive effect on Sri Lanka's medium-term sovereign risk.*"

Source: IHS Global Insight Daily Analysis, accessed via Factiva, 08.06.2017

Ukraine, 2015

"We view the risk of another default in the next two to three years as diminished due to the *Ukrainian authorities' commitment to the reforms set out in the International Monetary Fund (IMF) program,*" S&P analysts including Frank Gill said in the report."

Source: Bloomberg, <https://www.bloomberg.com/news/articles/2015-10-19/ukraine-rating-raised-to-b-by-s-p-on-debt-exchange-reform-plan>

Moody's:

Indonesia, 2002

"Moody's Investors Service changed the outlook [...] to positive from stable. The rating agency cited Indonesia's recent Paris Club memorandum of understanding and the country's *improved relationships with other foreign creditors, including the IMF, as bettering the country's liquidity position* in the coming two years. [...] Going forward, upward movement in the ratings will depend on, among other things, continued political stability, progress in disposing of IBRA assets, fiscal performance, and *the ability of the government to continue to meet the targets under its IMF program* and maintain good relations with foreign creditors generally. Moody's said that the positive outlook reflects progress made so far, but that *continued reforms were necessary to lift Indonesia's economic performance and improve investor confidence.*"

Source: Moody's Investor Service Press Release, accessed via Factiva, 08.06.2017

Pakistan, 2015

“Credit rating agency Moody’s has changed the outlook on Pakistan’s sovereign rating to Positive from Stable, affirming the rating itself at Caa1 [...]. The decision to change the outlook was prompted by Pakistan’s improving liquidity position, the *government’s continued efforts towards fiscal consolidation, and the steady progress with structural reforms under the International Monetary Fund (IMF)’s programme*. Pakistan’s external liquidity position has improved substantially in the past 12 months [...], supported by the narrowing current-account deficit, ongoing disbursements from the IMF, [...]. Meanwhile, fiscal discipline has also improved, as budget deficit and the government domestic borrowing have been gradually reduced. *On the structural reforms front, the agency pointed to the country’s successful completion of a number of IMF structural benchmarks, including those on the fiscal and debt management front and energy sector reforms.*”

Source: IHS Global Insight Daily Analysis, accessed via Factiva, 08.06.2017

Egypt, 2016

“Importantly, the rating agency views the staff-level agreement with the IMF which was announced on 11 August 2016 as credit-positive, because it will help alleviate some of Egypt’s external liquidity pressures. Under the Extended Fund Facility (EFF) Egypt would gain access to about \$12 billion of external funding through the IMF. The agreement is subject to approval by the IMF’s Executive Board, which Moody’s expects within 6-8 weeks. In Moody’s view, the agreement reached with the IMF is also important because it will unlock external funding from other multilateral and bilateral sources, and support the implementation of fiscal and economic reforms. These include the long-delayed introduction of a value-added tax and moves to a more flexible exchange rate regime.”

Source: Moody’s, https://www.moody.com/research/Moodys-affirms-Egypt-B3-rating-outlook-stable--PR_352656

Rwanda, 2016

“Moody's assigned Rwanda first-time local and foreign-currency issuer ratings of B2 last week, and gave the country a Stable outlook. [...] In Moody's view, a Stable outlook for Rwanda's sovereign credit is justified given access to USD204 million from the *International Monetary Fund (IMF) under the country's 18-month Standard Credit Facility (SCF) arrangement*. Additionally, it sees the government's policy implementation track record as strong, and *expects improvements in both its fiscal and external positions* to materialise over the medium term.”

Source: IHS Global Insight Daily Analysis, accessed via Factiva, 08.06.2017

Sri Lanka, 2016

“Therefore, in Moody's view, while the *IMF program* will alleviate Sri Lanka's external liquidity pressures, a more durable improvement in the macro-economic and balance of payments pressures will depend on the extent to which authorities can durably reverse the ongoing fiscal deterioration while improving Sri Lanka's international competitiveness and attractiveness to foreign investors. The study underpins *Moody's view that effective policy implementation determines the extent to which a country reaps the benefits of an IMF program*.”

Source: Moody's, https://www.moody's.com/research/Moodys-Sri-Lankas-benefits-from-its-IMF-program-depend-on--PR_350166

Ukraine, 2015

“The decision to upgrade the sovereign rating of Ukraine's government to Caa3 is based on the following key drivers: [...] 2. *Progress in political and economic reform under the auspices of the IMF-led programme, supporting a rebalancing of the economy and a meaningful reduction in public and external financial deficits*.”

Source: Moody's, https://www.moody's.com/research/Moodys-upgrades-Ukraines-sovereign-rating-to-Caa3-outlook-stable--PR_336283

Fitch:

Benin, 2004

“Fitch stated that successive IMF reform programmes have led to macro-economic stabilisation, including a reduction in the budget deficit and a stabilisation of the government’s debt burden through tight fiscal policies.”

Source: World Markets Research Centre Daily Analysis, accessed via Factiva, 08.06.2017

Ghana, 2005

“Fitch Ratings has upgraded Ghana's long-term foreign and local currency rating [...] The International Monetary Fund (IMF)/World Bank supported Poverty Reduction Strategy will be supported with higher aid funding, which should improve public investment, counteract a projected current-account deterioration and improve international reserves.”

Source: World Markets Research Centre Daily Analysis, accessed via Factiva, 08.06.2017

3.6.9 Appendix 3.I: Systematic Analysis of Statements by Rating Agencies

Based on the exploratory analysis, we selected FACTIVA as the more suitable database for a systematic analysis. In particular the feature to select an industry class improved the matching rate between search terms and statements significantly. Our final systematic approach was to

- 1.) **Open the database and login (library access or account is required).**
- 2.) **Issue search queries:**
 - “program” within three words distance to “IMF or International Monetary Fund”, Industry: Rating Agency, Language: English or German
 - “liquidity” within three words distance to “IMF or International Monetary Fund”, Industry: Rating Agency, Language: English or German
 - “reform” within three words distance to “IMF or International Monetary Fund”, Industry: Rating Agency, Language: English or German
 - “program,” “IMF or International Monetary Fund” and “rating” within a ten word corridor, Industry: All, Language: English or German
- 3.) **Manually skim over all statements and delete obviously false matches.**
- 4.) **Pool all remaining text in one text file.**
- 5.) Relevant text is often embedded in larger bodies of text irrelevant to our purpose. **Thus, we run a python script (see below) that searches the text for “IMF” or “International Monetary Fund” and extracts ten lines of text buffer prior and subsequent to a hit.** Moreover, we used regularities in text structure to extract the according publisher and date. Selecting the size of the buffer one faces a trade-off between reducing the volume of text and cutting potentially relevant information. A ten line buffer is a conservative choice towards minimizing the loss of information.
- 6.) Because these are still relatively large chunks of texts, we **manually read the remaining texts and delete irrelevant relevant parts**, and then copy the rest of the text and additional information (name of rating agency and country) to excel. If duplicates appear they are deleted. This left us with 126 statements.

We then developed the following codebook. Two student assistants were equipped with this codebook and went through all statements. In case of deviations in opinion, we always choose the choice biasing against our priors, i.e. the effects we hypothesize. Accordingly, in case of deviating opinions statements are grouped as “liquidity and reforms” instead of “reforms only” and are grouped as “mixed/neutral/negative” based on the more negative of two assessments.

Positive-Negative Dimension:

Positive = 1 iff the statement in question includes remarks which **indicate** that the IMF is seen in a positive light by the rating agency. Assumes background knowledge about basic economic processes and implications of measures for economy.

- **Indicators for IMF being seen in positive light by rating agency:**

- *Citing **actual** or **possible** implementation or continuation of an IMF program or measure or **actual** or **possible** positive developments due to an IMF program or measure as a reason for an **actual** or **possible** positive rating. Conversely, citing **actual** or **possible** lack of implementation or discontinuation of an IMF program or measure as a reason for an **actual** or **possible** negative rating.*

- ➔ **Example for actual continuation of program as reason for actual positive rating:**

ID5: "The ratings firm cited the country's improved performance under the European Union-International Monetary Fund program, falling near-term liquidity risk and a better fiscal track record for its upgrade"

- ➔ **Example for possible discontinuation of/ compliance problems with program as reason for possible negative rating:** ID10: "Greece's ratings could also be lowered for reasons unrelated to the proposed ESM, if the Greek government's ability to comply with the program is undermined by domestic political opposition or materially weakens for other reasons, increasing the likelihood of failure to fully comply with the IMF/EU program."

- ➔ **Example for possible discontinuation of program as reason for actual negative rating:** ID69: "The outlook is negative, reflecting what we view as ongoing social and political risks associated with deleveraging efforts by Portugal's highly indebted private and public sectors, as well as financing uncertainties related to Portugal's exit from the EU/IMF program, expected in May 2014. We believe this is symptomatic of diminishing political backing for further fiscal and structural reforms. The Constitutional Court's deliberations over further fiscal measures could coincide with Portugal's planned EU/IMF program exit in the second quarter of 2014."

- ➔ **Example for actual implementation of program as reason for possible positive rating:** ID20: "Turkey's economy has been improving and a continuation of the current positive trend could lead to higher credit ratings for the country, according to the general manager of Moody's Interbank Credit Service's regional Middle East office. [...] "We see lower inflation, the fiscal deficit relatively under control and the International Monetary Fund (IMF) targets seem to be achievable," he said. The IMF is helping Turkey through a stabilization package that sets macroeconomic targets and provides aid in return. [...] In Turkey, programs have been suggested by the IMF that are aimed at lifting its economy out of the debt trap and making it into a debt paying machine. "The IMF provides financing to Turkey through a macro-economic stabilization program. The program calls for the government to take certain actions to correct the macro-economic imbalances. These imbalances include various fiscal and economic reforms that would lead to improvement in the macro-economic conditions."

- Citing **actual** or **possible** implementation or continuation of an IMF program or measure as a factor for **actual** or **possible** positive economic developments in the country. Conversely, citing **actual** or **possible** lack of implementation or discontinuation of an IMF program or measure as a factor for **actual** or **possible** negative economic developments the country.
 - ➔ **Example for actual implementation/ compliance with program as factor for actual positive developments:** ID121: "As a result of the Chuan's cabinet's decisive policy to comply with the IMF program together with the disbursement of US\$10.282 billion as of March 30, 1998 out of the IMF rescue package for US\$17.2 billion, the present market situation is relatively stable and the market confidence seems to be recovered to some extent. (...)"
- Use of terms such as "successful completion" when talking about an IMF-program or measure.
 - ➔ **Example:** ID79: "Such political developments allowed to strengthen the fiscal management stability. The Latvian government also in late 2011 successfully completed the international assistance program with the European Commission and the International Monetary Fund (IMF), said the agency."

Negative = 1 iff the statement in question includes remarks which **indicate** that the IMF is seen in a negative light by the rating agency. Assumes background knowledge about basic economic processes and implications of measures for economy

- **Indicators for IMF being seen in negative light by rating agency:**

- Citing application for or implementation of IMF program or measures as a reason for an actual or possible downgrading
 - ➔ No examples
- Citing application or implementation of IMF program or measures as a reason to keep outlook at negative
 - ➔ **Example:** ID74 "Moody's Investors Service has today confirmed Egypt's B2 government bond ratings and maintained the rating outlook at negative. [...] The key drivers of today's confirmation of Egypt's B2 sovereign rating and negative outlook are: [...]4) The formal request by the new Egyptian government for IMF support"

Positive =0 and *Negative* =0 iff the statement in question neither includes remarks which indicate that the IMF is seen in a positive nor remarks which indicate that the IMF is seen in a negative light by the rating agency, or status of remark (positive/negative) is unclear.

- *Purely descriptive statements about IMF without evaluative content*
 - ➔ **Example:** ID59 "Pakistan is also moving forward on structural reforms under its program with the International Monetary Fund (IMF). These reforms focus primarily on fiscal consolidation, debt management, and addressing structural constraints in the energy sector."
- *Statements with not enough context to conclude status (e.g. because it is unclear if rating has changed in any way)*
 - ➔ **Example** ID93: ""However, policy adjustments and financial support under an 18-month IMF program agreed in April 2009 support a stable rating outlook," says Byrne."

Liquidity-Reform Dimension:

Liquidity Only = 1 iff the only feature addressed by the rating agency in their remarks in connection with the IMF is the liquidity of the country that is being rated (regardless of whether IMF is seen as donor or whether there might be consequences for liquidity resulting from e.g. implementation of IMF-program).¹⁴¹

- **Verbal indicators taken to address liquidity in statements about IMF:**

- *“financial assistance”*
- *“public wages (lowered)”*
- *“program to relieve the financial burden”*
- *“(future) disbursements”*
- *“financial support from the IMF”, etc.*

➔ **Example:** ID8 “(...) In our view, such improvements could be brought about by a positive conclusion to the negotiations with Gazprom on Ukraine's gas contract and/or a resumption of disbursements under Ukraine's IMF program,” the press release reads.”

Reform Only = 1 iff the only feature addressed by the rating agency in their remarks in connection with the IMF are reforms for the country that is being rated (regardless of whether IMF is seen as the one demanding reforms or the source of further IMF-unrelated reforms)¹⁴².

- **Verbal indicators taken to address reform in statements about IMF:**

- *“technical assistance”*
- *“(...) bolstering its institutional framework”*
- *“policy measures”*
- *“IMF assisted economic reform program”, etc.*

➔ **Example:** ID3 “(...) Moody's report explains that the Solomon Islands successfully graduated from an IMF program in 2016, with progress in bolstering its institutional framework.”

Reform and Liquidity = 1 iff the rating agency addresses both reforms and liquidity in their remarks in connection with the IMF (regardless of whether IMF is seen as the one demanding the reforms or the source of further IMF-unrelated reforms and regardless of whether IMF is seen as donor or whether there might be consequences for liquidity resulting from e.g. implementation of IMF-program).

➔ **Example:** ID2 “(...) The new IMF credit facilities (an Extended Credit Facility and an Extended Fund Facility (ECF/EFF)) approved in November unleashed official lending that had been withheld for more than a year. The second driver for stabilizing the outlook relates to the adoption of key structural reforms both in connection with the IMF program and in technical consultation with the IMF and other multilateral lenders and donors. (...)”

¹⁴¹ If there are consequences resulting from IMF-related liquidity, then statement is coded as 1. However, if statement only addresses circumstances or conditions which led to IMF-measures with regard to liquidity, statement is coded as 0.

¹⁴² If there are consequences resulting from IMF-related reforms, then statement is coded as 1. However, if statement only addresses circumstances or conditions which led to IMF-measures with regard to reform, statement is coded as 0.

Liquidity Only = 0, *Reform Only* = 0 and *Reform and Liquidity* = 0 iff either the rating agency neither addresses liquidity, nor reform nor both in their remarks about the IMF, or status of statement is unclear.

- *Use of the expressions “IMF program” or “IMF agreements” (or synonymous expressions) with no further specification with regard to what the program or agreement is about*

➔ **Example:** ID13 “(...) Under this scenario, the government can get the International Monetary Fund's program "back on track" and there is a strong prospect of positive ratings action, said Edward Parker, a senior Fitch analyst.”

4 The Economics of the Democratic Deficit: The Effect of IMF Programs on Inequality

Note: This paper is single-authored. An earlier version of it is available as *Heidelberg University Discussion Paper 617* (2016).

4.1 Introduction

International organizations (IOs) have become powerful political actors. In the increasingly globalized world, these specialized multilateral institutions often can address the growing amount of cross-border interdependencies more effectively than states individually (Keohane 1984). A main point of criticism on IOs, however, is the prominent claim that they suffer from a ‘democratic deficit.’ In the literature “[t]he main problem [...] of the democratic deficit is generally understood to be the relative lack of accountability of IOs to the individuals whose lives they directly affect” (Grigorescu 2013, 177).

The public and academic debate on the democratic deficit has often focused on the extent to which decision-making in IOs deviates from idealistic standards of democracy. The concept has thus primarily been used in a normative context in order to critique IOs on the basis of democratic ideals originally developed for nation-states. The fact that states and IOs cannot easily be compared against the same standards posed a problem for positive approaches to the topic.¹⁴³ More recently, however, the academic debate shifted to the question of how IOs *interact* with domestic democracy. Some scholars argue that if IOs are less responsive to voter preferences than democratic governments, their interference with domestic decision-making processes “involves the risk of making democracies less representative;” others disagree and point to the “democracy-enhancing potential” of IOs.¹⁴⁴ While this way of approaching the topic helps to derive testable hypotheses, both sides of the debate acknowledge that empirical evidence illuminating this controversy from a positive perspective remains scarce.¹⁴⁵ This study explicitly aims to provide such empirical evidence.

To do so, it focuses on the International Monetary Fund, an IO that is both particularly powerful and often criticized for being ‘democratically deficient.’¹⁴⁶ In the context of the IMF, the idea that IOs may undermine domestic democracy leads to a hypothesis that can be tested with the help of statistical techniques that enable inferences based on plausible

¹⁴³ See Moravcsik (2002) for a critique of the normative debate on the democratic deficit along these lines. See Grant and Keohane (2005) on the challenges involved in transferring standards of democratic accountability from nation-states to IOs.

¹⁴⁴ For the prominent debate in the journal *International Organization* see Keohane, Macedo, and Moravcsik (2009) and the rejoinders by Gartzke and Naoi (2011) and Keohane, Macedo, and Moravcsik (2011).

¹⁴⁵ In fact, Keohane, Macedo, and Moravcsik (2011, 600) state: “we recommend empirical analysis precisely to explain variation in democratic consequences of multilateralism.” Similarly, Gartzke and Naoi (2011, 596) note that “scholars should focus on how domestic politics and institutions ‘filter’ the distributional effects of policies adopted by multilateral organizations.” This paper is a direct response to their calls.

¹⁴⁶ Stone (2002, 1) calls the IMF “the most powerful international institution in history.” Stiglitz (2000, 4) criticizes the IMF for “undermin[ing] the democratic process by imposing policies.”

counterfactuals. I hypothesize that the influence of the IMF on domestic politics results in distributional outcomes that are observably different to the outcomes democracies would produce in the absence of such an influence. Potential differences in such distributional outcomes can then be interpreted as evidence for the idea that the IMF is capable of temporarily undermining the responsiveness of democratic governments to the distributional preferences of their citizens.

The considered outcome is the distribution of income within national societies. As an egalitarian distribution of political power is at the core of the concept of democracy, standard theoretical concepts suggest that *ceteris paribus* a more democratic form of governance is typically associated with a more equal distribution of economic goods. To the extent that the lending activities of the IMF have a distributional dimension, the relative lack of democratic control over these by citizens in program countries means that citizens lack mechanisms to guide and constrain (re)distributional policies according to their preferences. If the argument holds, the effect, I argue, should be driven particularly by democratic program countries: it is only in such countries that the interference of a democratically deficient IO in national policy-making can limit the functioning of existing domestic accountability mechanisms. Recent efforts to ‘democratize’ IO decision-making processes, which scholars have noted in recent years, could, in turn, mitigate this effect (Grigorescu 2015).

The most direct channel through which the IMF can affect economic outcomes in member countries is conditionality in loan programs. These policy reforms that the IMF demands in exchange for its loans aim to resolve balance-of-payment crises and correct underlying macroeconomic and structural problems. In practice, this often translates into conditions with potentially adverse distributional implications, such as cuts to public spending, labor market reforms as well as trade and capital account liberalization (Kentikelenis, Stubbs, and King 2016; Stone 2008).

At the core of this study is an empirical setting that makes it possible to construct a plausible counterfactual. The key methodological challenge is to address endogeneity since IMF programs are not randomly assigned but usually take place in countries with current economic crises. Extant approaches addressing this problem either control only for observables or rely on problematic instrumental variables (IVs) that are likely to be related to the outcome through channels other than IMF programs and thus violate the exclusion restriction. To fill this gap, I employ a novel identification strategy for IMF programs inspired by recent methodological

innovations (Nunn and Qian 2014; Werker, Ahmed, and Cohen 2009). I exploit exogenous variation over time in the IMF's liquidity and interact this variable with a country's probability of participating in IMF programs, thereby introducing variation across countries. When controlling for the levels of the interacted variables, this interaction is, arguably, excludable to country-specific variables such as income inequality and thus allows me to determine the causal effect of IMF programs. As the exclusion restriction of this new IV holds not only for inequality but also for other economic and political outcomes on the country-level, the methodological section of this paper can be considered as an attempt to provide the literature with a tool to investigate the causal effects of IMF programs at large.¹⁴⁷

Foreshadowing the main results, I find strong statistical evidence in support of the hypotheses. IMF programs increase inequality in democracies but have no such effect in non-democracies. In democracies, the effect is statistically significant, robust to a battery of additional tests and substantial in magnitude. On average, the rise in inequality induced by IMF programs is equivalent to a lump-sum transfer of four to eight percent of the poorer half's mean income by each person in the poorer half to each person in the richer half. Examining potential channels, I analyze IMF conditions and find evidence suggesting that inequality rises faster in program countries that face more extensive conditionality and that austerity and labor market reforms are among the policy conditions driving the main result. In light of these results, the study adds not only to the democratic deficit controversy and the causal identification of IMF program effects, but also to research stressing interaction effects between IMF programs and domestic politics (Caraway, Rickard, and Anner 2012; Nooruddin and Simmons 2006), and to the growing literature on the causes behind the continuing trend of rising inequality within many countries.

4.2 The Argument

4.2.1 *The IMF and the Democratic Deficit*

While the concept of the democratic deficit relates to IOs in a general sense, the focus of this study is on a single organization to make both the theoretical argument and the empirical

¹⁴⁷ Since this paper was published as a working paper in 2016, one study adopted its identification strategy (Nelson and Wallace 2016), one study applied a closely related strategy (Foster et al. 2018) for studying the IMF's effects, and two studies applied a related strategy based on using the World Bank's liquidity to studying the World Bank's effects (Dreher et al. 2017; Gehring, Kaplan, and Wong 2018).

analysis as concise as possible. The prime reason for the focus on the IMF, and more specifically on its loan programs, is that this setting arguably constitutes the easiest case for testing the theoretical predictions.¹⁴⁸ First, the IMF is sometimes considered “the most powerful international institution in history” (Stone 2002, 1). It has vast financial resources at its disposal and its loan arrangements can use conditionality as a potent instrument to affect economic policies in program countries. The fact that *almost all* developing countries have received IMF programs since the early 1970s additionally underlines the organization’s global relevance. Second, in the literature, the concept of the democratic deficit has been repeatedly applied to the IMF (Nye 2001; Stiglitz 2000). More generally, IMF decision-making has been subject to a considerable amount of research, making the organization one of the best-researched IOs to date. As I will argue, in this vast literature there is ample evidence indicating that the IMF is a prime example of an IO with a democratic deficit.¹⁴⁹

To do so, an analytical framework considering IOs as a set of “nested principal-agent relationships” helps to structure the argument (Nielson and Tierney 2003, 250; see also Vaubel 2006). From this perspective, the IMF is part of a delegation chain starting with voters in member countries (the ‘ultimate principal’), running via national governments and their representatives in the IMF’s executive board, and ending with the IMF’s staff. The problem for the ultimate principal arises because there may be ‘agency slack’ at each point in the delegation chain resulting from divergent preferences between agents and their ‘proximate principals’ as well as limits to the accountability of the former to the latter. Such limits to accountability result primarily from high costs of information and control, and the fact that agents can exploit preference heterogeneity among multiple and collective principals to pursue their own interests (Hawkins et al. 2006; Nielson and Tierney 2003; Vaubel 2006). The literature on IMF decision-making has mainly focused on the ‘agency slack’ that is due to the influence of individual major shareholder governments (see also chapter 1) and the IMF’s staff (see also chapter 2):

¹⁴⁸ Of course this also means that the findings are not automatically valid with regards to other IOs. Future research could apply variants of the argument to other IOs.

¹⁴⁹ For reviews of this literature see Dreher and Lang (2016) as well as Steinwand and Stone (2008).

Major Shareholders

Systematic empirical evidence for the disproportional influence that the United States and other major shareholders¹⁵⁰ exert on IMF decision-making abounds. It has repeatedly been shown that countries receive favorable treatment from the Fund if they are politically close or (economically) important to the United States. Such countries receive more IMF programs, larger loans, less stringent conditionality as well as more positively biased growth forecasts and debt sustainability ratings (Broz and Hawes 2006; Copelovitch 2010a; Dreher, Marchesi, and Vreeland 2008; Dreher, Sturm, and Vreeland 2009a, 2015; Gould 2003; Reynaud and Vauday 2009; Stone 2008; Thacker 1999; see also chapters 1 and 2). Theoretical models accounting for these empirical findings suggest that the governments of major shareholders intervene in the IMF's decision making process when their strategic interests are at stake (Stone 2008; 2011; 2013). While these governments also have the largest formal voting power in the Executive Board, scholars underline that much of their considerable impact on IMF policies runs through so-called "informal governance." The governments of the countries that typically receive IMF programs tend to lack both significant formal voting power and such informal channels of influence.¹⁵¹

In sum, this part of the literature suggest that the IMF is not unaccountable, but that its accountability to citizens in typical program countries is limited by the fact that their governments' influence on the organization is weak relative to that of the major shareholders. In the words of Grant and Keohane, "[t]he problem is not a lack of accountability as much as the fact that the principal lines of accountability run to powerful states" (Grant and Keohane 2005, 37). While citizens of these powerful countries can have an indirect impact on IMF policy via their governments, the influence of citizens whose countries receive IMF programs, and are thus directly affected, is relatively limited.

The Bureaucracy

Another line of research on IMF decision-making has examined the extent to which the particular interests of the bureaucracy are reflected in the Fund's policy decisions. According

¹⁵⁰ The other "G5" countries: United Kingdom, France, Germany, and Japan (Copelovitch 2010a).

¹⁵¹ An example is the fact that the vast majority of typical program countries has no individual representatives in the Board as minor shareholders are grouped and represented by joint Directors. In the World Bank, the IMF's similarly structured sister organization, countries with a Director have been shown to exploit informal channels to receive favorable treatment from the organization (Kaja and Werker 2010). For other examples see Stone (2008).

to public choice theory, IMF staff face bureaucratic incentives, making the maximization of power, budget, and prestige key determinants of the IMF's bureaucratic decision-making (Vaubel 1986). Several studies have observed such behavior within the IMF and argue that its officials are able to push for longer programs, larger loans and more far-reaching conditionality than what is economically optimal (Barnett and Finnemore 2004; Bird and Willett 2004; Copelovitch 2010a; Vaubel 1996; chapter 2) The expansion of the IMF's responsibilities and budget over time through "mission creep" is also linked to the bureaucracy's power and rent-seeking motives (Dreher and Lang 2016; Reinhart and Trebesch 2015). Other scholars show that the staff's ideological beliefs and policy preferences are also directly reflected in the organization's policy decisions (Barro and Lee 2005; Chwioroth 2013; Nelson 2014). These findings support the view that the ultimate principal's opportunity to hold IMF staff to account is severely limited. The high information asymmetry between the two actors resulting from high information costs, language barriers, distance, complexity, and limited transparency of IMF decision-making is considered to be a major obstacle for such accountability (Vaubel 2006).

In sum, from the perspective of the ultimate principal there is agency slack at least at two points in the IMF's delegation chain concerning the design of IMF programs. Both major shareholder governments and staff are able to influence this design according to their particular interests. The former are able to exploit their influence on the organization to further their own political and economic interests, while the latter can shape the IMF's policy decision in accordance with their material interests and ideological preferences. Both sets of particular interests are in many cases unlikely to align with the preferences of the ultimate principal. The 'relative lack of accountability' of IMF decision making to individuals affected by it thus makes it likely that the IMF's policy output will diverge from the policy preferences of citizens in typical program countries.

I argue that this may have distributional implications. If IMF programs have a distributional dimension, the influence of individual major shareholder governments and the bureaucracy on their design runs the risk of producing distributional outcomes that do not match the distributional preferences of citizens in program countries. The policy areas addressed by conditionality in IMF programs that according to economic research are likely to have a distributional dimension are thus the focus of the subsequent section.

4.2.2 IMF Conditionality and Inequality

Trade and Capital Account Liberalization

First, one of the key policy areas that conditionality in IMF programs has often addressed is the liberalization of trade and capital flows across borders and the promotion of economic globalization more broadly (Woods 2006). A recent analysis shows that since the late 1980s more than 70% of IMF arrangements set conditions for the external sector, including the liberalization of trade and the capital account (Kentikelenis, Stubbs, and King 2016). Multiple studies find effects of IMF programs on capital account liberalization¹⁵² while some – albeit not all – find direct positive effects on inflows of foreign direct investment (FDI).¹⁵³

These policy conditions are partly reflective of the preferences of the Fund's major shareholders and its staff. The former have an economic interest in the liberalization of trade and capital flows in developing countries. IMF programs were often followed by U.S. and European firms acquiring and investing in local companies; this was facilitated by IMF conditions stipulating liberalization in economies that were previously relatively shielded from international markets and foreign capital (Klein 2008; Stiglitz 2002). In addition to many anecdotes, quantitative evidence shows that following IMF programs countries receive more FDI from the United States (Biglaiser and DeRouen 2010). At the same time such liberalization policies also reflect staff preferences. Scholars have, for instance, argued that the IMF's strong emphasis on capital account liberalization was in part driven by bureaucratic competition between the IMF, the Organization for Economic Co-operation and Development and the European Commission. In line with public choice reasoning, each organization aimed to control this issue area (Steinwand and Stone 2008).

The distributional consequences of these various dimensions of economic globalization have been subject to a large amount of research. Studies employing composite measures of globalization have repeatedly found inequality-increasing effects (Bergh and Nilsson 2010; Dreher and Gaston 2008; Lang and Tavares 2018). Recent studies focusing on more narrow measures of capital account openness and financial liberalization also find adverse distributional consequences of such policies (Bumann and Lensink 2016; Furceri and Loungani

¹⁵² See Chwioroth (2007b; 2010) and Mukherjee and Singer (2010) but note that they point to important heterogeneities.

¹⁵³ See Bauer, Cruz, and Graham (2012) and literature cited therein.

2018; de Haan and Sturm 2016; Jaumotte, Lall, and Papageorgiou 2013).¹⁵⁴ Likewise, FDI inflows have been linked to increasing inequality (Feenstra and Hanson 1996; Jaumotte, Lall, and Papageorgiou 2013). The empirical evidence on the effect of trade liberalization on inequality is more mixed yet positive for at least some country groups (Goldberg and Pavcnik 2007). In sum, the ‘globalizing’ effect of IMF programs is a potential channel through which income inequality may be affected.

Public Spending

Second, many IMF programs include conditions on fiscal issues that aim to limit or reduce government expenditure. According to Stone’s analysis of conditionality data for the 1992-2002 period, “there is almost always some limit on public debt or government spending” (Stone 2008, 600). Gould (2006, 90) finds that in the 1983-1990 and the 1991-2000 period 91 percent and 100 percent respectively of IMF programs included fiscal targets. Hence, IMF programs are often considered tantamount to austerity.

One reason why IMF programs limit public spending is that they usually put a strong focus on the timely repayment of debt to creditors. Gould (2003, 2006) documents that many IMF programs include “bank-friendly conditions” that give priority to the repayment of principal and interest to commercial banks. Building on this finding, research has repeatedly supported the notion that governments of major shareholders use their influence on the institution to prevent financial losses for creditors from their country (Broz and Hawes 2006; Copelovitch 2010a; 2010b). While major Fund shareholders face few incentives to push for increasing public spending, conditions that prioritize debt repayment limit the amount of resources available for other public expenses. Other scholars link the policy preferences expressed in conditions that stipulate smaller governments (and liberalization) to the beliefs and political views of the IMF bureaucracy and its tendency to favor market-liberal ideas over government intervention in market processes and outcomes. They argue that its staff’s educational background promotes confidence in market-oriented policies and conclude that “the institution’s ideational culture is dominated by a set of ‘neoliberal’ economic beliefs” (Barnett and Finnemore 1999; Nelson 2014, 304; Woods 2006). Quantitative studies show that the share of “neoliberal” bureaucrats within the organization is positively associated with increases of

¹⁵⁴ Note, however, that there is no consensus in this literature and that important heterogeneities exist. De Haan and Sturm (2016) provide a comprehensive literature review.

capital account openness in IMF program countries (Chwieroth 2007a; 2007b; 2010) and that IMF program design depends on the degree to which IMF staff and program country officials share “neoliberal” beliefs (Nelson 2014).

The resulting effects of IMF programs on public spending have been subject to multiple studies. Nooruddin and Simmons (2006) find a significantly negative effect on social spending in democracies. While two studies by IMF staff find a positive effect, Stubbs, Kentikelenis, and King could not replicate this using the same data (Clements, Gupta, and Nozaki 2013; IMF 2003; Stubbs, Kentikelenis, and King 2016). Instead, the latter find positive effects on health expenditure only in Sub-Saharan Africa, and negative effects in other regions. Other studies provide support for the hypothesis that IMF programs reduce public spending by documenting negative effects on public wages (Nooruddin and Vreeland 2010; Rickard and Caraway 2018).

The reduction of government expenditure can increase inequality in net incomes either via reducing the amount of redistribution through progressive taxes and transfers or it can directly affect the distribution of gross income, i.e., income before taxes and transfers. The literature suggests that decreasing public spending on education can result in rising wage inequality to the extent that the level or the equality of educational attainment are adversely affected by it (Goldin and Katz 2010; De Gregorio and Lee 2002). Furthermore, the frequently included conditions demanding pension cuts or freezes are a potential channel. More generally, the fact that IMF programs restrict government expenditure during periods of economic liberalization limits the opportunities to ‘embed liberalism.’ As IMF conditionality typically combines liberalization and austerity, vulnerable segments of society, which usually benefit from insurance policies that governments often implement in order to provide protection against economic volatilities and risks resulting from increasing openness (“embedded liberalism”), may lack these “compensations” under IMF programs (see Cameron 1978; Rodrik 1998; Ruggie 1982; Walter 2010).

Labor Market Reforms

A third policy area with potentially distributional implications often covered by IMF conditionality is the labor market. Depending on the definition of labor market policies such conditions can include, but are not limited to, reductions in minimum wage levels, dismissals in the public sector, pension cuts, the legalization of nonpermanent labor, and the privatization

of state-owned enterprises (SOEs). According to one count, at least a quarter of IMF programs launched since 1987 have been found to include labor market conditions (Caraway, Rickard, and Anner 2012).¹⁵⁵ More than half include conditions that stipulate reforms of SOEs (Kentikelenis, Stubbs, and King 2016).

Again, the preference for such policies can be linked to the interests of major shareholder governments and the bureaucracy. Coupled with the liberalization policies discussed above, privatizations during IMF programs, on the one hand, often resulted in firms from the United States and Europe winning the bids for privatized SOEs. Labor market deregulations in program countries also benefitted investors from major Fund shareholder countries.¹⁵⁶ On the other hand, the gradual expansion of the scope of IMF conditionality into such areas as labor market policies has often been linked to the bureaucratic incentive to expand the organization's mission: While initially not part of the IMF's repertoire, conditions that address the details of structural policies in program countries ("microconditionality") became increasingly frequent. Many scholars argue that this was driven by staff aiming to increase their power, responsibilities, and budget (Barnett and Finnemore 2004; Dreher and Lang 2016; Kentikelenis, Stubbs, and King 2016; Polak 1991; Reinhart and Trebesch 2015).

Caraway, Rickard, and Anner (2012) argue that "[r]eforms of this nature are costly for workers in the short to medium term." The results of Vreeland (2002) indirectly support this by showing that IMF programs are associated with declines in the income share of labor manufacturing. Blanton, Blanton, and Peksen (2015) find IMF programs to reduce collective labor rights. Nooruddin and Vreeland (2010) find negative effects on public employment; Rickard and Caraway (2018) find IMF conditions targeting the public sector to reduce the public wage bill. Such cuts could both increase and reduce wage inequality, depending on the country-specific relation between public wages and median income. Other IMF labor conditions can adversely affect the distribution of gross income as minimum wage reductions and weaker collective labor rights have been found to increase wage inequality (Autor, Manning, and Smith 2016; Kerrissey 2015; Koeniger, Leonardi, and Nunziata 2007). Furthermore they can, at least in the short to medium term, increase unemployment resulting from layoffs in the public sector and in SOEs.

¹⁵⁵ Note that Kentikelenis, Stubbs, and King (2016, 14) apply a broader definition and report that "about 50% or more of lending programmes have carried at least one labour-related condition over the 1994-2007 period."

¹⁵⁶ For anecdotal evidence see Klein (2007) and Stiglitz (2002).

In sum, the considerations up to this point suggest a first testable hypothesis:

H1: IMF programs increase income inequality.

4.2.3 IMF Programs and Domestic Democracy

Testing H1, however, is only the first step of the empirical exercise. In order to shed more light on the theoretical question this paper poses, the interaction of IMF programs with democracy needs to be addressed.

The standard theories in the literature on democracy and inequality suggest that democratic governments, responding to voter preferences, promote a more egalitarian distribution of income across society than non-democratic governments.¹⁵⁷ For instance, Meltzer and Richard's (1981) prominent model suggests that the decisive voters earn less than mean income and therefore have a preference for redistribution from the rich to the poor and, thus, lower inequality. Here as well as in subsequent research, broadened access to political power is thus closely linked to broadened access to economic resources (Acemoglu and Robinson 2005; Boix 2003). While not all empirical studies find the effect the vast majority of the theoretical literature suggests, there is substantial empirical support for these arguments.¹⁵⁸

Going beyond this basic model and turning to the more specific policy areas discussed above, there is a well-established empirical link between democracy and the size of the government (Avelino, Brown, and Hunter 2005; Boix 2001; Jensen and Skaaning 2015). Democracies provide more public goods and spend more on education, health and social security. A prominent theoretical explanation for this finding is based on the idea of the "political survival" of leaders, according to which democratic leaders tend to provide public goods because they are accountable to a larger base of supporters, i.e., voters (Bueno de Mesquita et al. 2003). The evidence furthermore suggests that democracies increase public spending particularly when they liberalize. Democracies that open up their economies tend to combine such liberalization policies with more social spending, a larger public sector and other compensation measures and policies that protect the vulnerable from the risks associated with

¹⁵⁷ For a recent literature review see Acemoglu et al. (2015)

¹⁵⁸ Empirical support for the link between democracy and lower inequality is reported in, e.g., Blaydes and Kayser (2011), Reuveny and Li (2003), Rodrik (1999). Statistically insignificant or non-robust results for this relationship are, e.g., reported in Acemoglu et al. (2015), Mulligan, Gil, and Sala-i-Martin (2004), Scheve and Stasavage (2012). Dorsch and Maarek (2018) find an inequality-increasing effect when egalitarian (often communist) autocracies democratize and an inequality-reducing effect of democracy elsewhere.

higher exposure to global markets. Such policies include, for instance, public spending promoting human capital formation as well as social insurance against risks resulting from increased labor market volatilities like unemployment benefits (Avelino, Brown, and Hunter 2005). While democracies do not shy away from liberalizing trade and capital flows because of potentially adverse distributional consequences – if anything they tend to be more liberalized (Milner and Mukherjee 2009) – they usually ‘embed’ such liberalism. Adsera and Boix (2002) find democracy to be the key intervening variable explaining why some countries combine liberalization policies with larger public sectors while others do not (see also Rodrik 1998). Similarly, Rudra and Haggard (2005) find that autocratic governments spend less on welfare when they globalize than their democratic counterparts. Voters appear to prefer external liberalization to be ‘embedded’ in the provision of social insurance against the (distributional) risks resulting from it. Democratic institutions ensure that these preferences translate into policy.

These considerations suggest that democracies tend to pick different policies with distributional implications than what many IMF programs stipulate. Implementing typical IMF conditions concerning the public sector, social spending, compensation-for-liberalization policies and certain labor market policies are likely to mean a departure from a democracy’s policy path in the absence of an IMF program. Hence, an inequality-increasing effect of an IMF program in a democracy relative to the counterfactual of the same democracy not being under a program would mean that IMF programs reduce the democracy’s responsiveness to voter preferences. Such reduced responsiveness can result from rising costs for domestic political actors, e.g., parliaments, to block reforms under a program. The IMF itself states that its conditionality helps “strengthen [...] the hand of reformers” (IMF 2007, 8). In other words, IMF programs *ceteris paribus* can tip the balance within the domestic political process in favor of those who support distributional policies closer to the IMF’s preferences. Similarly, democratic leaders can also use the IMF as a scapegoat and “dilute accountability by blaming IMF conditionality” for reforms that are unpopular with the electorate (Smith and Vreeland 2006). For the empirical test of whether IMF programs increase inequality by limiting democratic responsiveness to voter preferences I propose a hypothesis that exploits the difference in counterfactuals when democratic and non-democratic countries are considered. In the democratic sample, the “non-treated” counterfactual is a democracy with the same characteristics as the program country. The only difference is the exogenously determined

presence of an IMF program that may undermine the democracy's responsiveness to the distributional preferences of its citizens. In the nondemocratic sample, the counterfactual is a political system in which the government's responsiveness to voter preferences is already weak. If the reduction of democratic responsiveness is indeed behind the increase in inequality, then the potentially inequality-increasing effect should mainly be driven by the democratic sample. In non-democracies the democratic tendency towards a more egalitarian distribution of income cannot be undermined. While the discussed equalizing policies and institutions that are typically stronger in democracies can be reversed under IMF programs in democracies, this is not the case in non-democracies. On average, these considerations suggest that IMF programs will increase inequality particularly when the counterfactual is a democracy without IMF intervention.

H2: The inequality-increasing effect of IMF programs is particularly driven by democratic program countries.

4.2.4 Heterogeneous Effects and Channels

To further investigate the 'democratic deficit' as a mechanism, I examine additional heterogeneous effects. Decision-making processes of the IMF have changed over time. Confronted with pressures for legitimization, international organizations have reacted by engaging in efforts to enhance accountability mechanisms with citizens (Grigorescu 2015). If the IMF's relative lack of accountability to affected citizens is behind the inequality-increasing effect of its loan programs, then improvements to its accountability should mitigate this effect. I use two proxies to quantify such variation in IMF accountability. The first is an index developed by Tallberg and co-authors measuring the degree to which non-governmental transnational actors (TNAs) have access to IO decision-making processes (Tallberg et al. 2014). As it has often been argued that "[n]on-governmental organizations can democratize [international governmental organizations] by expanding participation and increasing accountability," (Vabulas 2013, 194) their access is a potential mechanism that could increase the IMF's responsiveness to the preferences of citizens in IMF program countries and thereby its sensitivity to adverse distributional effects.

For the second measure I exploit the difference in IMF lending facilities under which programs are designed. In 1999, the Fund established the Poverty Reduction and Growth Facility

(PRGF).¹⁵⁹ According to the IMF, “foremost among them [the distinctive features of the new facility] is broad public participation and increased national ownership” (IMF 2001). In combination with the explicit focus on poverty reduction, the aims to let country authorities lead the process and involve civil society in the program design are an attempt to avoid extensive interference in domestic political systems and strengthen accountability. As PRGF programs thus aim to override the domestic democratic system to a lesser extent, the theoretical considerations suggest that their inequality-increasing effects should be smaller.¹⁶⁰ To explore the channels underlying the main effects I provide suggestive evidence on the correlation between IMF conditions and subsequent changes in inequality in the Appendix. Not all IMF programs are the same. Research has repeatedly highlighted important differences in the design of conditionality and challenged the claim that the IMF applies identical ‘cookie-cutter’ programs (Stone 2008). To empirically investigate conditionality as a mechanism I run two different tests. The first examines whether the scope of conditionality makes a difference. Given the above discussion a natural expectation is that the inequality-increasing effect increases with the scope of conditionality. The second disaggregates conditions by policy area and tests which of these are likely mechanisms. The analysis above suggests that various typical IMF policy conditions – in particular those addressing liberalization, public spending and the labor market – could be potential channels for distributional effects.

4.3 Method and Data

4.3.1 *Endogeneity*

There is no lack of anecdotal evidence linking IMF programs to rising inequality. Many Latin American, East Asian, and former Soviet countries experienced a divergence in income levels while IMF programs were in place (Klein 2008; Peet 2009; Stiglitz 2002). An illustrative example is the case of Argentina, which was under one of the economically largest and longest IMF programs of all time. Democratic since 1983, Argentina received financial assistance from the Fund for almost the entire 1983–2004 period. Over the course of these two decades the country’s Gini coefficient of net income rose from 38 to 45. Especially during the mass protests

¹⁵⁹ In 2010, the “Extended Credit Facility” replaced the PRGF.

¹⁶⁰ I admit that both measures are imperfect proxies for a phenomenon that is difficult to quantify. It is controversial whether TNAs indeed ‘democratize’ IOs, and PRGF programs exhibit several additional distinctive features. While these results should thus be taken with a grain of salt, it is reassuring that both results, based on two proxies with different shortcomings, point in the same direction.

at the turn of the millennium many blamed this trend, as well as widespread poverty and unemployment, on reforms with origins in IMF conditions implemented by Carlos Menem's government. The IMF had demanded and supported policies such as fiscal austerity that resulted in wage and pension cuts, the privatization of state-owned enterprises leading to mass layoffs, and during the 1998-2002 recession opposed social programs for the poor and government plans such as increasing teachers' salaries (Klein 2008; Paddock 2002; Rodrik 2003). When the program ended after Argentina's last purchase of IMF resources in 2004, inequality started to decline and in 2013 the country's Gini coefficient reached 38 again.

While it is plausible that IMF programs contributed to rising inequality in Argentina, other simultaneous processes may explain this development just as well: The same period was also characterized by years of hyperinflation, economic crises, and high levels of debt – which, in turn, had made continued participation in IMF programs more likely in the first place. It is furthermore not excludable that Menem's government would have implemented similar free-market liberal reforms by itself in the absence of IMF influence and that the trend of decreasing inequality after 2004 is linked to the more egalitarian policies under Néstor and Cristina Kirchner's governments rather than to the end of the IMF programs.

The case of Argentina illustrates that the central challenge for any study investigating the causal effects of IMF programs on economic outcomes is nonrandom selection (Przeworski and Vreeland 2000). The national economic and political conditions that drive selection into IMF programs are likely related to determinants of inequality and other economic and political outcomes. As IMF programs and inequality could thus be correlated in the absence of a causal effect, regression coefficients could be severely biased without a valid identification strategy. Problematically, not all of the potentially confounding variables are observable. While many key variables that explain IMF program participation¹⁶¹ suffer from missing data, the more limiting problem is that many relevant conditions are intrinsically difficult, if not impossible, to measure. Vreeland lists "political will" as an example (Vreeland 2002). Transferred to the focus of this study, this argument suggests that governments that favor IMF programs, e.g., due to a political preference for austerity policies, might also be more likely to implement policies leading to more inequality, irrespective of the presence of an IMF program. The lack of measurement of such variables as "political will" would thus bias the coefficient.

¹⁶¹ For an overview see Steinwand and Stone (2008); Berger, de Haan, and Sturm (2005); Moser and Sturm (2011).

In theory, there is a straightforward solution to this endogeneity problem, but to applied quantitative research on the IMF it presents a difficulty: “Instrumental variables can address this problem, but they are not easy to come by, especially since so much of what drives selection into IMF programs also influences IMF program effects” (Vreeland 2007, 82). So far one strand of this research has, for this reason, either limited itself to correct for selection-on-observables (e.g., Doyle 2012; Hartzell, Hoddie, and Bauer 2010), or additionally controlled for selection-on-unobservables by means of selection models without exclusion restrictions (e.g., Mukherjee and Singer 2010; Nooruddin and Simmons 2006; Przeworski and Vreeland 2000). The former do not control for unobserved confounders while the latter have to make strong assumptions on the joint distribution of the error term and the correct specification of the participation equation.¹⁶²

The other strand of research has incorporated exclusion restrictions in their empirical models (e.g., Atoyan and Conway 2006; Barro and Lee 2005; Dreher and Gassebner 2012). In these studies the share of votes cast in line with the United States or other G7 countries in the UNGA has become the ‘standard instrument’ for IMF programs.¹⁶³ However, as the other IVs used in this literature,¹⁶⁴ this variable is not clearly excludable to macroeconomic outcomes. It rests on the assumption that the only channel through which a country’s UNGA voting behavior is linked to macroeconomic outcomes in the same country is the presence of an IMF program. But it is likely that a government’s foreign policy preferences articulated in UNGA voting are related to a government’s preferences in domestic policy, which in turn are clearly linked to macroeconomic outcomes.¹⁶⁵ To paraphrase Moravcsik (1997), I argue that identification strategies should ‘take preferences seriously;’ especially since the authors of the most widely used UNGA voting data suggest that their data “can be interpreted as states’ positions towards

¹⁶² For further details regarding problems related to selection models without exclusion restrictions see Puhani (2000).

¹⁶³ Barro and Lee (2005) first proposed this IV.

¹⁶⁴ Alternatively, economic variables such as GDP, budget balance, inflation (Biglaiser and DeRouen 2010), growth, reserves (Bauer, Cruz, and Graham 2012), exchange rates (Clements, Gupta, and Nozaki 2013), trade with G5 countries (Barro and Lee 2005), have been used. But the assumption that these country-specific macroeconomic variables do not affect the respective dependent country-specific macroeconomic variable of interest other than through the presence of an IMF program is not plausible as more direct channels within the country’s economy cannot be excluded. The country’s share of IMF staff also used by Barro and Lee is not excludable to the country’s economic and political power, its views on the IMF, and – as IMF staff was found to be mainly recruited from U.S. universities and to lean towards certain political views – to the program country’s ideological and political orientation. A proposed alternative is the number of countries under an IMF program or the number of past IMF program years (Oberdabernig 2013; Atoyan and Conway 2006). However the former is correlated with global economic crises, while the latter captures country-specific characteristics such as weak economic governance.

¹⁶⁵ For a theory along these lines, see Moravcsik (1997); for a recent empirical confirmation, see Mattes et al. (2015).

the U.S.-led liberal order” (Bailey, Strezhnev, and Voeten 2017). The assumption that this political position is unrelated to domestic policies and the state of the domestic economy is not plausible. Hence, a new instrument is needed.¹⁶⁶

4.3.2 Identification Strategy

My identification strategy exploits exogenous variation in the IMF’s liquidity of financial resources. I apply a recent methodological innovation (Nunn and Qian 2014; Werker, Ahmed, and Cohen 2009) and interact this time-variant variable with a country-variant variable indicating the country’s probability of receiving an IMF program. The resulting interaction term varies over time and across countries and, after controlling for the levels of the two variables, introduces exogenous variation to the extent that the isolated interaction effect is excludable from alternative channels. Thus, even if there was endogeneity between the time-variant level variable and the outcome, the exclusion restriction would only be violated if the unobserved variables driving this endogeneity were affecting inequality differently in countries with different levels of the probability to be under an IMF program (for econometric details see Bun and Harrison 2018; Esarey 2015; Nizalova and Murtazashvili 2016). The identifying assumption that such a relationship does not exist resembles the identifying assumption of a *difference-in-differences* design (Dreher and Langlotz 2017).

Specifically, I use the natural logarithm of the IMF’s liquidity ratio (LQR_t) – defined as the amount of liquid IMF resources divided by its liquid liabilities¹⁶⁷ – and interact it with the fraction of years country i has been under an IMF program between 1973 and year t ($IMFprob_{it}$).

$$IV_{it} = \ln(LQR_t) \times IMFprob_{it} = \ln(LQR_t) \times \frac{\sum_{T=1973}^t I(IMFprogram_{iT}=1)}{t - 1973}$$

In the first-stage equation $IMFprogram_{it}$ is regressed on this interaction term and on all second-stage variables. While year fixed effects control for the level effect of the liquidity ratio, I also control for $IMFprob_{it}$ in both stages.¹⁶⁸ The identification can therefore be interpreted as a

¹⁶⁶ Of the existing studies on the IMF’s distributional effects Pastor (1987) conducts before-and-after comparisons, Garuda (2000) controls only for selection-on-observables, Vreeland (2002) addresses selection-on-unobservables without an exclusion restriction, and Oberdabernig (2013) relies on the excludability of UNGA voting. Foster et al. (2018), who cite an earlier version of this paper, apply an IV strategy that builds on the IV introduced in this paper.

¹⁶⁷ For further details on this variable and on all others see Appendix A.

¹⁶⁸ Note that Nunn and Qian (2014) use a time-invariant probability measure in their baseline and a time-variant measure in a robustness test, which produces very similar results. I use the time-variant measure in the baseline as

difference-in-difference approach: After controlling for the levels, the IV's coefficient indicates how the IMF's liquidity affects the likelihood of receiving an IMF program in year t differently in countries with different participation probabilities.

I expect this coefficient to be negative for the following reason: Multiple studies show that "recidivism" is a prime determinant of IMF programs. Countries that have a history of frequently participating in IMF programs are much more likely to do so again. The IMF, thus, has a regular clientele that is routinely supplied (Bird, Hussain, and Joyce 2004). In years, however, in which the Fund has abundant liquid resources, it has the means to extend its clientele. The organizational incentive to do so is an important explanation of the IMF's expansion since the 1970s (Barnett and Finnemore 2004). Arguably, the higher its liquidity, the more generous the Fund will be and the more actively it will look for new clients.¹⁶⁹ In times of high liquidity ratios, the Fund is thus more likely to grant loans to countries that would otherwise be less likely to receive IMF programs. This would be captured by a negative coefficient.

While the instrument's relevance can and will be tested empirically, its excludability is untestable and must be theoretically defended. Figure 4.1 shows the variation of the natural logarithm of LQR over time. The main sources of this variation are the IMF's Quota Reviews.¹⁷⁰ The Articles of Agreement (Article III, Section 2a) require the Board of Governors to review the amount of financial resources members commit to the Fund ("quotas") once every five years. In the observation period these reviews led to quota increases in all but three cases. In Figure 4.1 these jumps can be seen, for instance, in the late 1970s, early 1980s and late 1990s when member countries executed their respective payments of the 7th, 8th, and 11th General Review of Quotas. As the timings of the quota reviews follow the mentioned institutional rule

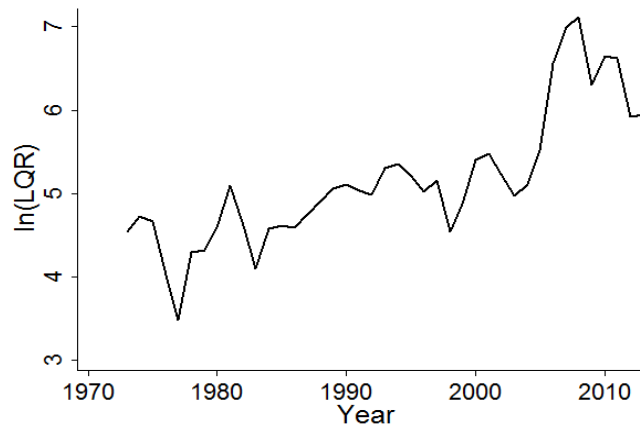
this increases the IV's relevance and avoids using future events to predict present realizations, but show in robustness tests that my results also hold with the time-invariant measure.

¹⁶⁹ This is an argument that my conversations with IMF staff in November 2017 support. Several IMF staff suggested that the IMF's re-designing and re-labelling of its lending facilities can be considered as a reaction to low demand for its programs and an attempt to make them more attractive for potential program countries. They supported the view that there is an incentive for the bureaucracy to ensure the IMF's global relevance by promoting participation in its loan programs, and that this incentive is particularly strong when the IMF's financial resources are relatively little used, which a high liquidity ratio indicates.

¹⁷⁰ A second source of variation is the fact that in some years repayments of individual, extraordinarily large loans affect liquidity liabilities. (In addition, minor changes in liquid liabilities can result from changes in IMF borrowing.) In the robustness section I show that this is unproblematic and that the results hold when this variation is excluded. A last source of variation is the fact that liquid resources additionally vary when the IMF adjusts the basket of currencies it considers "usable." The usability status, however, is highly stable over time, changes mostly for small economies and therefore has a very minor effect on the amount of liquid resources.

and are thus predetermined, it is very unlikely that they are linked to intra-state income inequality through unobserved channels. Even if this was the case, the correlation between such unobserved variables and the outcome would bias the result only if it was dependent on a country's probability of participating in IMF programs. In other words, a sceptic would have to find unobserved variables that affect the impact of the IMF's liquidity ratio on income inequality conditional on how regularly the country has received IMF programs in the past – after controlling for country and year fixed effects and a large vector of control variables. It is unlikely that such variables exist.¹⁷¹

Figure 4.1 – The IMF's Liquidity Ratio



4.3.3 Econometric Model and Data

Armed with this excludable instrument, I estimate 2SLS dynamic panel regressions to identify the causal effect of IMF programs on income inequality:

1st stage:

$$IMFprogram_{it-1} = \alpha \ln(LQR_{t-1}) \times IMFprob_{it-1} + \delta_1 IMFprob_{it-1} + \eta_1 Gini_{it-1} + \pi_1 X'_{it-1} + \xi_i + \varrho_t + u_{it}$$

2nd stage:

$$Gini_{it} = \beta \widehat{IMFprogram}_{it-1} + \delta_2 IMFprob_{it-1} + \eta_2 Gini_{it-1} + \pi_2 X'_{it-1} + \xi_i + \varrho_t + \varepsilon_{it}$$

The annual time-series cross-sectional data cover the 1973-2013 period and a maximum of 155 countries. As not all data are available for all countries and years, the panel is unbalanced and the number of observations depends on the explanatory variables used.

¹⁷¹ In addition, several robustness tests designed to challenge the identifying assumption fail to produce different results. See below and Appendix 4.G.

The dependent variable *Gini* is the Gini-coefficient of net income taken from the Standardized World Income Inequality Database (SWIID) (Solt 2016). The SWIID combines source data from multiple inequality databases and, in contrast to other panel datasets like All The Ginis (ATG) or the World Income Inequality Database, standardizes them to make the data comparable across countries and over time. Because of this standardization and its comprehensive coverage the SWIID is widely used in related research based on panel data (Acemoglu et al. 2015; Dorsch and Maarek 2018; Oberdabernig 2013). I follow this literature in the choice of this database, but also run robustness tests with ATG data.

The explanatory variable of interest, *IMFprogram*, is an indicator that equals 1 if country *i* was under an IMF program for at least five months in year *t* (Dreher 2006b, updated). In the baseline I follow the related literature on the effects of IMF programs and lag the variable by one year (Nooruddin and Simmons 2006). To look at longer-term effects I introduce different lags in additional regressions. It may take time for the effect to operate because of lagged consequences of economic reforms and it is relevant to see whether and when potential changes in inequality are undone.

Furthermore, to account for unobserved country-specific characteristics and time-specific trends, I include country and year fixed effects (ξ and ρ). As current levels of inequality are heavily dependent on previous levels it is standard to also include the lagged dependent variable (LDV).¹⁷² In addition, I include a lagged vector of covariates consisting of two variable sets.¹⁷³ The first comprises the standard covariates of inequality: *GDP per Capita* and its square to control for the country's level of economic development including a potential non-linear relationship à la Kuznets (Kuznets 1955) as well as *Education* measured by average years of schooling, *Trade (% GDP)*, *Life Expectancy* and *Regime Type*. The second set of covariates includes variables that the literature identified as key determinants of IMF programs: *Current Account Balance (% GDP)*, *Investments (% GDP)*, *GDP Growth*, *UNGA Voting*, and an indicator variable for the presence of a systemic *Banking Crisis*. I additionally interact the global total of both *Banking Crises* and *Global GDP Growth* with *IMFprob* to enhance the exclusion restriction's plausibility by demonstrating that global economic developments do not influence both the

¹⁷² Acemoglu et al. 2015. As in all regressions $T > 20$, a potential Nickell bias (Nickell 1981) is negligible, see Beck and Katz (2009). A Fisher-type augmented Dickey-Fuller unit-root test rejects the hypothesis that *Gini* has a unit root.

¹⁷³ For descriptive statistics, definitions and sources see Appendix 4.A.

IMF's liquidity ratio and inequality differently in countries with different levels of IMF participation probabilities.¹⁷⁴

4.4 Results

4.4.1 First-Stage Estimates

I begin by testing the relevance of the instrument. Table 4.1 shows the first-stage estimates of the 2SLS regressions with different sets of control variables, whose coefficients are omitted to reduce clutter (see Appendix 4.B for the full table). The results demonstrate that the instrument is relevant. They show a highly significant, negative conditional correlation between the IV and the presence of an IMF program. The Kleibergen-Paap LM statistic rejects the null hypothesis that the equation is underidentified at the 0.1 percent level. The cluster-robust Kleibergen-Paap F-statistics easily surpass conventional levels of weak identification tests, such as Staiger and Stock's (1997) threshold of ten as well as Stock and Yogo's (2005) most conservative critical value of 16.38, which tolerates a maximum 2SLS size distortion of 10 percent.

Table 4.1 – First Stage Regressions

	(1)	(2)	(3)
LQR × IMFprob	-0.276*** (0.052)	-0.308*** (0.063)	-0.356*** (0.067)
IMFprob	2.760*** (0.282)	2.673*** (0.315)	3.172*** (0.286)
Inequality controls	No	Yes	Yes
IMF controls	No	No	Yes
Observations	3766	3010	2625
Kleibergen-Paap (K.-P.) underidentification test LM-statistic	18.452	16.045	18.973
K.-P. underidentification test p-value	0.000	0.000	0.000
K.-P. weak identification test F-statistic	27.699	24.121	28.441

Notes: Dependent variable *IMFprogram*. All regressions include country and year fixed effects and the lagged dependent variable. Standard errors, robust to clustering at the country level, in parentheses.

Significance levels: * p<.10, ** p<.05, *** p<.01

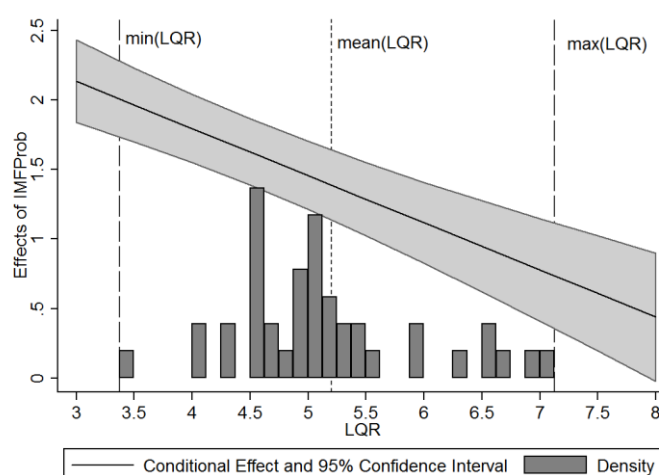
These results are robust across different specifications. In column 1, only the levels of the interaction term, the LDV (*Gini*_{*t*-1}), as well as country and year fixed effects are controlled for. Under the assumption that the IV is excludable conditional on these variables, this

¹⁷⁴ This is analogous to Nunn and Qian's (2014: 1643-5) choice of covariates.

specification without additional control variables already yields an unbiased coefficient of interest in the second stage. Nevertheless, in columns 2 and 3 I successively add the two sets of covariates described above. None of them significantly alters the relevant coefficient, its significance, the F-statistic or the underidentification test statistics.

As previous studies found, countries that participated in IMF programs in the past are significantly more likely to participate in such programs in the present (“recidivism”) (Bird, Hussain, and Joyce 2004). A new finding, however, is that this effect depends on the Fund’s liquidity. As the negative sign of the interaction term’s coefficient indicates and Figure 4.2 illustrates, in years with higher liquidity ratios the probability of past IMF participation is a weaker, but still significant, predictor of IMF programs. The Fund is, thus, not only more generous in years with higher liquidity ratios¹⁷⁵ but in these years it also implements more programs for countries beyond its more regular clientele. In sum, the IV is plausibly excludable to inequality levels in specific countries, proves to be highly relevant, and allows an intuitive interpretation of its linkage with the presence of IMF programs.

Figure 4.2 – Visualized Effect of the IV



4.4.2 Main Results

Table 4.2 presents the results of the second stage of the 2SLS regressions. Specifications 1-3 correspond to those reported in Table 4.1.¹⁷⁶ In line with H1, the results show that IMF programs significantly increase income inequality. Across all specifications the coefficient is

¹⁷⁵ The liquidity ratio – which is not included in the regressions because of perfect multicollinearity with year fixed effects – is positively correlated with the number of countries under IMF programs in a given year ($r = .3$).

¹⁷⁶ See Appendix 4.D for the full table and an interpretation of all other coefficients.

statistically significant at least at the 5 percent level (1 percent in specification 2) and substantial in size. Having an IMF program in year t on average increases the country's Gini coefficient of net income in year $t+1$ by a little more than one point. This result is robust both to different sets of control variables and to different samples, which vary because of missing data for the added controls.

To assess the magnitude of the effect, note that it is equivalent to an increase in the Gini coefficient by at least 34 percent and up to 51 percent of a within-country standard deviation. As inequality is slow to change, increases of this size within one year are rare events (8.6 percent of all observations in the sample). While this indicates a substantial effect size, differences in the Gini coefficient are difficult to interpret directly. Therefore, I elaborate on a method proposed by Blackburn (1989) to quantify the size of the effect in a more intuitive way. According to Blackburn's metric, an increase in the Gini by 1.1 (1.3) points is equivalent to a lump-sum transfer of 2.2 (2.6) percent of the country's mean income from the bottom half to the upper half. To view this from the perspective of an individual belonging to a country's poorer half, consider that in the sample's average country those below the median earn approximately 25 percent of the total national income. Hence, on average the change in inequality induced by having an IMF program in the previous year is equivalent to a transfer of four to five percent of the poorer half's mean income by each person in the poorer half to each person in the richer half (see Appendix 4.B).

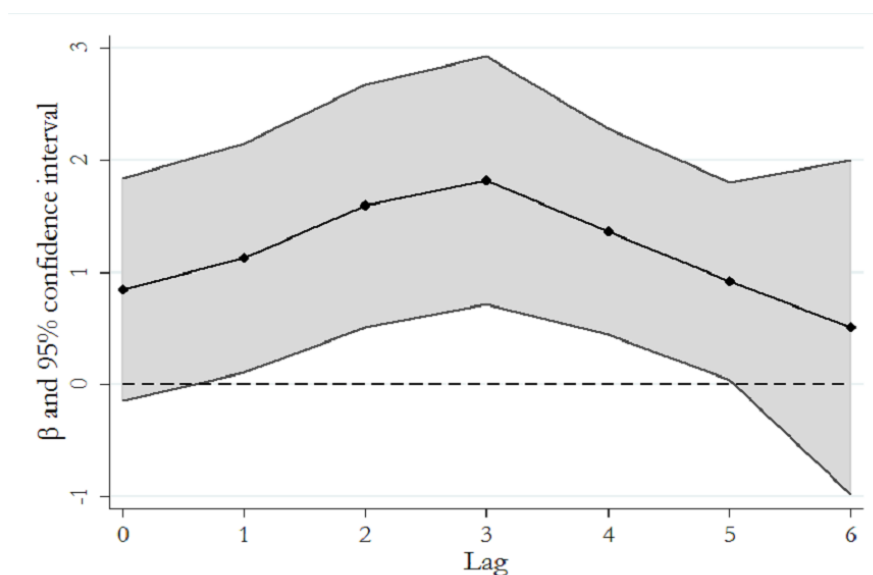
Table 4.2 – Main Results

	(1)	(2)	(3)
IMF Program _{t-1}	1.130** (0.521)	1.435*** (0.550)	1.288** (0.573)
Inequality controls (t-1)	No	Yes	Yes
IMF controls (t-1)	No	No	Yes
Observations	3766	3010	2625
Adjusted R ²	0.880	0.853	0.858
Notes: Dependent variable <i>Gini</i> . Second-stage regressions corresponding to Table 1. All regressions include country and year fixed effects, the lagged dependent variable, and <i>IMFprob</i> . Standard errors, robust to clustering at the country level, in parentheses. Significance levels: * p<.10, ** p<.05, *** p<.01			

Next, I test how IMF programs affect inequality in the longer term. Figure 4.3 illustrates and reports the estimates of the coefficient of interest for the baseline specification (1) with different levels of lags. It indicates that the effect is statistically significant during all of the following five years and strongest and most significant after three years. After six years the effect is no

longer significantly different from zero. Results are very similar when adding the control variables (see Appendix 4.E).

Figure 4.3 – Long Term Effects



	t	t-1	t-2	t-3	t-4	t-5	t-6
IMF program	0.847*	1.130**	1.593***	1.816***	1.363***	0.920**	0.511
	(0.506)	(0.521)	(0.552)	(0.564)	(0.468)	(0.450)	(0.758)
N	3766	3766	3726	3685	3643	3598	3556

Note: Reported are β -coefficients and standard errors for different lags of the variable *IMFprogram*. Specifications otherwise identical to specification 1 in Tables 1 and 2.

4.4.3 Heterogeneous Effects

To shed light on the underlying mechanisms I split the sample into democracies and non-democracies.¹⁷⁷ Note first the descriptive statistics in Table 4.3. They show the average of *Gini* depending on whether the observation is a democracy and on whether an IMF program was in place and reports *t*-tests comparing the respective means. As expected, inequality is significantly higher in non-democracies and in countries under an IMF program. It is furthermore interesting and in line with the hypotheses that the large and highly significant difference in inequality between democracies and non-democracies entirely disappears when only countries under IMF programs are compared.

¹⁷⁷ The definition of democracy follows the Polity IV index and treats observations with a Polity score of 6 and higher as democracies (Marshall, Jaggers, and Gurr 2011).

Table 4.3 – Conditional Means of the Gini Index

E(Gini)	not under IMF program	under IMF program	t-tests
Democracies	34.52	41.22	$t = 14.37; p < 0.001$
Non-democracies	39.61	41.05	$t = 3.28; p = 0.001$
t-tests	$t = 12.95; p < 0.001$	$t = -0.39; p = 0.68$	

As these descriptive statistics are obviously inadequate to isolate the IMF's causal effect, Table 4.4 presents the 2SLS regression results with the sample split into democracies and non-democracies to test H2. In columns 1-2 and 5-6 the sample is split on both stages, in columns 3-4 and 7-8 the fitted values of the variable of interest calculated by means of the entire sample are used.¹⁷⁸

Columns 1-4 show that IMF programs increase inequality in democracies.¹⁷⁹ The effect is robust to whether or not control variables are included and whether fitted values from the full or only the democratic sample are used. The instrument maintains its relevance despite the smaller sample size in columns 1 and 2. The point estimates, which are statistically significant across all specifications, range from 1.8 to 2.3 and are, thus, larger compared to the full sample. Applying the metric based on Blackburn, this is equivalent to a transfer of about eight percent of the average poor person's income to the average rich person. As soon as only non-democracies are considered (columns 5-8) the effect entirely disappears. The coefficients are close to zero and not statistically significant at conventional levels.¹⁸⁰

¹⁷⁸ The latter is a valid strategy to the extent that there is no systematic difference of the IV's effect on *IMFprogram* between democracies and non-democracies. Theoretically, there is no obvious reason why this should be the case. Empirically, the first-stage regressions for the split samples show that the coefficients of the IV are similar in both samples and only in column 5 do they not reach statistical significance at the 10%-level. This suggests that splitting the sample only on the second stage is also valid. Standard errors in these regressions are cluster bootstrapped to account for two-stage estimation.

¹⁷⁹ In accordance with the results for long-term effects reported in Table 4.3, in this and in the following table, *IMFprogram* is lagged by three years. The substance of the results, however, does not depend on this choice.

¹⁸⁰ Note that in column 5, the specification without control variables and with the sample split on both stages, the IV in the small non-democratic sample is not strong enough to reliably rule out weak instrument bias. Columns 6-8, however, show that the coefficient of interest remains insignificant when the IV's relevance is increased by adding control variables or by using fitted values from the full sample. In column 6 the F-statistic surpasses Stock and Yogo's (2005) critical value of 6.66 that tolerates a 2SLS size distortion of 20%.

Table 4.4 – Sample Split

	Democracies				Non-Democracies			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IMF Program _{t-3}	1.901** (0.739)	2.063** (0.810)	2.271*** (0.861)	1.829** (0.824)	-0.057 (2.260)	-0.363 (0.787)	-0.031 (1.462)	-0.407 (1.141)
controls	No	Yes	No	Yes	No	Yes	No	Yes
sample split	1 st & 2 nd stage	1 st & 2 nd stage	2 nd stage	2 nd stage	1 st & 2 nd stage	1 st & 2 nd stage	2 nd stage	2 nd stage
N	2094	1743	3766/2094	2526/1743	1317	878	3766/1317	2526/878
First Stage:								
IV	-0.315*** (0.077)	-0.367*** (0.088)	-0.276*** (0.052)	-0.356*** (0.067)	-0.142 (0.125)	-0.464*** (0.168)	-0.276*** (0.052)	-0.356*** (0.067)
K.-P. underid. p	0.001	0.001	0.000	0.000	0.256	0.026	0.000	0.000
K.-P. weak id. F	16.958	17.545	27.699	24.121	1.286	7.647	27.699	24.121

Note: Dependent variable *Gini*. All regressions include country and year fixed effects, the lagged dependent variable, and *IMFprob*. In columns 1-2 and 5-6 standard errors (in parentheses) are robust to clustering at the country level; in the remaining regressions standard errors are cluster bootstrapped. Significance levels: * p<.10, ** p<.05, *** p<.01

In sum, the inequality-increasing effect of IMF programs seems to be entirely driven by the democratic sample. In line with the hypotheses, IMF programs appear to weaken the inclination of democratic governments to more egalitarian income distributions. The size of the effect is equivalent to cutting the average difference of approximately four Gini points between democracies and autocracies in half.

Further examining the plausibility of this channel, I include the interaction $IMF \times TNA$ ($= IMF_{program} \times TNA_{access}$) as an additional regressor (Table 4.5 columns 1-2). To estimate its coefficient I employ the IV estimator proposed by Bun and Harrison (2018).¹⁸¹ As expected, the interaction enters with a significantly negative coefficient. This supports the expectation that the ‘opening up’ of the IMF towards societal actors makes the organization more sensitive to the distributional effects of its activities.

¹⁸¹ Bun and Harrison’s (2018) “IV3” estimator adds the IV multiplied by *TNAaccess* as well as $IMF \times TNA$ to the set of instruments for *IMFprogram* while treating $IMF \times TNA$ as exogenous. This identification is valid under the plausible assumptions that TNA access to the IMF is exogenous to inequality levels and that the degree of endogeneity of IMF programs and inequality does not depend on TNA access: $E(IMF \times Gini \mid TNA) = E(IMF \times Gini)$.

Table 4.5 – TNA Access and PRGF Programs

	(1)	(2)	(3)	(4)	(5)	(6)
IMF Program _{t-3}	0.717** (0.285)	0.985** (0.472)				
IMF×TNA _{t-3}	-2.799*** (1.075)	-3.824** (1.688)				
PRGF program _{t-3}			0.490 (0.961)	0.483 (1.120)		
Non-PRGF program _{t-3}					1.841*** (0.698)	2.089** (0.893)
Controls	No	Yes	No	Yes	No	Yes
N	2094	1743	2094	1743	2094	1743
Adj.R ²	0.838	0.829	0.841	0.834	0.782	0.759
K.-P. underid. LM	37.968	21.039	5.625	5.204	10.827	8.506
K.-P. underid. p	0.000	0.000	0.018	0.023	0.001	0.004
K.-P. weak id. F	77.962	32.011	14.900	14.443	19.953	12.307

Notes: See Table 2. Only democracies are considered. For columns 1-2 see text and footnote 119.

Columns 3-6 separately examine the effects of PRGF programs and other IMF programs in democracies and show a substantial difference. The coefficient for PRGF programs is small and not statistically significant at conventional levels. If all other types of IMF programs are considered, however, the effect on inequality is significantly positive. As discussed above, the IMF's emphasis on public participation and national ownership in its PRGF programs could (but does not necessarily have to) explain why these kinds of lending arrangements have no significant adverse distributional effects. In sum, these empirical tests suggest that variation in the decision-making processes concerning the design of IMF programs explain variation in the programs' effects on inequality. The more inclusive and 'democratized' these processes are, the smaller the adverse distributional consequences are.

In Appendix 4.F, I extend the core analysis by providing correlational evidence on the role that IMF conditionality plays for the link between IMF programs and increasing inequality. I find that inequality increases more in IMF program countries that receive more extensive conditionality than in program countries that receive fewer conditions. When examining specific policy areas, the results are consistent with the theoretical considerations on the 'public spending' and 'labor market reforms' channels discussed above: Conditionality addressing the social and pension sector as well as the labor market are associated with increasing inequality. I find, however, no evidence for the conjecture that IMF conditions targeting trade or capital account policies are associated to increasing inequality. While these results, which are discussed in more detail in the Appendix, do not provide causal evidence, they are consistent with the idea that conditionality is a plausible channel for the main effect

that this study identifies. Nevertheless, more research that focuses on specific IMF conditions as a channel for the distributional effects of IMF programs is needed.

4.4.4 Robustness

I run a series of additional tests to confirm the robustness of the main results.¹⁸² First, I address concerns regarding the exclusion restriction. The results are robust to excluding the country-year observations with large purchases and repurchases of IMF credit as well as to using only the IMF's liquid resources as the time-varying component of the IV (Table 4.11). Motivated by a recent finding by Christian and Barrett (2017), I then plot time trends in *LQR* and the cross-country average of *Gini* for sets of countries with different levels of *IMFprob* and find no evidence for time trends that could threaten the exclusion restriction. I also randomize the order of the *LQR* values for 1000 placebo first-stage regressions. The resulting IV coefficients are, as expected, normally distributed around zero and have all smaller *t*-statistics than the IV coefficient in the first-stage regression with the real *LQR* values. Furthermore, substituting the time-varying probability (*IMFprob*) by a time-constant probability that is absorbed by country fixed effects also does not affect the results (Table 4.12). Tests à la Altonji, Elder, and Taber (2005) show that selection-on-unobservables relative to selection-on-observables would have to be more than three times as large and go in the opposite direction if the true reduced form effect was in fact zero (Table 4.13). The same table also reports OLS and reduced form estimates of the baseline specification for comparison.

Second, to make the regressions comparable to previous studies I substitute my IV with UNGA voting (Table 4.14, columns 1-3). The results are similar. However, the UNGA voting IV is less relevant, with F-statistics below critical thresholds in two out of three specifications. Besides, the coefficients of interest are larger and the effect size that UNGA voting as IV identifies is somewhat doubtful (140 percent of a within-country standard deviation). Under the assumption that this study's IV is excludable, this finding and the fact that UNGA voting enters with a significantly positive sign as a control in the baseline (see Full Table 4.2 in Appendix 4.D) suggest that UNGA voting is linked to inequality through more channels than only IMF programs. When used as an IV this violates the exclusion restriction and biases the coefficient upwards.

¹⁸² For a detailed description of these tests see Appendix 4.G.

Third, I modify the main variables of interest. While my baseline definition of IMF programs follows the literature's standard, in a robustness test I use Barro and Lee's (2005) definition and only consider Stand-By Arrangements (SBA) and the Extended Fund Facility (EFF). The results are similar (Table 4.14, columns 4-6). When using the Gini coefficient of market income as an alternative outcome variable, the regressions again yield very similar results, suggesting that IMF programs tend to redistribute net income more via changes in gross income than by means of changes in taxes and transfers (Table 4.15, columns 1-3).¹⁸³ Additionally, I employ ATG data as an alternative to the SWIID (Table 4.15, columns 4-6). Even though the use of this dataset dramatically reduces the sample size, the results are again robust.

4.5 Conclusions

IMF loan programs increase income inequality. According to the evidence presented in this paper, this effect is causal and economically significant. The effect's heterogeneities – most importantly the finding that the effect is driven by the sample of democratic countries – are consistent with the theory that 'democratically deficient' international organizations that interfere in domestic politics are capable of temporarily restricting the responsiveness of democratic governments to the distributional preferences of their citizens.

For the IMF the results highlight an unintended consequence of its lending arrangements. It may encourage the Fund to revise its policy advice and conditionality with regards to their distributional implications. The findings furthermore suggest that giving greater autonomy to program countries' national political processes, as is the case in PRGF programs, and allowing societal actors to have more influence on decision-making processes, can mitigate adverse distributional effects. They thus support recent internal IMF reforms aimed at "country ownership" and "participatory processes" (IMF 2014b).

While future research is needed to examine whether these findings are transferable to other IOs, the study's main result suggest that the activities of at least one of the most powerful IOs can not only affect public goods provision and general welfare but can also have a distributional dimension. From a normative perspective, such distributional power points to the need for more effective accountability mechanisms between IOs and affected societies to ensure that the allocation of gains and losses is under democratic control. And apart from

¹⁸³ See the Appendix for caveats regarding this inference.

being normatively desirable, the analyses suggest that enhanced democratic accountability mechanisms in IOs can also produce better policy results.

Future research both on IMF program effects and on the drivers of income inequality can draw lessons from this study. First, the proposed identification strategy can be useful for scholars investigating the causal effects of IMF programs more broadly. The instrumental variable plausibly fulfills the exclusion restriction with regards to a range of additional political and economic outcomes. Second, the results add to the growing literature that stresses the role of policies and institutions as determinants of inequality. While their contribution to current trends of rising inequality across countries is well-established in this literature, it remains an open question as to why so many countries modify their national policies and institutions in a way that inequality increases. This study's results suggest that *inter*-national policies and institutions can play a significant role in this regard. Examining the underlying mechanisms linked to how international organizations interact with democratic processes at the national level is a promising area for future research.

4.6 Appendices to Chapter 4

4.6.1 Appendix 4.A: Variables

Table 4.6 – Descriptive Statistics and Data Sources

Variable	Mean	SD	Min	Max	Description and Source
<i>Gini</i>	38.08	9.23	17.96	68.16	Gini coefficient of net income according to the SWIID version 5.0 (Solt 2016).
<i>IMF Program</i>	0.32	0.47	0	1	Indicator 1 if IMF program in place for at least 5 months in year t , (Dreher 2006).
<i>LQR (ln)</i>	5.42	0.75	4.1	7.11	IMF liquidity ratio, which equals liquid resources (usable currencies plus Special Drawing Rights contributed) divided by liquid liabilities (total of members' reserve tranche positions plus outstanding IMF borrowing from members); own calculation based on data from the IMF's Annual Reports 1973-2013 and the IMF's International Financial Statistics
<i>IMFprob</i>	0.25	0.25	0	1	$\frac{\sum_{t=1973}^t I(IMFprogram_{it} = 1)}{t-1973}$ Own calculation based on (Dreher 2006b).
<i>GDP per capita (ln)</i>	8.21	1.61	4.92	11.38	Gross domestic product per capita in constant 2005 USD (World Bank 2016)
<i>Education</i>	7.56	2.85	0.89	13.18	Average years of schooling, linear interpolation of data for five-year periods (Barro and Lee 2013)
<i>Trade</i>	77.09	51.26	12.01	439.66	Trade (% GDP) (World Bank 2016)
<i>Life Expectancy</i>	68.71	9.61	27.08	82.93	Life expectancy at birth in years (World Bank 2016b)
<i>Democracy</i>	0.66	0.47	0	1	Indicator 1 if Polity IV index is 6 or higher (Marshall, Jaggers, and Gurr 2011)
<i>Current Account Balance</i>	-2.1	6.53	-47.21	26.77	Balance on current account (% GDP) (IMF 2016d).
<i>Investments</i>	23.11	7.14	-2.42	74.82	Gross capital formation (% of GDP) (World Bank 2016).
<i>GDP growth</i>	3.63	4.45	-50.25	35.22	GDP growth (annual %) (World Bank 2016).
<i>Banking Crisis</i>	0.11	0.31	0	1	Indicator 1 if systemic banking crisis in year t in country i , (Laeven and Valencia 2012).
<i>UNGA voting</i>	0.14	0.91	-2.14	3.01	Ideal point of voting behavior in the UNGA (Bailey, Strezhnev, and Voeten 2017).
<i>Global GDP Growth</i>	3.03	1.51	-2.08	6.98	Growth of global GDP; own calculations based on World Bank (2016).

<i>Banking Crises</i>	14.54	10.12	0	30	Global total of <i>Banking Crisis</i> in year t , based on Laeven and Valencia (2012)
<i>TNA index</i>	0.14	0.14	0	0.32	Index of TNA access to the IMF in year t (Tallberg et al. 2014)
<i>PRGF program</i>	0.15	0.36	0	1	Indicator 1 if IMF program under the PRGF in place for at least 5 months in year t (Dreher 2006)
<i>Liquid Resources (ln)</i>	11.3	0.66	9.84	12.96	IMF liquid resources (see <i>LQR</i>)
<i>Gross Gini</i>	45.36	7.12	20.25	71.13	Gini coefficient of market income according to the SWIID version 5.0 (Solt 2016)
<i>Gini (ATG)</i>	39.63	9.91	20	69.8	Gini coefficient (Gini _{all}) according to the ATG Dataset (Milanovic 2014)

Note: The sample of the full specification (Table 4.2, column 3) was used for calculating the values in this table.

Conditionality

<i>Scope of Conditionality</i>	5.46	2.26	0	9	Number of policy areas covered by IMF Conditionality
<i>Foreign Exchange Systems</i>	0.24	0.43	0	1	IMF condition addressing foreign exchange systems and restrictions (current and capital)
<i>Trade Policy</i>	0.44	0.50	0	1	IMF condition addressing international trade policy
<i>Central Bank</i>	0.13	0.33	0	1	IMF condition addressing the central bank
<i>Financial Sector</i>	0.78	0.42	0	1	IMF condition addressing the financial sector
<i>Government</i>	0.84	0.36	0	1	IMF condition addressing the general government
<i>Public Employment</i>	0.08	0.28	0	1	IMF condition addressing the civil service, public employment and wages
<i>Pensions and Social Sector</i>	0.10	0.30	0	1	IMF condition addressing pensions and other social sector reforms
<i>SOE reform</i>	0.77	0.42	0	1	IMF condition addressing reforms of public enterprises in the non-financial sector
<i>Labor Market</i>	0.03	0.17	0	1	IMF condition addressing labor market reforms in the private sector
<i>Residual Category</i>	0.63	0.48	0	1	IMF condition addressing other structural reforms

Note: The sample of the specifications 1 and 3 in Table 4.10 was used for calculating the values in this table. Source: Andone and Scheubel (2017) based on IMF (2016c)

4.6.2 Appendix 4.B: Interpreting Differences in Gini Coefficients

Following Blackburn (1989), a change in the Gini coefficient ($G \in [0, 100]$) by ΔG points is equivalent to a lump-sum transfer of L from all those below the median to all those above the median, given by

$$L = \frac{2\Delta G}{100} \times M, \text{ where } M \text{ is the country's mean income.}$$

Knowing M and the poorer half's share of total income S , the mean income of the poorer half P is given by

$$(P \times 0.5) + (P \times \frac{1-S}{S} \times 0.5) = M$$
$$P = 2MS$$

The lump-sum transfer relative to the poorer half's mean income is, hence, given by:

$$\frac{L}{P} = \frac{\Delta G}{100} \times \frac{1}{S}$$

4.6.3 Appendix 4.C: Full Table 4.1

Table 4.7 – Baseline, First Stage

	(1)	(2)	(3)
LQR × IMFprob	-0.276*** (0.052)	-0.308*** (0.063)	-0.356*** (0.067)
IMFprob	2.760*** (0.282)	2.673*** (0.315)	3.172*** (0.286)
Gini	0.003 (0.003)	-0.001 (0.004)	-0.003 (0.005)
GDP/Capita (ln)		-0.177 (0.280)	-0.135 (0.331)
GDP/Capita Squared (ln)		-0.002 (0.017)	-0.007 (0.020)
Education		-0.057** (0.025)	-0.058** (0.028)
Trade		-0.000 (0.001)	-0.001 (0.001)
Life Expectancy		0.008 (0.005)	0.010* (0.006)
Regime Type		-0.005 (0.049)	0.004 (0.053)
Current Account Balance			0.002 (0.003)
Investments			-0.006** (0.003)
GDP Growth			0.002 (0.002)
Banking Crisis			0.080** (0.038)
UNGA Voting			0.102*** (0.034)
Global GDP Growth × IMFprob			0.002 (0.027)
Banking Crises × IMFprob			0.006 (0.004)
Observations	3766	3010	2625
K.-P. underid. LM	18.452	16.045	18.973
K.-P. underid. p	0.000	0.000	0.000
K.-P. weak id. F	27.699	24.121	28.441

Notes: Dependent variable *IMFprogram*. All regressions include country fixed effects and year fixed effects. Standard errors, robust to clustering at the country level, in parentheses.

Significance levels: * p<.10, ** p<.05, *** p<.01

4.6.4 Appendix 4.D: Full Table 4.2 and Discussion of Control Variable Results

Table 4.8 – Baseline, Second Stage

	(1)	(2)	(3)
IMF Program _{t-1}	1.130** (0.521)	1.435*** (0.550)	1.288** (0.573)
IMFprob _{t-1}	-1.844** (0.841)	-1.901** (0.881)	-2.591** (1.087)
Gini _{t-1}	0.916*** (0.011)	0.915*** (0.013)	0.911*** (0.014)
GDP/Capita (ln) _{t-1}		2.224** (0.932)	2.728*** (0.868)
GDP/Capita squared (ln) _{t-1}		-0.077 (0.051)	-0.100** (0.050)
Education _{t-1}		-0.066 (0.079)	-0.056 (0.091)
Trade _{t-1}		-0.001 (0.002)	0.001 (0.003)
Life Expectancy _{t-1}		-0.025 (0.019)	-0.016 (0.023)
Regime Type _{t-1}		0.020 (0.110)	-0.044 (0.127)
Current Account Balance _{t-1}			0.003 (0.009)
Investments _{t-1}			0.013 (0.009)
GDP growth _{t-1}			-0.018** (0.008)
Banking Crisis _{t-1}			-0.235* (0.134)
UNGA Voting _{t-1}			0.231* (0.132)
Global GDP Growth × IMFprob _{t-1}			0.122** (0.050)
Banking Crises × IMFprob _{t-1}			-0.003 (0.012)
Observations	3766	3010	2625
Adjusted R ²	0.880	0.853	0.858

Notes: Dependent variable *Gini*. Second-stage regressions corresponding to Table 4.1. All regressions include country and year fixed effects. Standard errors, robust to clustering at the country level, in parentheses. Significance levels: * p<.10, ** p<.05, *** p<.01

As regards the coefficients of the control variables, the lagged dependent variable is, unsurprisingly, highly significant as inequality is a highly time-persistent phenomenon. The coefficient on *IMFprob* is not straightforward to interpret: It shows that – once the positive causal effect of a current IMF program is netted out – a more frequent past participation rate

in IMF programs is weakly associated with a lower Gini.¹⁸⁴ Note, however, that this coefficient cannot be interpreted in isolation. This variable only captures the variation in inequality that the predicted values of *IMFprogram*, which themselves include possibly endogenous variation of *IMFprob*, do not already capture. The sole purpose of controlling for *IMFprob* is to make sure that this possibly endogenous part of the variation in predicted values is controlled for and netted out.

GDP per capita is associated with higher inequality levels, while there is some weak evidence for the Kuznets curve hypothesis: Albeit consistently negative, the coefficient on the squared term is only significant in the third regression. As in previous studies, education is associated negatively with income inequality, even though the effect is not statistically significant. Systemic banking crises are also associated with decreasing inequality. As capital is usually distributed more unequally than income, the reduction of income from capital during such crises could explain this finding (see Piketty 2014).

¹⁸⁴ An IMF program with average length (3 years) would increase *IMFprob* in the average sample year by 0.136 and would thus be linked to a Gini coefficient that is lower by 0.259 points (column 2).

4.6.5 Appendix 4.E: Full Table on Long-Term Effects

Table 4.9 – Long-term Effects

Lag:	(1)	(2)	(3)
t	0.847* (0.506) [3766]	1.417*** (0.543) [3010]	1.525** (0.604) [2625]
t-1	1.130** (0.521) [3766]	1.435*** (0.550) [3010]	1.288** (0.572) [2625]
t-2	1.593*** (0.552) [3726]	1.633*** (0.561) [2977]	1.330** (0.702) [2625]
t-3	1.816*** (0.564) [3685]	1.773*** (0.554) [2943]	1.303** (0.664) [2625]
t-4	1.363*** (0.468) [3643]	1.683*** (0.543) [2909]	1.234** (0.632) [2625]
t-5	0.920** (0.450) [3598]	1.401** (0.605) [2871]	0.962** (0.675) [2625]
t-6	0.511 (0.758) [3556]	0.914 (0.911) [2834]	0.547 (0.907) [2625]

Note: The table reports β -coefficients for different lags of the variable *IMFprogram* in specifications (1)-(3), which are otherwise identical to the regressions in Tables 4.1 and 4.2. Standard errors in parentheses; number of observations in square brackets.

4.6.6 Appendix 4.F: IMF Conditionality

As an extension to the paper's core analysis I examine suggestive evidence on the role of IMF conditionality in the link between IMF programs and increasing inequality. I use data extracted from the IMF's Monitoring of Fund Arrangements (MONA) database with an algorithm developed by Andone and Scheubel (2017) in order to create an annualized and harmonized dataset from both the archived (1993-2002) and the current (2002-2013) MONA data.¹⁸⁵ First, I code the variable *Scope of Conditionality* defined as the number of policy areas conditionality covers.¹⁸⁶ Second, I code binary variables indicating whether any condition addressed one of nine policy areas.¹⁸⁷ For the analysis, I restrict the sample to country-years for which the MONA database indicates the start of an IMF program. Informed by the results of the main analysis I then regress the change in *Gini* over the subsequent three-year-period

¹⁸⁵ I thank Beatrice Scheubel for sharing these data with me.

¹⁸⁶ This approach follows Dreher, Sturm, and Vreeland (2015).

¹⁸⁷ See Appendix 4.A for a description of these policy areas.

on the conditionality variables at the time of the IMF program start. This sample restriction follows the approach by Rickard and Caraway (2018) to circumvent the selection-into-program problem, but allows inferences only for countries under IMF programs and provides correlational evidence only. Like Rickard and Caraway (2018) I was unable to find a relevant and excludable instrument for IMF conditions. To nevertheless mitigate the selection-into-conditions problem, I add the same set of control variables as before. The regressions, whose results are presented in Table 4.10, are run for both the full MONA sample (columns 1 and 3) as well as for the sample excluding the “archived” (1993-2002) version of the dataset, which Andone and Scheubel (2017) consider less reliable (columns 2 and 4).

In both samples the results show that inequality increases more in IMF program countries that receive more extensive conditionality than in program countries that receive fewer conditions. When examining specific policy areas, it becomes apparent that conditions targeting the labor market or the social and pension sector are associated with rising inequality. In program countries in which IMF conditions address the labor market the Gini rises by almost three points more than in countries whose programs do not cover this policy area. In observations in which conditionality addresses the social and pension sector income inequality in the subsequent three-year period rises, on average, by about two Gini points more than otherwise. While they provide no causal evidence, these results are consistent with the idea that conditionality is a plausible channel for the main effect. They are also consistent with the theoretical considerations on ‘public spending’ and ‘labor market reforms’ discussed above. Regarding the ‘liberalization’ channel, however, conditions addressing trade policy or the financial sector are not significantly associated with rising inequality. Interestingly, there is a statistically significant, negative coefficient on *Public Employment*. Given that such conditions usually demand cuts in public sector wages (Rickard and Caraway 2018), this suggests that such reforms tend to be equalizing because public sector wages are above the median wage in typical IMF program countries.

While these results tend to support the theoretical argument, a word of caution regarding their interpretation is in order. First, the data on conditionality is limited to a much shorter time period than the data used for the main analysis. Second, its structure does not allow a direct test of all channels discussed above, as the disaggregation by policy areas in the MONA database is not quite in line with the scholarly literature’s theoretical considerations on determinants of inequality. While social spending and labor market reforms can be captured,

the effects of more general spending cuts and capital account liberalization cannot be isolated. Third, the information that is included only provides the policy area and not the exact content of the condition. It does neither cover its stringency nor the extent of compliance. Fourth, while restricting the sample to IMF program countries circumvents the selection-into-program problem, potential endogeneity bias resulting from selection-into-conditions cannot be ruled out. For these reasons this evidence should be considered as suggestive and correlational rather than as definitive and causal. While this study's focus is on causally identifying the aggregate effect, future research should zero in on the underlying channels.¹⁸⁸

Table 4.10 – Conditionality

	(1)	(2)	(3)	(4)
Scope of Conditionality	0.163** (0.073)	0.441** (0.206)		
Foreign Exchange Systems			0.519 (0.426)	0.799 (0.648)
Trade Policy			0.133 (0.381)	0.568 (0.808)
Central Bank			0.622 (0.625)	0.305 (0.755)
Financial Sector			0.459 (0.460)	0.813 (1.925)
Government			0.128 (0.838)	0.822 (3.580)
Public Employment			-1.032 (0.661)	-1.594** (0.677)
Pensions and Social Sector			1.727*** (0.577)	1.953** (0.762)
SOE reform			0.022 (0.519)	-0.707 (0.711)
Labor Market			2.614*** (0.761)	2.897*** (0.943)
Residual Category			-0.089 (0.371)	0.843 (0.747)
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Period	1993-2013	2002-2013	1993-2013	2002-2013
Observations	273	71	273	71
R-squared	0.218	0.289	0.262	0.500

Note: OLS Regressions. Unit of observation: county-year with a start of an IMF arrangement. Dependent variable: change in Gini in subsequent three-year-period. Standard errors robust to clustering at the country level in parentheses. Significance levels: * p<.10, ** p<.05, *** p<.01

¹⁸⁸ In ongoing work with Alexander Kentikelenis we work on identifying a source of exogenous variation in IMF conditions.

4.6.7 Appendix 4.G: Robustness

This section describes the robustness tests summarized in the results section in more detail.

First, I address concerns regarding the exclusion restriction. Some readers might worry that the denominator of the liquidity ratio, i.e., the amount of the Fund's liquid liabilities, threatens the excludability of the instrument. While most variation in the liquidity ratio is induced by the changing amount of liquid resources, to a significantly lesser extent it also depends on the liquid liabilities.¹⁸⁹ These vary when economically large members obtain and repay loans that are large relative to total IMF resources ("purchase" and "repurchase" in IMF jargon).¹⁹⁰ In Figure 4.1 this is visible, for instance, in the mid-2000s when Brazil and Turkey repaid extraordinarily large loans. In general, I argue that this does not undermine the excludability of the IV: First, the vast majority of these flows are not sizable enough to significantly affect the liquidity ratio. As in most cases the amount of resources transferred is significantly less than 1 percent of total IMF quotas, any concern regarding excludability would relate to very few observations. Second, the timing of such transactions is agreed upon years in advance. Given also that explanatory variables are lagged, it is unlikely that the schedule of large transactions developed with economically large countries is correlated with future levels of inequality in specific countries. Third, even if there was a correlation it would have to be conditional on *IMFprob* because of the difference-in-differences style model the interacted IV estimates.

Nevertheless, to be cautious I run a robustness test in which I exclude the 100 observations that exhibit the largest flows from and to the IMF.¹⁹¹ As the first three columns in Table 4.11 show, the results do not differ substantially. To address these concerns in the most cautious way possible, I also run regressions using only liquid resources as the time-variant factor of the IV. This variable is, by construction, not determined by the Fund's liquid liabilities. By refraining from dividing the variable by liquid liabilities, I only exploit variation in liquid resources, whose only substantial source of variation is the exogenous timing of quota reviews. These results are presented in the last three columns of Table 4.11. While the instrument's

¹⁸⁹ The logged liquidity ratio's correlation with logged liquid resources is $r = 0.83$, while with logged liquid liabilities it is $r = 0.23$.

¹⁹⁰ The liquid liabilities' second source of variation is the Fund's borrowing from its members. While total borrowing by the Fund is zero in many years, its average share of the liquid liabilities is approximately 15%.

¹⁹¹ This leaves only observations with a (re)purchase to total quota ratio of less than 0.57% (0.37%) in the sample. Regressions with 50 and 200 excluded observations produce virtually the same results.

relevance naturally decreases because some valuable variation is lost, it is still strong enough to confirm the robustness of the result to this alternative specification.

In addition, I examine whether there are general trends of inequality in sets of countries with different levels of *IMFprob* that could be correlated with the IMF's liquidity ratio. Christian and Barrett (2017) show that the findings by Nunn and Qian (2014) could be driven by a spurious correlation between the time-varying constituent term of their interacted IV and a particular time trend in their outcome variable for a set of countries with a specific level of their probability measure. This is why in Figure 4.4, I plot year-specific cross-country averages of *Gini* for countries with different levels of *IMFprob* over time. I find no evidence for trends that could threaten the exclusion restriction. Instead, the *Gini* trends seem to be parallel across these groups and substantially different as compared to the *LQR* time series. As Christian and Barrett (2017) show, a problem in Nunn and Qian (2014) arises from the fact that the time series of the time-varying constituent term of their interacted IV is remarkably similar to a simple (inverse-U shaped) trend and does not vary strongly from one period to the next. As *LQR* exhibits no obvious similarity to any such simple trend and is subject to several idiosyncratic shocks (see Figure 4.1), it is much less likely to be correlated with a similar trend in the outcome variable.

In addition, and to increase the confidence that the first stage does not pick up an artefact, I run placebo regressions in which I randomize the values of *LQR*. I run 1000 iterations of such regressions, which are based on a randomized order of the actual values of *LQR*, and find that the resulting first-stage IV coefficients are normally distributed around zero, and that the coefficient's *t*-statistics are all smaller than in the first-stage regression based on the actual values of *LQR*. This increases confidence in the mechanism driving the first-stage and suggests that it is unlikely that in the first stage the IV picks up an artefact.

Another modification concerns the second factor of the interacted instrument (Table 4.12). Like Nunn and Qian (2014) I also report results employing an IV based on a country-specific probability that does not vary over time, substituting *IMFprob_{it}* by *IMFprob_const_i*, which is given by

$$IMFprob_const_i = \frac{\sum_{t=1973}^{2013} I(IMFprogram_{it} = 1)}{41}$$

I thereby make the probability multicollinear with the country fixed effects. While I am more convinced by the time-varying probability because it avoids using future realizations to explain the present, the results are robust to this modification.

In the next table I report OLS and reduced form estimates (Table 4.13). First, I run OLS and OLS-fixed effect (FE) models (columns 1-2) and then calculate the OLS estimates for the baseline model, i.e., I do not instrument for IMF programs, *ceteris paribus* (columns 3-5). As the results show, IMF programs are correlated with higher inequality in OLS and OLS-FE regressions without control variables but there is no correlation when endogeneity is only insufficiently addressed in OLS-FE models with different sets of control variables. Together with the statistically significant effect found in the 2SLS regressions these results suggest that the proposed IV is able to eliminate the selection bias the OLS coefficients suffer from. In columns 6-8 I report the results of reduced form regressions of the baseline specifications. They show that the IV has a statistically significant effect on inequality. This relationship is not significantly affected when a large vector of control variables is added to the regression. Following Altonji, Elder, and Taber (2005) this enhances the plausibility of the exclusion restriction: The comparison of the β -coefficients of the models with and without these covariates (6 vs. 8) shows that the so-called “selection ratio” is 3.12. This means that if the effect, which I claim is causal, was in reality driven by unobserved variables, this selection on unobservables would have to be more than three times as large as the selection on observed variables, and it would have to go in the opposite direction.

To compare the results to studies using the current standard instrument for IMF programs, I substitute the IV with UNGA voting behavior *ceteris paribus* (Table 4.14, columns 1-3). These regressions estimate IMF programs to cause rises in inequality of approximately four Gini points, comparable to Oberdabernig (2013), who uses the same IV. Considering that the estimated coefficients are equivalent to a change of up to 140 percent of a within-country standard deviation, this effect is strikingly large. One reason why these coefficients may be biased is that the instrument is not relevant enough; in specifications 2 and 3 the Kleibergen-Paap F-statistics fall below Stock and Yogo’s (2005) lowest critical value of 5.53 that tolerates a 2SLS size distortion of 25 percent. A second reason could be that the instrument is not excludable. As argued above, plausible alternative channels are governments’ political and ideological preferences. Under the assumption that my IV strategy identifies the true causal effect of IMF programs, the baseline regressions reported in Table 4.2 (see also Appendix 4.D) provide empirical evidence for the violation of the exclusion restriction of UNGA voting: In the full baseline specification voting similarity with the United States in the UNGA is associated with higher levels of inequality when controlling for the causal effect of IMF

programs. This finding suggests that UNGA voting is linked positively to inequality through channels other than IMF programs and is, thus, an invalid instrumental variable when the outcome of interest is inequality.¹⁹²

As a last step, I modify the main variables of interest. Regarding the independent variable, the paper so far followed the conventional practice of the literature on IMF program effects by jointly considering Stand-By Arrangements (SBA), the Extended Fund Facility (EFF), the Structural Adjustment Facility (SAF) and the Poverty Reduction and Growth Facility (PRGF) (e.g., Oberdabernig 2013). Barro and Lee (2005, 1248), however, argue that only SBA and EFF programs should be considered while the others “should be viewed more as foreign aid, rather than lending or adjustment programs”. In Table 4.14 (columns 4-6) I follow their approach and find that the results hold when only SBA and EFF programs are considered.

Regarding the dependent variable, I first substitute the Gini coefficient of net income by that of gross income (*Gross Gini*), which is also taken from the SWIID. The fact that the results are very similar, could indicate that IMF programs affect inequality mainly by leading to changes in the distribution of wages in contrast to affecting the extent of redistribution. This could, for instance, be driven by labor market reforms such as minimum wage reductions, cuts in pensions or by rising short-term unemployment after privatizations. An important caveat of these findings, however, is that the *differences* between market and net inequality that the SWIID indicates are not reliable for all countries (Solt 2016, 1274-5). Future research could investigate the exact channels in more detail. As a final robustness test, I change the inequality dataset. Until here I followed the related literature (Acemoglu et al. 2015; Dorsch and Maarek 2018; Oberdabernig 2013) in choosing the SWIID as the source for panel data on Gini coefficients. Jenkins (2015) however, voices concerns about the SWIID’s methodology and recommends the World Income Inequality Database (WIID), on which the SWIID builds, over the SWIID.¹⁹³ The WIID, however, offers multiple Gini coefficients for many country-year observations. Since there is no commonly accepted procedure for choosing the respective values, the use of the WIID for regression analyses necessitates highly arbitrary decisions. This is presumably also why the SWIID is used much more frequently than the WIID. An

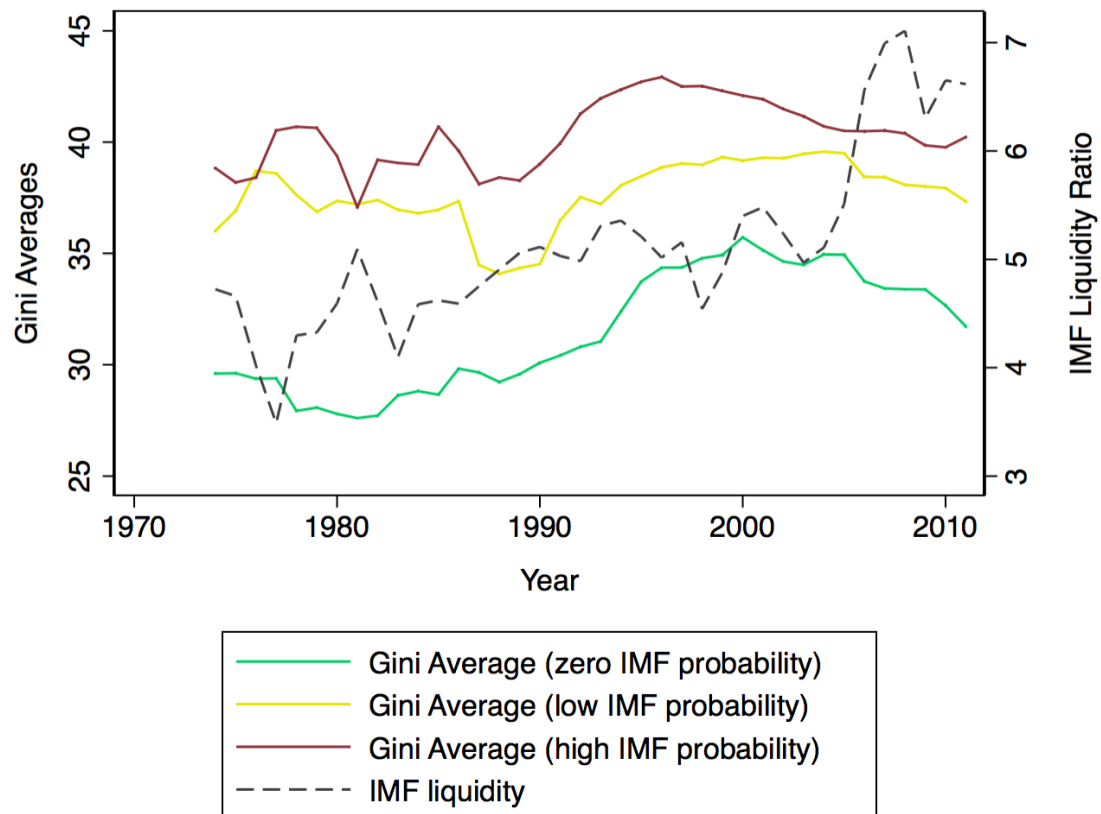
¹⁹² As inequality is clearly linked to other economic conditions, analyses of IMF program effects on other economic outcomes are likely to suffer from the same problem when UNGA voting is used as an IV.

¹⁹³ Jenkins’ (2015) concerns, however, relate to an older version of the SWIID and Solt (2015) is able to overcome many of these concerns. The reader is referred to the entire special issue of the *Journal of Economic Inequality* (December 2015, Volume 13, Issue 4) for details on this debate.

alternative is offered by Milanovic (2014), who derives the final Gini value if multiple observations exist through “choice by precedence.” While this approach makes sure that in each case the observation of the highest possible quality is chosen, it combines data from nine different sources with different methodologies without further standardization. Milanovic himself advises caution when using the resulting variable *Gini_{all}* in regressions as the concepts underlying the calculation of the Gini coefficients are based on income and consumption, net and gross, as well as household and individual levels. Unfortunately, too few observations remain if the sample is restricted to one concept. Nevertheless, to address this issue I control for dummy variables that indicate the respective concepts interacted with country fixed effects. Columns 4-6 in Table 4.15 report the results. Note that, compared to the baseline, the sample size is severely limited. Nevertheless, the coefficient of interest is positive and statistically significant in the specifications that include control variables.

I conclude that the results are robust to these modifications.

Figure 4.4 – Spurious Correlations Between Inequality and IMF Liquidity?



Note: The figure plots the year-specific cross-country averages of the Gini index of net income for three sets of countries with different levels of *IMFprob*. For the bottom line all countries that never received an IMF program are included for calculating these averages. For the middle line all countries whose *IMFprob* is in the bottom half of this variable's distribution are included. For the top line all countries whose *IMFprob* is in the upper half of this variable's distribution are included. The dashed line plots the IMF liquidity time-series for comparison.

Table 4.11 – Robustness 1: Challenging the Liquidity Variable

	Excluding large (re)purchases			IV with liquid resources		
	(1)	(2)	(3)	(4)	(5)	(6)
2nd stage results						
IMF Program _{t-1}	1.244** (0.546)	1.727*** (0.610)	1.403** (0.568)	1.172* (0.668)	2.037** (0.970)	1.351** (0.615)
1st stage results						
IV _{t-1}	-0.277*** (0.053)	-0.299*** (0.064)	-0.357*** (0.069)	-0.168*** (0.043)	-0.177*** (0.056)	-0.387*** (0.087)
K.-P. underid. LM	17.480	14.320	17.423	12.652	9.325	18.762
K.-P. underid. p	0.000	0.000	0.000	0.000	0.002	0.000
K.-P. weak id. F	27.240	21.671	26.727	15.599	10.058	19.818
Inequality Controls (t-1)	No	Yes	Yes	No	Yes	Yes
IMF Controls (t-1)	No	No	Yes	No	No	Yes
N	3654	2901	2536	3766	3010	2625
Adjusted R ²	0.878	0.840	0.853	0.879	0.823	0.855

Note: Dependent variable *Gini*. All regressions control for *IMFprob*, country fixed effects and year fixed effects as well as the lagged dependent variable. Standard errors, robust to clustering at the country level, are in parentheses. Significance levels: * p<.10, ** p<.05, *** p<.01

Table 4.12 – Robustness 2: Challenging the Probability Variable

	(1)	(2)	(3)
2nd stage results			
IMF Program _{t-1}	1.901* (1.145)	1.691** (0.746)	1.634** (0.788)
1st stage results			
IMFprob_const x LQR _{t-1}	-0.173** (0.071)	-0.271*** (0.072)	-0.300*** (0.075)
K.-P. underid. LM	4.877	10.832	12.933
K.-P. underid. p	0.027	0.001	0.000
K.-P. weak id. F	5.876	14.241	15.906
Inequality Controls	No	Yes	Yes
IMF Controls	No	No	Yes
N	3766	3010	2625
Adjusted R ²	0.851	0.838	0.841

Note: Dependent variable *Gini*. All regressions control for country fixed effects and year fixed effects as well as the lagged dependent variable. Note that *IMFprob_const* does not need to be controlled for because it is fully absorbed by country fixed effects. Standard errors, robust to clustering at the country level, are in parentheses. Significance levels: * p<.10, ** p<.05, *** p<.01

Table 4.13 – Robustness 3: Selection on Unobservables

	OLS	OLS-FE	OLS (Baseline)			Reduced Form (Baseline)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2nd stage results								
IMF Program _{t-1}	5.113*** (0.903)	0.651** (0.270)	0.016 (0.071)	0.082 (0.073)	0.122 (0.076)			
1st stage results								
IV _{t-1}						-0.312** (0.142)	-0.442*** (0.168)	-0.459** (0.201)
Selection Ratio $\beta_8 / (\beta_8 - \beta_6)$						3.12		
Country & Year fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
LDV & IMFprob _{t-1}	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Inequality Controls (t-1)	No	No	No	Yes	Yes	No	Yes	Yes
IMF Controls (t-1)	No	No	No	No	Yes	No	No	Yes
N	3963	3963	3768	3012	2627	3768	3012	2627
Adjusted R ²	0.057	0.120	0.898	0.885	0.884	0.898	0.885	0.884

Note: Dependent variable *Gini*. Standard errors, robust to clustering at the country level, are in parentheses. Significance levels: * p<.10, ** p<.05, *** p<.01

Table 4.14 – Robustness 4: Alternative IV, Alternative Treatment

	UNGA voting as IV			Only SBA & EFF Programs		
	(1)	(2)	(3)	(4)	(5)	(6)
2nd stage results						
IMF	4.644***	4.258**	4.259*			
Program _{t-1}	(1.773)	(2.002)	(2.247)			
SBA/EFF				0.762***	0.805***	0.722*
Program _{t-1}				(0.286)	(0.312)	(0.389)
1st stage results						
UNGA	0.061***	0.073**	0.075**			
voting _{t-1}	(0.023)	(0.033)	(0.034)			
IV _{t-1}				-0.558***	-0.576***	-0.531***
				(0.059)	(0.068)	(0.071)
K.-P.						
underid. LM	6.084	2.546	3.005	18.574	19.987	19.963
K.-P.						
underid. p	0.014	0.111	0.083	0.000	0.000	0.000
K.-P.						
weak id. F	7.139	4.999	4.961	89.310	72.677	56.272
Inequality						
Controls (t-1)	No	Yes	Yes	No	Yes	Yes
IMF						
Controls (t-1)	No	No	Yes	No	No	Yes
N	3520	3001	2682	3766	3010	2625
Adjusted R ²	0.626	0.635	0.620	0.890	0.875	0.874

Note: Dependent variable *Gini*. All regressions include country fixed effects and year fixed effects as well as the lagged dependent variable. In columns 4-6 only SBA and EFF programs are used to calculate the variable *IMFprob*, which the regressions also control for. Standard errors, robust to clustering at the country level, are in parentheses. Significance levels: * p<.10, ** p<.05, *** p<.01

Table 4.15 – Robustness 5: Alternative Inequality Data

	Gross Gini (SWIID)			ATG Data		
	(1)	(2)	(3)	(4)	(5)	(6)
2nd stage results						
IMF Program _{t-1}	1.666*** (0.561)	1.544*** (0.579)	1.222** (0.575)	1.220 (1.155)	1.974** (0.927)	2.043** (1.014)
1st stage results						
IV _{t-1}	-0.276*** (0.053)	-0.314*** (0.063)	-0.363*** (0.067)	-0.557*** (0.113)	-0.699*** (0.130)	-0.624*** (0.135)
K.-P. underid. LM	18.804	16.732	19.608	12.093	12.857	10.483
K.-P. underid. p	0.000	0.000	0.000	0.001	0.000	0.001
K.-P. weak id. F	27.637	24.864	29.173	24.112	28.779	21.447
Inequality Controls (t-1)	No	Yes	Yes	No	Yes	Yes
IMF Controls (t-1)	No	No	Yes	No	No	Yes
ATG Controls (t-1)	No	No	No	Yes	Yes	Yes
N	3765	3009	2624	928	814	758
Adjusted R ²	0.870	0.862	0.859	0.493	0.512	0.491

Note: Dependent variables *Gross Gini* (columns 1-3) and *Gini_{all}* (columns 4-6). All regressions control for *IMFprob*, country fixed effects and year fixed effects as well as the lagged dependent variable. Standard errors, robust to clustering at the country level, are in parentheses.

Significance levels: * p<.10, ** p<.05, *** p<.01

References

- Abbott, Kenneth W., and Duncan Snidal. 1998. "Why States Act through Formal International Organizations." *Journal of Conflict Resolution* 42(1): 3–32.
- Acemoglu, Daron, Suresh Naidu, Pascual Restrepo, and James Robinson. 2015. "Democracy, Redistribution, and Inequality." In *Handbook of Income Distribution*, eds. Anthony B. Atkinson and Francois Bourguignon. Oxford: North-Holland, 1885–1966.
- Acemoglu, Daron, and James Robinson. 2005. *Economic Origins of Dictatorships and Democracy*. Cambridge: Cambridge University Press.
- Adsera, Alicia, and Carles Boix. 2002. "Trade, Democracy, and the Size of the Public Sector: The Political Underpinnings of Openness." *International Organization* 56(2): 229–62.
- Afonso, António. 2003. "Understanding the Determinants of Sovereign Debt Ratings: Evidence for the Two Leading Agencies." *Journal of Economics and Finance* 27(1): 56–74.
- Afonso, António, Davide Furceri, and Pedro Gomes. 2012. "Sovereign Credit Ratings and Financial Markets Linkages: Application to European Data." *Journal of International Money and Finance* 31(3): 606–38.
- Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113(1): 151–84.
- Andersen, Thomas Barnebeck, Thomas Harr, and Finn Tarp. 2006. "On US Politics and IMF Lending." *European Economic Review* 50(7): 1843–62.
- Andone, Irina, and Beatrice Scheubel. 2017. "Memorable Encounters? Own and Neighbours' Experience with IMF Conditionality and IMF Stigma." *CESifo Working Paper* 6399.
- Angrist, Joshua D., and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.
- . 2010. "The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics." *NBER Working Paper* 15794.
- Archer, Candace C., Glen Biglaiser, and Karl DeRouen. 2007. "Sovereign Bonds and the 'Democratic Advantage': Does Regime Type Affect Credit Rating Agency Ratings in the Developing World?" *International Organization* 61(2): 341–65.
- Atoyan, Ruben, and Patrick Conway. 2006. "Evaluating the Impact of IMF Programs: A Comparison of Matching and Instrumental-Variable Estimators." *Review of International*

- Organizations* 1: 99–124.
- Autor, David, Alan Manning, and Christopher Smith. 2016. “The Role of the Minimum Wage in the Evolution of U.S. Wage Inequality over Three Decades: A Modest Re-Assessment.” *American Economic Journal: Applied Economics* 8(1): 58–99.
- Avelino, George, David S. Brown, and Wendy Hunter. 2005. “The Effects of Capital Mobility, Trade Openness, and Democracy on Social Spending in Latin America, 1980-1999.” *American Journal of Political Science* 49(3): 625–41.
- Babb, Sarah L. 2009. *Behind the Development Banks: Washington Politics, World Poverty, and the Wealth of Nations*. Chicago: University of Chicago Press.
- Bailey, Michael A., Anton Strezhnev, and Erik Voeten. 2017. “Estimating Dynamic State Preferences from United Nations Voting Data.” *Journal of Conflict Resolution* 61(2): 430–456.
- Baker, James. 1995. *The Politics of Diplomacy: Revolution, War and Peace: 1989–1992*. Putnam.
- Bandow, Doug. 1992. “Avoiding War.” *Foreign Policy* 89: 156–74.
- Bank for International Settlements. 2016. “Consolidated Banking Statistics.”
- Barnett, Michael, and Martha Finnemore. 1999. “The Politics, Power, and Pathologies of International Organizations.” *International Organization* 53(4): 699–732.
- — —. 2004. *Rules for the World*. Cornell: Cornell University Press.
- Barro, Robert J., and Jong Wha Lee. 2005. “IMF Programs: Who Is Chosen and What Are the Effects?” *Journal of Monetary Economics* 52: 1245–69.
- Bartels, Bernhard, and Beatrice Weder di Mauro. 2013. “A Rating Agency for Europe - A Good Idea?” *Centre for Economic Policy Research Working Paper*.
- Bas, Muhammet, and Randall W. Stone. 2014. “Adverse Selection and Growth under IMF Programs.” *Review of International Organizations* 9(1): 1–28.
- Bauer, Molly, Cesi Cruz, and Benjamin Graham. 2012. “Democracies Only: When Do IMF Agreements Serve as a Seal of Approval?” *Review of International Organizations* 7(1): 33–58.
- Beaudry, Paul, and Tim Willems. 2018. “On the Macroeconomic Consequences of Over-Optimism.” *unpublished manuscript*.
- Beck, Nathaniel, and Jonathan N. Katz. 2009. “Modelling Dynamics in Time-Series Cross-Section Political Economy Data.” *Social Science Working Paper 1304, California Institute of Technology*.

- Beck, Thorsten et al. 2001. "New Tools in Comparative Political Economy: The Database of Political Institutions." *World Bank Economic Review* 15(1): 165–76.
- Becker, Raphael N., Arye L. Hillman, Niklas Potrafke, and Alexander H. Schwemmer. 2014. "The Preoccupation of the United Nations with Israel: Evidence and Theory." *Review of International Organizations* 10(4): 413–37.
- Ben-Artzi, Ruth. 2005. "Political Aid: Is Multilateral Lending Strategic?" *unpublished manuscript*.
- Berensmann, Kathrin, Florence Dafe, and Ulrich Volz. 2015. "Developing Local Currency Bond Markets for Long-Term Development Financing in Sub-Saharan Africa." *Oxford Review of Economic Policy* 31(3–4): 350–78.
- Berg, Andrew et al. 2014. "Assessing Bias and Accuracy in the World Bank- IMF's Debt Sustainability Framework for Low-Income Countries." *IMF Working Paper* 14/48.
- Berger, Helge, Jakob de Haan, and Jan Egbert Sturm. 2005. "Which Variables Explain Decisions on IMF Credit? An Extreme Bounds Analysis." *Economics and Politics* 17(2): 177–213.
- Bergh, Andreas, and Therese Nilsson. 2010. "Do Liberalization and Globalization Increase Income Inequality?" *European Journal of Political Economy* 26(4): 488–505.
- Biglaiser, Glen, and Karl DeRouen. 2010. "The Effects of IMF Programs on U.S. Foreign Direct Investment in the Developing World." *Review of International Organizations* 5(1): 73–95.
- Bird, Graham. 1996. "The International Monetary Fund and Developing Countries: A Review of the Evidence and Policy Options." *International Organization* 50: 477–511.
- Bird, Graham, Mumtaz Hussain, and Joseph P. Joyce. 2004. "Many Happy Returns? Recidivism and the IMF." *Journal of International Money and Finance* 23(2): 231–51.
- Bird, Graham, and Dane Rowlands. 2002. "Do IMF Programmes Have a Catalytic Effect on Other International Capital Flows." *Oxford Development Studies* 30(3): 229–49.
- — —. 2009. "The IMF's Role in Mobilizing Private Capital Flows: Are There Grounds for Catalytic Conversion?" *Applied Economics Letters* 16(17): 1705–8.
- Bird, Graham, and Thomas D. Willett. 2004. "IMF Conditionality, Implementation and the New Political Economy of Ownership." *Comparative Economic Studies* 46(3): 423–50.
- Blackburn, McKinley. 1989. "Interpreting the Magnitude of Changes in Measures of Income Inequality." *Journal of Econometrics* 42(1): 21–25.
- Blanchard, Olivier J., and Daniel Leigh. 2013. "Growth Forecast Errors and Fiscal Multipliers." *American Economic Review* 103(3): 117–20.

- Blanton, Robert, Shannon Lindsey Blanton, and Dursun Peksen. 2015. "The Impact of IMF and World Bank Programs on Labor Rights." *Political Research Quarterly* 68(2): 324–36.
- Blaydes, Lisa, and Mark A. Kayser. 2011. "Counting Calories: Democracy and Distribution in the Developing World." *International Studies Quarterly* 55(4): 887–908.
- Boix, Carles. 2001. "Democracy, Development, and the Public Sector." *American Journal of Political Science* 45(1): 1–17.
- . 2003. *Democracy and Redistribution*. Cambridge: Cambridge University Press.
- Borensztein, Eduardo, Kevin Cowan, and Patricio Valenzuela. 2013. "Sovereign Ceilings 'Lite'? The Impact of Sovereign Ratings on Corporate Ratings." *Journal of Banking and Finance* 37(11): 4014–24.
- Broz, J. Lawrence. 2005. "Congressional Politics of International Financial Resources." *American Journal of Political Science* 49(3): 479–96.
- . 2008. "Congressional Voting on Funding the International Financial Institutions." *Review of International Organizations* 3(4): 351–74.
- . 2011. "The United States Congress and IMF Financing, 1944-2009." *Review of International Organizations* 6(3): 341–68.
- Broz, J. Lawrence, and Michael Brewster Hawes. 2006. "Congressional Politics of Financing the International Monetary Fund." *International Organization* 60(2): 367–99.
- Bueno de Mesquita, Bruce, and A. Smith. 2010. "The Pernicious Consequences of UN Security Council Membership." *Journal of Conflict Resolution* 54(5): 667–86.
- Bueno de Mesquita, Bruce, Alastair Smith, Randolph Siverson, and James D. Morrow. 2003. *The Logic of Political Survival*. Cambridge: MIT Press.
- Bumann, Silke, and Robert Lensink. 2016. "Capital Account Liberalization and Income Inequality." *Journal of International Money and Finance* 61: 143–62.
- Bun, Maurice, and Teresa Harrison. 2018. "OLS and IV Estimation of Regression Models Including Endogenous Interaction Terms." *Econometric Reviews*.
- Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82(6): 2295–2326.
- Cameron, David R. 1978. "The Expansion of the Public Economy: A Comparative Analysis." *American Political Science Review* 72(4): 1243–61.
- Cantor, Richard, and Frank Packer. 1996. "Determinants and Impact of Sovereign Credit

- Ratings." *Economic Policy Review* 2(2): 37–53.
- Caraway, Teri, Stephanie J. Rickard, and Mark Anner. 2012. "International Negotiations and Domestic Politics: The Case of IMF Labor Market Conditionality." *International Organization* 66(1): 27–61.
- Carter, David B., and Randall W. Stone. 2015. "Democracy and Multilateralism: The Case of Vote Buying in the UN General Assembly." *International Organization* 69(1): 1–33.
- Casper, Brett A. 2017. "IMF Programs and the Risk of a Coup d'etat." *Journal of Conflict Resolution* 61(5): 964–96.
- Cattaneo, Matias D, Michael Jansson, and Xinwei Ma. 2016a. "Rddensity: Manipulation Testing Based on Density Discontinuity." *The Stata Journal*: 1–18.
- . 2016b. "Simple Local Regression Distribution Estimators with an Application to Manipulation Testing." *unpublished manuscript*.
- Chapman, Terrence, Songying Fang, Xin Li, and Randall W. Stone. 2015. "Mixed Signals: IMF Lending and Capital Markets." *British Journal of Political Science* 47: 329–49.
- Chapman, Terrence L., and Dan Reiter. 2004. "The United Nations Security Council and the Rally 'Round the Flag Effect." *Journal of Conflict Resolution* 84(6): 886–909.
- Cho, Hye Jee. 2014. "Impact of IMF Programs on Perceived Creditworthiness of Emerging Market Countries: Is There a 'Nixon-Goes-to-China' Effect?" *International Studies Quarterly* 58(2): 308–21.
- Christian, Paul, and Christopher B. Barrett. 2017. Revisiting the effect of food aid on conflict: A methodological caution. *World Bank Policy Research Working Paper* 8171.
- Chwieroth, Jeffrey M. 2007a. "Neoliberal Economists and Capital Account Liberalization in Emerging Markets." *International Organization* 61(02): 443–63.
- . 2007b. "Testing and Measuring the Role of Ideas: The Case of Neoliberalism in the International Monetary Fund." *International Studies Quarterly* 51: 5–30.
- . 2010. *Capital Ideas: The IMF and the Rise of Financial Liberalization*. Princeton: Princeton University Press.
- . 2013. "'The Silent Revolution': How the Staff Exercise Informal Governance over IMF Lending." *Review of International Organizations* 8(2): 265–90.
- CIA. 2018. "CIA World Factbook." <https://www.cia.gov/library/publications/resources/the-world-factbook/> (May 10, 2018).
- Clements, Benedict, Sanjeev Gupta, and Masahiro Nozaki. 2013. "What Happens to Social

- Spending in IMF-Supported Programmes?" *Applied Economics* 48: 4022–33.
- Congressional Research Service. 2002. "U.S. Assistance to the Former Soviet Union 1991-2001: A History of Administration and Congressional Action." *CRS Report for Congress*.
- Copelovitch, Mark S. 2010a. "Master or Servant? Common Agency and the Political Economy of IMF Lending." *International Studies Quarterly* 54: 49–77.
- — —. 2010b. *The International Monetary Fund in the Global Economy: Banks, Bonds, and Bailouts*. Cambridge: Cambridge University Press.
- Corsetti, Giancarlo, Bernardo Guimaraes, and Nouriel Roubini. 2006. "International Lending of Last Resort and Moral Hazard: A Model of IMF's Catalytic Finance." *Journal of Monetary Economics* 53(3): 441–71.
- Dabla-Norris, Era et al. 2015. "Causes and Consequences of Income Inequality: A Global Perspective." *IMF Staff Discussion Note* 15/13.
- Deaton, Angus S. 2009. "Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development." *NBER Working Paper* 14960.
- Dell'Ariccia, Giovanni, Isabel Schnabel, and Jeromin Zettelmeyer. 2006. "How Do Official Bailouts Affect the Risk of Investing in Emerging Markets?" *Journal of Money, Credit and Banking* 38(7): 1689–1714.
- Derek, Headey. 2008. "Geopolitics and the Effect of Foreign Aid on Economic Growth: 1970–2001." *Journal of International Development* 20(2): 161–80.
- Dieter, Heribert. 2006. "The Decline of the IMF: Is It Reversible? Should It Be Reversed?" *Global Governance* 12(4): 343–49.
- Dietrich, Simone. 2013. "Bypass or Engage? Explaining Donor Delivery Tactics in Foreign Aid Allocation." *International Studies Quarterly* 57(4): 698–712.
- Dorsch, Michael T., and Paul Maarek. 2018. "Democratization and the Conditional Dynamics of Income Distribution." *American Political Science Review*, forthcoming.
- Doyle, David. 2012. "Pressures to Privatize? The IMF, Globalization, and Partisanship in Latin America." *Political Research Quarterly* 65(3): 572–85.
- Dreher, Axel. 2004. "Does the IMF Cause Moral Hazard? A Critical Review of the Evidence." *unpublished manuscript, Heidelberg University*.
- — —. 2006a. "Does Globalization Affect Growth? Evidence from a New Index of Globalization." *Applied Economics* 38(10): 1091–1110.
- — —. 2006b. "IMF and Economic Growth: The Effects of Programs, Loans, and Compliance

- with Conditionality." *World Development* 34(5): 769–88.
- — —. 2009a. "IMF Conditionality: Theory and Evidence." *Public Choice* 141: 233–67.
- — —. 2009b. "The Development of IMF and World Bank Conditionality." In *The Law and Economics of Globalisation: New Challenges for a World in Flux*, ed. L Yueh. Cheltenham: Edward Elgar.
- Dreher, Axel, Vera Eichenauer, and Kai Gehring. 2016. "Geopolitics, Aid, and Growth." *World Bank Economic Review*.
- Dreher, Axel, Andreas Fuchs, Bradley Parks, Austin M. Strange, and Michael J. Tierney. 2017. Aid, China, and Growth: Evidence from a New Global Development Finance Dataset. *AidData Working Paper* 46.
- Dreher, Axel, and Martin Gassebner. 2012. "Do IMF and World Bank Programs Induce Government Crises? An Empirical Analysis." *International Organization* 66(2): 329–58.
- Dreher, Axel, and Noel Gaston. 2008. "Has Globalization Increased Inequality?" *Review of International Economics* 16(3): 516–36.
- Dreher, Axel, Noel Gaston, and Pim Martens. 2008. *Measuring Globalization – Gauging Its Consequences*. Berlin: Springer.
- Dreher, Axel, Matthew Gould, Matthew D. Rablen, and James Raymond Vreeland. 2014. "The Determinants of Election to the United Nations Security Council." *Public Choice* 158(1–2): 51–83.
- Dreher, Axel, and Nathan M. Jensen. 2007. "Independent Actor or Agent? An Empirical Analysis of the Impact of U.S. Interests on International Monetary Fund Conditions." *The Journal of Law and Economics* 50: 105–24.
- Dreher, Axel, Stephan Klasen, James Raymond Vreeland, and Eric Werker. 2013. "The Costs of Favoritism: Is Politically Driven Aid Less Effective?" *Economic Development and Cultural Change* 62(1): 157–91.
- Dreher, Axel, and Valentin F. Lang. 2016. "The Political Economy of International Organizations." *CESifo Working Paper* 6077.
- Dreher, Axel, and Sarah Langlotz. 2017. "Aid and Growth. New Evidence Using an Excludable Instrument." *Heidelberg University Discussion Paper* 635.
- Dreher, Axel, Silvia Marchesi, and James Raymond Vreeland. 2008. "The Political Economy of IMF Forecasts." *Public Choice* 137(1–2): 145–71.
- Dreher, Axel, Heiner Mikosch, and Stefan Voigt. 2015. "Membership Has Its Privileges - The

- Effect of Membership in International Organizations on FDI." *World Development* 66: 346–58.
- Dreher, Axel, Peter Nunnenkamp, and Rainer Thiele. 2008. "Does US Aid Buy UN General Assembly Votes: A Disaggregated Analysis." *Public Choice* 136(1–2): 139–69.
- Dreher, Axel, and Katharina Richert. 2017. "The Political Economy of International Finance Corporation Lending." *CEPR Discussion Paper* 12290.
- Dreher, Axel, Peter B. Rosendorff, and James Raymond Vreeland. 2013. "Buying Votes and International Organizations." *unpublished manuscript, Heidelberg University*.
- Dreher, Axel, Jan-Egbert Sturm, and James Raymond Vreeland. 2009a. "Development Aid and International Politics: Does Membership on the UN Security Council Influence World Bank Decisions?" *Journal of Development Economics* 88(1): 1–18.
- — —. 2009b. "Global Horse Trading: IMF Loans for Votes in the United Nations Security Council." *European Economic Review* 53(7): 742–57.
- — —. 2015. "Politics and IMF Conditionality." *Journal of Conflict Resolution* 59(1): 120–48.
- Dreher, Axel, and Jan Egbert Sturm. 2012. "Do the IMF and the World Bank Influence Voting in the UN General Assembly?" *Public Choice* 151: 363–97.
- Dreher, Axel, and Stefan Voigt. 2011. "Does Membership in International Organizations Increase Governments' Credibility? Testing the Effects of Delegating Powers." *Journal of Comparative Economics* 39(3): 326–48.
- Dreher, Axel, and James Raymond Vreeland. 2011. "Buying Votes and International Organizations." *Courant Research Centre, Discussion Paper* May(78): 1–39.
- Dreher, Axel, and Stefanie Walter. 2010. "Does the IMF Help or Hurt? The Effect of IMF Programs on the Likelihood and Outcome of Currency Crises." *World Development* 38: 1–18.
- Easterly, William. 2005. "What Did Structural Adjustment Adjust? The Association of Policies and Growth with Repeated IMF and World Bank Adjustment Loans." *Journal of Development Economics* 76(1): 1–22.
- Edwards, Martin S. 2006. "Signalling Credibility? The IMF and Catalytic Finance." *Journal of International Relations and Development* 9(1): 27–52.
- Eichengreen, Barry, Kenneth Kletzer, and Ashoka Mody. 2006. "The IMF in a World of Private Capital Markets." *Journal of Banking and Finance* 30(5): 1335–57.
- Eldar, Ofer. 2008. "Vote-Trading in International Institutions." *European Journal of International*

- Law* 19(1): 3–41.
- Esarey, Justin. 2015. “Using Interaction Terms as Instrumental Variables for Causal Identification: Does Corruption Harm Economic Development?” *Rice University Working Paper*.
- Faye, Michael, and Paul Niehaus. 2012. “Political Aid Cycles.” *American Economic Review* 102(7): 3516–30.
- Feenstra, Robert C., and Gordon H. Hanson. 1996. “Globalization, Outsourcing, and Wage Inequality.” *American Economic Review* 86(2): 240–45.
- Fink, Gerhard, Peter Haiss, and Sirma Hristoforova. 2003. “Bond Markets and Economic Growth.” *IEF Working Paper* 49.
- Fleck, Robert K., and Kilby Christopher. 2006. “How Do Political Changes Influence US Bilateral Aid Allocations? Evidence from Panel Data.” *Review of Development Economics* 10(2): 210–23.
- Foster, Timon et al. 2018. “How Structural Adjustment Programs Affect Inequality: A Disaggregated Analysis of IMF Conditionality, 1980–2014.” *unpublished manuscript, University of Cambridge*.
- Fratianni, Michele, and John Pattison. 2005. “Who Is Running the IMF: Critical Shareholders or the Staff.” In *Multidisciplinary Economics: The Birth of a New Economics Faculty in the Netherlands*, eds. Peter De Gijssel and Hans Schenk. Springer, 279–92.
- Fratzscher, Marcel, and Julien Reynaud. 2011. “IMF Surveillance and Financial Markets-A Political Economy Analysis.” *European Journal of Political Economy* 27(3): 405–22.
- Fuchs, Andreas, Axel Dreher, and Peter Nunnenkamp. 2014. “Determinants of Donor Generosity: A Survey of the Aid Budget Literature.” *World Development* 56: 172–99.
- Fuchs, Andreas, and Kai Gehring. 2017. “The Home Bias in Sovereign Ratings.” *Journal of the European Economic Association* 6: 1386–1423.
- Furceri, Davide, and Prakash Loungani. 2018. “The Distributional Effects of Capital Account Liberalization.” *Journal of Development Economics* 130: 127–44.
- Gartzke, Erik, and Megumi Naoi. 2011. “Multilateralism and Democracy: A Dissent Regarding Keohane, Macedo, and Moravcsik.” *International Organization* 65(3): 589–98.
- Garuda, Gopal. 2000. “The Distributional Effects of IMF Programs: A Cross-Country Analysis.” *World Development* 28(6): 1031–51.
- Gehring, Kai, Lennart Kaplan, and Melvin Wong. 2018. “Aid and Conflict at the Local Level.”

unpublished manuscript, University of Zurich.

- Gehring, Kai, and Valentin Lang. 2018. "Stigma or Cushion? IMF Programs and Sovereign Creditworthiness." *ETH and University of Zurich, CIS Working Paper* 98.
- Gelman, Andrew, and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.
- Gerster, Richard. 1993. "Accountability of Executive Directors in the Bretton Woods Institutions." *Journal of World Trade* 27: 88–116.
- Goldberg, Pinelopi, and Nina Pavcnik. 2007. "Distributional Effects of Globalization in Developing Countries." *Journal of Economic Literature* 45(1): 39–82.
- Goldgeier, James M., and Michael McFaul. 2005. *Power and Purpose: U.S. Policy Toward Russia After the Cold War*. Washington, DC: Brookings Institution Press.
- Goldin, Claudia, and Lawrence F. Katz. 2010. *The Race between Education and Technology*. Harvard: Belknap Press.
- Gould, Erica R. 2003. "Money Talks: Supplementary Financiers and International Monetary Fund Conditionality." *International Organization* 57(3): 551–86.
- . 2006. *Money Talks: The International Monetary Fund, Conditionality, and Supplementary Financiers*. Stanford: Stanford University Press.
- Grant, Ruth W., and Robert O. Keohane. 2005. "Accountability and Abuses of Power in World Politics." *American Political Science Review* 99(1): 29–43.
- Gray, Julia. 2009. "International Organization as a Seal of Approval: European Union Accession and Investor Risk." *American Journal of Political Science* 53(4): 931–49.
- . 2013. *The Company That States Keep: International Economic Organizations and Investor Perceptions*. Cambridge: Cambridge University Press.
- . 2018. "Life, Death, or Zombie? The Vitality of International Organizations." *International Studies Quarterly* 62(1): 1–13.
- De Gregorio, José, and Jong-Wha Lee. 2002. "Education and Income Inequality: New Evidence from Cross-Country Data." *Review of Income and Wealth* 48(3): 395–416.
- Grigorescu, Alexandru. 2007. "Transparency of Intergovernmental Organizations: The Roles of Member States, International Bureaucracies and Nongovernmental Organizations." *International Studies Quarterly* 51(3): 625–48.
- . 2013. "International Organizations and Their Bureaucratic Oversight Mechanisms: The Democratic Deficit, Accountability, and Transparency." In *Routledge Handbook of*

- International Organization*, ed. Bob Reinalda. London: Routledge, 176–88.
- — —. 2015. *Democratic Intergovernmental Organizations? Normative Pressure and Decision-Making Rules*. Cambridge: Cambridge University Press.
- Grossman, Gene M, and Elhanan Helpman. 2015. "Globalization and Growth." *American Economic Review* 105(5): 100–104.
- de Haan, Jakob, and Jan-Egbert Sturm. 2016. "Finance and Income Inequality: A Review and New Evidence." *CESifo Working Paper* 6079.
- Haberler, Gottfried. 1974. *Economic Growth and Stability*. Westminster: Prometheus.
- Hagen, Rune Hansen. 2006. "Samaritan Agents? On the Strategic Delegation of Aid Policy." *Journal of Development Economics* 79(249–263).
- Harrison, Ann, John McLaren, and Margaret S. McMillan. 2010. "Recent Findings on Trade and Inequality." *NBER Working Paper Series* 16425.
- Hartley, Keith, and Todd Sandler. 1999. "NATO Burden-Sharing: Past and Future." *Journal of Peace Research* 36(6): 665–80.
- Hartzell, Caroline, Matthew Hoddie, and Molly Bauer. 2010. "Economic Liberalization via IMF Structural Adjustment: Sowing the Seeds of Civil War?" *International Organization* 64(2): 339–56.
- Hawkins, Darren, David A. Lake, Daniel Nielson, and Michael J. Tierney. 2006. *Delegation and Agency in International Organizations*. Cambridge: Cambridge University Press.
- Heckman, James J. 2005. "Rejoinder: Response to Sobel." *Sociological Methodology* 35(1): 135–62.
- Heinrich, Tobias, and Yoshiharu Kobayashi. 2018. "How Do People Evaluate Foreign Aid To 'Nasty' Regimes?" *British Journal of Political Science*.
- Hernandez, Diego, and Krishna Chaitanya Vadlamannati. 2017. "Politics of Religiously Motivated Lending: An Empirical Analysis of Aid Allocation by the Islamic Development Bank." *Journal of Comparative Economics* 45(4): 910–29.
- Hicks, Robert L., Bradley C. Parks, J. Timmons Roberts, and Michael J. Tierney. 2010. *Greening Aid?: Understanding the Environmental Impact of Development Assistance*. Oxford University Press.
- Hill, Paula, Robert Brooks, and Robert Faff. 2010. "Variations in Sovereign Credit Quality Assessments across Rating Agencies." *Journal of Banking and Finance* 34(6): 1327–43.
- Hillman, Arye L., and Niklas Potrafke. 2015. "The UN Goldstone Report and Retraction: An

- Empirical Investigation." *Public Choice* 163(3–4): 247–66.
- Hoeffler, Anke, and Verity Outram. 2011. "Need, Merit, or Self-Interest—What Determines the Allocation of Aid?" *Review of Development Economics* 15(2): 237–50.
- Hurd, Ian. 2007. *Legitimacy and Power in the UN Security Council*. Princeton University Press.
- Hurd, Ian, and Bruce Cronin. 2008. *The UN Security Council and the Legitimacy of International Authority*. London: Routledge.
- IMF. 1998. "IMF Financing Helps Members Pursue Sound Policies." <https://www.imf.org/external/pubs/ft/survey/sup0998/08.htm> (May 11, 2018).
- — —. 2001. "IMF Lending to Poor Countries: How Does the PRGF Differ from the ESAF?" <https://www.imf.org/external/np/exr/ib/2001/043001.htm> (July 17, 2016).
- — —. 2003. *Fiscal Adjustment in IMF-Supported Programs*. Independent Evaluation Office.
- — —. 2007. "Structural Conditionality in IMF-Supported Programs - Background Documents."
- — —. 2009. *IMF Policy Paper A Review of Some Aspects of the Low-Income Country Debt Sustainability Framework*. Washington, DC.
- — —. 2013. *Staff Guidance Note on the Application of the Joint Bank-Fund Debt Sustainability Framework for Low-Income Countries*. Washington, DC.
- — —. 2014a. "Review of Flexible Credit Line, the Precautionary and Liquidity Line, and the Rapid Financing Instrument." *IMF Policy Paper*.
- — —. 2014b. "Revised Operational Guidance to IMF Staff on the 2002 Conditionality Guidelines."
- — —. 2015. *Staff Guidance Note on the Implementation of Public Debt Limits in Fund-Supported Programs*. Washington, DC.
- — —. 2016a. "IMF Lending." <https://www.imf.org/external/np/exr/facts/howlend.htm> (April 26, 2016).
- — —. 2016b. "Monitoring of Fund Arrangements (MONA) Database."
- — —. 2016c. "Monitoring of Fund Arrangements (MONA) Database." <https://www.imf.org/external/np/pdr/mona/index.aspx> (September 16, 2016).
- — —. 2016d. "World Economic Outlook."
- — —. 2017. "Reforms of Lending and Conditionality Frameworks." <https://www.imf.org/external/np/exr/faq/facfaqs.htm#q11> (February 7, 2018).
- — —. 2018a. "IMF Lending Arrangements."

- — —. 2018b. "The IMF and the Sustainable Development Goals." <http://www.imf.org/en/About/Factsheets/Sheets/2016/08/01/16/46/Sustainable-Development-Goals> (May 14, 2018).
- IMF, and World Bank. 2012. *Revisiting the Debt Sustainability Framework for Low-Income Countries. Policy Paper, IMF and The World Bank.*
- Ito, Takatoshi. 2012. "Can Asia Overcome the IMF Stigma?" *American Economic Review* 102(3): 198–202.
- Jaumotte, Florence, Subir Lall, and Chris Papageorgiou. 2013. "Rising Income Inequality: Technology, or Trade and Financial Globalization?" *IMF Economic Review* 61(2): 271–309.
- Jenkins, Stephen P. 2015. "World Income Inequality Databases: An Assessment of WIID and SWIID." *Journal of Economic Inequality* 13(4): 629–71.
- Jensen, Carsten, and Svend-Erik Skaaning. 2015. "Democracy, Ethnic Fractionalisation, and the Politics of Social Spending: Disentangling a Conditional Relationship." *International Political Science Review* 36(4): 457–72.
- Jensen, Nathan. 2004. "Crisis, Conditions, and Capital: The Effect of International Monetary Fund Agreements on Foreign Direct Investment Inflows." *Journal of Conflict Resolution* 48(2): 194–210.
- Jorra, Markus. 2012. "The Effect of IMF Lending on the Probability of Sovereign Debt Crises." *Journal of International Money and Finance* 31(4): 709–25.
- Kaja, Ashwin, and Eric Werker. 2010. "Corporate Governance at the World Bank and the Dilemma of Global Governance." *World Bank Economic Review* 24(2): 171–98.
- Kanbur, Ravi. 2013. "Globalization and Inequality." In *Handbook of Income Distribution*, eds. A Atkinson and Francois Bourguignon. North-Holland.
- Kegley Jr., Charles W., and Steven W. Hook. 1991. "U.S. Foreign Aid and U.N. Voting: Did Reagan's Linkage Strategy Buy Deference or Defiance?" *International Studies Quarterly* 35(3): 295–312.
- Kentikelenis, Alexander E., Thomas H. Stubbs, and Lawrence P. King. 2016. "IMF Conditionality and Development Policy Space, 1985–2014." *Review of International Political Economy*.
- Keohane, Robert O. 1984. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton: Princeton University Press.
- Keohane, Robert O., Stephen Macedo, and Andrew Moravcsik. 2009. "Democracy-Enhancing

- Multilateralism." *International Organization* 63(1): 1–31.
- — —. 2011. "Constitutional Democracy and World Politics: A Response to Gartzke and Naoi." *International Organization* 65(3): 599–604.
- Keohane, Robert O., and Joseph S. Nye. 2000. "Introduction." In *Governance in a Globalizing World*, eds. Joseph S. Nye and John D. Donahue. Washington DC: Brookings Institution.
- Kerrissey, Jasmine. 2015. "Collective Labor Rights and Income Inequality." *American Sociological Review* 80(3): 626–53.
- Kersting, Erasmus, and Christopher Kilby. 2016. "With a Little Help from My Friends: Global Electioneering and World Bank Lending." *Journal of Development Economics* 121: 153–65.
- Kilby, Christopher. 2000. "Supervision and Performance: The Case of World Bank Projects." *Journal of Development Economics* 62(1): 233–59.
- — —. 2009. "The Political Economy of Conditionality: An Empirical Analysis of World Bank Loan Disbursements." *Journal of Development Economics* 89(1): 51–61.
- — —. 2013a. "An Empirical Assessment of Informal Influence in the World Bank." *Economic Development and Cultural Change* 61(2): 431–64.
- — —. 2013b. "The Political Economy of Project Preparation: An Empirical Analysis of World Bank Projects." *Journal of Development Economics* 105: 211–25.
- — —. 2015. "Assessing the Impact of World Bank Preparation on Project Outcomes." *Journal of Development Economics* 115: 111–23.
- Klein, Naomi. 2008. *The Shock Doctrine*. New York: Metropolitan Books.
- Koeniger, Winfried, Marco Leonardi, and Luca Nunziata. 2007. "Labor Market Institutions and Wage Inequality." *Industrial and Labour Relations Review* 60(3): 340–56.
- KOF. 2016. "KOF Index of Globalization." <http://globalization.kof.ethz.ch/> (April 11, 2018).
- Kolo, Philipp. 2012. "Measuring a New Aspect of Ethnicity – The Appropriate Diversity Index." *Universität Göttingen Discussion Paper* 221.
- Koremenos, Barbara. 2013. "What's Left out and Why? Informal Provisions in Formal International Law." *Review of International Organizations* 8(2): 137–62.
- Kuziemko, Ilyana, and Eric Werker. 2006. "How Much Is a Seat on the Security Council Worth? Foreign Aid and Bribery at the United Nations." *Journal of Political Economy* 114(5): 905–30.
- Kuznets, Simon. 1955. "Economic Growth and Income Inequality." *American Economic Review* 45(1): 1–28.

- Laeven, Luc, and Fabian Valencia. 2012. "Systemic Banking Crises Database: An Update." *IMF Working Paper* 12(163).
- Lang, Valentin F. 2016. "The Economics of the Democratic Deficit: The Effect of IMF Programs on Inequality." *Heidelberg University Discussion Paper* 617.
- Lang, Valentin F., and Andrea F. Presbitero. 2018. "Room for Discretion? Biased Decision-Making in International Financial Institutions." *Journal of Development Economics* 130: 1–16.
- Lang, Valentin F, and Marina Mendes Tavares. 2018. "The Distribution of Gains from Globalization." *IMF Working Paper* 18/54.
- Lee, Jong Wha, and Kwanho Shin. 2008. "IMF Bailouts and Moral Hazard." *Journal of International Money and Finance* 27: 816–30.
- Lim, Daniel Yew Mao, and James Raymond Vreeland. 2013. "Regional Organizations and International Politics: Japanese Influence over the Asian Development Bank and the UN Security Council." *World Politics* 65(1): 34–72.
- Lumsdaine, David Halloran. 1993. *Moral Vision in International Politics: The Foreign Aid Regime, 1949-1989*. Princeton: Princeton University Press.
- Malik, Rabia, and Randall W Stone. 2017. "Corporate Influence in World Bank Lending." *Journal of Politics* 80(1): 103–18.
- Malone, David. 1998. *Decision-Making in the UN Security Council: The Case of Haiti, 1990–1997*. Oxford University Press.
- Marchesi, Silvia, and Emanuela Sirtori. 2011. "Is Two Better than One? The Effects of IMF and World Bank Interaction on Growth." *Review of International Organizations* 6(3): 287–306.
- Marchesi, Silvia, and Jonathan P. Thomas. 1999. "IMF Conditionality as a Screening Device." *Economic Journal* 109: 111–25.
- Marshall, Monty, Keith Jaggers, and Ted Robert Gurr. 2011. "Polity IV Project: Dataset Users' Manual." *Centre for Systemic Peace: Polity IV Project*.
- Mattes, Michaela, Brett Ashley Leeds, and Royce Carroll. 2015. "Leadership Turnover and Foreign Policy Change: Societal Interests, Domestic Institutions, and Voting in the United Nations." *International Studies Quarterly* 59: 280–90.
- McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142(2): 698–714.
- McDonnell, Ida, and Liam Wegimont, Henri-Bernhard Solignac Lecomte. 2003. *Public Opinion*

- and the Fight against Poverty*. Paris: OECD Publishing.
- McKeown, Timothy J. 2009. "How U.S. Decision-Makers Assessed Their Control of Multilateral Organizations, 1957-1982." *Review of International Organizations* 4(3): 269–91.
- Mearsheimer, John J. 1994. "The False Promise of International Institutions." *International Security* 19(3): 5–49.
- Meltzer, Allan H., and Scott F. Richard. 1981. "A Rational Theory of the Size of Government." *Journal of Political Economy* 89(5): 914–27.
- Mikulaschek, Christoph. 2017a. "Cueing Foreign Elite Consensus or Divisions : The Effect of Unanimity in International Organizations on Public Opinion." *Paper presented at the Political Economy of International Organization Conference, 2018*.
- — —. 2017b. "Issue Linkage across International Organizations: Does European Countries' Temporary Membership in the UN Security Council Increase Their Receipts from the EU Budget?" *Review of International Organizations*: 1–28.
- Milanovic, Branko. 2014. "All The Ginis Dataset." <http://go.worldbank.org/9VCQW66LA0> (April 12, 2016).
- Milner, Helen V. 2006. "Why Multilateralism: Foreign Aid and Domestic Principal-Agent Problems." In *Delegation and Agency in International Organizations*, eds. Darren Hawkins, David A. Lake, Daniel L. Nielson, and Michael J. Tierney. Cambridge University Press.
- Milner, Helen V., and Bumba Mukherjee. 2009. "Democratization and Economic Globalization." *Annual Review of Political Science* 12(1): 163–81.
- Milner, Helen V., and Dustin Tingley. 2013. "The Choice for Multilateralism: Foreign Aid and American Foreign Policy." *Review of International Organizations* 8(3): 313–41.
- Milner, Helen V., and Dustin H. Tingley. 2010. "The Political Economy of U.S. Foreign Aid: American Legislators and the Domestic Politics of Aid." *Economics and Politics* 22(2): 200–232.
- — —. 2015. *Sailing the Water's Edge The Domestic Politics of American Foreign Policy*. Princeton University Press.
- Mody, Ashoka, and Diego Saravia. 2006. "Catalysing Private Capital Flows: Do IMF Programmes Work as Commitment Devices?" *Economic Journal* 116(513): 843–67.
- Moravcsik, Andrew. 1997. "Taking Preferences Seriously: A Liberal Theory of International Politics." *International Organization* 51(4): 513–53.
- — —. 2002. "In Defence of the 'Democratic Deficit': Reassessing Legitimacy in the European

- Union." *Journal of Common Market Studies* 40(4): 603–24.
- Morris, Stephen, and Hyun Song Shin. 2006. "Catalytic Finance: When Does It Work?" *Journal of International Economics* 70(1): 161–77.
- Moser, Christoph, and Jan Egbert Sturm. 2011. "Explaining IMF Lending Decisions after the Cold War." *Review of International Organizations* 6: 307–40.
- Mukherjee, Bumba, and David Andrew Singer. 2010. "International Institutions and Domestic Compensation: The IMF and the Politics of Capital Account Liberalization." *American Journal of Political Science* 54(1): 45–60.
- Mulligan, Casey, Ricard Gil, and Xavier Sala-i-Martin. 2004. "Do Democracies Have Different Public Policies than Non-Democracies?" *Journal of Economic Perspectives* 18(1): 51–74.
- Nelson, Stephen C. 2014. "Playing Favorites: How Shared Beliefs Shape the IMF's Lending Decisions." *International Organization* 68(2): 297–328.
- Nelson, Stephen C., and Geoffrey P. R. Wallace. 2017. "Are IMF Lending Programs Good or Bad for Democracy?" *The Review of International Organizations* 12(4): 523–58.
- Nielson, Daniel L., and Michael J. Tierney. 2003. "Delegation to International Organizations: Agency Theory and World Bank Environmental Reform." *International Organization* 57(2): 241–76.
- Nizalova, Olena, and Irina Murtazashvili. 2016. "Exogenous Treatment and Endogenous Factors: Vanishing of Omitted Variable Bias on the Interaction Term." *Journal of Econometric Methods* 5(1): 71–77.
- Nooruddin, Irfan, and Joel W. Simmons. 2006. "The Politics of Hard Choices: IMF Programs and Government Spending." *International Organization* 60(4): 1001–33.
- Nooruddin, Irfan, and James Raymond Vreeland. 2010. "The Effect of IMF Programs on Public Wages and Salaries." In *Global Governance, Poverty and Inequality*, eds. Jennifer Clapp and Rorden Wilkinson. London: Routledge.
- Noy, Ilan. 2008. "Sovereign Default Risk, the IMF and Creditor Moral Hazard." *Journal of International Financial Markets, Institutions and Money* 18(1): 64–78.
- Nunn, Nathan, and Nancy Qian. 2014. "US Food Aid and Civil Conflict." *American Economic Review* 104(6): 1630–66.
- Nye, Joseph S. 2001. "Globalization's Democratic Deficit: How to Make International Institutions More Accountable." *Foreign Affairs* (July/August).
- O'Neill, Barry. 1996. "Power and Satisfaction in the United Nations Security Council." *Journal*

- of Conflict Resolution* 40(2): 219–37.
- Oatley, Thomas, and Jason Yackee. 2004. “American Interests and IMF Lending.” *International Politics* 41(3): 415–29.
- Oberdabernig, Doris A. 2013. “Revisiting the Effects of IMF Programs on Poverty and Inequality.” *World Development* 46: 113–42.
- OECD. 2011. *Divided We Stand: Why Inequality Keeps Rising*. Paris: OECD Publishing.
- — —. 2018a. “DAC List of ODA Recipients.” <http://www.oecd.org/dac/financing-sustainable-development/development-finance-standards/daclist.htm> (May 8, 2018).
- — —. 2018b. “OECD.Stat.” www.stats.oecd.org (May 8, 2018).
- Paddock, John V. 2002. “IMF Policy and the Argentine Crisis.” *The University of Miami Inter-American Law Review* 34(1): 155–87.
- Panizza, Ugo, Federico Sturzenegger, and Jeromin Zettelmeyer. 2009. “The Economics and Law of Sovereign Debt and Default.” *Journal of Economic Literature* 47(3): 651–98.
- Papi, Luca, Andrea F. Presbitero, and Alberto Zazzaro. 2015. “IMF Lending and Banking Crises.” *IMF Working Papers* 63(3): 644–91.
- Pastor, Manuel. 1987. “The Effects of IMF Programs in the Third World: Debate and Evidence from Latin America.” *World Development* 15(2): 249–62.
- Paul Christian, and Christopher B Barrett. 2017. “Revisiting the Effect of Food Aid on Conflict: A Methodological Caution.” *World Bank Policy Research Working Paper* 8171. [http://barrett.dyson.cornell.edu/files/papers/Revisiting Food Aid and Conflict 28 April 2017\(b\).pdf](http://barrett.dyson.cornell.edu/files/papers/Revisiting_Food_Aid_and_Conflict_28_April_2017(b).pdf).
- Peet, Richard. 2009. *Unholy Trinity: The IMF, World Bank and WTO*. London: Zed.
- Piketty, Thomas. 2014. *Capital in the Twenty-First Century*. Harvard: Belknap Press.
- Polak, Jacques. 1991. Essays in International Finance *The Changing Nature of IMF Conditionality*.
- Potrafke, Niklas. 2015. “The Evidence on Globalisation.” *World Economy* 38(3): 509–52.
- Presbitero, Andrea F., Dhaneshwar Ghura, Olumuyiwa Adededeji, and Lamin Nije. 2016. “Sovereign Bonds in Developing Countries: Drivers of Issuance and Spreads.” *Review of Development Finance* 6(1): 1–15.
- Przeworski, Adam, and James Raymond Vreeland. 2000. “The Effects of IMF Programs on Economic Growth.” *Journal of Development Economics* 62: 385–421.
- Puhani, Patrick. 2000. “The Heckman Correction for Sample Selection and Its Critique.” *Journal of Economic Surveys* 14(1): 53–68.

- Rai, Kul B. 1980. "Foreign Aid and Voting in the UN General Assembly." *Journal of Peace Research* 17(3): 269–77.
- Reinhart, Carmen M. 2002. "Default, Currency Crises, and Sovereign Credit Ratings." *World Bank Economic Review* 16(2): 151–70.
- Reinhart, Carmen M., and Christoph Trebesch. 2015. "The International Monetary Fund: 70 Years of Reinvention." *Journal of Economic Perspectives* 30(1): 3–28.
- Reinsberg, Bernhard, Katharina Michaelowa, and Stephen Knack. 2017. "Which Donors, Which Funds? Bilateral Donors' Choice of Multilateral Funds at the World Bank." *International Organization* 71(4): 767–802.
- Reuveny, Rafael, and Quan Li. 2003. "Economic Openness, Democracy, and Income Inequality: An Empirical Analysis." *Comparative Political Studies* 36(5): 575–601.
- Reynaud, Julien, and Julien Vauday. 2009. "Geopolitics and International Organizations: An Empirical Study on IMF Facilities." *Journal of Development Economics* 89(1): 139–62.
- Reynolds, Evangeline M., and Matthew S. Winters. 2016. "Foreign Aid Funnel? A Placebo-Based Assessment of Aid Flows to Non-Permanent United Nations Security Council Members." *Research & Politics* 3(1): 1–9.
- Rickard, Stephanie, and Teri Caraway. 2018. "International Demands for Austerity: Examining the Impact of the IMF on the Public Sector." *Review of International Organizations*.
- Rieffel, Lex. 2003. *Restructuring Sovereign Debt: The Case for Ad-Hoc Machinery*. Brookings Institution Press.
- Rocabert, Jofre, Frank Schimmelfennig, Thomas Winzen, and Lorian Crasnic. 2017. "The Rise of International Parliamentary Institutions: Authority and Legitimation." *unpublished manuscript, University of Zurich*.
- Rodrik, Dani. 1998. "Why Do More Open Economies Have Bigger Governments?" *Journal of Political Economy* 106(5): 997–1032.
- — —. 1999. "Democracies Pay Higher Wages." *Quarterly Journal of Economics* 114(3): 707–38.
- — —. 2003. "Argentina: A Case of Globalisation Gone Too Far or Not Far Enough?" In *The Crisis That Was Not Prevented: Argentina, the IMF, and Globalisation*, eds. Jan Joost Teunissen and Age Akkerman. FONDAD.
- Rowlands, Dane. 2001. "The Response of Other Lenders to the IMF." *Review of International Economics* 9(3): 531–46.
- Rudra, Nita, and Stephan Haggard. 2005. "Globalization, Democracy, and Effective Welfare

- Spending in the Developing World." *Comparative Political Studies* 38(9): 1015–49.
- Ruggie, John Gerard. 1982. "International Regimes, Transactions, and Change: Embedded Liberalism in the Postwar Economic Order." *International Organization* 36(2): 379–415.
- Scheve, Kenneth, and David Stasavage. 2012. "Democracy, War, and Wealth: Lessons from Two Centuries of Inheritance Taxation." *American Political Science Review* 106(1): 81–102.
- Schneider, Christina, and Branislav L. Slantchev. 2018. "The Domestic Politics of International Cooperation: Germany and the European Debt Crisis." *International Organization* 72(1): 1–31.
- Schneider, Christina, and Jennifer L. Tobin. 2016. "Portfolio Similarity and the Diversification of Multilateral Aid." *International Studies Quarterly* 60(4): 647–64.
- Smith, Alastair, and James Raymond Vreeland. 2006. "The Survival of Political Leaders and IMF Programs." In *Globalization and the Nation State: The Impact of the IMF and the World Bank*, eds. Gustav Ranis, James Raymond Vreeland, and Stephen Kosack. Routledge.
- Solt, Frederick. 2016. "The Standardized World Income Inequality Database." *Social Science Quarterly* 97(5): 1267–81.
- Staiger, Douglas, and James H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica* 65(3): 557–86.
- Standard and Poor's. 2016. "Understanding Ratings." http://www.spratings.com/en_US/understanding-ratings (September 12, 2016).
- Stasavage, David. 2004. "Open-Door or Closed-Door? Transparency in Domestic and International Bargaining." *International Organization* 58(4): 667–703.
- Steinwand, Martin C., and Randall W. Stone. 2008. "The International Monetary Fund: A Review of the Recent Evidence." *Review of International Organizations* 3(2): 123–49.
- Stiglitz, Joseph E. 2000. "What I Learned at the World Economic Crisis." *New Republic* 222(16/17): 56–60.
- — —. 2002. *Globalization and Its Discontents*. New York: Norton.
- Stock, James H., and Motohiro Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression." In *Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg*, eds. James H. Stock and D. Andrews. New York: Cambridge University Press.
- Stone, Randall W. 2002. *Lending Credibility: The International Monetary Fund and the Post-Communist Transition*. Princeton: Princeton University Press.

- — —. 2004. "The Political Economy of IMF Lending in Africa." *American Political Science Review* 98(4): 577–91.
- — —. 2008. "The Scope of IMF Conditionality." *International Organization* 62(4): 589–620.
- — —. 2011. *Controlling Institutions: International Organizations and the Global Economy*. Cambridge: Cambridge University Press.
- — —. 2013. "Informal Governance in International Organizations: Introduction to the Special Issue." *Review of International Organizations* 8(2): 121–36.
- Strand, Jonathan R., and Tina M. Zappile. 2015. "Always Vote for Principle, Though You May Vote Alone: Explaining United States Political Support for Multilateral Development Loans." *World Development* 72: 224–39.
- Stubbs, Thomas H., Alexander E. Kentikelenis, and Lawrence P. King. 2016. "Catalyzing Aid? The IMF and Donor Behavior in Aid Allocation." *World Development* 78: 511–28.
- Sturm, Jan Egbert, Helge Berger, and Jakob de Haan. 2005. "Which Variables Explain Decisions on IMF Credit? An Extreme Bounds Analysis." *Economics and Politics* 17(2): 177–213.
- Tallberg, Jonas, Thomas Sommerer, Theresa Squatrito, and Christer Jönsson. 2014. "Explaining the Transnational Design of International Organizations." *International Organization* 68(4): 741–74.
- Thacker, Strom C. 1999. "The High Politics of IMF Lending." *World Politics* 52(1): 38–75.
- Tomz, Michael, and Mark L. J. Wright. 2007. "Do Countries Default in 'Bad Times'?" *Journal of the European Economic Association* 5: 352–60.
- Uppsala Conflict Data Program. 2015. "UCDP Battle-Related Deaths Dataset v.5-2015." <http://www.pcr.uu.se/research/ucdp/ucdp-data/> (October 17, 2017).
- US Department of State. 1985. "Report to Congress on Voting Practices in the United Nations."
- — —. 2000. "Report to Congress on Voting Practices in the United Nations."
- USAID. 2016. "U.S. Overseas Loans and Grants (Greenbook)."
- Vabulas, Felicity. 2013. "Consultative and Observer Status of NGOs in Intergovernmental Organizations." In *Routledge Handbook of International Organization*, ed. Bob Reinalda. London: Routledge, 189–202.
- Vaubel, Roland. 1986. "A Public Choice Approach to International Organization." *Public Choice* 51(1): 39–57.
- — —. 1996. "Bureaucracy at the IMF and the World Bank: A Comparison of the Evidence." *World Economy* 19(2): 195–210.

- — —. 2006. "Principal-Agent Problems in International Organizations." *Review of International Organizations* 1(2): 125–38.
- Voeten, Erik. 2001. "Outside Options and the Logic of Security Council Action." *American Political Science Review* 95(4): 845–58.
- — —. 2005. "The Political Origins of the UN Security Council's Ability to Legitimize the Use of Force." *International Organization* 59(3): 527–57.
- Voigt, Stefan, and Eli M. Salzberger. 2002. "Choosing Not To Choose: When Politicians Choose To Delegate Powers." *Kyklos* 55(April): 289–310.
- Vreeland, James Raymond. 2002. "The Effect of IMF Programs on Labor." *World Development* 30(1): 121–39.
- — —. 2003. *The IMF and Economic Development*. Cambridge: Cambridge University Press.
- — —. 2007. "The Politics of IMF Conditional Lending." *World Economics* 8(3): 185–93.
- Vreeland, James Raymond, and Axel Dreher. 2014. *The Political Economy of the United Nations Security Council*. Cambridge: Cambridge University Press.
- Walter, Stefanie. 2010. "Globalization and the Welfare State: Testing the Microfoundations of the Compensation Hypothesis." *International Studies Quarterly* 54(2): 403–26.
- Wang, T. Y. 1999. "US Foreign Aid and UN Voting." *International Studies Quarterly* 43(1): 199–210.
- Werker, Eric, Faisal Z. Ahmed, and Charles Cohen. 2009. "How Is Foreign Aid Spent? Evidence from a Natural Experiment." *American Economic Journal: Macroeconomics* 1(2): 225–44.
- Wittkopf, Eugene R. 1973. "Foreign Aid and United Nations Votes: A Comparative Study." *American Political Science Review* 67(3): 868–88.
- Woo, Byungwon. 2013. "Conditional on Conditionality: IMF Program Design and Foreign Direct Investment." *International Interactions* 39: 292–315.
- Woods, Ngaire. 2003. "The United States and the International Financial Institutions: Power and Influence within the World Bank and the IMF." In *US Hegemony and International Organizations*, eds. N. Foot, N. McFarlane, and M. Mastanduno. Oxford University Press, 92–114.
- — —. 2006. *The Globalizers: The IMF, the World Bank, and Their Borrowers*. Ithaca: Cornell University Press.
- World Bank. 2016a. *Poverty and Shared Prosperity 2016: Taking On Inequality*. Washington, DC:

World Bank Group.

— — —. 2016b. *World Development Indicators*. Washington, DC: World Bank Group.

— — —. 2018. *World Development Indicators*. Washington, DC: World Bank Group.

World Inequality Lab. 2017. *World Inequality Report 2018*. Paris: World Inequality Database.

Acknowledgements

Chapter 1: This chapter builds on Dreher and Vreeland (2011) and Dreher, Rosendorff, and Vreeland (2013). It is part of an ongoing project with Dreher, Rosendorff, and Vreeland. We thank Andreas Fuchs, Tobias Rommel, participants of the Annual Meeting of the European Public Choice Society (EPCS, Rome 2018), the Center for the Study of African Economies (CSAE) Conference (Oxford 2018), seminars at Heidelberg University and the University of Zurich for helpful comments and Simon Hug for valuable support regarding the analysis of UNSC meeting minutes.

Chapter 2: The chapter is co-authored with Andrea F. Presbitero. We thank two anonymous referees, Luca Bandiera, Andrew Berg, Tito Cordella, Axel Dreher, Andreas Fuchs, Christopher Humphrey, Christopher Kilby, Sarah Langlotz, Gerard Padro i Miquel, Juan Pradelli, Katharina Richert, Eyal Robinson, Tim Willems, several IMF and World Bank staff, and participants at the conference on the Political Economy of International Organizations (PEIO, Bern 2017) at the EPCS (Budapest 2017), the Development Economics and Policy conference (AEL, Göttingen 2017), at the International Political Economy Society (IPES) Conference (Austin 2017), the DIAL Conference (Paris 2017), the Conference of the Verein für Socialpolitik (Vienna 2017), and at seminars at the IMF and at Heidelberg University for useful comments and discussions. We also thank numerous staff at the IMF for assistance in accessing the data on the debt sustainability analyses. The views expressed herein are those of the authors and should not be attributed to the IMF, its Executive Board, or its management.

Chapter 3: The chapter is co-authored with Kai Gehring. We thank Lawrence Broz, Giovanni Dell’Ariccia, Dominique Desruelles, Axel Dreher, Christopher Humphrey, David A. Singer, Randall Stone, Marina Tavares, seminar participants at Heidelberg University and the University of Cambridge, as well as participants at the PEIO (Bern 2017), the EPCS (Budapest 2017), the ZEW Public Finance Conference (Mannheim 2017), and the Transformations in Global Economic Governance conference (Cambridge 2017). We thank Jamie Parsons for proof-reading. Excellent research assistance was provided by Philip Kerler, Gina Messerli, Lukas Willi, and Johannes von Mandach.

Chapter 4: The paper is single-authored. I thank Andrew Berg, Enrico Bertacchini, Lawrence Broz, Axel Dreher, Andreas Fuchs, Kai Gehring, Shom Mazumder, Liam McGrath, Andrew Moravcsik, Jonathan Ostry, Florian Rühl, Randall Stone, Mike Tierney, Agnès Zabsonré, participants at the PEIO (Salt Lake City 2016), the CSAE (Oxford 2016), the EPCS (Freiburg 2016), the AEL (Heidelberg 2016), the Conference of the European Political Science Association (EPSA, Rome 2017), and at seminars at Heidelberg University, the University of Zurich, the University of Cambridge, and the University of Hamburg for helpful comments and Jamie Parsons for proof-reading.

•

I would not have been able to write this dissertation without the wonderful people around me. First of all, I thank Axel for being simply the best PhD supervisor I could have asked for. It is incredible how much I have learnt from him. I thank Christopher Kilby for agreeing to be my second supervisor, for making such a long trip for my defense, and for many valuable comments. I thank Jale Tosun for agreeing to be the third member of my PhD committee and for creating this cordial atmosphere in all E&P Seminars that she attended.

I thank my colleagues at Heidelberg University. I thank Sarah for – I am not exaggerating – hundreds of suggestions on the regressions that made it into this dissertation and for endless discussions on interacted instrumental variables. I thank Katha, whose nodding and smiling during my presentations got me even through the toughest internal seminars. I thank Andi, whose thoughtful comments improved every single chapter of this dissertation substantially. I thank Lenny for being the world's best Chair Drinks intern and I thank him and Sebastian for great comments, particularly during EPSA, which takes place only once a year. I thank Gerdy for her contagiously positive attitude. I thank Angelika for enlightening discussions in the lounge. I thank John for many great debates on the IMF over beer. I thank Sven for Pausenexpress. I thank Vera for many great comments on chapter 4. I thank Kai for endless discussions on parallel-trends assumptions and for his work on chapter 3. I thank Alexandra and Diego for their support during the early days of my PhD.

I thank my new colleagues at the University of Zurich, Stefanie, Tobi, Raphael, Lori, Tabea, Ari, and Nils for welcoming me so warmly in Zurich during the final days of my PhD. I thank Nils, for many helpful suggestions during the very final days before I submitted this thesis, one of which was the extremely helpful warning to not write these acknowledgements without a glass of wine.

I thank Alex for inviting me to Cambridge for a research visit in September 2016. There I received great support, and many eye-opening comments from him, Bernhard, and Thomas. I thank my co-author Andrea for his work on chapter 2, for all his support and for inviting me to a research visit at the IMF in November 2016. This visit and the many conversations with IMF staff that he arranged helped me to gain a much deeper understanding of the institution that I have now been studying for almost four years.

I thank my friends and family from whom I learnt more than in any university. I thank Tim, my oldest friend, for never-ending friendship and never-ending discussions. I thank Jakob for his present that eventually led to the topic of this dissertation. I thank Johannes for fantastic conversations on the same topic. I thank David for Cambridge and Paris. I thank Flo H. for Paris and Heidelberg. I thank Flo R. for Heidelberg and Destille. I thank Paul, without whom I would have never survived Mannheim. I thank Felix and Lukas for wonderful conversations in the *théâtre*. I thank Philipp for the same conversations and for some of the best literature suggestions I have ever received. I thank Carolin and Leonie, the best little sisters an older brother could hope for. I thank my parents, who are always there for me and whom I dedicate this dissertation to. I thank, most importantly and wholeheartedly, Julia.

Heidelberg and Zurich, May 2018