Unimplementable by design? A theory of compliance with International Monetary Fund policy conditionality

Bernhard Reinsberg,[1*] Alexander Kentikelenis,[2] Thomas Stubbs[3]

1 – University of Cambridge

2 – University of Oxford

3 – Royal Holloway—University of London

* Correspondence to: br385@cam.ac.uk, 12 Trumpington Street, Cambridge CB2 1QA, United Kingdom

September 30, 2018

Abstract

International agreements often fail due to lack of compliance by the countries that agree to them. While scholars have scrutinized domestic factors as obstacles to compliance, we argue that agreements may be unimplementable by design. We develop our argument in relation to International Monetary Fund programs, in which borrowing countries must commit to far-reaching economic policy reforms for access to credit. We provide a framework for understanding why the Fund initially insists on over-ambitious policy design and subsequently waives unimplemented conditions to keep programs afloat. Specifically, we model program design as a common-pool resource problem among the various departments of the IMF bureaucracy, in which each department manages to have policy conditions included that it cares about, without taking the impact on program sustainability into account. To test our argument, we collect detailed compliance data on individual policy conditions. We deploy these data to assess the determinants of waived conditions across Fund programs from 1980 to 2009. Controlling for a host of initial conditions, features of the loan, and unexpected shocks during implementation, we find a robust positive correlation between the number of conditions and the percentage of conditions waived, and an even stronger correlation when considering only structural conditions, which are widely considered as more difficult to implement. These findings have important implications for theories of compliance as well as for policy-making at international organizations.

Highlights

- We present new data to quantify compliance with IMF conditionality
- We argue that lack of compliance is due to program design
- We use graphical analysis, bivariate statistics, and multivariate analyses to corroborate our argument

Keywords: International Monetary Fund; conditionality; compliance; effectiveness

JEL codes: F33, F34, F53

Acknowledgements: We would like to thank Toke Aidt for helpful comments.

1. Introduction

Lack of compliance is widespread in international relations (Chayes and Chayes 1993; Findley, Nielson, and Sharman 2013; Hathaway 2002; Underdal 1998)—not only for international agreements between governments (such as the GATT, the Paris climate accord, or human rights conventions) but also for contractual agreements between individual states and international organizations. One prominent example of the latter is implementation failure of borrowing countries with the policy conditions of the International Monetary Fund (IMF)—an international organization mandated to uphold global financial stability and provide loans to countries in economic trouble. While researchers concur that the ability to tie disbursements to the fulfillment of policy conditions makes the IMF one of the most powerful international organizations (Stone 2004; Vreeland 2003; Woods 2006), its ability to compel actual policy reform in borrowing countries remains often limited (Abbott, Andersen, and Tarp 2010; Haggard and Kaufman 1992; Rickard and Caraway 2018). This may be problematic to the extent that full implementation of IMF programs might be necessary to achieving macroeconomic stabilization. For instance, IMF researchers have established that better compliance with its conditions lead to better macroeconomic outcomes (Nsouli, Atoyan, and Mourmouras 2004). Even where IMF programs adversely affect socioeconomic outcomes, such as economic growth, their adverse effects are mitigated where countries have a better record of compliance (Dreher 2006). Over-ambitious programs may also drain state capacity of developing countries without bringing about welfare gains (Reinsberg et al. 2018). Lack of compliance with IMF programs raises transaction costs by causing frequent re-negotiation to program content, while delays in disbursement may cause loss of market confidence and increased financial instability (Bird 2002, 838–39).

Against this background, it is important to understand why countries do not comply with their international policy commitments. Most of the literature focuses on recipient-country factors, such as a lack of political will and deficits in capacity, but neglects why political will is lacking and why countries entered programs with which they anticipated not to comply in the first place (Joyce 2006, 340). We argue that most IMF programs are unimplementable by design—that is, they include too many and excessively complex conditions that cause implementation problems, in turn requiring the Fund to scale back on its ambitions at a later stage of the program. For example, in 1999—the heyday of the structural adjustment era—countries under IMF programs faced an average of 26.3 conditions per year, but implemented only 24.2 of them.

Our theoretical argument explains why unimplementable programs persist. We conceive the Fund as a collective agent in which several departments have a stake in program design. Each of them wants to champion a particular set of policies that it considers to be critical to the achievement of its wider objectives—the repayment of the loan and, more importantly, the enactment of policy reforms by the borrowing government. However, any given department fails to consider the negative externality that its insistence on specific conditions imposes on the overall likelihood that the government will be able to fully implement program conditions. In other words, we consider the design of IMF programs as a ‘common pool resource problem’ (Ostrom 1990) among different bureaucratic actors.

Empirically, we introduce a new dataset which extends earlier data on IMF conditionality (Kentikelenis, Stubbs, and King 2016) by including information on the implementation record of all policy conditions agreed between the Fund and its borrowers between 1980 and 2009. The data

are based on original coding of the Letters of Intent and associated Memoranda of Economic and Financial Policies to identify individual program conditions, as well as Staff Progress Reports and Executive Board Specials to assess the extent of country compliance. The country-year coverage of the data well extends earlier datasets that have been used in IMF research (Ivanova et al. 2001; Mercer-Blackman and Unigovskaya 2004; Nsouli, Atoyan, and Mourmouras 2004).

As the first comprehensive researcher dataset on compliance with IMF policies, we explore our data descriptively to illustrate the most important trends in compliance, drawing on widely discussed measures in the relevant literature. We show that countries regularly implement fewer binding conditions than they originally commit to. The share of waivers—IMF decisions exempting countries from the need to implement specific conditions—has been increasing over the past three decades, particularly for so-called 'structural conditions' that require far-reaching reforms touching upon the role of the state in the economy and the institutional arrangements of the borrowing country (Easterly 2005). Moreover, of all 668 programs during 1980 and 2009, 414 programs were interrupted due to implementation problems, and 225 programs never resumed. Overall, these figures suggest that non-implementation of IMF programs is widespread.

We then deploy this data to test our argument that non-compliance may be the result of over-ambitious policy design. Using bivariate analyses, we find that the percentage of waived conditions positively correlates with the total number of binding conditions, the number of structural conditions, and the scope of conditionality in a given IMF program. Specifically the programs with the most controversial policy measures—privatization of state-owned enterprises, fiscal retrenchment, and public-sector reform—are more likely to fail. To provide a definitive test of our argument, we examine the determinants of implementation problems using multivariate analysis. Our primary measure of implementation difficulties is the proportion of waived conditions within a program. It is IMF staff who propose to the IMF Executive Board to waive conditions so that program review can be formally completed and the country can continue to obtain IMF funds. Waivers hence represent a 'best-case scenario' of implementation difficulties. Our secondary measure is temporary program interruptions, which indicate that a program has not even been sent to the Executive Board by IMF staff—oftentimes the result of more severe lack of compliance. Controlling for alternative explanations such as initial conditions, unexpected shocks, and global factors, we obtain a robust correlation between implementation problems and the number of conditions.

Our article also offers useful contributions to current debates. By demonstrating that non-compliance occurs even after considering initial conditions and unexpected shocks, we identify program design—which is at the discretion of the IMF bureaucracy and its powerful shareholders—as another determinant of compliance in addition to borrower-specific factors such as political will. Our findings challenge some previous work regarding the salience of different factors in compliance decisions (Ivanova et al. 2001, 4). For instance, contrary to earlier studies, political instability, bureaucratic quality, and trade openness are not robustly associated with compliance. While domestic politics, unexpected shocks, and geopolitics may determine compliance, these factors jointly do not account fully for the extent of non-compliance. Our work also addresses some under-researched questions, notably whether compliers differ from non-compliers and whether non-compliers face tougher initial circumstances (Bird 2002, 846).

Theoretically, by emphasizing over-ambitious program design, we explicate an under-researched channel for why countries often fail to comply with policy commitments. Our argument differs from the view that countries accept programs with which they can comply and the Fund assigns conditionality profiles that it expects to be complied with, but as economic circumstances deteriorate, borrowers may fail to implement some conditions (Bird 2002; Joyce 2006; Stone 2004). We are not the first to note that programs design matters for compliance decisions (Baqir, Ramcharan, and Sahay 2005; Bird 2002; Dawson 2002). However, we offer more refined predictions and systematic tests of the unimplementability hypotheses. By considering variation in non-compliance due to different types of IMF conditions, we also respond to calls for a disaggregated analysis of compliance (Vreeland 2006), addressing shortcomings of earlier work that relied on rough proxies.

Finally, our work contributes to a generalist IR literature on compliance, which invariably consider the state perspective for compliance (Chayes and Chayes 1993; Girod and Tobin 2016; Underdal 1998). While realists consider compliance as purely epiphenomenal, constructivists emphasize lack of capacity as a source of non-compliance, while rational design scholars intuit the existence of temporary derogations, information exchange, and dispute settlement as deliberate mechanisms that allow countries to comply again (Bagwell and Staiger 2005; Koremenos 2005; Rosendorff and Milner 2001). We complement this perspective by emphasizing the independent agency of international organizations to affect compliance decisions and rationalize over-ambitious policy design by emphasizing organizational incentives, notably the need for visibility and program success that are requested by their authorizing environments (Barnett and Coleman 2005; Selznick 1948; Weaver 2008). By emphasizing the substance of policies mandated by international organizations, our work also complements prominent IR research that scrutinizes organizational structure as drivers of compliance when state interests are fundamentally opposed to the prescriptions of international agreements (Mitchell and Hensel 2007).

2. A primer on IMF program compliance

Our point of departure is a country turning to the IMF to alleviate a balance-of-payments crisis. When a government requests IMF assistance, the respective area department prepares a blueprint of an adjustment program that considers key characteristics of the country and entails the timetable of disbursements and policy measures to which financial support is tied. The staff circulates this information as a briefing paper to other (non-area) departments to invite comments from them. A revised blueprint incorporating such comments is then submitted to IMF management for clearance, who evaluates the blueprint and authoritatively decides on the **prior actions** that borrowing countries need to implement before getting funding access (Mussa and Savastano 2000, 86–87).

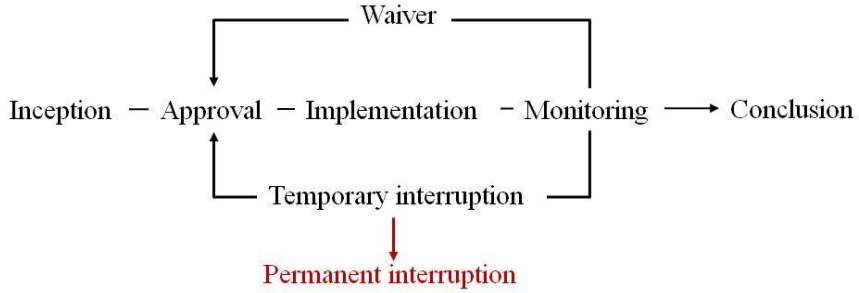
After clearance from management, an IMF mission visits the country to enter into negotiations with the government over the terms of agreement. The terms of a tentative agreement are enshrined in the Letter of Intent, in which the government details all policy measures that it commits to undertake in exchange for access to IMF credit. Mission staff circulates the letter of intent along with a report summarizing the technical features of the proposed agreement—program duration, phasing of tranches, and performance clauses ascribed to the tranches—to the

(non-area) departments again. A revised staff report is then submitted to management for clearance. Eventually, any tentative agreement—as reflected by a Letter of Intent and the staff report—requires approval by the Executive Board, which represents the collective interests of the member states. Board approval triggers disbursement of the first tranche, followed by additional tranches as the government demonstrates progress toward reform (Mussa and Savastano 2000, 88-90).

The subsequent phase in which IMF staff monitor compliance with the terms of an agreement is the most important one. The IMF staff assesses the progress of implementation periodically through so-called **program reviews** (Goldstein et al. 2003). These reviews are conducted quarterly for most lending facilities. Governments must comply with binding conditions to complete program reviews and trigger subsequent loan disbursements. Binding conditions include prior actions, structural performance criteria, and quantitative performance criteria. In contrast, failure to implement non-binding conditions—indicative targets and structural benchmarks—does not automatically lead to program interruptions (IMF 2001a).

Facing difficulties with implementing a binding condition, the government may approach IMF staff, which can then propose to the Executive Board to waive the respective condition. A waiver decision is the only way to avoid program interruption in the face of non-implementation of a binding condition. Failure to waive conditions—often the result of substantial deviations from performance criteria—leads to temporary interruption and the renegotiation of agreement conditions. If the government and the IMF cannot come to an agreement about how to get the program back on track, the program is permanently interrupted (Arpac, Bird, and Mandilaras 2008).

Figure 1: Flow chart of IMF program compliance process



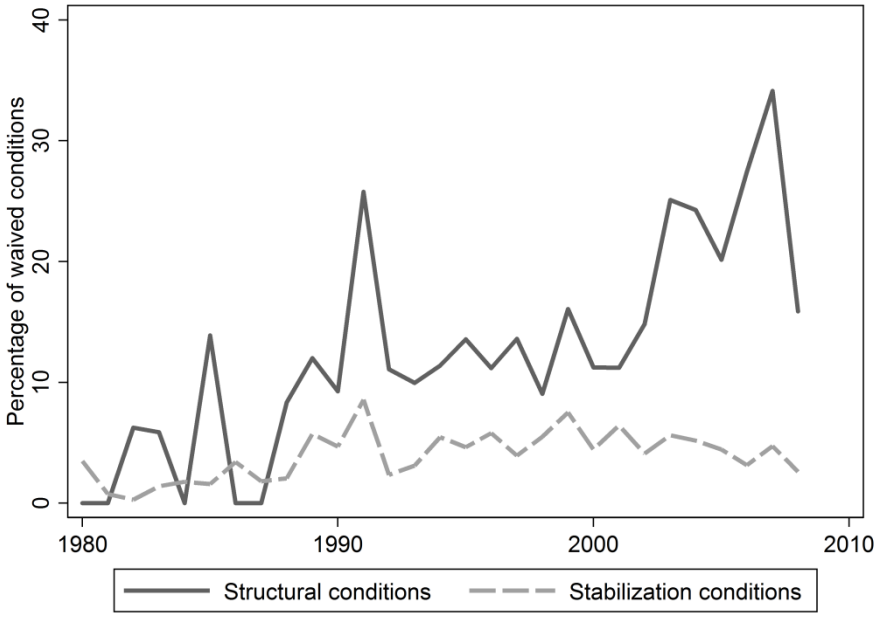
Waivers are one proxy for compliance difficulties. Occurrence of a waiver implies that the country failed to implement a binding condition—although absence of a waiver could still mean that the country did not implement a binding condition (in which case the program would be interrupted). In other words, the percentage of waived conditions is a lower bound of the true implementation deficit.

Figure 2 shows the percentage of waived conditions per program year has been steadily increasing over time.[1] There are also pronounced differences in waiving between structural conditions and stabilization conditions. The percentage of waived structural conditions has increased

[1] We calculate the percentage of conditions waived with respect to all originally applicable conditions to guard against the possibility that more waivers per program simply reflect a growing number of conditions in such programs.

tremendously, from as low as 5% in the mid-1980s to about 30% in the mid-2000s. Conversely, the percentage of waived stabilization conditions has been hovering at less than 5% over the 1980-2009 period, with slightly higher waiving intensity in the early and late 1990s.

Figure 2: Average percentage of waived conditions over time



Yet, waivers are a partial measure of compliance because only binding conditions have such data available, where non-implementation necessitates a waiver to be granted by the Executive Board, thereby ensuring there is always a trace of this in the relevant Board decisions. In contrast, non-binding conditions are modified by IMF staff without requiring approval from the Board and no systematic data on the implementation of these is available. In addition to inevitably discarding non-binding conditions, this method does not capture the case where an IMF program is off-track and reviews are not concluded because it does not even reach the Board to complete the review. Subtracting waivers from total conditions may also be an inappropriate inflection of the burden a condition carries given that waivers reflect not only non-implementation of a condition but also partial or delayed implementation.

Therefore, we also consider program interruptions as a proxy for non-compliance. While waivers reflect a mild form of non-compliance given that they allow IMF disbursements to continue, program interruptions are more serious because they indicate that the Board has not waived a missed condition. Program interruptions can be directly measured by examining borrowing countries' failure to complete reviews. Interruptions can be temporary, lasting a few months, or permanent. They are measured as the time lag between the initially agreed-upon review dates and the actual review dates (Mecagni 1999). However, a limitation of this approach is that it tells us little about what actually caused the review delay. While program interruptions most often occur as a result of failing to meet conditions, they can also be due to events that are extraneous to conditionality, such as administrative delays or changes in political leadership.

Following the approach adopted by IMF staff analyses (Ivanova et al. 2001; Mecagni 1999; Nsouli, Atoyan, and Mourmouras 2004), we coded temporary interruptions—a deviation from

program implementation that is subsequently corrected (i.e., the country gets back on-track with the program)—and permanent interruptions. An interruption is formally defined as a program review for a Stand-By Agreement delayed by more than 90 days or not completed at all; or a program review for an Extended Fund Facility, (Enhanced) Structural Adjustment Facility, or Poverty Reduction and Growth Facility program delayed by more than 180 days or not completed at all. The exception to this rule refers to programs that are cancelled and replaced with another, in which case non-completed reviews are not counted as interruptions. A permanent interruption occurs if the program does not resume (i.e., there is no subsequent review after the interruption event).

Table 1 shows that program interruptions are prevalent. Using IMF programs as unit of analysis, we find that about 62% of all programs become interrupted over their lifetime. This figure includes 28% of programs with at least one temporary interruption but that eventually get back on track, and 34% permanently interrupted programs (Table 1). These figures are consistent with previous research which, however, is based on much shorter time series (Bird 2002, 838; Ivanova et al. 2001, 7).

Table 1: Program interruptions.

	Total number	Percentage
All programs	668	100%
Interrupted programs	414	62%
… permanently interrupted	225	34%
… temporarily interrupted	189	28%

Notes: The set of programs excludes programs that cannot be interrupted, for example because they do not include conditionality; augmentations of existing programs; and programs after 2009. Programs can be interrupted either permanently or temporarily. Hence, a program that becomes temporarily interrupted but does not resume is considered as permanently interrupted here.

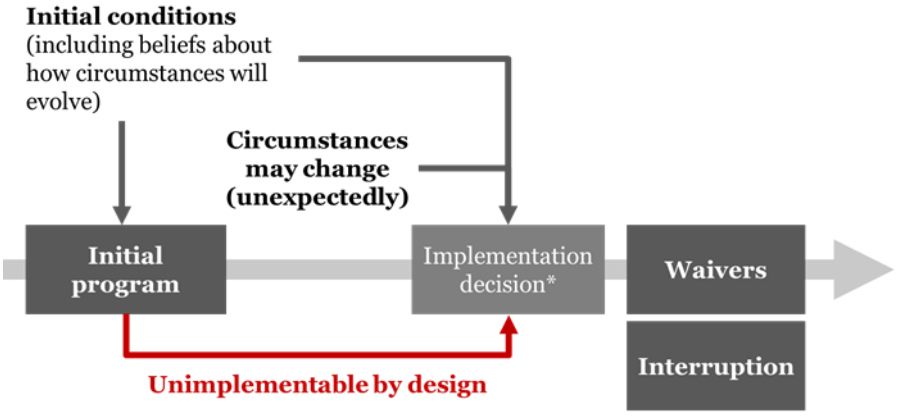
In sum, implementation failure is the norm rather than the exception—echoing findings of an earlier IMF staff report (Mussa and Savastano 2000, 94). We also find that implementation failure vis-à-vis waivers is becoming more prevalent – hence the pressing need to uncover what is driving non-compliance with IMF programs.

3. Initial hypotheses on non-compliance: Economic shocks, political constraints, and information asymmetries

Our conceptual framework depicts compliance as a three-stage process involving two actors (Figure 3). First, an IMF program is designed that provides a specified lending amount and a range of policy conditions that borrowing countries need to fulfill. Second, external economic shocks are realized that may affect country compliance with the IMF program, especially when such shocks were unanticipated. If the government fails to implement a binding policy condition, it will liaise with IMF staff who will then decide whether to request the Board to waive it so that the program can continue. Third, the Board decides whether or not to approve the waiver. If the Board waives all conditions that the government failed to implement, the program review is successfully completed and the program does not get interrupted. Failing that, the program is

(temporarily) interrupted, which halts the disbursement of loans. If actors cannot re-negotiate the terms of the agreement, the program gets permanently interrupted. In practice, the game is potentially infinite, because not only are programs paid out in several tranches, but also programs follow one another (Stone 2002). However, each round of the compliance game can be broken down into these three stages, and we ignore the long-term dynamic implications in this simple graph.

Figure 3: The compliance process in three stages.



This simple framework allows us to account for prior theories on compliance with IMF programs. For instance, some researchers have modeled compliance as a decision-making problem by the borrowing government, arguing that a government will comply as long as the marginal benefits exceed the marginal costs from doing so (Bird 2008; Joyce 2006; Stone 2002). This view takes as given the initial program design as the outcome of prior bargaining between the Fund and the borrower government. In this model, non-compliance arises for three reasons. First, as the program evolves, circumstances may change such that realized conditions are less favorable than initially expected. An adverse shock may cause governments to fail to implement certain conditions even if they intended to comply with them (Bird 2002). Other sources of implementation failure are economic crises, regime change, executive elections, unexpected levels of domestic opposition, and aid catalysis below prior expectations (Bird 2008; Dreher 2003). In essence, compliance depends on how well actors correctly predict the cost-benefit schedule over the lifetime of the program. Second, there is a possibility that a country never intended to comply fully, but only up to point where the net benefits of compliance become negative. Motivation is unobservable to the Fund but private information of the government. Third, implementation failure may also be the result of different cost-benefit evaluations by both actors. In particular, while the IMF takes into account compliance-related benefits beyond the program country, an individual country does not internalize these benefits (but faces the full costs of adjustment (Joyce 2006).

Some scholars have highlighted 'political constraints'—special interest groups with anti-reform preferences (Coate and Morris 2006; Drazen 2002; Mayer and Mourmouras 2005)—that often upset program implementation. In the above framework, however, such political constraints must be unexpected at the program design stage to really affect compliance. Some constraints are known to both parties and can be factored into program design. For instance, a democratic government that faces more domestic opposition may obtain a more lenient deal to begin with (Caraway, Rickard, and Anner 2012). Similarly, a country with little capacity might get fewer

conditions. This view inadvertently stresses the agency of the Fund in designing programs that provide enough benefits for countries to comply in view of anticipated implementation challenges.

The Fund as a 'strategic lender' will observe all these factors in its initial offer to enable borrowers to fully comply with the program (Stone 2002). However, borrower compliance is only one of the objectives that the Fund pursues. To be sure, the IMF values implementation of conditionality as it values market-friendly economic reforms. IMF staff also considers conditionality as a means to ensure that countries pay back their loans and to ration funding when many countries request access to its loans (Dreher 2009). When imposing conditionality, the IMF wants to appear credible in that it actually punishes non-implementation. This objective conflicts with the need to please powerful shareholders, who may demand more lenient treatment in terms of compliance (Stone 2002). Finally, the Fund also needs to prevent spillover of economic instability in a crisis-affected country to the global economy, especially when the crisis country is systemically important (Copelovitch 2010; Dreher 2009, 244; Joyce 2013).

In contrast to the Fund, governments want access to credit at the lowest possible price—that is, with few conditions attached. Due to the presence of domestic veto players, even a program with moderate level of conditionality faces risk of implementation failure—assuming that not all factors can be considered at the design stage (Stone 2004). Under asymmetric information, the government may have an additional advantage by being able to blame unobserved shocks for a non-compliance outcome. The IMF then is challenged to propose a welfare-maximizing program that satisfies a number of constraints: the incentive compatibility constraint (the government must be incentivized to implement reforms in exchange for lending), the political stability constraint (the program must ensure living standards that prevent civil violence even when economic conditions deteriorate), and the financial constraint (the Fund must be able to afford the loan from its budget and repayment should be likely). For a range of parameters, there may not be a program which simultaneously satisfies these constraints. The IMF then faces a tradeoff between a tightly budgeted program that the government implements fully but that does not deliver substantive reforms and a generous program that yields progressive reforms but which the Fund finds financially unsustainable (Mayer and Mourmouras 2005). Taking this argument further, we suggest there may be a program that is financially viable and that sets ambitious reform targets but the recipient fails to implement some of them, as further discussed below.

Following implementation failure of some conditions, the Fund thus faces a decision of how to punish such failure. For geopolitically important countries, the IMF faces a credibility issue in that it is less likely to punish poor implementation, for instance by granting waivers and by allowing countries to resume borrowing rather quickly (Stone 2002). In view of threats by powerful member states, the IMF accepts to suffer reputational damage by being more lenient to some recipients. As these recipients anticipate laxer enforcement, they are more likely to renege on their commitments. Hence, waivers are a double-edged sword: They are important for program flexibility in situations where exogenous circumstances make compliance difficult, but they are also potential mechanisms for political interference by allowing unpunished deviations from IMF conditionality (Pop-Eleches 2009, 793). Hence, in light of donor politics, it is difficult to uphold the claim that most countries negotiate in good faith. In the argument we develop below, we do not make this assumption.

4. An alternative hypothesis: Unimplementability by design

Complementing the above views, we argue that the Fund adopts over-ambitious policy designs—in terms of both the policy substance and the complexity of programs—that even a well-intended government may not fully implement under favorable circumstances. The view that the Fund sets over-ambitious targets is not new (Baqir, Ramcharan, and Sahay 2005; Bird 2008; Bird and Willett 2004). Even the IMF does not deny that its programs have long suffered from unrealistic expectations. For example, the former IMF External Relations Director explained that "though the expansion of structural conditionality was a largely appropriate response to changing circumstances, there is a sense that we may have gone a bit too far," asking for "too much, too soon" (Dawson 2003). Likewise, in its conditionality review, the IMF stated that "[…] conditionality may have been established on policies that were unlikely to be delivered, calling into question the realism of program design" (IMF 2001a). It has also recognized the dangers of over-ambition for country compliance and socioeconomic outcomes, given that "conditionality covering a broad range of policy areas may also place an enormous burden on limited administrative capacities in borrowing countries and may make it more difficult to focus on getting the most important things done" (Ahmed, Lane, and Schulze-Ghattas 2001). The same study also warns "that conditionality that is unrealistically ambitious [...] may result in repeated failure to meet agreed targets and foster a culture of nonperformance."

Official policy documents offer no convincing explanation for excessive conditionality, but attempt to provide assurance that the Fund has been successfully streamlining its conditionality practice after the Asian Financial Crisis (IMF 2001a). Most official statements treat periods of excessive conditionality such as around the Asian Financial Crisis as an historical accident. James Boughton—IMF historian reflecting on the evolution of conditionality—argued "there was an ambiguity that had crept into the whole process over time that was, I think people realize now, beginning to get out of hand" (IMF 2001b).

We do not find idiosyncratic explanations too convincing though. Adopting a rational choice perspective, we conceive the Fund as a collective agent in which several departments have a stake in program design. Each of them wants to champion a particular set of policies that it considers to be critical for 'program success', which entails the repayment of the loan and, more importantly, the borrowing government enacting policy reforms that the department cares most about. However, any given department fails to consider the negative externality that its insistence on specific conditions imposes on the overall likelihood that the government will be able to fully implement program conditions. In other words, we consider the design of IMF programs as a 'common pool resource problem' (Hallerberg and Marier 2004; Ostrom 1990; Potts 2017) among different bureaucratic actors. The administrative protocol by which IMF programs come about is consistent with this dilemma situation: Non-area departments—given their specific sector expertise for instance on legal affairs, tax matters, and monetary issues—hold *de facto* veto power over the design of IMF programs, and thereby can easily amend a proposed program as they see fit (Mussa and Savastano 2000, 89).

The outcome of the common pool resource problem is an IMF program with both too many conditions as IMF departments fail to coordinate on the number of conditions. Many of such conditions—reflecting the specific expertise and policy preferences of IMF sector departments—

are likely to face domestic opposition to reform. But departments have no incentives to water down the ambition of their policy conditions. They are aware that only ambitious programs will shift the balance toward one group in a divided society to unleash reform (Drazen 2002). In other words, the IMF is often said to 'tip the balance' in favor of reform (Vreeland 2003). This is also the prevailing view among IMF departments, which note that "establishing conditions covering a wide variety of policies and the detailed steps involved in each policy reform may effectively short-circuit national decision-making and undermine political institutions within the country" (Ahmed, Lane, and Schulze-Ghattas 2001). Given the desire of departments to be seen as effective promoters of reform, they are less likely to consider the external effects of their own policy advice that they include in IMF-supported programs.

Our collective-agency view differs from previous work that adopted a unitary-actor framework. To explain over-ambitious policy designs, the latter view must make different assumptions on what drives IMF staff and their principals. For example, an alternative view holds that IMF staff realize that (over-)ambitious programs are necessary to get the program approved by the Executive Board although their targets may be missed (Bird and Willett 2004). Over-ambitious design is thus linked to the myopia of powerful donor countries. However, most observers would not consider this to be a plausible assumption. Yet others emphasize that bureaucratic self-interest drives program design (Babb and Buira 2005; Dreher and Vaubel 2004; Vaubel 2006). A unitary Fund might be interested in mandating too many conditions because doing so provides it with increased discretion to waive conditions subsequently. To the extent that over-ambitious programs increase the number of performance criteria that a country will miss, staff gain discretionary authority to keep a program afloat (Babb and Buira 2005). Moreover, as the Board rubber-stamps virtually all IMF staff requests for waivers (Mussa and Savastano 2000, 90), this also implies more discretion of IMF staff vis-à-vis the Board.

A remaining puzzle we need to address in our argument is why governments knowingly accept over-ambitious programs. One might think that governments consider over-ambitious programs as too politically costly in relation to their expected benefits. Furthermore, overly ambitious programs do not serve as credible commitment device because market actors do not believe they will help the country resolve the crisis. In reality, however, countries have limited influence on the terms of an agreement. This is because all tentative agreements effectively are subject to ultimate approval by non-area departments, IMF management, and the Executive Board (Mussa and Savastano 2000, 88). Even if some countries can influence the terms of their agreements—for example because they are geopolitically important—they find it rational to agree on over-ambitious terms. Specifically, such countries will anticipate implementation failure of some conditions but expect the Fund will anyway waive such conditions at a later stage (Pop-Eleches 2009). This renders the net benefits from program participation positive again. Evidence that such strategy is consistent with beliefs abounds. James Boughton—talking about an IMF program with Indonesia that contained 117 conditions—said that "obviously, nobody expected Indonesia to fulfill all 117 of these promises. It was impossible, and everybody recognized it was impossible" (IMF 2001b). The Indonesian case is no outlier but seems to reflect a general policy approach. Timothy Geithner—then-director of the Policy Development and Review Department—asked almost rhetorically: "Has the Fund been too tough or too accommodating? Or [...] has it been both, by setting unrealistic aspirations for policy reform and then acquiescing to the inevitable failure of even relatively well intentioned governments to meet the bar? There is something to this [...]" (Goldstein et al. 2003, 442).

In sum, we have offered a rational explanation for why programs are over-ambitious and consequently run into implementation problems that the Fund must remedy by granting waivers. We relate this behavior to the incentives of the IMF staff and the operational protocol they follow. In particular, staff need to present their department as an effective engine of policy reform by insisting on specific policy conditions, but they do not internalize the effect of doing so on the overall implementability of a program. This leads to an excessive number of conditions that need to be waived later on to keep programs afloat. Our argument is based on the simple assumption of full information regarding borrowing country circumstances, rather than asymmetric information, where the Fund may use over-ambitious designs to screen out the uncooperative type of borrowers.

Our argument has several observable implications. Specifically, it implies that non-compliance—as a result of rational behavior—is the norm rather than the exception. The figures in the previous section are consistent with this expectation. In addition, we would expect that programs with many conditions are more likely to run into implementation problems. We thus should observe a positive correlation between the number of program conditions and the share of waived conditions. If the main alternative—that programs are designed to be implementable unless unexpected shocks occur after initiation—were to hold, we would find no effect of program design variables on waivers after controlling for such shocks.

5. Empirical evidence

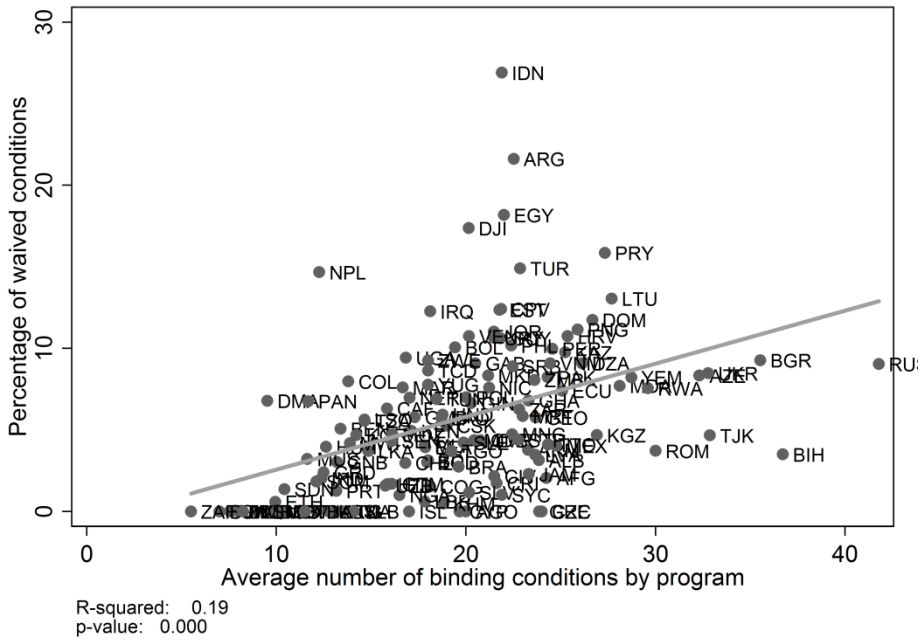
In the following, we first conduct bivariate analysis to test some empirical observations of our theoretical argument. To adjudicate between our theory and its main alternative, we then conduct multivariate analysis using well-specified models.

Co-evolution of adjustment burden and implementation failure

Following a disaggregated approach to compliance, we first examine the incidence of waivers as a proxy for compliance difficulties. Occurrence of a waiver implies that the country failed to implement a binding condition—although absence of a waiver could still mean that the country did not implement a binding condition (in which case the program would be interrupted). In other words, the number of waived conditions is a lower bound of the true implementation deficit. As further discussed in the appendix, the waiver percentage is the most pertinent measure of implementation failure.

Are more ambitious programs less implementable judging by the number of waivers to remedy implementation failure? Figure 4 plots the percentage of waived conditions against the average number of conditions that a country had to implement across its programs. In support of our argument, we find a positive correlation. This result is consistent with the view that more burdensome programs generate subsequent implementation problems, which in turn require a formal waiver to keep programs alive. The figure is qualitatively similar when using the average number of structural conditions rather than the total number.

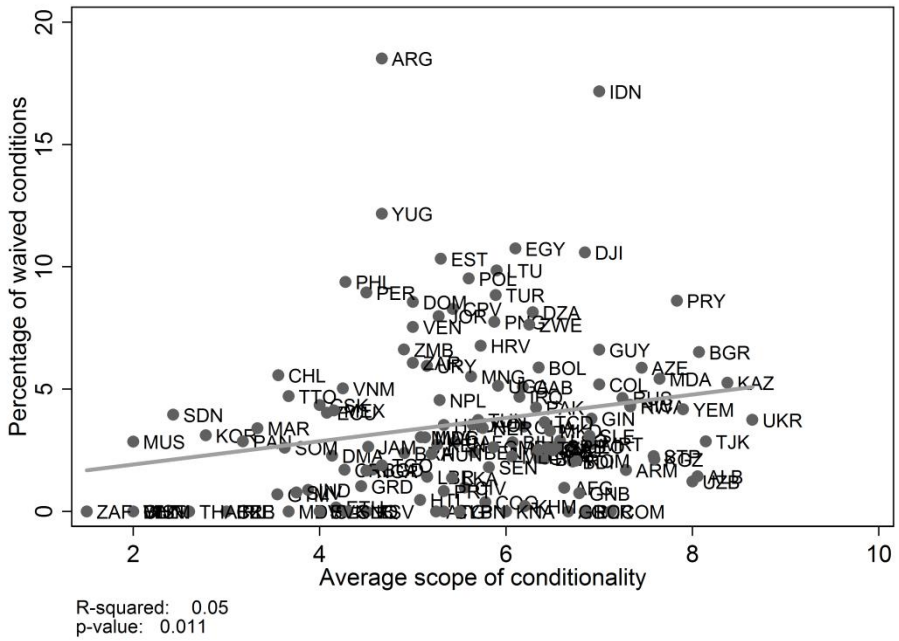
Figure 4: Waiver percentage and burden of adjustment.



Our use of percentages of waived conditions guards against the possibility that the increased incidence of waivers is simply an artifact of increasing numbers of conditions. Drawing on our earlier result on the growing use of waivers—notably for structural conditions—we find that the IMF has become more penitent over time, in awareness of the failures of its various conditions. The IMF has come to waive 25% of all structural conditions, in contrast to stabilization conditions for which the waiver percentage has been stable at around 5% (Figure 2). This shows that structural conditions, rather than stabilization conditions, make IMF programs unimplementable, and the Fund—having long underestimated implementation difficulties—has increasingly embraced waiving as a means to avert implementation failure.

We obtain similar results when plotting waivers against the average scope of conditionality—the number of distinct issue areas in which countries are required to implement reforms. IMF programs with larger scope are more complex and thus more likely to run into implementation problems, which in turn may lead to waivers. We again find a positive relationship between how burdensome the average country program is and the extent to which its conditions are waived (Figure 5).

Figure 5: Waivers and scope of conditionality.



We also scrutinize issue areas, for which the emerging picture is not as clear as expected. For most issue areas, waiver percentages were highest in the period immediately following the Cold War, but for some issues, this percentage has increased further in the post-Millennium era, including public-sector conditions, privatization, and price deregulation. Furthermore, the most controversial issue areas—fiscal policy, privatization, and public-sector reform—tend to be related to a higher waiver percentage, but the differences are small.

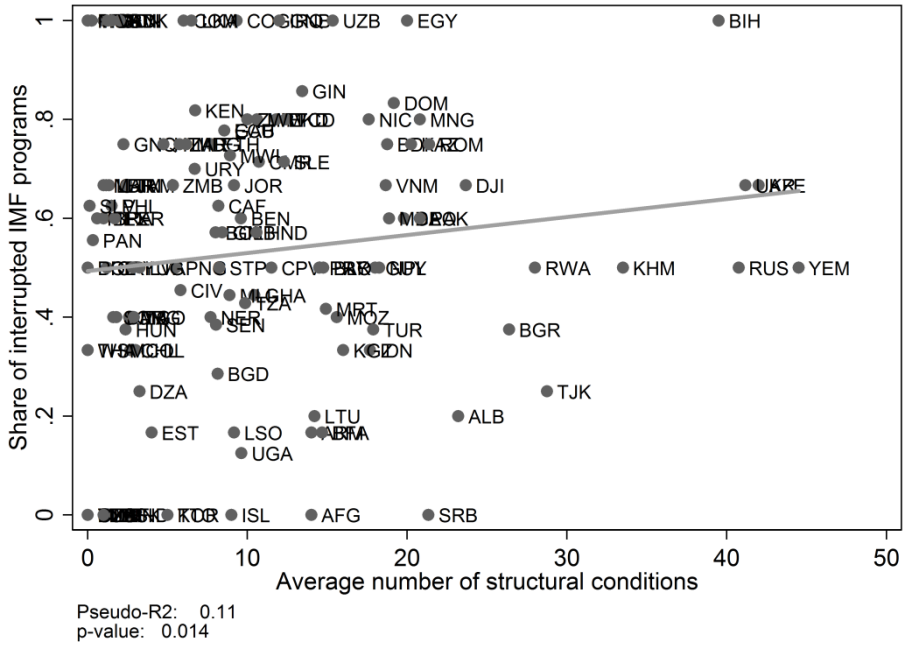
Table 2: Waiver percentages by issue area and time period.

	1980-1989	1990-1999	2000-2009
Debt issues	2.06	5.41	3.18
External sector	3.65	6.04	3.34
Financial sector	2.49	5.71	4.39
Fiscal policy	2.77	6.65	4.40
Institutional conditions	--	2.12	3.40
Labor conditions	--	3.84	5.35
Privatization	--	5.30	6.51
Revenue issues	1.84	3.64	3.70
Price deregulation	0.83	4.64	5.66

Do these result hold when looking at program interruptions? As discussed before, interruptions also reflect the willingness of the Fund to keep programs alive despite missed targets. An interruption can only be averted if the Fund waives all binding conditions that the recipient country failed to implement. In Figure 6, we find a positive correlation between the number of structural conditions and the share of interrupted programs for a given country. Similarly, we find positive associations between adjustment burden and frequency of interruptions for key sectors of structural reform such as privatization of state-owned enterprises, price liberalization, and public-

sector employment. In contrast, there is no statistical relationship between program interruptions and the total number of conditions, and even a negative relationship with stabilization conditions.

Figure 6: Structural conditions and program interruptions.



A closer look at the types of programs that fail shows that implementation failure is related to the prevalence of policy conditions in controversial areas. Table 3 conducts t-tests on the differences in the types of conditions mandated across interrupted programs and non-interrupted programs. The fact that failure is unrelated to the number of conditions in areas like revenue generation, fiscal policy, and financial services, but strongly related to structural conditions more generally, and prior actions, privatization, and public-sector reform more specifically, suggests that unimplementability is a direct consequence of over-ambitious program design involving politically costly reform measures.

Table 3: Features of interrupted programs.

	Interrupted programs	Non-interrupted programs	p-value for difference
Binding conditions	55.5	51.9	0.270
Structural conditions	10.8	8.3	0.029
Stabilization conditions	44.6	43.5	0.649
Prior actions	7.4	5.0	0.059
Privatization	2.4	1.6	0.008
Public sector	1.4	0.7	0.004
Fiscal policy	7.5	7.2	0.745
Financial sector	14.9	13.9	0.242
Revenue issues	2.4	2.2	0.610

Notes: t-test with unequal variances. Interrupted programs last longer by about 2 months ($p<0.024$) but are otherwise not different, for example in terms of SDR amount ($p<0.160$).

Our inference arguably is limited by relying on bivariate correlations. Furthermore, unlike the existing theories, our theory makes no predictions regarding variation of compliance across different types of countries and different political-economic circumstances. In the following, we assess to what extent our outcomes of interest vary across different circumstances. Are compliant countries different from non-compliant countries? And does the design of IMF programs vary with initial country circumstances? While these questions still remain to be addressed in the literature (Bird 2002, 846), the answers to them allow us not only to probe existing theories but also to identify important control variables for our subsequent multivariate test of our argument.

Our simple comparisons yield interesting insights (Table 4). In particular, we find that economic fundamentals at the time of program initiation do not matter significantly for program design. In fact, they matter more at the compliance stage. For instance, countries with FDI inflows above the sample mean—which receive somewhat more structural conditions in their programs—have significantly more waivers and interrupt programs more frequently. Conversely, program countries with a more favorable balance of payments are less likely to interrupt, although this is obviously endogenous to IMF treatment.

In terms of domestic politics, we find that democracies obtain more conditions, but also more waivers. Furthermore, while more veto players are positively related to the number of conditions, there is no significant relationship with compliance outcomes. We also highlight that state capacity—proxied by ICRG bureaucratic quality—is unrelated to all outcome variables studied here. The same result holds when using the State Capacity Index (not shown here). Interestingly, countries do not get more lenient programs when a civil war is raging but obtain more waivers to remedy implementation failure.

International politics matters a great deal for program lending, as suggested by previous seminal work (Dreher, Sturm, and Vreeland 2009; Stone 2004; Thacker 1999). For instance, temporary UNSC members who serve the institution at the time of IMF program initiation face around eight conditions less—a sizeable effect. However, they do not obtain more favorable treatment at the implementation stage—as probably their service has ended then. This interpretation is confirmed by the positive effect of UNGA voting alignment with the G7: While G7 allies face more conditions, they also obtain significantly more waivers during implementation. Further, foreign aid has no impact on design variables and most compliance variables but helps countries avert permanent interruptions.

Finally, IMF bureaucratic politics significantly affect lending programs. Bigger loans in terms of country income contain more conditions but also are positively related to interruptions. Longer programs are positively associated with all lending stages. Interestingly, first-time borrowers receive substantively fewer conditions, but they are not different in terms of compliance. Finally, when IMF resources are relatively stretched—when the Fund is assisting many countries at the same time—it requests more conditions but also waives more.

In sum, these simple comparisons strengthen the view of the Fund as a ‘strategic lender’, which must cater to the interests of important shareholders but which also enjoys discretion in designing programs (Lang and Presbitero 2017). The findings for waivers indicate that the Fund responds to a mix of factors including political instability, geopolitical alignment, and program characteristics.

Table 4: Outcomes of interests under varying country circumstances.

	Design stage		**Compliance stage**		
	All conditions	Structural conditions	Waiver percentage	Temporary interruptions	Permanent interruptions
Economic fundamentals					
GDP per capita	-0.782	0.601	-0.027	-0.052	0.070
Current account	2.810	0.421	-0.676	-0.095*	-0.023
Debt service	0.809	-0.171	0.310	-0.039	-0.020
Reserves	-0.020	1.040	0.951	0.060	0.020
FDI/GDP	1.733	1.826*	2.950*	0.153*	0.038
Trade openness	1.173	0.827	0.450	0.008	0.039
Domestic politics					
Democracy	3.691*	0.958	1.385*	0.064	-0.015
Veto players	3.884*	1.410*	1.258	0.043	0.002
Left-wing government	-0.563	-0.594	-1.273	-0.016	0.009
Fractionalization	0.685	2.080*	0.970	-0.068	0.040
Bureaucratic quality	-0.497	0.663	-1.847	-0.056	-0.005
Civil war	0.861	-0.755	2.688*	-0.010	-0.057
International politics					
UNSC member	-7.796*	-1.692*	2.857	-0.136	0.006
UNGA alignment	6.611*	3.897*	2.325*	-0.019	0.019
ODA per capita	0.571	-0.382	0.140	-0.005	-0.188*
Political openness	1.629	0.828	0.373	0.004	-0.015
Bureaucratic politics					
IMF quota/GDP	3.595*	1.912*	0.133	0.095*	0.002
Program duration	1.965	2.081*	1.803*	0.109*	0.116*
First-time borrower	-8.919*	-1.787*	-0.788	-0.064	-0.068
Countries under Fund	11.888*	5.581*	4.543*	0.057	0.010

Notes: Cell entries show the difference in means for above-average group and the below-average group with respect to the criterion shown in the row header. All comparisons are based on individual programs (N=668). All continuous variables are measured in the year of program initiation and are dichotomized for group mean comparisons. Binary variables that did not need to be dichotomized include democracy, left-wing government, civil war, UNSC member, and first-time borrower.

Significance level: * $p<0.05$.

Multivariate analysis

Previous studies have examined the determinants of compliance using multivariate analysis. Their results are difficult to compare due to different samples, measures of compliance, and econometric methods. Most studies consider compliance across entire programs. For instance, Killick (1995) conducts probit analysis of failed programs proxied by a disbursement rate threshold, finding that higher debt, lower exports, and bigger loans are related to failure. In contrast, Joyce (2006) finds disbursement ratios to be related to democracy, trade openness, length of tenure, and ethnic fractionalization. Others consider program interruptions a more valid measure of implementation problems and find that these are more likely under circumstances of trade openness, the number of veto players, and the loan-to-quota ratio (Arpac, Bird, and Mandilaras 2008). Only two studies use a different unit of analysis, primarily to accommodate shorter time horizons. Dreher (2003) finds that programs break down specifically before elections, while Stone (2004) studies the length of punishment intervals after programs have gone off-track. Table A1 in the supplemental appendix provides an overview of previous studies.

Multivariate analysis provides a definitive test of our argument because it allows us to include covariates that represent competing accounts of compliance. Our primary dependent variable is the percentage of conditions waived in a given program. Complementary outcomes are binary indicators respectively for temporary program interruption and permanent interruption. Our two main predictors relating to program design include the total NUMBER OF CONDITIONS and the number of STRUCTURAL CONDITIONS in a given program. As previously discussed, we only consider binding conditions, because failure to implement binding conditions interrupts scheduled disbursements of IMF loans (Copelovitch 2010; Reinsberg et al. 2018; Woo 2013).

When using waiver percentages as dependent variables, we employ Ordinary Least Squares regression. To model the likelihood that a program gets interrupted, we use logistic regression. We cluster standard errors on program countries. As our unit of analysis is the **IMF program**, we pool all observations of a given program. In alternative model estimations, we also include country-fixed effects to control for unobserved heterogeneity.

We rely on well-specified models to arbitrate between alternative models of compliance. Our first set of control variables gauge (unexpected) changes in the environment which occur between program initiation and the first incidence of program failure.[2] Considering the global financial environment, we include the percentage change in the US interest rate given that re-financing becomes more difficult under higher US rates. The data is from the Global Financial Development Database (World Bank 2018). We also include a dummy indicating a financial crisis (Laeven and Valencia 2013). In addition, we include percentage changes in standard macroeconomic variables that the Fund closely monitors—current account balance, foreign aid inflows, reserves, and debt service as of GNI (World Bank 2015). Regarding political variables, we measure the absolute difference in the Polity IV score (Marshall, Jaggers, and Gurr 2010), a dummy for executive elections (Beck et al. 2001), and the percentage change in the number of veto players (Henisz 2002). Data sources and descriptive statics for all variables can be found in the appendix (Table A2).

[2] If the program is not interrupted, we measure changes between the year of program initiation and the year in which it was concluded.

Second, based on our comparison of compliant countries and non-compliant countries, we also need to control for observable confounders which induce spurious correlation between program design and subsequent compliance, for instance FDI inflows, UNGA alignment, program size, program duration, and countries under IMF assistance. In principle, we do not need to include country features that are correlated with compliance but not with design as these do not confound our relationship of interest, but doing so is no harm and can improve efficiency.

Third, the thorniest inferential challenge may be omitted-variable bias—the same factors that cause non-compliance may also underlie bad designs and thus cause a spurious relationship. To use a classical example, 'political will' is unaccounted for by the Fund when designing a program but affects country compliance (Vreeland 2003). For our main analysis, we do not consider this potential source of bias for two reasons. First, our models are well-specified as they include a wide array of observable characteristics (which correlate with the remaining unobservables). Second, omitted variables that also confound the relationship of interest are unlikely to exist: the Fund is in control of program design (and unable to consider unobserved variables), while the recipient makes a decision to comply. It is thus unlikely that there is a direct relationship between design and compliance induced by unobservables. In robustness tests, however, we examine an instrumental-variable design to consider potential endogeneity of IMF conditions, but as the results are unaffected, we favor a simpler approach for our main analysis.

Our regressions provide overall support for our argument. Table 5 shows results for regressions of waiver percentages on the number of conditions. We find a robust positive correlation between program conditions and compliance failure that is unaffected by varying sets of controls and the inclusion of country-fixed effects. In substantive terms, Model 1 suggests that an additional 42 conditions (the standard deviation of this variable) are related to an increase in waivers by 2.35 percentage points ($p<0.01$). When adding the most likely confounders in model 2, this effect drops to 1.64 percentage points ($p<0.01$). When controlling for other potential determinants of waivers, we recover roughly the original coefficient size. Finally, changing circumstances do not mute this effect. Across all models, only a few variables are comparably robust than the number of conditions. In particular, countries under programs at the time of approval is positively related to waiver intensity, suggesting a more lenient IMF behavior toward borrowers when global circumstances are dire. Furthermore, longer agreements are related to a lower waiver share, which may reflect stronger IMF resolve in the shadow of the future. Turning to circumstances changing during the agreement, we find that higher US interest rates imply lower waiving percentages (contrary to expectations). An improving current account balance reduces the need for waivers, while debt service growth is encountered with more generous waiving.

Table 6 repeats the analysis for waiver percentages but examines the relationship with structural conditions. Reflecting the common wisdom that these conditions are less implementable, we find a substantive stronger correlation with non-compliance. An increase by 15 conditions (the standard deviation of this variable) is related to an increase in waivers by at most 2.67 percentage points ($p<0.01$). The effect is fairly robust across all models shown. Control variables generally behave as in the previous analysis but are less robust in terms of statistical significance.

Table 7 scrutinizes the determinants of temporary interruptions of IMF programs. Again we find robust positive associations with conditionality, most of them highly statistically significant ($p<0.01$). When adding a set of initial conditions, we find only IMF quota to be positively related,

indicating that countries in bigger trouble (and requiring more lending per capita income) interrupt more often. When adding changing circumstances, some initial conditions become weakly statistically significant in the anticipated directions; and higher debt service reduces interruptions (as the IMF is granting more waivers then), as does political liberalization. The latter theory is consistent with theories emphasizing the commitment value of democratic institutions.

We have also examined the determinants of permanent interruptions, but do not find any robust association with conditionality-related variables. We are not surprised by this outcome. Permanent interruptions are a less appropriate—if not invalid—measure of compliance problems because they imply that the Fund and the authorities have not reached an understanding of how to get the program back on track. These special situations are not representative of the compliance decisions that we wish to examine here. In fact, we find a negative relationship between the number of conditions and permanent interruptions, which is due to our coding decision that interruptions can be either temporary *or* permanent but not both (and we found a positive effect on temporary ones). When we condition on the waiver share—itself negatively related to permanent interruptions—the number of conditions become insignificant.

We conduct two additional tests to probe the robustness of our main findings. First, we explore how different sets of control variables affect the relationship of interest. We find our results to be remarkably robust. Second, we attempt to instrument the number of conditions as third variables (such as 'political will') might induce a spurious relationship between the percentage of waivers and the number of conditions. Our first-stage model includes the number of countries simultaneously under IMF programs, which provides a plausibly exogenous source of variation for the number of conditions with respect to waiving. The underlying logic is that under times of high demand for its resources, the Fund imposes more conditions, thus raising the 'price' of obtaining access to a scarcer resource. We indeed find this variable to be highly correlated. To further improve the accuracy of prediction, we include variables such as first-time borrower, prior interruptions, IMF quota, loan-to-quota ratio, and the logged number of US troops. Using a two-stage model does not affect our main results, as there remains a strong correlation between the number of conditions and the waiver percentage.

Overall, our various tests warrant the conclusion that over-ambitious programs are unimplementable by design. Hence, there is a robust relationship between the number of conditions and the degree to which countries fail to comply with them. This finding does not invalidate other explanations for non-compliance—such that important countries obtain more lenient treatment (Stone 2004) and unforeseen events may lead to adaptations in the program design (Bird 2008). Our argument is complementary to these views as it does not hinge on cross-country comparisons but the overall tendency for programs to have many conditions along with many waivers indicating implementation failure. Hence, our failure to find robust support for alternative theories does not necessarily suggest that these are wrong, but rather that our chosen research design makes robust findings unlikely.

Table 5: Waiver percentage and total number of conditions.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of conditions	0.056***	0.039***	0.059***	0.042***	0.055***	0.042***	0.073***	0.054***
	(0.006)	(0.009)	(0.012)	(0.012)	(0.009)	(0.012)	(0.017)	(0.015)
Democracy		0.334	1.199	-0.786		-0.972	-0.373	-1.578
		(0.819)	(0.903)	(0.861)		(0.919)	(1.390)	(1.050)
UNGA alignment		-0.765	-7.153	-6.918		-3.848	-13.742*	-3.464
		(4.724)	(6.002)	(7.421)		(5.602)	(7.645)	(8.595)
Countries under programs		0.167***	0.139***	0.205***		0.144**	0.097	0.200***
		(0.044)	(0.046)	(0.053)		(0.060)	(0.061)	(0.056)
FDI inflows			0.004				0.016	
			(0.008)				(0.032)	
Civil war			2.107				-3.534	
			(2.183)				(2.191)	
Agreement duration			-0.097*				-0.178**	
			(0.053)				(0.071)	
US interest rate (Δ)				-1.974**				-1.986*
				(0.907)				(1.058)
Financial crisis (Δ)				0.756				1.033
				(1.137)				(1.426)
Current account balance (Δ)				-0.023*				-0.046***
				(0.013)				(0.016)
ODA inflows (Δ)				0.056				0.087
				(0.089)				(0.065)
Reserves (Δ)				0.256				0.280
				(0.273)				(0.372)
Debt service (Δ)				0.039				0.090***
				(0.033)				(0.032)
Polity IV (Δ)				0.119				-0.002
				(0.129)				(0.106)
Executive elections				-0.411				-0.098
				(0.585)				(0.833)
Political constraints (Δ)				-1.001				-0.356
				(2.204)				(2.491)
Fixed effects	no	no	no	no	yes	yes	yes	yes
r2	0.09	0.12	0.14	0.23	0.31	0.35	0.44	0.47
N	640	569	390	318	640	569	390	318

Notes: * p<.1 ** p<.05 *** p<.01.

Table 6: Waiver percentage and structural conditions.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Structural conditions	0.148***	0.120***	0.127***	0.178***	0.166***	0.163***	0.161***	0.162**
	(0.022)	(0.034)	(0.035)	(0.049)	(0.030)	(0.049)	(0.049)	(0.065)
FDI inflows		0.009	0.007	0.063		0.035	0.034	0.039
		(0.009)	(0.009)	(0.127)		(0.039)	(0.039)	(0.150)
UNGA alignment		-4.971	-4.479	-6.966		-9.447	-9.599	-6.033
		(5.968)	(5.908)	(10.014)		(7.388)	(7.400)	(11.188)
Agreement duration		-0.051	-0.037	-0.094*		-0.112*	-0.107*	-0.161**
		(0.042)	(0.043)	(0.049)		(0.058)	(0.057)	(0.068)
Countries under programs		0.150***	0.147***	0.203***		0.106	0.107	0.223***
		(0.050)	(0.049)	(0.065)		(0.067)	(0.068)	(0.077)
Democracy			1.425				-0.195	
			(0.927)				(1.528)	
Civil war			2.564				-3.382*	
			(2.288)				(1.985)	
US interest rate (Δ)				-2.358**				-2.755**
				(1.070)				(1.288)
Financial crisis (Δ)				0.861				1.008
				(1.721)				(2.292)
Current account balance (Δ)				-0.037				-0.068**
				(0.032)				(0.027)
ODA inflows (Δ)				0.082				0.062
				(0.077)				(0.051)
Reserves (Δ)				0.217				0.380
				(0.398)				(0.590)
Debt service (Δ)				-0.109				0.114
				(0.193)				(0.286)
Polity IV (Δ)				0.181				0.152
				(0.150)				(0.131)
Executive elections				-0.440				-0.440
				(0.784)				(1.166)
Political constraints (Δ)				-0.190				1.086
				(2.810)				(3.208)
Fixed effects	no	no	no	no	yes	yes	yes	yes
r2	0.08	0.11	0.13	0.25	0.32	0.42	0.43	0.53
N	640	395	390	207	640	395	390	207

Notes: * p<.1 ** p<.05 *** p<.01.

Table 7: Determinants of temporary interruptions.

	(1)	(2)	(3)	(4)	(5)	(6)
Number of conditions	0.011***	0.013***	0.014**			
	(0.002)	(0.004)	(0.006)			
Structural conditions				0.023***	0.027***	0.026**
				(0.006)	(0.010)	(0.013)
IMF quota / GDP		0.062***	0.028		0.064***	0.041
		(0.019)	(0.073)		(0.015)	(0.070)
Current account balance		-0.016	-0.031*		-0.013	-0.026
		(0.016)	(0.016)		(0.014)	(0.017)
FDI inflows		-0.016	-0.051*		-0.011	-0.036
		(0.017)	(0.031)		(0.016)	(0.028)
Agreement duration		-0.012	-0.018		0.000	-0.005
		(0.014)	(0.019)		(0.012)	(0.017)
US interest rate (Δ)			-0.268			-0.273
			(0.380)			(0.360)
Financial crisis (Δ)			-0.822			-0.798
			(1.130)			(1.141)
Current account balance (Δ)			-0.009			-0.011
			(0.009)			(0.009)
ODA inflows (Δ)			0.004			0.001
			(0.021)			(0.022)
Reserves (Δ)			-0.097			-0.110
			(0.119)			(0.112)
Debt service (Δ)			-0.102*			-0.103**
			(0.053)			(0.050)
Polity IV (Δ)			-0.098**			-0.093**
			(0.042)			(0.042)
Executive elections			0.025			0.098
			(0.263)			(0.250)
Political constraints (Δ)			1.364			1.059
			(1.323)			(1.339)
Fixed effects	no	no	no	no	no	no
r2_p	0.04	0.06	0.10	0.02	0.05	0.08
N	658	331	166	658	331	166

Notes: * p<.1 ** p<.05 *** p<.01.

6. Conclusions

This article developed a framework for analyzing the politics of compliance with IMF-mandated policy conditions. We argued that IMF programs to a considerable degree are 'unimplementable by design'. In other words, they contain too many conditions that require far-reaching reforms that most countries find themselves incapable of implementing. Therefore, the IMF often authorizes to waive conditions to keep programs afloat.

We explain over-ambitious program design as the result of coordination failure among various IMF departments involved in the design of adjustment programs: Each sector department adds a maximally-inclusive set of policy conditions that it cares about to the final loan package, failing to account the negative externality that it imposes on the overall implementability of the program. The result of this common-pool resource problem is a program that contains too many conditions and the Fund must waive some conditions to address the inevitable implementation difficulties that countries face once program reviews are due. Individual departments have no incentive to back down on their policy demands as their insistence can help underpin their reputation as effective engine for reform in the borrowing country. But the reputational costs of inconsistent application of IMF conditionality is borne by the Fund as a whole—rather than individual departments—which implies that being over-ambitious in the design of conditionality is individually rational. Borrowing governments—in the rare cases in which they can influence program design—choose not to do so as they anticipate this behavior and hence initially commit to program conditions that they expect not to fully comply with at a later stage.

Our argument complements existing views arguing that countries negotiate in good faith and the Fund insists on full implementation initially, but unexpected negative shocks require modification of program conditions and thus the granting of waivers. Similarly, the rational lender view suggests that all observable characteristics—macroeconomic circumstances, domestic politics, global factors, and expected future realizations of all these variables—determine the initial program design, which implies that no adjustments to program conditions need to be made if these expectations are true. While this view suggests that non-compliance is a rare event, the opposite is true—most programs break down and the Fund waives at least some portion of conditions in most programs. This also implies that researchers need to be cautious to base their findings not only on the formal terms of agreements alone. Neglecting compliance gives an incomplete—if not misleading—picture of global politics and may overestimate the authority of international organizations in promoting domestic policy change. Our article thus provides a fairly accurate picture of the extent to which international organizations affect national policies, given that governments do not always implement all conditions to which they initially committed.

Some researchers—particularly organizational sociologists—have long noted excessive conditionality in IMF programs, arguing that these programs to follow a boilerplate model which seeks to roll out market-liberalizing policy prescriptions across all major sectors and in all borrowing countries—without due consideration of local circumstances (Babb 2013; Easterly 2005; Stiglitz 2002). Constructivist accounts of IMF program designs would emphasize the preference among IMF staff for ideological consistency (Chorev and Babb 2009) and—relatedly—organizational culture and a homogenous staff corps (Chwieroth 2014; Nelson 2017; Seabrooke 2010). We complement these views with a rational-choice approach that demonstrates why over-ambitious programs under collective agency are practically inevitable.

We also introduce new data on compliance with IMF conditionality. It extends existing data sources—such as the IMF MONA database (Arpac, Bird, and Mandilaras 2008)—with respect to programs (668 programs) and time frames (from 1980 to 2009) covered. We use the data to cast light on compliance patterns over three decades. We also examine bivariate correlations between program design and compliance performance to show that over-ambitious programs more often lead to program interruptions and the waiving of unimplementable conditions. Our definitive test of the unimplementability hypothesis relies on multivariate analysis, in which we control for initial conditions and unexpected changes in external circumstances. Our main result is unaffected by the inclusion of these variables, suggesting that our theory is complementary to existing theoretical arguments on compliance.

Before discussing the implications of our research, we note a few limitations of our study. First, some aspects of the compliance process remain unobserved. In particular, we do not know exactly the set of conditions that governments failed to implement. While the presence of waivers implies that the government must have failed to implement the respective conditions, we do not know the conditions that governments failed to implement but for which the Fund granted no waiver (while program interruption as the necessary outcome again is observable). Second, our results remain subject to omitted-variable bias due to unobservables. While we have controlled for a battery of observed confounders—country characteristics and environmental factors affecting program design and compliance decisions—unobserved factors may still be present. We leave it to future research to address this challenge. In addition, future research should study more closely the relationship between various compliance decisions. For instance, waivers—if granted exhaustively to cover non-implemented conditions—can avert program interruption. But when does the Fund waive all conditions in full, and when does it only waive some conditions? In addition, there is also limited research into the issue of what factors help programs get back on track once they have been interrupted (Stone 2004).

In addition to its theoretical contribution, our results also update the existing empirical record of compliance research with IMF conditions. Much of the earlier work is based on disbursement ratios, which is a rather poor proxy of country compliance. Therefore, we may not expect that correlates of disbursement—short-term debt, loan size, political constraints, initial GDP per capita, government consumption, democracy, and upcoming elections (Dreher 2003; Joyce 2006; Killick 1995)—also relate significantly to waivers. Indeed, except for democracy, we find none of these factors to affect compliance consistently.

Finally, our results have implications for policy-makers who aim to enhance the effectiveness of conditional lending assistance. Specifically, we find that programs are interrupted due to over-ambitious conditionality. If non-interruption is indeed crucial for program success, it would be important to reduce the incidence of program failure. Since over-ambitious program designs are the result of coordination failure among individual departments, our theory suggests that member states need to strengthen the veto power of those parts in the IMF bureaucracy that articulate the collective interests of the Fund and its borrowing countries. Organizational reform may therefore be an important step toward enhancing the effectiveness of IMF lending.

References

Abbott, Philip, Thomas Barnebeck Andersen, and Finn Tarp. 2010. "IMF and Economic Reform in Developing Countries." *Quarterly Review of Economics and Finance* 50(1): 17–26.

Ahmed, Masood, Timothy Lane, and Marianne Schulze-Ghattas. 2001. "Refocusing IMF Conditionality." *Finance and Development* 38(4): 40–43.

Arpac, Ozlem, Graham Bird, and Alex Mandilaras. 2008. "Stop Interrupting: An Empirical Analysis of the Implementation of IMF Programs." *World Development* 36(9): 1493–1513.

Babb, Sarah L. 2013. "The Washington Consensus as Transnational Policy Paradigm: Its Origins, Trajectory and Likely Successor." *Review of International Political Economy* 20(2): 268–97. http://dx.doi.org/10.1080/09692290.2011.640435.

Babb, Sarah L, and Ariel Buira. 2005. "Mission Creep, Mission Push and Discretion: The Case of IMF Conditionality." *The IMF and the World Bank at sixty*: 59–84. http://lccn.loc.gov/2005472666.

Bagwell, Kyle, and Robert W Staiger. 2005. "Enforcement, Private Political Pressure, and the General Agreement on Tariffs and Trade/World Trade Organization Escape Clause." *Journal of Legal Studies* 34(2): 471–513.

Baqir, Reza, Rodney Ramcharan, and Ratna Sahay. 2005. "IMF Programs and Growth: Is Optimism Defensible?" *IMF Staff Papers* 52(2): 260–86.

Barnett, Michael N, and Liv Coleman. 2005. "Designing Police: Interpol and the Study of Change in International Organizations." *International Studies Quarterly* 49(4): 593–619.

Beck, T et al. 2001. "New Tools in Comparative Political Economy: The Database of Political Institutions." *World Bank Economic Review* 15(1): 165–76.

Bird, Graham. 2002. "The Completion Rate of IMF Programmes: What We Know, Don't Know and Need to Know." *The World Economy* 25(6): 833–47. http://doi.wiley.com/10.1111/1467-9701.00465.

———. 2008. "The Implementation of IMF Programs: A Conceptual Framework." *Review of International Organizations* 3(1): 41–64.

Bird, Graham, and Thomas D Willett. 2004. "IMF Conditionality, Implementation and the New Political Economy of Ownership." *Comparative Economics Studies* 46(3): 423–50. http://www.palgrave-journals.com/doifinder/10.1057/palgrave.ces.8100060.

Caraway, Teri L, Stephanie J Rickard, and Mark S Anner. 2012. "International Negotiations and Domestic Politics: The Case of IMF Labor Market Conditionality." *International Organization* 66(1): 27–61. http://journals.cambridge.org/abstract_S0020818311000348.

Chayes, Abram, and Antonia Handler Chayes. 1993. "On Compliance." *International Organization* 47(2): 175–205.

Chorev, Nitsan, and Sarah L Babb. 2009. “The Crisis of Neoliberalism and the Future of International Institutions: A Comparison of the IMF and the WTO.” *Theory and Society* 38(5): 459–84. http://www.springerlink.com/index/10.1007/s11186-009-9093-5.

Chwieroth, Jeffrey M. 2014. “Professional Ties That Bind: How Normative Orientations Shape IMF Conditionality.” *Review of International Political Economy* 22(4): 757–87. http://www.tandfonline.com/doi/full/10.1080/09692290.2014.898214.

Coate, Stephen, and Stephen Morris. 2006. “Policy Conditionality.” In *Globalization and the Nation State: The Impact of the IMF and the World Bank*, eds. Stephen Kosack, Gustav Ranis, and James R Vreeland. , 36–50.

Copelovitch, Mark S. 2010. “Master or Servant? Common Agency and the Political Economy of IMF Lending.” *International Studies Quarterly* 54(1): 49–77.

Dawson, Thomas C. 2002. “Stiglitz, the IMF and Globalization.” http://www.imf.org/external/np/speeches/2002/061302.HTM.

———. 2003. “On the IMF’s Role and Activities in Africa.”

Drazen, Allan. 2002. “Conditionality and Ownership in IMF Lending: A Political Economy Approach.” *IMF Staff Papers* 49(Special Issue): 36–67.

Dreher, Axel. 2003. “The Influence of Elections on IMF Programme Interruptions.” *Journal of Development Studies* 39(6): 101–20. http://www.tandfonline.com/doi/abs/10.1080/0022038031233129359 7.

———. 2006. “IMF and Economic Growth: The Effects of Programs, Loans, and Compliance with Conditionality.” *World Development* 34(5): 769–88. http://www.sciencedirect.com/science/article/pii/S0305750X06000210.

———. 2009. “IMF Conditionality: Theory and Evidence.” *Public choice* 141(1–2): 233–67. http://www.springerlink.com/index/10.1007/s11127-009-9486-z.

Dreher, Axel, Jan-Egbert Sturm, and James R Vreeland. 2009. “Global Horse Trading: IMF Loans for Votes in the United Nations Security Council.” *European Economic Review* 53(7): 742–57. http://www.sciencedirect.com/science/article/pii/S0014292109000312.

Dreher, Axel, and Roland Vaubel. 2004. “The Causes and Consequences of IMF Conditionality.” *Emerging Markets Finance and Trade* 40(3): 26–54.

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. http://www.sciencedirect.com/science/article/pii/S0304387804000872.

Findley, Michael G., Daniel L. Nielson, and Jason C. Sharman. 2013. 67 International Organization *Using Field Experiments in International Relations: A Randomized Study of Anonymous Incorporation*. http://www.journals.cambridge.org/abstract_S0020818313000271.

Girod, Desha M., and Jennifer L. Tobin. 2016. “Take the Money and Run: The Determinants

of Compliance with Aid Agreements." *International Organization* 70(1): 209–39. http://www.journals.cambridge.org/abstract_S0020818315000326.

Goldstein, Morris, Timothy F Geithner, Paul Keating, and Yung Chul Park. 2003. "IMF Structural Programs." In *Economic and Financial Crises in Emerging Market Economies*, ed. Martin Feldstein. University of Chicago Press, 363–457. http://www.nber.org/chapters/c9778.pdf.

Haggard, Stephan, and Robert R Kaufman. 1992. "The Politics of Economic Adjustment: International Constraints, Distributive Conflicts, and the State." http://www.amazon.com/Politics-Economic-Adjustment-Stephan-Haggard/dp/0691003947%3FSubscriptionId%3D1V7VTJ4HA4MFT9XBJ1R2%26tag%3Dmekentosjcom-20%26linkCode%3Dxm2%26camp%3D2025%26creative%3D165953%26creativeASIN%3D0691003947.

Hallerberg, Mark, and Patrik Marier. 2004. "Executive Authority, the Personal Vote, and Budget Discipline in Latin American and Caribbean Countries." *American Journal of Political Science* 48(3): 571–87.

Hathaway, Oona A. 2002. "Do Human Rights Treaties Make a Difference?" *The Yale Law Journal* 111(8): 1935–2042.

Henisz, Witold J. 2002. "The Political Constraint Index (POLCON) Dataset."

IMF. 2001a. *Conditionality in Fund-Supported Programs*. Washington D.C.

———. 2001b. "IMF Conditionality: How Much Is Enough?" In *IMF Economic Forum*, Washington D.C.: International Monetary Fund.

Ivanova, Anna, Wolfgang Mayer, Alex Mourmouras, and George Anayiotos. 2001. *What Determines the Success or Failure of Fund-Supported Programs?* Washington D.C.

Joyce, Joseph P. 2013. "The IMF and Global Financial Crises: Phoenix Rising?" http://books.google.co.uk/books?id=yDYqDRQNHAoC&pg=PA120&dq=The+IMF+and+Global+Financial+Crises+inauthor:joyce&hl=&cd=1&source=gbs_api.

Joyce, Joseph P. 2006. "Promises Made, Promises Broken: A Model of IMF Program Implementaion." *Economics and Politics* 18(3): 339–65.

Kentikelenis, Alexander E, Thomas H Stubbs, and Lawrence P King. 2016. "IMF Conditionality and Development Policy Space, 1985-2014." *Review of International Political Economy* (Online first).

Killick, Tony. 1995. *IMF Programmes in Developing Countries – Design and Impact*. London: Routledge.

Koremenos, Barbara. 2005. "Contracting around International Uncertainty." *American Political Science Review* 99(4): 549–65.

Laeven, Luc, and Fabian Valencia. 2013. "Systemic Banking Crises Database." *IMF*

Economic Review 61(2): 225–70.

Lang, Valentin F, and Andrea F Presbitero. 2017. "Room for Discretion? Biased Decision-Making in International Financial Institutions." *Journal of Development Economics*.

Marshall, M, Keith Jaggers, and T R Gurr. 2010. "Polity IV Project: Political Regime Characteristics and Transitions 1800− 2010: Polity Level 2004− 2008."

Mayer, Wolfgang, and Alex Mourmouras. 2005. "The Political Economy of IMF Conditionality: A Common Agency Model." *Review of Development Economics* 9(4): 449–66. http://siteresources.worldbank.org/INTRAD/Resources/Breen.pdf.

Mecagni, Mauro. 1999. "The Causes of Program Interruptions." In *Economic Adjustment in Low-Income Countries*, Washington D.C.: International Monetary Fund, 215–76.

Mercer-Blackman, V, and A Unigovskaya. 2004. "Compliance with IMF Program Indicators and Growth in Transition Economies." *Emerging Markets Finance and Trade* 40(3): 55–83. http://mesharpe.metapress.com/index/xpjpkepq0tny9wge.pdf.

Mitchell, Sara McLaughlin, and Paul R Hensel. 2007. "International Institutions, Cooperation, and Compliance with Agreements." *American Journal of Political Science* 51(4): 721–37.

Mussa, Michael, and Miguel Savastano. 2000. 14 NBER Macroeconomics Annual 1999, Volume 14 *The Approach to Economic Stabilization*.

Nelson, Stephen C. 2017. *The Currency of Confidence: How Economic Beliefs Shape the IMF's Relationship with Its Borrowers*. Ithaca, NY: Cornell University Press.

Nsouli, Saleh M, Ruben Atoyan, and Alex Mourmouras. 2004. 04/184 IMF Working Paper *Institutions, Program Implementation, and Macroeconomic Performance.* http://papers.ssrn.com/abstract=879009.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press. http://books.google.com/books?id=4xg6oUobMz4C.

Pop-Eleches, Grigore. 2009. "Public Goods or Political Pandering: Evidence from IMF Programs in Latin America and Eastern Europe." *International Studies Quarterly* 53(3): 787–816. http://doi.wiley.com/10.1111/j.1468-2478.2009.00556.x.

Potts, Jason. 2017. "Governing the Innovation Commons." *Journal of Institutional Economics* (May): 1–23.

Reinsberg, Bernhard, Alexander Kentikelenis, Thomas Stubbs, and Lawrence P King. 2018. "The World System and the Hollowing-out of State Capacity: How Structural Adjustment Programs Affect Bureaucratic Quality in Developing Countries." *American Journal of Sociology* forthcomin.

Rickard, Stephanie J., and Teri Caraway. 2018. "International Demands for Austerity: Examining the Impact of the IMF on the Public Sector." *Review of International Organizations* forthcomin.

Rosendorff, B Peter, and Helen V Milner. 2001. “The Optimal Design of International Trade Institutions: Uncertainty and Escape Clauses.” *International Organization* 55(4): 829–57.

Seabrooke, Leonard. 2010. “Bitter Pills to Swallow: Legitimacy Gaps and Social Recognition of the IMF Tax Policy Norm in East Asia.” In *Owning Development*, eds. Susan Park and Antje Vetterlein. Cambridge University Press, 137–60.

Selznick, Philip. 1948. 13 American sociological review *Foundations of the Theory of Organization*. American Sociological Association. http://www.jstor.org/stable/2086752.

Stiglitz, Joseph E. 2002. “Globalization and Its Discontents.” : 282. http://books.google.co.uk/books?id=075MS-ZBsswC&dq=stiglitz+2002&ie=ISO-8859-1&cd=1&source=gbs_api.

Stone, Randall W. 2002. “Lending Credibility: The International Monetary Fund and the Post-Communist Transition.” : 304. http://books.google.com/books?hl=en&lr=&id=lXXSeWsZE7wC&oi=fnd&pg=PR15&dq=Lending+Credibility&ots=3lBgrPigeU&sig=mK_dqn6-WPAJFJ1cKhv45N38WXQ.

———. 2004. “The Political Economy of IMF Lending in Africa.” *American Political Science Review* 98(4): 577–92. http://journals.cambridge.org/production/action/cjoGetFulltext?fulltextid=265353.

Thacker, Strom C. 1999. “The High Politics of IMF Lending.” *World Politics* 52(1): 38–75.

Underdal, Arild. 1998. “Explaining Compliance and Defection: Three Models.” *European Journal of International Relations* 4(1): 5–30.

Vaubel, Roland. 2006. “Principal-Agent Problems in International Organizations.” *The Review of International Organizations* 1(2): 125–38.

Vreeland, James R. 2003. *The IMF and Economic Development*. Cambridge University Press.

———. 2006. “IMF Program Compliance: Aggregate Index versus Policy Specific Research Strategies.” *Review of International Organizations* 1(4): 359–78.

Weaver, Catherine E. 2008. *The World Bank’s Hypocrisy and the Poverty of Reform*.

Woo, Byungwon. 2013. “Conditional on Conditionality: IMF Program Design and Foreign Direct Investment.” *International Interactions* 39(3): 292–315.

Woods, Ngaire. 2006. *The Globalizers: The IMF, the World Bank, and Their Borrowers*. Cornell University Press.

World Bank. 2015. “World Development Indicators.” http://data.worldbank.org/.

———. 2018. “Global Financial Development Database (GFDD).” https://datacatalog.worldbank.org/dataset/global-financial-development (July 31, 2018).

Supplemental appendix

Table A1: Previous studies on the determinants of compliance.

Study	Compliance measure	Unit of analysis	Sample	Specification	Significant determinants
Compliance as outcome					
Killick 1995	Disbursement rate threshold	Program	305 programs	Probit analysis	Export base, debt level, size of IMF loan
Mecagni 1999	Interruption	Program	36 programs	Case study	External shocks, domestic political economy
Dreher 2003	Disbursement rate below 75% in a year	Country-year	104 countries over 1975-98	Probit analysis	Programs break down more likely before election; past compliance, low government consumption, a low share of short-term debt, high per capita GDP at beginning of program
Stone 2004	Length of interruption	Country-year	53 countries over 1990-2000	Weibull model of state transitions	Suspensions more likely when reserves low, exchange rate, and short-term debt; duration of punishment: short-term debt; suspensions are based on technical criteria, but resumption of lending is political; probability of government failure predicts shorter punishment; US aid receipts;
Ivanova et al. 2006	Interruption	Program	170 programs (95 countries over 1992-98)	IV probit	MIMIC results: ELF, political instability, bureaucratic quality and executive leadership; Random-effect IV analysis: No effect of IMF effort;
	Implementation index			IV tobit	
	Disbursement ratio			IV tobit	
Joyce 2006	Disbursement ratio	Program	352 programs (77 countries from 1975-99)	Tobit analysis	Democracy, trade openness, length of tenure, ethnic fractionalization
Arpac, Bird, and Mandilaras 2008	Interruption	Program	218 programs (95 countries over 1992–2004)	Probit analysis	Trade openness, number of veto players, loan-quota ratio
Compliance as predictor (all except one examining economic growth)					
Mercer-Blackman and Unigovskaya 2004	Implementation index	Country	84 countries over 1993-97	Linear regression	Economic growth: Inflation, reform indicator, implementation index
Nsouli, Atoyan, and Mourmouras 2006	Non-interruption dummy	Program	195 programs (95 countries over 1992-2002)	Before-after using 2SLS and 3SLS	Economic growth: Large programs, longer engagement, better implementation leads to better macroeconomic outcomes
	Share of disbursed funds	Program			Economic growth: Intense civil war and weak law enforcement before program approval
Dreher 2006	Disbursement share; drawn funds below threshold; program	Country-period	67 countries over 1975-98	3SLS	Economic growth: Democracy, elections, government consumption, short-term debt, GDP per capita, budget deficit

	suspension				
Dreher and Walter 2010	Disbursement below 75% in any year over five-year period	Country-period	68 countries over 1975-2002	2SLS and GMM	Currency crisis less likely when IMF program active in past five years, no effect of compliance beyond that

Discussion of different compliance measures

Previous research examining the determinants of compliance with IMF conditionality have used three measures (Bird 2002, 638; Arpac et al; Dreher 2009). Table A2 summarizes the main studies using these measures and discusses the relative (de)merits of the measures. While each measure has its own merits, there has been a tendency toward disaggregation of compliance patterns, which suggests that waivers are the most pertinent measure (Vreeland 2006). We therefore consider waivers as our preferred measure.

First, the **disbursement ratio** refers to the amount of funds drawn down by the recipient in relation to the originally committed funds. A low disbursement ratio—typically below a threshold of 75%—may indicate non-implementation of some essential targets. While Killick (1995) considers disbursement ratios of entire IMF programs, Dreher (2003) proposes a refinement that considers the percentage of funds left undrawn in any program year, assuming equal-size tranches. However, disbursement-based measures are noisy proxies of compliance because countries may not need to draw on the funds, for example if economic circumstances have improved.

Second, **program interruptions** are a more valid measure as they directly relate to non-implementation of binding conditions. However, a program is interrupted if the Fund fails to grant a waiver for the unimplemented conditions. The Fund might need to do so if program implementation is severely off-track, but it has more discretion to continue programs if only a few conditions are unimplemented. To the extent that only some countries benefit from such discretion, the measure does not actually reflect the implementation performance of borrowing countries. The above measures are disadvantageous for another reason. They both ignore that "[c]ompliance is a spectrum, not a binary variable. A country that has complied with many conditions may subsequently have disbursements suspended because of the failure of others; program non-completion is not equivalent to total failure. Conversely, program completion is not necessarily equivalent to success because many governments are granted waivers despite unmet targets" (Babb and Carruthers 2008, 22).

The third measure thus follows a disaggregated approach by tracking the implementation record of each individual condition in any given program. In particular, it considers whether the IMF board approved **waivers** for unimplemented conditions, which ensures that countries complete program reviews to have the next tranche disbursed. A disaggregated approach allows us to identify the policy conditions that were responsible for (imminent) implementation failure. If IMF programs follow a boilerplate model, not all reforms are equally easy to implement across borrowing countries (Polak 1991; Vreeland 2006; Babb and Carruthers 2008).

Table A2: Comparison of compliance measures

Measure	Studies and findings	Advantages	Disadvantages
Disbursement ratio	Killick (1995): At least 20% of program undrawn specifically under high debt, and small loan size Dreher (2003): At least 25% of resources available each year undrawn specifically before elections and under democracy Ivanova et al. (2006): Disbursement share lower under ethnic fractionalization, when domestic interests stronger, under political instability, and number of conditions (IV regressions); lower bureaucratic quality and political cohesion (OLS regressions) Joyce (2006): Disbursement ratio is higher under democracy, higher trade openness, longer tenure of government, and ethnic fractionalization	Simplicity	Aggregate measure of compliance Other factors beyond country compliance may explain why funds remain undrawn (for instance an unexpected improvement in economic circumstances)
Waiver percentage	Ivanova et al. (2006): Implementation index lower under ethnic fractionalization and when special interests are stronger (IV regressions); more program conditions per year (OLS regression)	Disaggregated measure of compliance	Proxy measure for actual implementation (since the IMF may decide not to waive some conditions)
Program interruption	Mecagni (1999): Breakdown related to slippage on conditionality (in 33 cases, often due to fiscal deficit), far fewer (8 cases) due to disagreement about future actions; elections played a role in 12 interruptions Ivanova et al. (2006): Non-interruption more likely under lower political instability, ethnic fractionalization, weaker special interests, and better political cohesion (IV regression);	Simplicity	Aggregate measure of compliance Other factors beyond country compliance may explain delays in scheduled reviews IMF discretion to resume program (some conditions may be waived but if essential ones are not waived a program gets interrupted)

	bureaucratic quality (OLS regression)
	Arpac et al. (2008): Interruption less likely under high trade openness, low number of veto players, higher loan-quota ratio

Table A3: Variable definitions and descriptive statistics.

Variable	Description and sources	Obs	Mean	Sd	Min	Max
Key variables						
Waiver percentage	Percentage of conditions waived in the total number of conditions applicable to the program (Kentikelenis, Stubbs, and King 2016)	640	5.57	8.09	0.00	62.50
Temporary interruption	Dichotomous indicator of temporary program interruption; an interruption is temporary if the program review is delayed but there is a subsequent review in the data (Kentikelenis, Stubbs, and King 2016)	668	0.28	0.45	0.00	1.00
Permanent interruption	Dichotomous indicator of permanent program interruption; an interruption is permanent if the program does not resume after the last scheduled review was not completed (Kentikelenis, Stubbs, and King 2016)	668	0.42	0.49	0.00	1.00
Number of conditions	Number of binding conditions applicable over the duration of the program (Kentikelenis, Stubbs, and King 2016)	658	54.21	42.75	0.00	253.00
Structural conditions	Number of binding structural conditions applicable over the duration of the program (Kentikelenis, Stubbs, and King 2016)	658	9.79	15.27	0.00	123.00
Covariates measured at program initiation						
First-time borrower	Dichotomous measure of borrowing government turning to the IMF for the first time since 1979 (Kentikelenis, Stubbs, and King 2016)	648	0.17	0.37	0.00	1.00
Left-wing government	Left-wing government (Beck et al. 2001)	589	0.25	0.44	0.00	1.00
IMF quota	IMF quota of borrowing country (Scheubel and Stracca 2017)	603	3.57	3.48	0.19	37.78
Loan-to-quota ratio	Loan-to-quota ratio (Scheubel and Stracca 2017)	508	104.88	157.94	4.99	1938.47
Veto player index	Veto player index (Henisz 2002)	635	0.21	0.21	0.00	0.68
Democracy	Dichotomous measure of democracy (Alvarez et al. 2000)	628	0.45	0.50	0.00	1.00
Previous interruptions	Number of permanent or temporary interruptions of programs before program initiation year (Kentikelenis, Stubbs, and King 2016)	648	0.94	3.58	0.00	26.00
Countries under programs	Number of countries under IMF programs (Kentikelenis, Stubbs, and King 2016)	648	57.65	10.36	41.00	75.00
Agreement duration	Length of agreement in months (Kentikelenis, Stubbs, and King 2016)	641	23.43	10.73	3.00	36.00
Changing circumstances						
Inflation (Δ)	Percentage change in inflation rate (World Bank 2015), occurring between program initiation and the year a program is first interrupted or the last program year if no interruption	543	2.21	28.45	-47.66	583.41
Current account balance (Δ)	Percentage change in current account balance (World Bank 2015), occurring between program initiation and the year a program is first interrupted or the last	453	1.39	22.18	-89.06	411.14

	program year if no interruption					
ODA per capita (Δ)	Percentage change in ODA per capita (World Bank 2015), occurring between program initiation and the year a program is first interrupted or the last program year if no interruption	517	1.56	24.37	-61.38	419.15
Reserves (Δ)	Change in reserve months of imports (World Bank 2015), occurring between program initiation and the year a program is first interrupted or the last program year if no interruption	466	0.61	1.93	-7.70	11.54
Debt service (Δ)	Percentage point change in debt service (as of GNI) (World Bank 2015), occurring between program initiation and the year a program is first interrupted or the last program year if no interruption	497	-0.61	11.29	-132.66	127.48
Polity index (Δ)	Absolute change in Polity IV index (Marshall, Gurr, and Jaggers 2010), occurring between program initiation and the year a program is first interrupted or the last program year if no interruption	560	1.43	4.20	-14.00	17.00
Government fractionalization (Δ)	Absolute change in government fractionalization (Beck et al. 2001), occurring between program initiation and the year a program is first interrupted or the last program year if no interruption	481	0.01	0.26	-0.88	0.88
Political constraints (Δ)	Absolute change in veto player index (Henisz 2002), occurring between program initiation and the year a program is first interrupted or the last program year if no interruption	600	0.04	0.19	-0.67	0.67
Ideological shift	Change in executive party orientation (Beck et al. 2001), between program initiation and the year a program is first interrupted or the last program year if no interruption	267	0.14	0.86	-2.00	2.00
US interest rate (Δ)	Percentage change in US interest rate (World Bank 2015), occurring between program initiation and the year a program is first interrupted or the last program year if no interruption	613	-0.16	0.50	-0.90	3.94
II rating (Δ)	Absolute change in Institutional Investor rating (Scheubel and Stracca 2017), occurring between program initiation and the year a program is first interrupted or the last program year if no interruption	414	3.71	10.06	-28.10	37.45
Public debt (Δ)	Percentage change in public debt (Scheubel and Stracca 2017), occurring between program initiation and the year a program is first interrupted or the last program year if no interruption	493	0.09	0.72	-0.92	7.43
Central bank assets (Δ)	Percentage change in central bank assets (World Bank 2018), occurring between program initiation and the year a program is first interrupted or the last program year if no interruption	488	0.11	1.74	-1.00	21.29
Financial assets (Δ)	Percentage change in financial assets (World Bank 2018), occurring between	515	0.22	1.46	-1.00	24.24

	program initiation and the year a program is first interrupted or the last program year if no interruption					
UNGA alignment with G7 (Δ)	Absolute change of G7 voting alignment in the UNGA (Bailey, Strezhnev, and Voeten 2015), occurring between program initiation and the year a program is first interrupted or the last program year if no interruption	592	0.01	0.06	-0.26	0.30
UNSC member	Temporary UNSC member (Dreher, Sturm, and Vreeland 2009), any time occurring between program initiation and the year a program is first interrupted or the last program year if no interruption	654	0.19	0.39	0.00	1.00
G5 bank exposure (Δ)	Percentage change of G5 bank exposure (Bank for International Settlements 2018), occurring between program initiation and the year a program is first interrupted or the last program year if no interruption	513	1.77	5.36	-1.00	57.00
Capital stop	Any stop of capital inflow to the country (Scheubel and Stracca 2017), occurring between program initiation and the year a program is first interrupted or the last program year if no interruption	160	0.43	0.50	0.00	1.00
Capital flight	Any capital flight out of the country (Scheubel and Stracca 2017), occurring between program initiation and the year a program is first interrupted or the last program year if no interruption	161	0.52	0.50	0.00	1.00
Financial crisis	Any financial crisis (Laeven and Valencia 2013), occurring between program initiation and the year a program is first interrupted or the last program year if no interruption	653	0.42	0.49	0.00	1.00
Executive election	Any executive election (Beck et al. 2001), occurring between program initiation and the year a program is first interrupted or the last program year if no interruption	646	0.60	0.49	0.00	1.00
Legislative election	Any legislative election (Beck et al. 2001), occurring between program initiation and the year a program is first interrupted or the last program year if no interruption	646	0.83	0.38	0.00	1.00
Civil war	Any civil war (Gleditsch et al. 2002), occurring between program initiation and the year a program is first interrupted or the last program year if no interruption	667	0.10	0.29	0.00	1.00

Table A3: 2SLS estimation taking endogeneity of conditions into account.

	(1)	(2)	(3)	(4)	(5)	(6)
Waiver percentage						
Number of conditions	0.082***	0.087***	0.079***			
	(0.024)	(0.019)	(0.027)			
Structural conditions				0.226***	0.187***	0.239***
				(0.087)	(0.057)	(0.070)
UNGA alignment (Δ)	-11.637**	-12.910**	-13.048	-6.853	-8.838	-10.206
	(5.521)	(5.505)	(8.106)	(6.088)	(5.497)	(8.184)
Democracy	0.166	0.598	0.039	0.533	1.359	0.800
	(1.010)	(0.955)	(1.179)	(1.089)	(0.990)	(1.160)
Agreement duration	-0.138***	-0.124***		-0.058	-0.042	
	(0.045)	(0.047)		(0.039)	(0.039)	
FDI inflows		-0.106			-0.048	
		(0.099)			(0.101)	
Civil war (Δ)		-2.532*			-3.271**	
		(1.319)			(1.580)	
US interest rate (Δ)			-1.336			-1.224
			(0.890)			(0.895)
Financial crisis			2.476			1.994
			(1.510)			(1.573)
Current account balance (Δ)			-0.057***			-0.052***
			(0.015)			(0.015)
ODA per capita (Δ)			0.075			0.037
			(0.069)			(0.058)
Reserves (Δ)			0.341			0.259
			(0.347)			(0.337)
Debt service (Δ)			0.112***			0.057**
			(0.037)			(0.028)
Polity index (Δ)			-0.006			0.080
			(0.091)			(0.094)
Upcoming executive election			0.212			0.334**
			(0.179)			(0.161)
Political constraints (Δ)			0.383			0.221
			(2.198)			(2.187)
Number of conditions						
First-time borrower	-19.704***	-19.867***	-19.896***	-2.458	-2.444	-2.424
	(6.815)	(6.771)	(6.797)	(2.148)	(2.181)	(2.161)
IMF quota	2.063	2.071	2.062	0.469	0.466	0.438
	(1.841)	(1.825)	(1.838)	(0.548)	(0.553)	(0.556)
Loan-to-quota ratio	0.039***	0.039***	0.038***	0.015***	0.015***	0.014***
	(0.015)	(0.015)	(0.015)	(0.005)	(0.005)	(0.005)
Prior interruptions	3.011	3.055	2.865	0.935*	0.898*	0.818
	(2.155)	(2.105)	(2.170)	(0.496)	(0.495)	(0.501)
Countries under programs	1.577***	1.580***	1.590***	0.656***	0.656***	0.661***
	(0.256)	(0.252)	(0.259)	(0.090)	(0.091)	(0.092)
Log(US troops)	0.001	0.001	0.001	0.000	0.000	0.000
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Fixed effects	yes	yes	yes	yes	yes	yes

Within-R2	0.21	0.26	0.27	0.19	0.24	0.25
Observations	590	560	524	590	560	524

Notes: * p<.1 ** p<.05 *** p<.01.

Figure F1: The implementation gap of IMF conditionality.

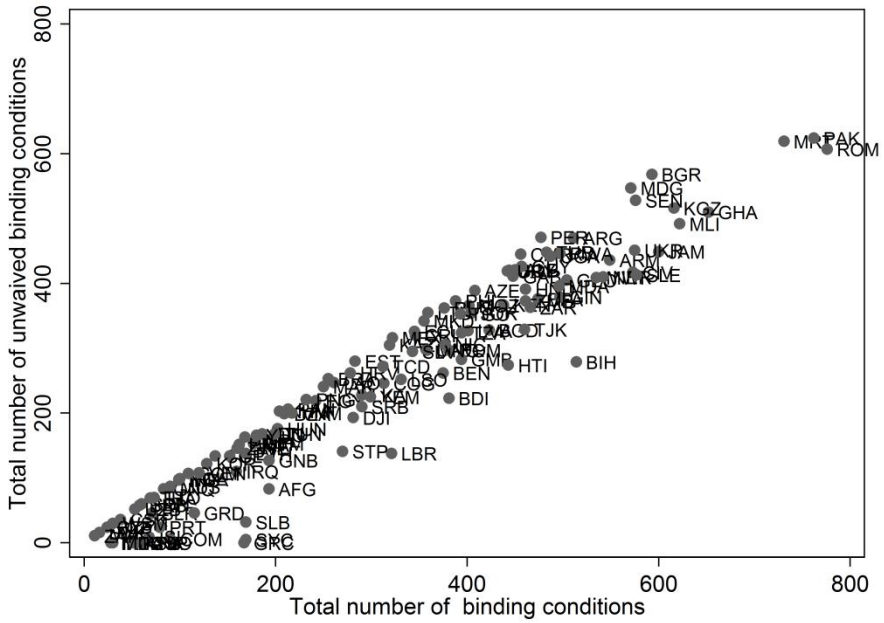


Figure 1 compares the total number of (binding) conditions initially agreed in Letters of Intent with conditions net of those waived, showing that almost all countries implemented fewer conditions than originally envisaged over the 1980-2009 period. For some countries the gap is considerably larger, regardless of how many conditions they faced at the onset of a given program. For instance, Afghanistan cumulatively faced around 200 conditions but implemented only half of them. Burundi was waived almost 200 conditions out of the 400 conditions that were initially agreed.