INTERNATIONAL MONETARY FUND

IMF-Supported Programs in Low-Income Countries: Fragile versus Non-Fragile States

Kaihao Cai, Édouard Martin, and Felix Várdy

WP/24/221

2024
OCT

WORKING PAPER

 WP/24/221

IMF Working Paper
African Department

IMF-Supported Programs in Low-Income Countries: Fragile versus Non-Fragile States

Prepared by Kaihao Cai, Édouard Martin, and Felix Várdy

Authorized for distribution by Charalambos Tsangarides
October 2024

ABSTRACT: This paper examines the macroeconomic frameworks of IMF-supported programs with low-income countries from 2009 to 2022, focusing on how macroeconomic targets and their achievement differ between fragile and conflicted-affected states (FCS) and non-FCS. Key findings include similar program targets for FCS and non-FCS, optimism in all dimensions considered other than inflation, and no significant correlation between targets and outcomes. For variables other than inflation, country-independent targets equal to the mean or median outcomes of other programs outperform program projections as predictors of actual outcomes. This underscores the challenges in setting realistic, country and program-specific targets in IMF-supported programs with low-income countries. Finally, we discuss potential caveats, including GDP rebenchmarking, non-linear relationship between initial conditions and targets, and repeat programs. We do not study, and make no claims about, causality.

RECOMMENDED CITATION: Cai, Kaihao, Édouard Martin, and Felix Várdy. 2024. IMF-Supported Programs in Low-Income Countries: Fragile versus Non-Fragile States. IMF Working Papers WP/24/221.

JEL Classification Numbers:	F33, F34, F35, O11, O19.
Keywords:	IMF Programs; LICs; Fragility; Tailoring; Optimism
Authors E-Mail Addresses:	kcai@imf.org, emartin@imf.org, fvardy@imf.org

WORKING PAPERS

IMF-Supported Programs in Low-Income Countries: Fragile versus Non-Fragile States

Prepared by Kaihao Cai, Édouard Martin, and Felix Várdy[1]

[1] The authors would like to thank, without implicating, Sever Can, Aliona Cebotari, Max Dovì, Mai Farid, Chris Hinchcliffe, Frank Hespeler, Carlos van Hombeeck, Daehaeng Kim, Vitaliy Kramarenko, Toh Kuan, Tomo Mineyama, Mico Mrkaic, Chris Papageorgiou, Luca Ricci, Axel Schimmelpfennig, Haris Tsangarides, Olaf Unteroberdoerster, Issouf Samake, and seminar participants at the IMF, for their comments and suggestions. Roberta Guarnieri provided excellent research assistance. Joanna Delcambre's support with editing the manuscript is gratefully acknowledged.

Contents

TABLES

I. Introduction

This paper studies the macroeconomic frameworks of IMF-supported programs with low-income countries (LICs), examining how macroeconomic objectives and the achievement thereof differ between programs with fragile and conflicted-affected states (FCS) and non-FCS. The IMF designates about one fifth of its 191 members as FCS. Most of them are also LICs.[3] Conversely, almost half of LICs are FCS (see, e.g., IMF, 2022). Here, we only consider LICs, which may or may not be FCS. FCS are home to nearly one billion people who face long-standing challenges such as weak governmental institutions, limited public services, and ongoing conflicts. Fragility and conflict severely impact economic stability, disrupting the balance of payments, trade, and financial flows, while also stalling economic development (*ibid.*).

Our comparative study covers the period 2009-2022 and is motivated, in part, by the recommendations of the recent IMF Strategy for FCS, which has called for a better "tailoring" of IMF engagement with FCS to the fragility and conflict context. The Strategy was launched in 2022. However, calls for the tailoring of IMF-supported programs go back to at least 2008, when the IMF published its first Review of Experience in Fragile States and Post-Conflict Countries. While tailoring has many dimensions, an important element is "avoiding overly optimistic assumptions" and applying "more realistic macro frameworks" (see, e.g., IMF, 2008, 2011, 2015, 2022, and Independent Evaluation office (IEO), 2018). This aspect, which we refer to as "quantitative tailoring," is the focus of the current paper.

We seek to assess the extent of quantitative tailoring, in terms of the targets embedded in the macro frameworks of IMF-supported programs. We also study optimism and the degree of correlation between targets and outcomes. Past papers on tailoring mainly looked at differences, or lack thereof, in how programs were designed (e.g., the number and kind of prior actions, quantitative performance criteria, and structural benchmarks), rather than at potential differences in projections of macroeconomic variables. These variables are the focus of the present paper. They include growth, inflation, fiscal consolidation, and external adjustment variables. In this context, optimism refers to the difference between projections and outcomes, while correlation refers to the degree of association between the two.

In this paper, we shall use the terms program objectives, projections, and targets interchangeably. Objectives, projections, and targets all refer to three-year quantitative forecasts taken from the macro frameworks of IMF-supported programs at the time of program approval. These variables are not necessarily formal performance criteria that must be satisfied for IMF financing to be disbursed. Rather, they reflect IMF staff's expectations as to how the economy will develop under the program. While targets may be adjusted during the program, we focus on their original values.

Main Findings

Our main findings are as follows:

***Limited Quantitative Tailoring*:** Program targets do not appear to differ significantly between FCS and non-FCS LICs, be it in absolute terms or relative to initial conditions.

Considerable Optimism: Targets tend to be missed in all dimensions other than inflation.[4]

[2] Following common IMF practice, we consider a country to be a LIC if it qualifies for financing under the IMF's Poverty Reduction and Growth Trust (PRGT), the IMF's main source of concessional financing. For a list of LICs, as well as of FCS, see Appendix H.

[4] Among on-track programs, targets are missed in all dimensions other than inflation and reserves.

Weak Correlations: For variables other than growth and inflation, we cannot reject the null hypothesis that targets and outcomes are statistically independent. Country and program-independent targets equal to the mean or median outcomes of all other programs would have outperformed program projections as predictors of actual outcomes in dimensions other than inflation.

While it is interesting to speculate about the reasons for our findings, we purposefully stay away from questions of causality. Pervasive endogeneity of targets, outcomes, program implementation, and other potentially explanatory variables makes credible identification daunting. In the absence of good instruments, we prefer to stay non-committal on questions of causality.

The rest of the paper is structured as follows: A brief review of the literature is presented in the remainder of the Introduction. Section II details the research question, methodology, and data. Sections III to V study quantitative tailoring, optimism, and the correlation between targets and outcomes, respectively. Section VI discusses potential caveats, while Section VII concludes. Additional figures and tables are included in appendices, while the data and code are accessible online.

Literature Review

While there exists a substantial literature on IMF-supported programs with LICs (see, e.g., Bal Gunduz (2013) and references therein), studies focusing on IMF-supported programs with FCS are much scarcer. Noteworthy exceptions include Kunduz (2018), Lane *et al.* (2021), and Collyns *et al.* (2021), which builds on the IMF's Independent Evaluation Office's report on The IMF and Fragile States (IEO, 2018). These studies highlight that: (i) the borrowing frequency from the IMF of FSC LICs is like that of non-FCS LICs; (ii) however, completion rates for lending programs with FCS are markedly lower—30 versus 75 percent; (iii) the level of conditionality, in terms of the number and nature of prior actions, quantitative performance criteria, and structural benchmarks, shows little variation between FCS and non-FCS.

While most studies concentrate on how IMF-supported programs were designed, Kunduz (2018) assesses macroeconomic outcomes—specifically, aid flows and growth. He documents that IMF-supported programs positively influence official development assistance in FCS compared to non-FCS. He also finds more volatility and marginally lower growth outcomes in FCS—findings we confirm. Finally, he shows that growth in FCS tends to increase by approximately one percentage point (p.p.) after program approval.

Regarding optimism, our research is most closely related to Kim *et al.* (2021). The evidence on optimism in program design presented here is broadly consistent with the earlier work of Kim *et al.* They juxtapose GRA and PRGT programs, with a particular focus on growth and fiscal multipliers. We concentrate on the PRGT and divide them into FCS and non-FCS. Owing to the passage of time, our dataset benefits from a somewhat extended temporal scope, furnishing us with approximately three years of additional data. Notable earlier work on optimism in IMF-supported programs includes Baqir *et al.* (2005a, 2005b).

Regarding correlation of targets and outcomes, our paper is related to Celasun *et al.* (2021). Celasun *et al.* evaluate the accuracy of WEO growth forecasts from 2004 to 2017. They find that short-term real GDP growth forecasts show minimal bias. However, longer-term WEO growth forecasts (two to five years ahead) tend to be overly optimistic and, in many cases, less accurate than a simple forecast based on the average growth rate of the recent past. We replicate this finding for LICs in a program context and extend it to variables other than growth.

Outside the context of IMF-supported programs, the literature on the relationship between economic performance and fragility is vast (see, e.g., Leepipatpiboon *et al*., 2023, and references therein). Leepipatpiboon *et al*. distinguish between two veins in this literature, depending on the direction of causality:

The first vein examines the detrimental effects of fragility and conflict on economic performance. This negative impact is transmitted through various channels, including heightened economic uncertainty, worsening fiscal balances, reduction in consumption, and the deterioration and destruction of physical capital, as illustrated in the works of Collier (1999), Rodrik (1999), Cerra and Saxena (2008), and others. Additionally, the literature points to negative spillover effects of fragility and conflict on the economies of neighboring nations, exacerbated by challenges such as the impact of large refugee populations on infrastructure and fiscal accounts, highlighted by researchers like Anselin and O'Loughlin (1992) and Buhaug and Gleditsch (2008).

The second vein reverses the causality, investigating how economic performance can affect the likelihood and intensity of fragility and conflict. Initial studies in this area identified a statistically significant link between poverty and fragility incidence. However, subsequent research has cast doubt on the direct causal relationship between the two, suggesting that the observed correlation often diminishes when accounting for country-specific characteristics and utilizing more recent datasets. This body of work has grappled with methodological challenges such as endogeneity and omitted variable bias, leading to the use of global commodity prices as plausibly exogenous variables to explore the economic-conflict nexus. Yet, the findings from these inquiries have been mixed, with studies like Brückner and Ciccone (2010) and Bellemare (2015) finding that fluctuations in commodity prices can either mitigate or exacerbate fragility and conflict, depending on numerous factors including the type of commodities and their impact on household income versus state revenue.

To our knowledge, the present paper is the first to systematically investigate quantitative tailoring, optimism, and the relationship between targets and outcomes in IMF-supported programs with low-income FCS.

II. Questions, Methodology, and Data

Research Questions

In the context of IMF-supported programs with LICs, we explore the following questions.

1. **Quantitative Tailoring:** To what extent do macroeconomic objectives of programs with FCS differ from those with non-FCS?
2. **Optimism:** How optimistic are these objectives? Is there a difference between FCS and non-FCS, and between programs that remain on track and those that do not?
3. **Correlation:** How correlated are program targets and outcomes?

Methodology

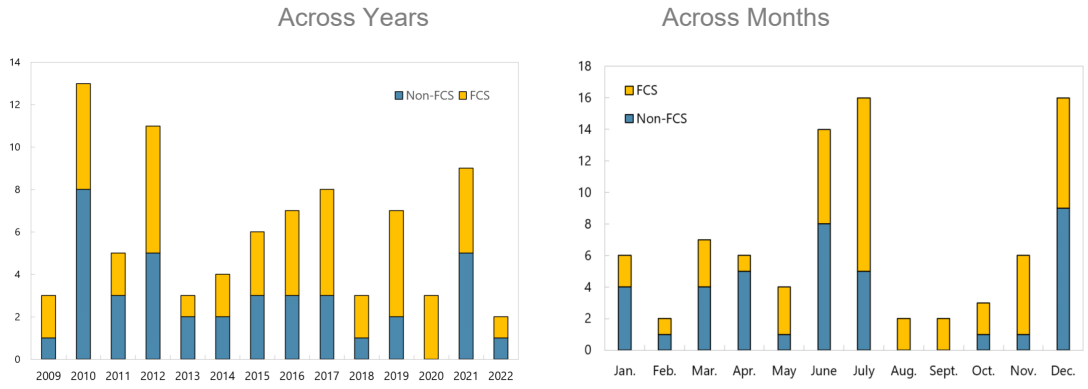
Program Inclusion Criteria

To be included in our sample, programs had to meet the following criteria. They needed to be with a LIC, i.e., eligible for financing from the PRGT, approved during fiscal years 2009 to 2022, and have a planned duration of 1.5 years or more.[5] These criteria yielded 84 programs across 43 countries for the ambitiousness assessment. They comprised 75 ECFs and 9 Standby Credit Facilities. Using a time horizon of three years—the standard program duration for IMF-supported programs with LICs—outcome variables for programs

[5] To avoid duplication, blended programs in which countries received financial support from both the Extended Credit Facility (ECF) and the Extended Fund Facility (EFF) were handled by retaining the PRGT-financed ECF arrangement and excluding the EFF.

initiated after mid-2019 were not yet available at the time of data collection.[6] This reduced the number of observations for studying optimism and correlation to, maximally, 62 programs across 37 countries.

Figure 1. Histogram of Program Approval Dates



Source: MONA Database and IMF staff calculations.

The left panel shows the distribution of program approvals across years. The right panel shows it across the months of the year.

Programs with planned durations under 1.5 years were excluded. First, their short length—less than 50 percent of the modal length—might render them less effective. Second, even if effective, they could not be "held accountable" for outcomes occurring more than 18 months after their completion. Lastly, they could lead to multiple and similar programs within the three-year forecast horizon, potentially yielding highly correlated observations.[7] However, including programs with durations less than 1.5 years does not materially affect the results.

When assessing optimism and correlation, we distinguish between the full sample of programs and the subsample of programs that remained on track. When comparing targets to outcomes, our time horizon was three years, equal to the planned duration of most programs. Since programs typically were on a six-month review cycle, we say that a program was on track if it had a successful reviews 2.5 years or more after its start. Notice that a program could fail to make the subsample of on-track programs either because it failed to pass Board review or because its planned duration was less than 2.5 years to begin with.

FCS were less likely to remain on track than non-FCS. [8] We have outcomes for 62 out of the 84 programs in our sample. Of these, 8 had a planned duration of less than 2.5 years. Of the remaining 54, 29 programs were

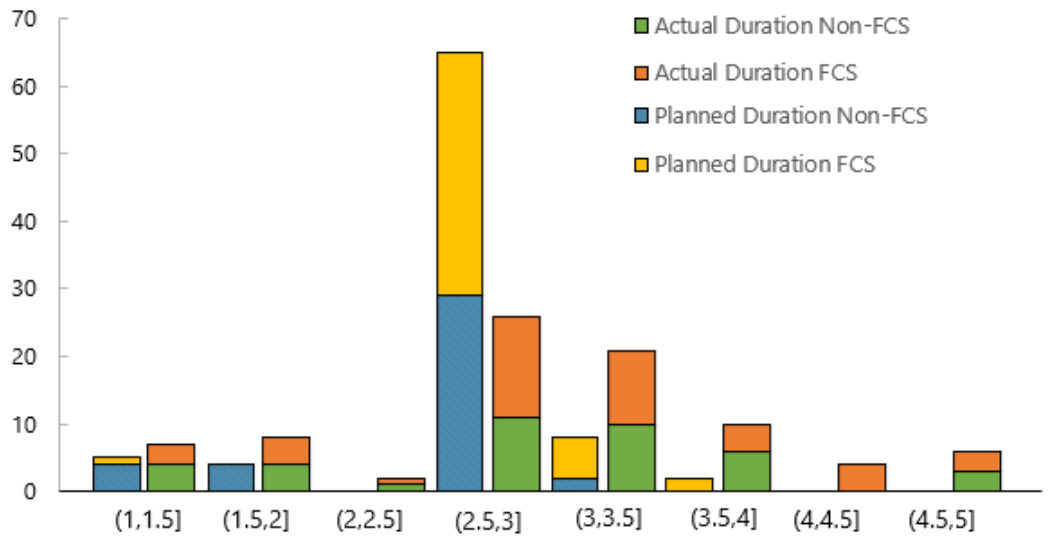
[6] Data were collected during the fall of 2023. To ensure that the most recent WEO reflected realizations rather than preliminary estimates or forecasts, we only considered programs for which, by mid-2022, three years had passed since adoption.

[7] See Section VI Potential Caveats for a discussion of dependence of observations and its implication for the validity of statistical tests.

[8] The IMF introduced its FCS classification in FY2008. For details, see the Data Section, below.

with FCS, 12 of which went off track (i.e., 41 percent). 25 programs were with non-FCS, 4 of which went off track (i.e., 16 percent). Hence, in our sample, FCS were 2.6 times more likely to go off track than non-FCS.[9]

Figure 2. Histogram of In-sample Program Durations



Distribution of planned and actual program durations.
Source: IMF Financial Data Query and IMF staff calculations.

Variables

Our dataset comprises variables that capture countries' initial conditions, program targets, and outcomes. Our dataset focuses on the key areas of growth, inflation, government finances, and balance of payments. The variables, listed in Table 1, have been chosen for their significance in the design of IMF-supported programs and their reasonably consistent availability throughout the study period. They include real growth ("Real GDP Growth"); inflation ("CPI Inflation"); general government revenue excluding oil and grants ("Revenues XOXG"); primary current expenditure ("PCE"); the public wage bill ("Wage Bill"); the primary fiscal balance ("PB"); PB excluding oil and grants ("PB XOXG"), general government debt ("Debt"), the external current account balance ("CAB"), and international reserve coverage ("Reserves"). While reserve coverage is expressed in months of imports, growth and inflation are in average percentage changes. All other variables are in percentage of (total or non-oil) GDP. [10]

For each variable, we have defined an initial condition, a target, and an outcome (see Table 1). These definitions depend on whether the variable itself represents a percentage change, a flow, or a stock. Growth and inflation represent percentage changes (of flow variable GDP and stock level CPI, respectively). All

[9] Cf., IEO (2018), who find a ratio of 2.8. While these ratios are essentially identical, in their sample, the fractions of programs going off track is greater, namely, 70 and 25 percent for FCS and non-FCS. The difference could be explained by us limiting attention to the first 2.5 years of programs.

[10] Some variables are expressed as a percentage of GDP, others as a percentage of non-oil GDP. For each variable, we have chosen what we believe to the more natural choice. However, our conclusions do not depend on this.

revenues, government expenditures, and fiscal and current account balances represent flows, while debt and foreign reserves represent stocks.

Initial conditions, targets, and outcomes of variables representing percentage changes, namely, growth and inflation, are measured as geometric averages—i.e., average growth rates, over three-year periods. Let *T* denote the year of program approval. The initial condition is defined as the average growth rate during years *T*-2 to *T*, as estimated at the time of the program adoption.[11] The target is the average growth rate programmed for years *T*+1 to *T*+3, while the outcome is the realization of this average. Put differently, targets and outcomes of change variables correspond to the forecasted and realized outcomes at *T*+3, respectively. For example, the initial condition of Growth is equal to the estimated average growth rate achieved during *T*-2 to *T*, the target is the average growth rate envisaged for the period *T*+1 to *T*+3, while the outcome is the ex-post realization thereof.

The initial condition of a flow variable is equal to its level during year *T*, while the initial condition of a stock variable is equal to its level at the end of that year. Unlike for change-variables, targets and outcomes of flows and stocks are defined as the projected and realized *difference* of the outcome at *T*+3 and the initial condition at *T*. Our signing convention for this difference is that a positive number corresponds an adjustment in the 'right' direction. Thus, for Revenue XOXG, PB (XOXG), CAB, and Reserves, a positive number for a target or an outcome corresponds to an *increase*, while for PCE, Wage Bill, and Debt a positive number corresponds to a *decrease*. For example, for Revenue XOXG, the initial condition is equal to the amount of non-oil non-grant revenue collected during *T*, expressed as a percentage of non-oil GDP. The target is the increase in the revenue-to-non-oil-GDP ratio envisaged for year *T*+3, compared to *T*. For Debt, the initial condition is the debt-to-GDP ratio at the end of year *T*. The target is the *de*crease of that ratio, as projected for the end of year *T*+3.

[11] We decided to include year *T* in the pre-program period, because there may be delay in programs becoming effective. Also, more programs were approved in the second half of the year than in the first half (45 versus 39—see Figure 1).

Table 1. Variables

Sector	Variable	Variable Type	Units	Initial Condition	Target	Outcome
Real	Real GDP Growth	Change	%Δ per annum	Avg growth *T*-2 to *T*	Avg growth *T*+1 to *T*+3	Avg growth *T*+1 to *T*+3
Inflation	CPI Inflation	Change	%Δ per annum	Avg growth *T*-2 to T	Avg growth *T*+1 to *T*+3	Avg growth *T*+1 to *T*+3
Gov. Finances	Revenue XOXG	Flow	% of non-oil GDP	Flow during *T*	Δ btwn. *T* and *T+3*	Δ btwn. *T* and *T+3*
	PCE	Flow	% of non-oil GDP	Flow during *T*	Δ btwn *T* and *T+3*	Δ btwn *T* and *T+3*
	Wage Bill	Flow	% of non-oil GDP	Flow during *T*	Δ btwn *T* and *T+3*	Δ btwn *T* and *T+3*
	PB	Flow	% of GDP	Flow during *T*	Δ btwn *T* and *T+3*	Δ btwn *T* and *T+3*
	PB XOXG	Flow	% of GDP	Flow during *T*	Δ btwn *T* and *T+3*	Δ btwn *T* and *T+3*
	Debt	Stock	% of GDP	Stock, end of *T*	Δ btwn *T* and *T*+3	Δ btwn *T* and *T*+3
Balance of Payments	CAB	Flow	% of GDP	Flow during *T*	Δ btwn *T* and *T+3*	Δ btwn *T* and *T+3*
	Reserves	Stock	Months of imports	Stock, end of *T*	Δ btwn *T* and *T*+3	Δ btwn *T* and *T*+3

The table defines initial conditions, targets, and outcomes for each variable.

Using these data, we assess:

Quantitative Tailoring: We compare the ambitiousness of targets for FCS versus non-FCS and study how targets relate to initial conditions.

Optimism: We estimate average differences between targets and outcomes. We also study to what extent optimism about growth can explain optimism in other dimensions.

Correlation: We study the correlation between targets and outcomes and test for independence between the two. We calculate the R^2 of Targets as predictors of Outcomes, as well as Theil's U_2, where we use mean or median outcomes for the "naïve" model.

Data Sources

Overview

Our main data sources were the IMF's Financial Data Query Tool, the World Economic Outlook database, and Country Staff Reports. We used these databases to determine country classifications and construct variables related to initial conditions, program objectives, and outcomes. During the cleaning process, we had to address several data issues, including the handling of GDP re-benchmarking and rebasing, the substantial impact of

HIPC debt relief on some countries' government revenue and debt stocks, and the incompleteness of some data series in the WEO.

FCS Classification

Our FCS classification used the IMF list in place at the time of program approval. Between 2008 and 2023, the IMF had its own FCS list, which differed somewhat from the World Bank's. It can be found in Appendix H. During our sample period (2009-2022), the list was updated four times. The number of FCS went from 41 in 2008; 33 in 2011; 39 in 2015; 43 in 2017; to 42 in 2019.[12] The number of countries changing category decreased from 18 in 2011; 12 in 2015; 10 in 2017; to only 1 in 2019, indicating increased stability in FCS classification.

LICs

Our LIC classification used the IMF's PRGT-eligibility list that was in place at the time of program approval. PRGT-eligibility was periodically reevaluated, with a focus on per capita GNI and sustainable access to international financial markets. At the beginning of our sample period in 2009, 77 countries were PRGT-eligible. During subsequent reviews, in 2010, 2013, 2015, and 2020, a total of 13 countries graduated, while 5 countries became newly eligible. This resulted in 69 PRGT-eligible countries at the end of our sample period in 2022, out of a total of 191 IMF members. Of these 69, 45 were Sub-Saharan African, while 30 were FCS.

IMF-supported Programs

For information on IMF-supported programs, we used the publicly available IMF Financial Data Query Tool. The database provided information on program approval dates, projected durations, Board reviews, and the amounts of financial support.

For macro variables, we used various vintages of the WEO, supplemented by publicly available Staff Reports. For initial conditions and program objectives, we used the first WEO vintage after program approval, which reflected, or only differed slightly from, the program's macroeconomic framework. For program outcomes, we used the latest version of the WEO database (i.e., October 2023). Missing data were supplemented by publicly available Staff Reports.

Data Issues

To ensure the consistency of data across countries and between initial conditions, targets, and outcomes, we proceeded with a few adjustments and verifications.

For program outcomes, we used the data from the October 2023 WEO. At times, this meant that GDP re-benchmarking had to be addressed to restore comparability of targets and outcomes. After consulting with statistical experts, we assumed that re-benchmarking changed GDP levels—sometimes substantially as in the case of, e.g., Ghana and Nigeria—but that the ratios between the old and new data series remained constant across years. To the extent possible, we verified this assumption and found that it held up quite well. See Appendix G for details.

When data were missing from the WEO database, we collected them from Staff Reports or country teams' macroeconomic frameworks.

We adjusted debt stock and revenue data for the impact of HIPC debt relief. The rationale was that, for some countries, HIPC contributed to substantial differences between program targets and outcomes that did not

[12] Since the 2023 harmonization of the IMF and World Bank FCS lists the number is 35.

directly reflect government policies or other macroeconomic developments. In the absence of comparable cross-country data, we did not adjust for other debt relief, whose magnitude was, in any case, substantially smaller than that of HIPC. Deleting programs subject to HIPC debt relief does not materially change the results.

The outcomes of a few programs were removed from the analysis, because of the eruption of civil war during the program period (CAF2012, YEM2014, SDN2021). These realizations were discarded as non-representative outliers.

Some variables that were initially considered, including social spending, and domestic versus foreign currency debt, were dropped from the analysis, because a lack of reliable and comparable cross-country data resulted in too few observations.

We now turn to the analysis. First, we consider quantitative tailoring. Then we study optimism. Finally, we look at the correlation between targets and outcomes.

III. Quantitative Tailoring

Tailoring IMF engagement and instruments to macro-critical manifestations of fragility and conflict is a key principle of the IMF's Strategy for FCS (2022). Calls for the tailoring of IMF-supported programs go back to at least 2008, when the IMF published its first Review of Experience in Fragile States and Post-Conflict Countries. While tailoring has many dimensions, an important element is "avoiding overly optimistic assumptions" and applying "more realistic macro frameworks" (see, e.g., IMF, 2008, 2011, 2015, 2022, and Independent Evaluation office (IEO), 2018). To assess this aspect of IMF-supported programs, which we refer to as "quantitative tailoring," we compare the targets of programs with FCS to those with non-FCS, in terms of 'absolute' and 'relative' ambitiousness. Since the FCS Strategy was adopted as recently as 2022, we do not provide an assessment thereof. Rather, we provide a benchmark against which the Strategy could assessed in the future.

Absolute ambitiousness refers, simply, to the magnitude of a target (e.g., the growth rate, inflation, or targeted fiscal or external adjustments). Relative ambitiousness also considers a country's initial condition. In absolute terms, programs for FCS might appear to be as ambitious as those for non-FCS, or even more so, but this could be due, wholly or in part, to worse initial conditions. In that case, arguably, FCS programs would be quantitatively tailored, in the sense of being less ambitious than programs with non-FCS facing similar conditions. Conversely, FCS programs could appear to have tailored adjustment targets, but this apparent tailoring might disappear once we control for initial conditions.

Absolute Ambition

To assess absolute ambitiousness, we compare targets for FCS versus non-FCS. The right side of Table 2 summarizes our findings. The "Target" column in Table 2 reports first and second moments of program targets for FCS and non-FCS—denoted by F and NF, respectively. The symbol "≈" means that we cannot reject the null hypothesis of equal means at the 10 percent significance level. Inequality signs with one, two, or three stars refer to statistically significant differences at the 10, 5, and 1 percent levels.

Table 2. Absolute Ambitiousness

Arithmetic Averages (Standard Deviations)

Variable	Initial Condition				Target			
	#Obs (F, NF)	**F**	2-sided test	**NF**	**#Obs (F, NF)**	**F**	2-sided test	**NF**
Growth	(45, 39)	3.4 (3.3)	≈ (>**)	4.1 (2.3)	(45, 39)	5.4 (2.5)	≈ (>*)	5.7 (1.9)
Inflation	(44, 39)	7.0 (5.8)	≈ (>**)	5.6 (3.9)	(44, 39)	5.5 (3.4)	≈ (>*)	5.0 (2.6)
Revenue XOXG	(45, 39)	**14** (5.8)	<*** (<*)	**21** (7.8)	(45, 39)	1.5 (1.6)	≈ (≈)	1.3 (1.5)
PCE	(44, 38)	17 (9.2)	≈ (≈)	19 (8.4)	(42, 38)	.7 (1.8)	≈ (≈)	1.0 (1.7)
Wage Bill	(44, 39)	7.4 (2.9)	≈ (≈)	7.7 (3.3)	(43, 39)	.3 (.7)	≈ (≈)	.4 (.9)
PB	(45, 39)	**-1.0** (3.6)	>** (≈)	**-2.7** (4.6)	(45, 39)	**.7** (2.7)	<** (≈)	**1.9** (2.5)
PB XOXG	(45, 39)	**-12** (7.8)	<*** (>***)	**-7.4** (2.8)	(45, 39)	2.0 (4.2)	≈ (>***)	2.5 (3.0)
Debt	(45, 39)	59 (45)	≈ (>***)	52 (21)	(45, 39)	**7.1** (13)	>** (>***)	**2.2** (7.5)
CAB	(38, 33)	-7.2 (11)	≈ (>***)	-11 (6.4)	(38, 33)	**-.3** (5.6)	<** (>***)	**2.5** (3.5)
Reserves	(44, 39)	4.3 (3.3)	≈ (>***)	3.7 (1.7)	(44, 39)	.4 (1.2)	≈ (≈)	.3 (1.1)

The table compares initial conditions and targets across FCS (F) and non-FCS (NF), in terms of arithmetic averages and standard deviations. The number of observations is shown under #Obs (F, NF) for F and NF, respectively. Inequality signs with one, two, or three stars indicate statistically significant differences at 10, 5, and 1 percent confidence, respectively, under a two-sided t-test. The equivalence sign, ≈, indicates no significant difference at 10 percent confidence. Bolding indicates a significant difference between F and NF.

We will discuss in some detail the entries for one flow variable, Growth, and for one stock variable, Debt. The first row of Table 2 reveals that the average growth target in programs with FCS is 5.4 percent per annum, with a standard deviation of 2.5 p.p. This is statistically indistinguishable at the 10 percent significance level from the 5.7 percent target, with standard deviation of 1.9, in programs with non-FCS. The column "#Obs (F, NF)" reports that these numbers are based on 45 programs with FCS and 39 programs with non-FCS. The "Debt" row in Table 2 reveals that the average debt-reduction target in programs with FCS is 7.1 percent of GDP, with a standard deviation of 13 p.p. At five percent confidence, this is greater than the 2.2 percent of GDP average debt-reduction target, with a standard deviation of 7.5, in programs with non-FCS. This means that, in absolute terms, programs in FCS have, in fact, more ambitious debt reduction targets than those in non-FCS.

More broadly, Table 2 shows that programs with FCS and non-FCS differ little in terms of absolute ambitiousness. There is no significant difference in terms of average targeted growth and inflation, nor in programmed adjustments for revenue, primary current expenditure, the wage bill, primary balance excluding oil and grants, and reserve coverage. Debt targets are more ambitious for FCS, while primary balance and current

account targets are less ambitious. While the averages are broadly similar, standard deviations in programs with FCS tend to be greater than or equal to those with non-FCS.

Using medians rather than means does not change this assessment. One might be concerned that the averages in Table 2 are unduly influenced by a few outliers. To check for this possibility, we have calculated median rather than mean targets and tested for a difference in medians. The results are shown in Table 1 in Appendix B. The conclusion about absolute ambitiousness remains unchanged: there is little difference between FCS and non-FCS. Relative to Table 2, the only change is that the difference between median targeted debt reduction of 4.7 percent for FCS and 1.6 percent for non-FCS is no longer statistically significant.

Relative Ambition

Differences in absolute ambitiousness are, of course, a rather crude measure of quantitative tailoring, because they ignore economies' initial conditions. For example, if FCS face worse initial conditions on average than non-FCS at the time of program design, equal absolute ambitiousness can in fact reflect quantitative tailoring, in the sense of lower *relative* ambitiousness. To investigate this issue, the column "Initial Condition" in Table 2 reports first and second moments of (measures of) countries' initial conditions, and it compares them across FCS and non-FCS. In terms of averages, there appears to be no systematic ordering: Growth, Inflation, PCE, Wage Bill, Debt, CAB, and Reserves are similar in the two types of countries. The PB is better in FCS, while the PB XOXG, as well as Revenues XOXG, are worse. The difference in Revenues XOXG is striking: 14 versus 21 percent of non-oil GDP for FCS and non-FCS, respectively. As for targets, the initial conditions of FCS appear more dispersed than those of non-FCS.

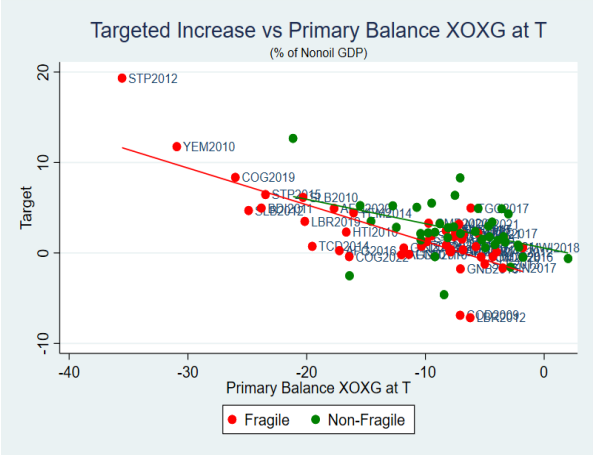
We now derive a quantitative measure of relative ambitiousness. First, we regress targets on initial conditions for each variable and for FCS and non-FCS separately. That is,

$$t_i = \beta_0 + \beta_1 s_i + \varepsilon_i \,.$$

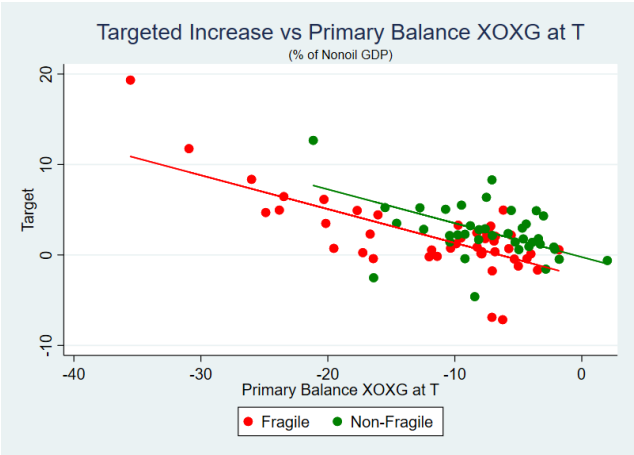
Here, t_i is the target for program *i*, s_i is the initial condition (state) at time *T*, β_0 is the intercept, β_1 is the regression coefficient, and ε_i is the error term. Then we check whether the slopes, β_1, are statistically different between the two groups of countries. If we cannot reject the null hypothesis that they are the same, we rerun the regression for FCS and non-FCS jointly, requiring the same slope but allowing for different intercepts. By construction, this yields two parallel lines, one for FCS, the other for non-FCS. The distance between the parallel lines is our measure of relative ambitiousness. Thus, a target is on average relatively less ambitious for FCS if its regression line lies significantly below the one for non-FCS, and vice versa. If we can reject the hypothesis of equal slopes, then we take that as potential evidence of tailoring. However, the exact interpretation would depend on the context. (Note that comparing intercepts when the slopes differ may not be meaningful.)

Figure 3. Initial condition versus Targeted Increase

Different slopes and different intercepts



Same slope but different intercepts



The left panel shows regression lines of initial conditions versus targets for FCS and non-FCS separately. The right panel shows the regression lines when we impose equality of slopes but allow for different intercepts. The distance between the parallel lines is our measure of relative ambitiousness.

Figure 3 illustrates that targeted increases of PB XOXG are, on average, less relatively ambitious for FCS than for non-FCS (See Appendix A, Figure 1, for variables other than PB XOXG). The initial condition, PB XOXG at *T*, is on the x-axis, while the targeted increase is on the y-axis. Each dot represents a program, with red dots for FCS and green dots for non-FCS. The left panel depicts the regression lines for FCS and non-FCS separately. At -.41 and -.27, the slopes are statistically and economically significant, and their signs are sensible: greater initial deficits are associated with higher adjustment targets. However, with a p-value of .22, they are not significantly different from each other. The right panel shows the regression lines when we require identical slopes for FCS and non-FCS but allow for different intercepts. The line for FCS lies a sizable 2.2 p.p. of GDP below that for non-FCS, which is significant at the 1 percent confidence level. Thus, while targeted PB XOXG increases for FCS are, on average, not less ambitious in absolute terms than those for non-FCS (see Table 2), they are less *relatively* ambitious, by 2.2 p.p. of GDP. Lower relative ambition for FCS may reflect tailoring to FCS' characteristics, such as limited capacity and considerable infrastructure and social spending needs. However, it might also be explained by FCS receiving more oil and grant revenues than non-FCS: indeed, programmed average annual oil revenues are 4.7 versus 3.0 percent of GDP for FCS and non-FCS, while grant revenues are 6.4 percent versus 3.1 percent of GDP, respectively. If we account for these revenues by focusing on the regular primary balance, the targeted relative adjustments for FCS and non-FCS become statistically indistinguishable. This is shown in Table 3.

Regression coefficients between initial conditions and targets are statistically and economically significant for most variables, and their signs are sensible. The only exceptions are Reserves for non-FCS and Revenues XOXG for both FCS and non-FCS. Their slopes are indistinguishable from zero. Signs are in line with IMF staff considering the initial conditions of an economy when formulating an IMF-supported program and targeting larger adjustments for countries facing worse initial conditions.

The slopes for FCS and non-FCS are mostly indistinguishable. For all variables other than the Wage Bill, we cannot reject the null hypothesis that the slopes between initial conditions and targets are the same for FCS and non-FCS. (Even for the wage bill, significance is only at the 10 percent level.) This allows us to calculate

our measure of relative ambitiousness, which is listed in the last column of Table 3. For the wage bill, the result has been bracketed out since its interpretation may be questionable.

Table 2. Relative Ambitiousness

(Standard errors in brackets)

Variable	#Obs (F, NF)	Slope btwn Initial Condition & Target				Δ Relative Ambition
		F	2-sided	NF	All	
Growth	(45, 39)	.25*** (.09)	≈	.43*** (.15)	.31*** (.08)	-.03 (.46)
Inflation[13]	(44, 39)	.52*** (.04)	≈	.52*** (.07)	.51*** (.04)	-.18 (.36)
Revenue XOXG	(45, 39)	-.05 (.04)	≈	-.00 (.03)	-.02 (.03)	.10 (.39)
PCE	(42, 38)	.15*** (.03)	≈	.11*** (.03)	.13*** (.02)	.04 (.39)
Wage Bill	(43, 39)	**.14*** ** (.03)	>*	**.21*** ** (.03)	.18*** (.02)	$\begin{bmatrix}-.11\\(.12)\end{bmatrix}$
PB	(45, 39)	-.47*** (.08)	≈	-.64*** (.11)	-.52*** (.07)	-.39 (.45)
PB XOXG	(45, 39)	-.41*** (.05)	≈	-.27*** (.10)	-.37*** (.07)	**-2.2*** ** (.64)
Debt	(45, 39)	.21*** (.03)	≈	.22*** (.06)	.21*** (.05)	**3.5** ** (1.7)
CAB	(38, 33)	-.39*** (.04)	≈	-.34*** (.09)	-.38*** (.04)	**-1.5*** (.79)
Reserves	(44, 39)	-.11** (.05)	≈	-.17 (.11)	-12** (.05)	.20 (.25)

"Δ Relative Ambition" measures the difference in relative ambitiousness of targets for FCS versus non-FCS. A negative number means that programs with FCS are, on average, less relatively ambitious. The result for the Wage Bill is bracketed because the slopes differ for FCS versus non-FCS, at least at10 percent significance. This complicates the interpretation.

Table 3 reveals that, as was the case for absolute ambitiousness, programs with FCS and non-FCS also differ little in terms of relative ambitiousness. In addition to PB XOXG, only CAB has a relatively less ambitious targeted increase for FCS than for non-FCS, while the targeted debt decrease is *more* ambitious. For all other variables, there is no significant difference between the degree of relative ambitiousness for FCS and non-FCS. Also, as argued above, the less ambitious target for PB XOXG may be attributable in part to the FCS's greater oil and grant revenues. In the absence of a greater overall PB target increase, FCS's more ambitious debt reduction targets might be attributable to greater expected non-HIPC debt relief.[14]

Other than possibly PB XOXG, this leaves the CAB as a potential case of quantitative tailoring. Even there, statistical significance is only at the 10 percent level. We conclude that, in terms of the ambitiousness of

[13] Excludes Sudan, which had hyperinflation preceding its 2021 program.

[14] Recall that debt and revenue data have been adjusted to remove the impact of HIPC debt relief, which in some instances contributed to major differences in projections and outcomes not directly attributable to government policies. We did not adjust for other, smaller kinds of debt relief, owing to the unavailability of comparable data for all countries.

targets, there is at most weak evidence of quantitative tailoring in program design once we account for initial conditions.[15]

IV. Optimism

We now turn to assessing optimism in program design.

Optimism is defined as the difference between targets and outcomes. Our signing convention is such that a positive number corresponds to an outcome falling short of its target. (For Inflation, an outcome exceeding its target is interpreted as a "shortfall.")

The tendency of IMF growth projections to be optimistic is a recognized problem, as highlighted by Kim *et al.* (2021), among others. Here, we confirm this trend in the context of IMF-supported programs with LICs and find that sizable optimism extends to nearly all variables under consideration. This holds true for the full sample as well as for the subsample of programs that remained on track. (Recall that programs were on track relative to our three-year time horizon if and only if they had a successful review 2.5 years or more after program approval.) The sole exception is inflation, where program targets appear accurately calibrated. We also verify that dropping programs that may have been negatively affected by COVID does not change our conclusions.

Average optimism is reported in Tables 4 and 5. Table 4 reflects the full sample, while Table 5 only considers on-track programs. Target values in these tables are somewhat different from those in Table 1. First, to evaluate optimism, we only consider *paired* observations of targets and outcomes for the same program, thus excluding the 18 programs that started less than four years prior to when the data were collected (three years for the time horizon, plus one year to ensure that the data reported in the WEO reflect actual outcomes rather than preliminary estimates). In practice, this means that only programs approved before 2019 are included in our assessment of optimism. Second, for variables other than growth and inflation, we only include programs that targeted a non-negative adjustment. This is to avoid findings driven by targets that might be rather too easy to achieve, such as fiscal expansions or running down reserves. Including programs with negative adjustment targets tends to reduce optimism but most often does not eliminate it. See Appendix D. Finally, in Table 5, we also exclude programs that were inactive or went off track, further reducing the number of observations. To avoid excessively small samples, in Table 5, we have consolidated FCS and non-FCS and only report optimism for the group of LICs as a whole. We have done the same in Table 4, when the number of observations fell below 20 for either category of countries.

The tables reveal that IMF-supported programs were optimistic in all dimensions other than inflation, irrespective of whether they remained on track. Optimism is reported in the last columns of Tables 4 and 5, as the average difference between target and outcome, and as a percentage of the original target. For example, Table 4 reveals that the difference between targeted and realized annual growth in FCS is 2.4 p.p., or 39 percent of target, and that it is significantly greater than for non-FCS, where the difference is 0.9 p.p., or 16 percent of target. Combining FCS and non-FCS, average growth optimism in the full sample is 1.6 p.p. (This is not shown in the table.) This compares to growth optimism of 1.2 p.p. when considering only programs that stayed on track, as reported in Table 5.[16] In other dimensions, overoptimism is at least as large in the

[15] Notice that our measure of relative ambitiousness crucially depends on the cardinal values of initial conditions and targets. Therefore, it does not lend itself easily to an analysis using quantiles, such as the median.

[16] Recall that only 4 programs with FCS went off track versus 12 with non-FCS. This suggests that the .4 p.p. the difference in average optimism between all programs versus on-track programs is at least partially explained by the latter subsample containing a disproportionate number of non-FCS, that tended to exhibit less growth optimism than FCS.

subsample of on-track programs. This is the case for PCE, PB, PB XOXG, and CAB. The tables also reveal that some variables deteriorated rather than improved, even on average, once we excluded off-track programs. For both samples, this is the case for PCE, Wage Bill, Primary Balance, and Debt. CPI Inflation constitutes an exception to almost uniform optimism: for both samples, inflation targets are well-calibrated, with only minor and statistically insignificant optimism.

Using medians rather than means does not materially change the picture. Both for the full sample and for the subsample of on-track programs, Tables 2 and 3 in Appendix B show that programs continue to be optimistic in almost all dimensions other than inflation. The only exceptions are PB XOXG for FCS in the full sample and Reserves in the subsample. We can no longer reject the null that these two are well calibrated.

In dimensions other than growth, optimism for FCS and non-FCS is either statistically identical, or we lack sufficient observations. In the full sample, realized annual growth for FCS is as much as 2.4 p.p. below target, yielding a GDP that is 7.4 percent lower than expected over the three-year horizon. At 10-percent confidence, this is significantly greater than the .9 p.p. optimism bias for non-FCS. For Inflation, Revenue XOXG, and PB XOXG, there is no differential optimism between FCS and non-FCS. Finally, for the remaining variables and for the subsample of on-track programs, we have less than the prespecified minimum of 20 observations for FCS or non-FCS, precluding a comparison. As shown in Appendix B, differential over-optimism for growth disappears if we rely on the median rather than the mean. Otherwise, our conclusions remain unchanged.

Table 3. Optimism

Arithmetic Averages (Standard Deviations)

Variable		Country Type	#Obs (Paired)	Target	1-sided	Outcome	Optimism 2-sided
Growth		F	30	6.0 (2.7)	>***	3.7 (3.5)	2.4*** [39%]
		NF	32	5.6 (1.9)	>***	4.7 (1.7)	>* .9*** (16%)
Inflation		F	30	5.5 (3.3)	≈	6.0 (4.1)	.5 (8%)
		NF	32	4.9 (2.3)	≈	5.3 (5.4)	≈ .5 (10%)
Revenue XOXG	Target ≥ 0	F	25	1.9 (1.2)	>**	.7 (2.2)	1.1** (62%)
	Target ≥ 0	NF	26	1.6 (1.2)	>**	.9 (1.6)	≈ .7** (44%)
PCE	Target ≥ 0	All	41	1.7 (1.4)	>***	-.6 (2.2)	2.3*** (135%)
Wage Bill	Target ≥ 0	All	44	.7 (.7)	>***	-.0 (.9)	.7*** [102%]
PB	Target ≥ 0	All	42	2.6 (2.4)	>***	-.2 (3.2)	2.8*** (107%)
PB XOXG	Target ≥ 0	F	22	3.8 (4.6)	>*	2.2 (4.5)	1.6* (43%)
	Target ≥ 0	NF	27	3.5 (2.7)	>***	.9 (2.9)	≈ 2.6*** (75%)
Debt	Target ≥ 0	All	36	9.6 (8.4)	>***	-3.2 (14)	13*** (133%)
CAB	Target ≥ 0	All	31	3.6 (3.7)	>***	.9 (4.8)	2.7*** (75%)
Reserves	Target ≥ 0	All	43	.8 (.9)	>**	.3 (1.4)	.5** (67%)

The table compares targets to outcomes in terms of arithmetic averages and standard deviations. Optimism is defined as the difference between the two. "Target ≥ 0" indicates that we limit attention to programs targeting a positive adjustment. If the number of paired observations, #Obs (paired), is less than 20 for F or NF, then programs are pooled ("All"). Otherwise, in the last column, average optimism is compared across FCS and non-FCS. Numbers in square brackets express optimism as a percentage of target. To compare Targets with Outcomes, we tested for a difference in means using a paired t-test. To compare optimism for F versus NF, we tested for a difference in average optimism using an independent t-test.

Table 5. Optimism—On-Track Only

Arithmetic Averages (Standard Deviations)

Variable		#Obs (Paired)	Target	1-sided	Outcome	Optimism
Growth		39	5.7 (2.0)	>***	4.5 (2.1)	1.2*** [21%]
Inflation		39	4.7 (2.5)	≈	4.8 (4.6)	.1 [2%]
Revenue XOXG	Target ≥ 0	32	1.7 (1.2)	>**	1.0 (1.9)	.6** [38%]
PCE	Target ≥ 0	26	1.9 (1.4)	>***	-.7 (2.5)	2.6*** [139%]
Wage Bill	Target ≥ 0	29	.7 (.7)	>***	.1 (.9)	.6*** [86%]
PB	Target ≥ 0	26	3.1 (2.7)	>***	-.8 (3.2)	3.9*** [126%]
PB XOXG	Target ≥ 0	30	3.3 (2.9)	>**	.7 (4.1)	2.6*** [79%]
Debt	Target ≥ 0	21	7.7 (5.3)	>***	-.5 (12)	8.2*** [106%]
CAB	Target ≥ 0	20	4.3 (4.3)	>***	.7 (5.0)	3.6*** [83%]
Reserves	Target ≥ 0	28	.8 (.8)	≈	.5 (1.2)	.32 [41%]

One may worry that our conclusions are driven, wholly or in part, by the negative shock of COVID-19. To check for this, we dropped all programs approved since 2017. This reduced the number of observations by 11, but it ensured that each remaining program had a full three years to come to fruition before COVID hit in the Spring of 2020. The results are shown in Appendix C, Tables 1 and 2. In all dimensions other than revenue and primary balances results are essentially unchanged, even quantitatively. For example, for FCS in the full sample, the difference between targeted and realized annual growth continues to amount to as much as 2.4 p.p., while average realized debt reduction remains negative. Optimism about revenues and primary balances is reduced. This is intuitive and even reassuring. Still, statistically, and economically significant optimism remains, ranging from, e.g., 36 percent of target for PB XOXG and 41 percent for Revenues XOXG, to 98 percent for PB. The subsample results for on-track programs are very similar.

Finally, we investigate to what extent optimism about growth could explain optimism in other dimensions (see Appendix F for details). One may wonder to what extent widespread optimism has a single, common predictor, namely, optimism about growth. All else equal, a shortfall in growth boosts variables expressed as a percentage of GDP. Notice, however, that growth deviations tend to affect not only the denominator (GDP), but also the numerator—revenues, say. This makes the net effect on the ratio ambiguous. Under unit elasticity, targeted GDP ratios and their outcomes are in fact growth independent. While an elasticity close to unity may hold for revenues, it is less likely to hold for expenditures, at least in the short run. Based on elasticities reported in the literature, in Appendix F we formulate hypotheses about the expected signs of the correlations between forecast errors for growth and those for other variables. These hypotheses are largely borne out by

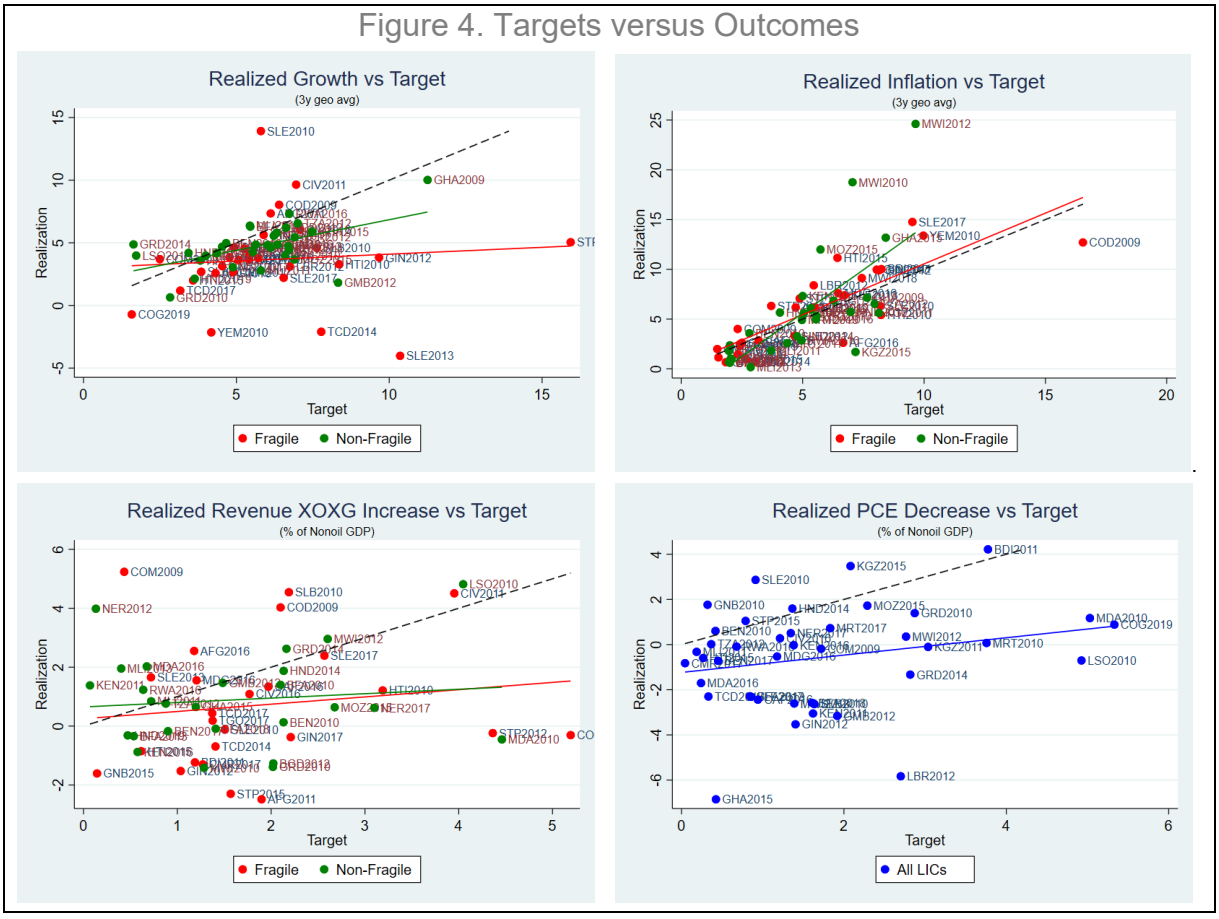
the data. Overall, we conclude that growth optimism helps explain optimism about Debt reduction. To a lesser extent, it also helps with optimism about Inflation and PCE—but not with other variables.

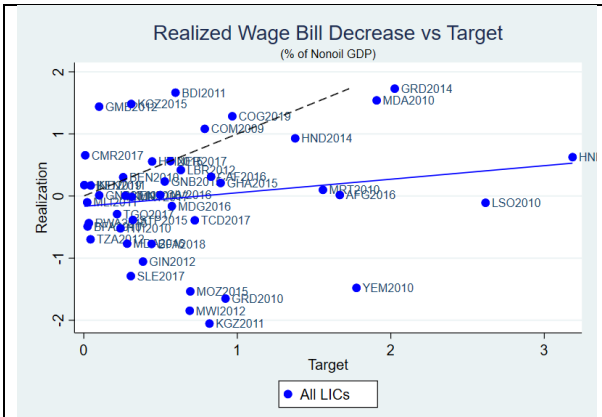
V. Correlation Between Targets and Outcomes

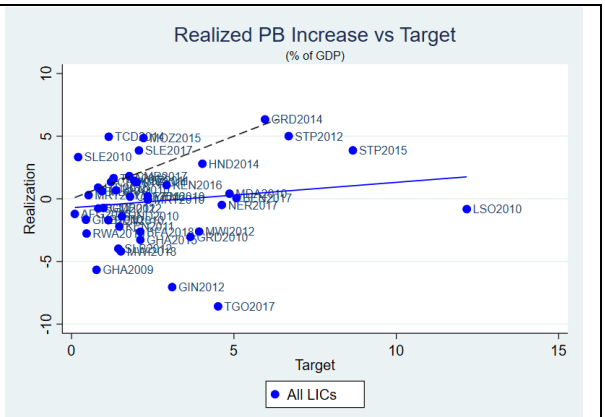
As we have seen, the macroeconomic assumptions underlying IMF-supported programs in LICs are generally optimistic, including for on-track programs. That is, on average, outcomes tend to fall short of targets. Notice, however, that this might not preclude targets from being accurate "on the margin." By this we mean that, even though their levels are off, a 1 p.p. higher target could still be associated with a 1 p.p. higher outcome, in expectation. More broadly, the question we pose is to what extent higher targets correlate with higher outcomes. As with optimism, we focus on programs with non-negative adjustment targets. Otherwise, our conclusions might be driven by countries achieving—or even systematically undershooting—projected deteriorations. Also, we distinguish between the full sample and the subsample of on-track programs.

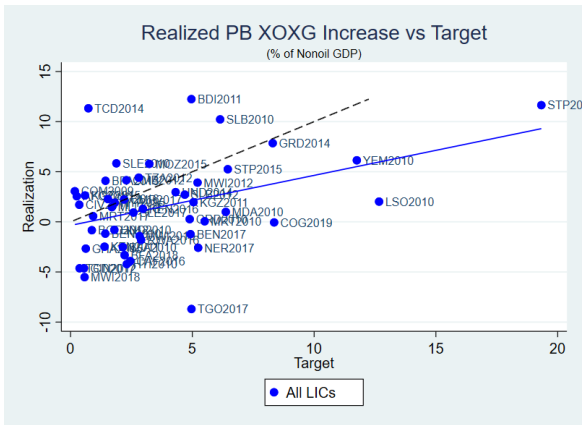
Figure 4 illustrates the relationship between targets and outcomes. The black dashed line represents the 45-degree line, while the solid lines denote the regression lines for the full sample. When we have more than 20 observations for each type of country, we differentiate between FCS and non-FCS, using red for the former and green for the latter. In cases with fewer observations, we combine the two types of countries and display a single regression line in blue.

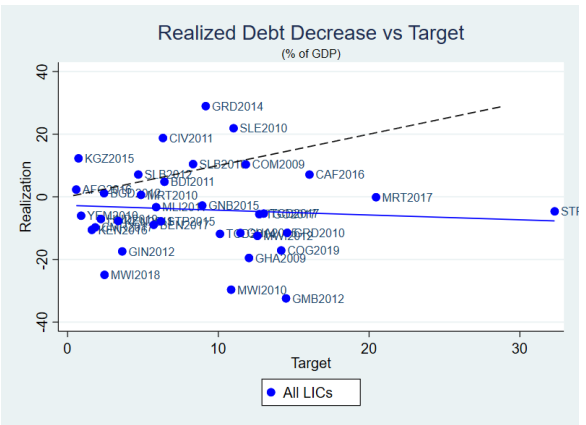
Figure 4. Targets versus Outcomes

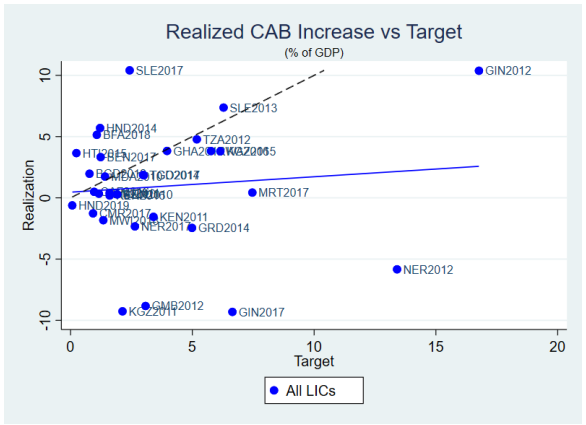


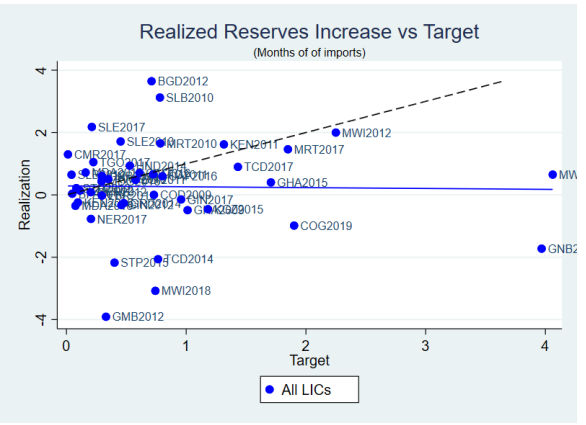












Regression lines are in red for FCS and in green for non-FCS. A dashed black line denotes the 45-degree line. If the number of paired observations is less than 20 for either type of country, then programs are pooled, and the regression line is in blue.

Also on the margin, targets are overly optimistic in all dimensions other than inflation. That is, the slope between target and outcome is strictly less than 1. Regression coefficients are given in the "Slope" columns of Tables 6 and 7. For all variables other than Inflation and in both samples, we can confidently reject the null hypothesis of unit slope.[17] The *p*-values usually lie below .01 and always below .05. Hence, also on the margin, targets tend to be optimistic.[18]

Except for growth and inflation, targets and outcomes appear to be statistically independent, even for on-track programs. We now turn to the broader question of the correlation between targets and outcomes. We have tabulated Pearson's correlation coefficient, Spearman's ρ, and Kendall's τ statistics.[19] The values of ρ and τ reveal that, for variables other than growth and inflation, we cannot reject independence of targets and outcomes, even at 10 percent significance. This holds for the full sample as well as for the subsample of on-track programs. With two exceptions for the full sample, the Pearson correlation and regression coefficients paint essentially the same picture.[20] Notice that zero correlation implies "optimism on the margin." However, the reverse does not hold.

[17] The stars in the tables do ***not*** refer to this test. They indicate whether the slopes are statistically different from 0.

[18] For CPI Inflation in non-FCS, the slope is strictly greater, instead of smaller, than 1. This also corresponds to optimism, however, in the sense that a 1 p.p. higher inflation forecast translates in more than 1 p.p. higher inflation, on average.

[19] Spearman's ρ and Kendall's τ are non-parametric measures of rank correlation, meaning they assess the relationship between two variables without making assumptions about the frequency distribution of those variables.

[20] In the full sample, the simultaneous rejection of independence and the non-significance of the correlation between targets and outcomes for growth in FCS suggest that the two may be non-linearly related. The simultaneous non-rejection of independence and the positive correlation between targets and outcomes for PB XOXG in FCS are somewhat harder to explain. They could be due either to a Type II error, i.e., false negative, for the independence tests, or to a Type I error, i.e., false positive, for the correlation test, where the latter test might also be affected by non-Normality of the data.

Table 6. Correlations

Variable		Type	#Obs (Paired)	Slope	Pearson Corr.	Spearman ρ	Kendall τ	R^2	Theil U_2		U_2 Handicap	
									Mean	Median	Mean	Median
Growth		F	30	.11	.08	**.34***	**.28****	-.95	1.14	1.38	1.21	1.18
		NF	32	**.49***	**.54*****	**.42****	**.31****	-.32	.57	1.09		
Inflation		F	30	**1.0*****	**82*****	**.84*****	**66*****	.68	.41	.56	.55	.58
		NF	32	**1.7*****	**.74*****	**.81*****	**.64*****	.43	.63	.75		
Revenue XOXG	Target ≥ 0	F	25	.25	.14	.19	.12	-.44	1.16	1.20	.98	1.01
		NF	26	.15	.11	.02	.02	-.57	1.15	1.23		
PCE	Target ≥ 0	All	41	.38	.24	.19	.11	-1.14	1.43	1.42	1.35	1.30
Wage Bill	Target ≥ 0	All	44	.23	.12	.14	.08	-.91	1.35	1.38	1.12	1.14
PB	Target ≥ 0	All	42	.20	.15	.13	.09	-1.11	1.42	1.43	1.15	1.17
PB XOXG	Target ≥ 0	F	22	**.56*****	**.43****	.31	.19	-.01	.96	.98	1.08	1.08
		NF	27	.34	.31	.27	.19	-1.13	1.41	1.39		
Debt	Target ≥ 0	All	35	-.15	-.07	-.12	-.07	-1.21	1.44	1.47	1.43	1.46
CAB	Target ≥ 0	All	31	.13	.10	.01	-.01	-.75	1.28	1.26	1.19	1.26
Reserves	Target ≥ 0	All	43	-.03	-.02	.05	.02	-.56	1.22	1.23	1.26	1.30

The table reports on the strength of the monotone relationship between targets and outcomes, as well as on how well the former forecasts the latter. "Slope" refers to the regression coefficient of a simple linear regression between outcomes and targets, while "Pearson Corr." represents the associated correlation coefficient. "Spearman ρ" and "Kendall τ" test for statistical independence between targets and outcomes. "Theil U_2" measures how well outcomes are forecasted by the mean and median outcomes of other programs compared to using program targets. A number greater than 1 means that the naive forecasts do better on average. "Handicap" only uses those programs whose outcomes were known at the time a program was initiated. "R^2"gives the proportion of the variance in outcomes that is forecasted by program targets. A negative value means that targets tend to deviate more from outcomes than the uniform average.

Table 7. Correlations—On-Track Only

Variable		Type	#Obs (Paired)	Slope	Pearson Corr.	Spear-man ρ	Ken-dall τ	R^2	Theil U_2		U_2 Handicap	
									Mean	Mean	Mean	Mean
Growth		All	39	.08	.08	**.32****	**.24****	-1.09	1.41	1.43	1.29	1.32
Inflation		All	39	1.4	**.77*****	**.83*****	**.64*****	.54	.66	.64	.53	.56
Revenue XOXG	Target ≥ 0	All	32	.27	.16	.04	.06	-.28	1.09	1.10	.69	.72
PCE	Target ≥ 0	All	26	.51	.29	.29	.20	-1.14	1.41	1.41	1.22	1.21
Wage Bill	Target ≥ 0	All	29	.39	.27	.29	.17	-.60	1.22	1.22	1.17	1.13
PB	Target ≥ 0	All	26	.26	.22	.18	.11	-1.82	1.61	1.58	1.29	1.38
PB XOXG	Target ≥ 0	All	30	.38	.27	.16	.09	-.54	1.20	1.15	1.18	1.17
Debt	Target ≥ 0	All	21	-.09	-.04	-.09	.06	-.72	1.25	1.24	1.26	1.30
CAB	Target ≥ 0	All	20	.12	.10	-.10	-.07	-1.08	1.37	1.31	1.29	1.40
Reserves	Target ≥ 0	All	28	-.12	-.09	.08	.04	-.73	1.27	1.29	1.30	1.36

Next, we calculate Theil's U_2 statistic to measure the accuracy of targets as forecasts of outcomes. Theil's U_2 is a relative accuracy measure that compares the performance of a forecast to a naïve model, by calculating the ratio of the mean squared errors. For the naive model we use either the mean or the median of all *other* program outcomes, while our forecast consists of the program Target. That is,

$$U_2 = \frac{MSE_{Targets}}{MSE_{Naive}}.$$

The results are listed in Tables 6 and 7. A U_2 greater than 1 signifies that the naïve model outperforms the forecast model, because it has lower mean squared error.

The analysis reveals that, even for on-track programs, the naïve model using the median tends to outperform the target, often by large margins. Tables 6 and 7 show that Theil's U_2 using the median is significantly below 1 only for Inflation. In the full sample only, it is close to 1 for PB XOXG in FCS. In all other cases, U_2 strictly exceeds 1, with most values ranging between 1.2 and 1.4. Thus, forecast errors using targets are significantly greater than those using the median-based model. This has the striking implication that, except for Inflation, using country-independent targets might raise the realism of macro frameworks in IMF-supported programs with LICs.

With one exception, the naïve model using the mean perform equally well. While both naïve models perform equally well for almost all variables, for Growth in non-FCS in the full sample, U_2 is 1.09 using the median but only 0.57 using the mean. This suggests that the median is the more robust forecasting choice, perhaps due to its insensitivity to outliers.

Except for inflation, targets generate negative R^2s. As we have seen, Theil's U_2-using-means tends to be strictly greater than 1, indicating that targets have worse forecasting performance than means. Since the R^2s of the means are zero by construction, the R^2s of targets tend to be strictly negative. That is, Targets deviate more from outcomes than the uniform average. On the one hand, optimism adds to the MSE of Targets as forecasts of outcomes. On the other hand, Targets have the (potential) advantage of functioning as conditional forecasts, that differ for, and can be tailored to, each program. The latter allows for a, potentially large, reduction of the MSE. The fact that $R^2 < 0$ or, equivalently, $U_2 > 1$, means that the former effect dominates the latter.

Our conclusions remain largely unchanged if, in the naïve model, we calculate means and medians using only outcomes known at the time a program was approved. Our naïve forecasts for each program consisted of the median or mean outcomes of all other programs. One might object that not all outcomes were known at the time of a program's approval. To address this concern, we modify our naïve forecasts for a program approved in year T to be the mean or median outcome of programs approved more than 3 years earlier, i.e., in $T - 4$ and before. Thus, for programs starting in 2013, our naïve forecast consists of the mean or median outcome of programs approved in 2009 only, while for programs starting in 2014, it consists of the mean or median of programs approved in 2009 and 2010, and so forth. Observe that programs approved during 2009-2012 drop out of our sample, while the naïve forecasts for programs approved in 2013 and the first few years thereafter are based on very few outcomes. The results are shown under the heading "U_2 Handicap" where also in the full sample we no longer distinguish between FCS and non-FCS, due to the relative paucity of observations. Despite this handicapping, our conclusions remain essentially unchanged. With one exception, means and medians continue to outperform targets in dimensions other than inflation. Now, Targets outperform the naïve forecasts for Revenue XOXG in the subsample of on-track programs.

Likewise, if we include programs with negative adjustment targets, means and medians continue to outperform targets in terms of forecasting, in all dimensions other than inflation. Outperformance is strict except for PB XOXG. This is shown in Appendix D, Table 3 and 4.

To the extent that the data allow for such an assessment, there is limited difference between FCS and non-FCS in terms of correlations between targets and outcomes. We already saw that there is not much significant difference between IMF-supported programs with FCS and non-FCS in terms of average ambitiousness and optimism. Table 6 reveals that the same is true in terms of correlations between program targets and outcomes. Both the Spearman's ρ and Kendall's τ are essentially identical for the two country groups for all variables where we have sufficient observations to make the comparison. These are Growth, Inflation, Revenue XOXG, and PB XOXG. The Pearson correlations and slopes do differ for growth and PB XOXG. Growth targets and outcomes are more correlated for non-FCS, while PB XOXG targets and outcomes are more correlated for FCS.

Dropping programs affected by COVID does not materially affect our conclusions. We continue to confidently reject the null hypothesis of marginal accuracy of forecasts, i.e., a unit slope between targets and outcomes. Second, Table 3 and 4 in Appendix C show that, except for growth and inflation, the null of independence between targets and outcomes still cannot be rejected, even at 10 percent confidence. In fact, we can no longer reject it for growth in FCS either. Furthermore, almost all p-values remain in their confidence "buckets," i.e., $p>10\%$, $5\%<p<10\%$, $1\%<p<5\%$, or $p<1\%$. For the full sample, the only exception is PB XOXG, whose correlation point estimate gets an extra star. Here, we no longer distinguish between FCS and non-FCS, because the number of observations for FCS has fallen below 20. Finally, in terms of forecasting, means and medians continue to outperform targets in all dimensions other than inflation. Outperformance is strict except for Revenue XOXG.

VI. Caveats

Our analysis is subject to several caveats.

Independence of Observations

Unlike non-parametric tests, the t-tests reported above rely on the Central Limit Theorem (CLT), which says that sample means are asymptotically Normally distributed. The standard CLT assumes that observations—IMF-supported programs in this case—are ex-ante independent and identically distributed. In our context, this assumption may well be violated for programs adopted around the same time. For example, their growth forecasts are likely to have been influenced by the same forecast for world growth, thus violating independence. However, we can rely on more general versions of the CLT that only require near-independence for programs temporally far apart (see, e.g., Dedecker, 2007). Notice that, in our context, this considerably weaker property is much more likely to be satisfied. To assuage remaining concerns, in Appendices B, C, and E, we report on the results of Wilcoxon signed-rank and Mann-Whitney U tests for paired and non-paired observations, respectively. These non-parametric tests for differences in median do not rely on the CLT or asymptotics and, reassuringly, lead to very similar conclusions.

Non-linearity

To evaluate relative ambitiousness, we postulated a linear relationship between initial conditions and targets. If this assumption fails and FCS and non-FCS have systematically different initial conditions, it could yield Type I or Type II errors for the detection of quantitative tailoring. With 14 versus 21 percent of GDP, average Revenues XOXG at *T* were significantly lower for FCS than for non-FCS (see Table 2). Regressing targets on initial conditions revealed that the difference between the regression coefficients for the two types of countries was not statistically significant at the 10 percent confidence (see Table 3 and Figure 4). This allowed us to calculate the distance between the two parallel regression lines and conclude that there was little evidence of tailoring. Suppose, however, that the true relationship between initial conditions and targets was convex rather than linear. In that case, finding a similar slope for FCS and non-FCS when the initial level of Revenues XOXG was much lower for the former than for the latter could, in fact, be an indication of quantitative tailoring. To see this, notice that without tailoring, convexity of the underlying relationship implies a more negative slope at lower initial revenue levels, while identical slopes indicate tailoring. Conversely, our tentative conclusion of quantitative tailoring for the Wage Bill might be a Type I error, i.e., a false positive. Even though the difference does not reach statistical significance, the average Wage Bill at *T* is around half a p.p. of GDP lower for FCS than for non-FCS. Therefore, rather than tailoring, the flatter slope for FCS could be a consequence of a highly convex relationship between initial conditions and targets.

Un-re-benchmarking

The process of un-re-benchmarking GDP data introduces noise and potentially affects our findings regarding optimism and correlations between targets and outcomes. The process involves adjusting realized GDP data for the program years as reported in the October 2023 WEO to align with what they would have been if re-benchmarking had not taken place. While necessary to compare outcomes with initial targets, it can create inconsistencies and inaccuracies and may introduce additional variability and distortions into the data, potentially obscuring true relationships between targets and outcomes and undermining the reliability of our analysis. Consequently, it complicates the interpretation of results and may lead to less accurate conclusions.

No "Why"

We do not provide causal explanations for our empirical findings, nor an analysis of counterfactuals. It is interesting to speculate about the economic, political, and institutional reasons for our findings, especially those regarding optimism and the independence of targets and outcomes. However, in the absence of convincing instrumental variables, credible causal inference and counterfactual analyses are not possible. Therefore, we prefer not to make any claims in this regard.

Omitted Variables

Unless properly controlled for, differences in ambitiousness, optimism, or correlation attributed to FCS status might well reflect omitted variables, like per capita GDP, rather than FCS status. However, since we have found almost no such differences between the two classes of countries, we need not worry too much about incorrect attribution (of differences that are not present). A related caveat is that the number of observations (84 in total) limits the number of controls we can accommodate.

COVID-19

Somewhat surprisingly, the effect of the COVID pandemic on our findings appears to be limited. Notice that the COVID shock, which occurred in 2020, is irrelevant to the question of quantitative tailoring, because it does not rely on forecasts. Tailoring only asks whether, at the time of program approval, program targets or the relationship between initial conditions and targets differed for FCS versus non-FCS. Unexpected shocks are immaterial to this question. Perhaps more surprisingly, our other findings appear not affected much by COVID either. That is, dropping the 11 programs whose outcomes have been influenced by the epidemic does not alter our conclusions of widespread optimism and weak correlation between outcomes and targets. Some details do change, however. On the one hand, for FCS, we can no longer reject the null that targeted and realized growth are independent. On the other hand, while targets continue to significantly exceed realizations for PB, now, the two are at least positively correlated—but only if we exclude oil and grant revenues.

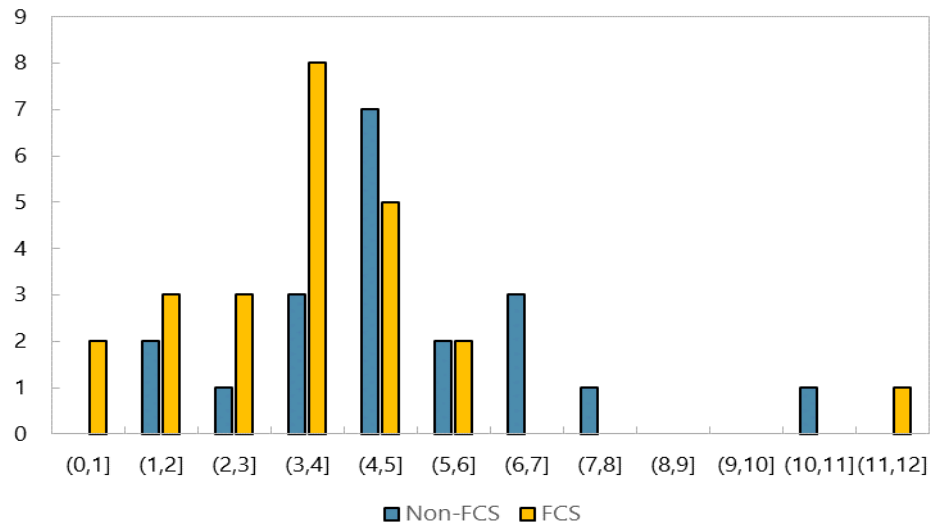
"Repeat" Customers

Many countries appear more than once in our sample. However, this should not unduly affect our conclusions. For the assessment of tailoring, our sample consists of 84 programs with 43 countries. Fourteen countries appear once, 18 twice, 10 three times and 1 country, Sierra Leone, four times. For our other assessments, we have 62 programs with 37 countries. Fourteen appear once, 19 twice, while 4 countries appear three times. Despite about half of countries having multiple programs, we treat each program as a separate observation. One might worry about correlation in the error terms. However, since we drop all programs with planned duration of less than 1.5 years, different programs for the same country tend to be spaced out in time. This reduces the probability of counting twice what might amount to highly correlated, or even essentially the same, observations and error terms. Furthermore, IMF country teams have rather high turnover, with typical staff member tenure of 2-3 years. This further reduces potential problems with correlated error terms across (sufficiently spaced-out) programs with the same country.

Ultimately, the possibility of repeat customers modifies the way 'nature' draws from the sample space, from one without replacement to one with replacement. This may influence the interpretation of our results, but it does not invalidate them. To see this, assume we had drawn 84 programs with one and the same country. Provided these programs were sufficiently spaced out, the generalized CLT ensures that the inference would not change. However, our conclusions would apply only to the sample space at hand, namely, programs with this particular country.

Still, one may wonder whether there is evidence of learning, in the sense of more realistic program designs for repeat programs. In Appendix E we consider this possibility but do not find material support for the hypothesis.

Figure 5. Repeat Customers: Time Between Program Approvals



Source: IMF Financial Data Query and IMF staff calculations.

Burden of Proof

Finally, an obvious but important caveat concerns the difference in certainty that results from not being able to reject a null hypothesis versus accepting an alternative hypothesis. Our results about (average and marginal) optimism fall in the latter, stronger category, while those on tailoring and independence fall in the former, weaker one. That is, we do not prove that IMF-supported programs with FCS are non-FCS equally ambitious, nor that most targets and outcomes are statistically independent from each other. We merely show that in most dimensions, we cannot reject these hypotheses at conventional confidence levels. However, we do positively establish that making targets country-independent would improve their accuracy for all variables other than inflation.

VII. Conclusion

Our findings support the rationale for the IMF's Strategy for Fragile and Conflict-Affected States. The strategy's focus on avoiding overly optimistic assumptions and applying more realistic macro frameworks is a positive development that could be systematically extended across all LIC programs.

Our analysis of IMF-supported programs in LICs, with a particular focus on FCS, offers the following insights. First, it is evident that, before the adoption of the IMF's Strategy for Fragile and Conflict-Affected States in 2022, there was limited differentiation in the ambition of program targets between FCS and non-FCS. It suggests that, in its program design, the IMF aimed to maintain a high degree of consistency. However, this uniformity also strengthens the rationale for the recent IMF's FCS Strategy, which emphasizes the opportunity

for refinement in tailoring programs to the specific needs and conditions of FCS, ensuring that they address the unique challenges these countries face.

Second, our findings reveal a tendency towards optimism in the macroeconomic projections of IMF-supported programs. This optimism is observed in several variables, though inflation projections are generally accurate. Recognizing this optimistic bias provides a valuable insight for future program designs. By setting more conservative and achievable targets, programs can enhance their credibility and effectiveness, fostering greater confidence among stakeholders, including governments and development partners.

Third, the observed independence of outcomes and targets for most variables suggests an area for enhancing the predictive accuracy of program targets. This insight highlights the complexity of economic forecasting and the challenges inherent in setting precise targets in dynamic environments. The finding that country-independent targets based on median program outcomes could improve accuracy underscores the potential for refining target-setting methodologies. By incorporating more robust analytical tools and a deeper understanding of country-specific factors, the IMF can further strengthen its support to LICs.

Potential caveats for our findings include the issue of GDP rebenchmarking, a non-linear relationship between initial conditions and targets, and repeat programs. While we have purposefully shied away from making causal claims, this may be a fruitful, though difficult, topic for future research.

References

Anselin L., and J. O'Loughlin. 1992. Geography of International Conflict and Cooperation: Spatial Dependence and Regional Context in Africa. In New Politics. Routledge.

Bal Gunduz, Y., Ebeke, C., Hacibedel, B., Kaltani, L., Kehayova, V., Lane, C., et al. 2013. The Economic Impact of IMF-Supported Programs in Low-Income Countries (IMF Occasional Paper, 13/277). Washington, DC: International Monetary Fund.

Baqir, R., R. Ramcharan, and R. Sahay. 2005. IMF Programs and Growth: Is Optimism Defensible? IMF Staff Papers 52(2): 260-286.

Baqir, R., R. Ramcharan, and R. Sahay. 2005. IMF Program Design, in Globalization and the Nation State: The Impact of the IMF and the World Bank, ed. By Gustav Ranis, James Vreeland, and Stephen Kosack (New York: Routledge).

Bellemare, M. F. 2015. Rising Food Prices, Food Price Volatility, and Social Unrest. American Journal of Agricultural Economics, 97(1):1–21.

Brückner, M. and A. Ciccone. 2010. International Commodity Prices, Growth and the Outbreak of Civil War in Sub-Saharan Africa. The Economic Journal, 120(544):519–534.

Buhaug, H. and K. S. Gleditsch. 2008. Contagion or Confusion? Why Conflicts Cluster in Space. International Studies Quarterly, 52(2): 215–233

Celasun, O., J. Lee, M. Mrkaic, and A. Timmermann. 2021. An Evaluation of World Economic Outlook Growth

Forecasts—2004-17, IMF Working Paper WP/21/216

Cerra, V. and S. C. Saxena. 2008. Growth Dynamics: The Myth of Economic Recovery. American Economic Review, 98(1): 439–457.

Collier, P. (1999). On the Economic Consequences of Civil War. Oxford Economic Papers, 51(1):168–183.

Collyns, C., Kuruc, K., and Takagi, S. 2021. "Assessing the Role of the IMF in Fragile States" In: *Macroeconomic Policy in Fragile States*. Edited by: Ralph Chami, Raphael Espinoza, and Peter Montiel, Oxford University Press.

Dedecker, J., P. Doukhan, G. Lang, L. Rafael, S. Louhichi, and C. Prieur Weak. 2007. Dependence: With Examples and Applications. Springer.

Easterly, W., and Rebelo, S. 1993. "Fiscal Policy and Economic Growth: An Empirical Investigation," NNBER Working Paper No. 4499.

Independent Evaluation Office of the International Monetary Fund. 2018. The IMF and Fragile States. Washington, DC: International Monetary Fund.

International Monetary Fund (IMF), 2008, "The Fund's Engagement in Fragile States and Post-Conflict Countries—A Review of Experience—Issues and Options," IMF Policy Paper, International Monetary Fund, Washington, DC.

International Monetary Fund (IMF). 2011. Macroeconomic and Operational Challenges in Countries in Fragile Situations. Washington, DC: International Monetary Fund.

International Monetary Fund (IMF). 2012. Staff Guidance Note on the Fund's Engagement with Countries in Fragile Situations. Washington, DC: International Monetary Fund.

International Monetary Fund (IMF). 2015. IMF Engagement with Countries in Post-Conflict and Fragile Situations-Stocktaking. Washington, DC: International Monetary Fund.

International Monetary Fund (IMF). 2019. Review of Conditionality and Program Design. Washington, DC: International Monetary Fund.

International Monetary Fund (IMF). 2022. The IMF Strategy for Fragile and Conflict-Affected States. Washington, DC: International Monetary Fund.

Kim, Jun I., Jean-Marc B. Atsebi, Kwang Y. Lee, Hasan H. Toprak, and Jiakun Li. 2021. Cross Country Analysis of Program Design and Growth Outcomes: 2008–19. Background Paper. Independent Evaluation Office of the International Monetary Fund.

Kuruc, Kevin. 2018. The IMF and Fragile States: Assessing Macroeconomic Outcomes. IEO Background Paper, Independent Evaluation Office of the International Monetary Fund.

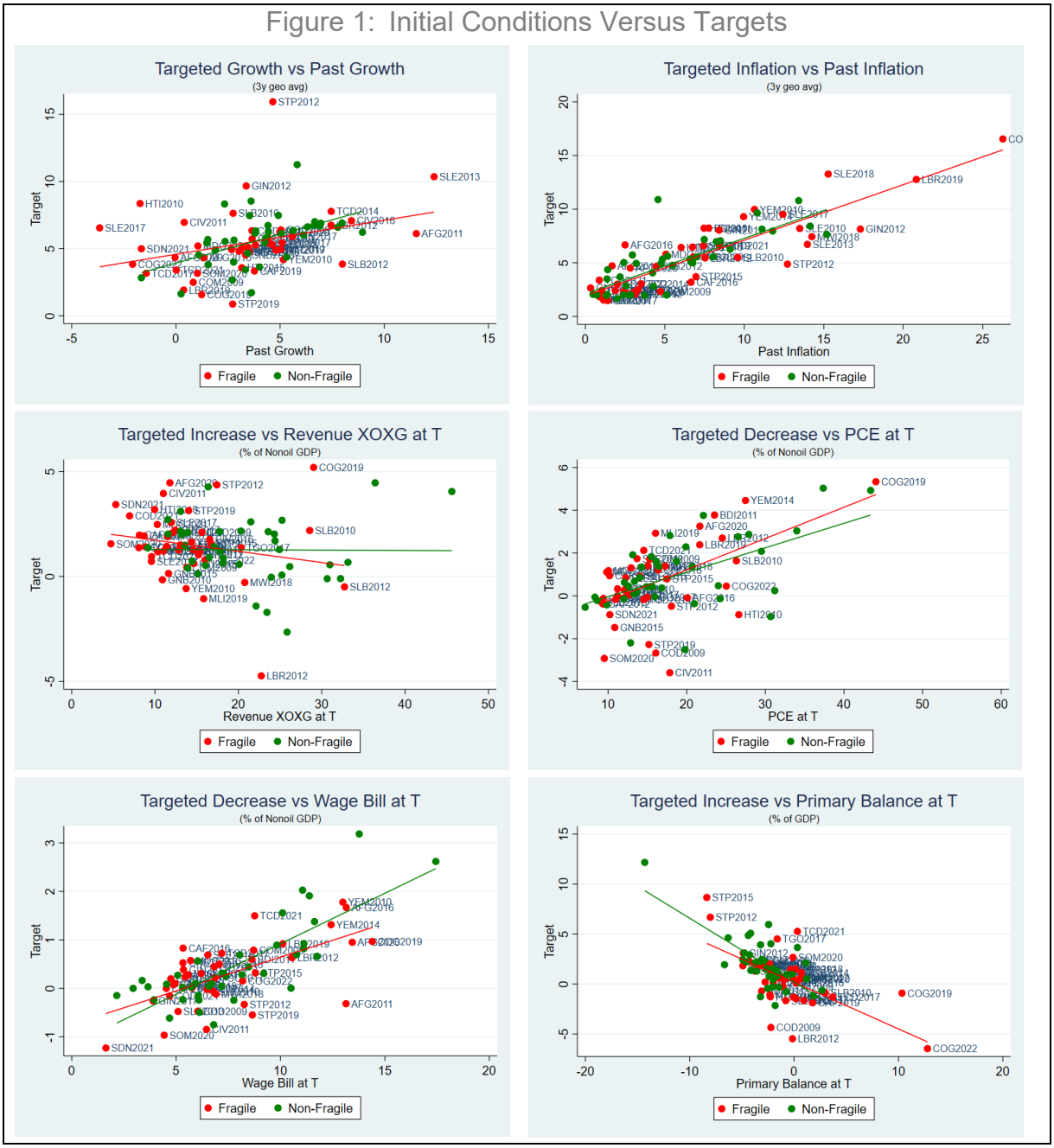
Lane, C., Turk, R., Hitaj, E., and Mehta, P. 2021. "Tailoring IMF-Supported Programs to Fragile and Conflict-Affected States' Needs: What Works?" In: *Macroeconomic Policy in Fragile States*. Edited by: Ralph Chami, Raphael Espinoza, and Peter Montiel, Oxford University Press.

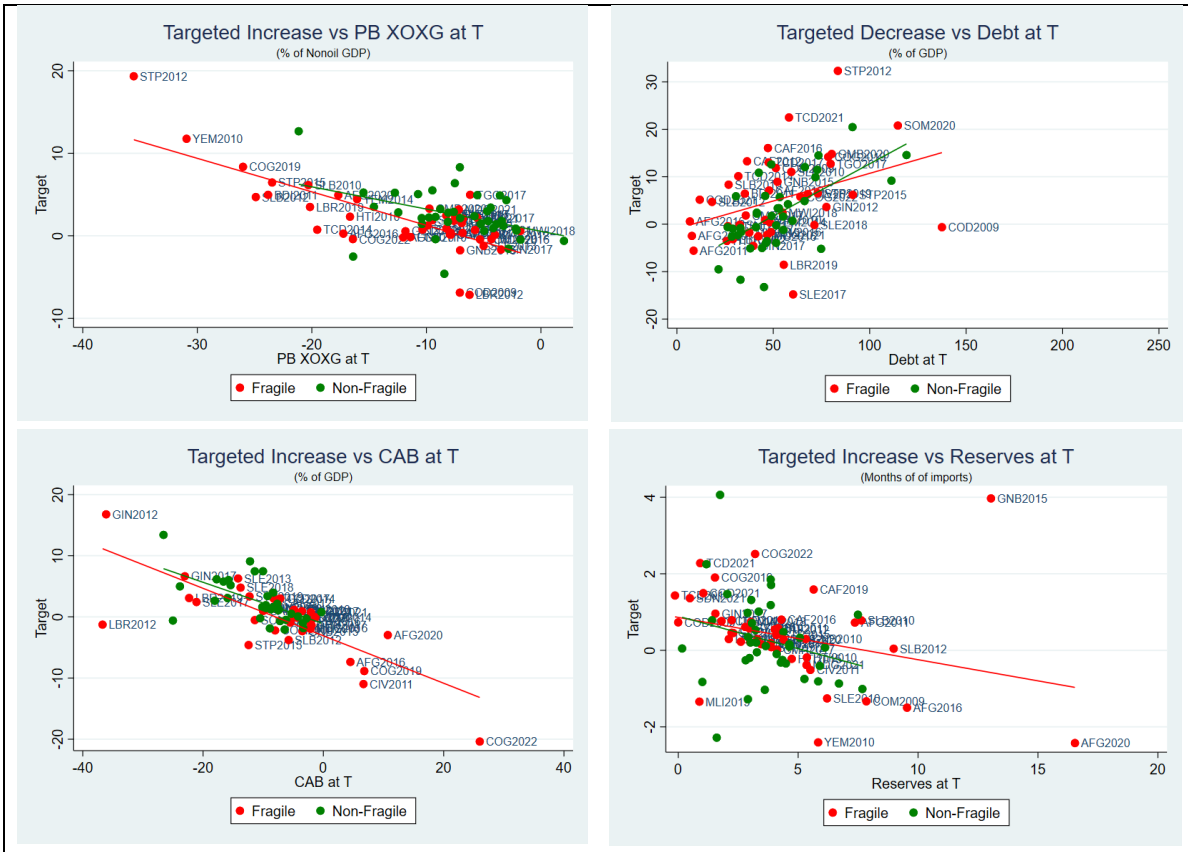
Leepipatpiboon, Patcharaporn, Chiara Castrovillari, and Tomohide Mineyama. 2023. Macroeconomic Shocks and Conflict. IMF Working Paper WP/23/68.

Rodrik, D. 1999. Where Did All the Growth Go? External Shocks, Social Conflict, and Growth Collapses. Journal of Economic Growth, 4(4):385–412.

Appendix A. Initial conditions versus Targets

Figure 1: Initial Conditions Versus Targets





The left panel shows regression lines of targets versus outcomes for FCS and non-FCS separately. The right panel shows the regression lines when we impose equality of slopes but allow for different intercepts. The distance between the parallel lines is our measure of relative ambitiousness.

Appendix B. Medians Instead of Means

Table 1: Absolute Ambitiousness

Medians [Inter-quartile Range]

Variable	Initial Condition				Target			
	#Obs (F, NF)	F	2-sided	NF	#Obs (F, NF)	F	2-sided	NF
Growth	(45, 39)	3.5 [1.1, 5.0]	≈	3.8 [2.4, 5.8]	(45, 39)	**5.1** [4.2, 6.3]	≈	**6.0** [4.5, 6.7]
Inflation	(44, 39)	6.3 [2.5, 9.8]	≈	4.8 [2.4, 8.0]	(44, 39)	5.2 [2.5, 7.7]	≈	5.0 [2.1, 7.0]
Revenue XOXG	(45, 39)	**13** [10, 16]	<***	**20** [14, 25]	(45, 39)	1.4 [.9, 2.2]	≈	1.3 [.5, 2.1]
PCE	(44, 38)	15 [11, 19]	≈	16 [13, 24]	(42, 38)	.6 [-.2, 1.6]	≈	.6 [-.1, 1.9]
Wage Bill	(44, 39)	6.6 [5.3, 8.7]	≈	7.2 [5.3, 11]	(43, 39)	.2 [-.1, .7]	≈	.2 [.0, .7]
PB	(45, 39)	-1.4 [-2.8, .2]	>**	-2.4 [-4.1, -1.6]	(45, 39)	.6 [-.9, 1.8]	<**	1.5 [.5, 2.9]
PB XOXG	(45, 39)	-8.2 [-17, -6.8]	<***	-7.1 [-9.8, -3.9]	(45, 39)	1.2 [.1, 3.3]	≈	2.2 [.9, 4.3]
Debt	(45, 39)	48 [35, 71]	≈	49 [38, 60]	(45, 39)	4.7 [-1.7, 11]	≈	1.6 [-2.6, 6.0]
CAB	(38, 33)	**-5.6** [-11, -2.2]	>**	**-9.3** [-15, -7.2]	(38, 33)	**.1** [-1.9, 2.2]	<**	**1.6** [-.2, 5.0]
Reserves	(44, 39)	3.9 [2.2, 5.4]	≈	3.6 [2.9, 4.6]	(44, 39)	.4 [.0, .8]	≈	.1 [-.3, .9]

To compare F with NF, we used a Mann-Whitney U Test, i.e., a nonparametric 2-sample test for the equality of medians of unmatched data.

Table 2. Optimism

Medians [Inter-quartile Range]

Variable		Country Type	#Obs (Paired)	Target	1-sided	Outcome	Optimism 2-sided
Growth		F	30	5.6 [4.5, 7.0]	>***	3.7 [2.6, 4.6]	1.9*** ≈ 1.2***
		NF	32	5.9 [4.4, 6.7]	>***	4.7 [4.0, 5.7]	
Inflation		F	30	5.2 [2.5, 8.0]	≈	6.2 [2.5, 9.1]	1.0 ≈ -.0
		NF	32	5.0 [2.4, 6.6]	≈	4.9 [1.7, 6.3]	
Revenue XOXG	Target ≥ 0	F	25	1.5 [1.2, 2.2]	>**	.2 [-.9, 1.7]	1.3** ≈ .6*
		NF	26	1.5 [.6, 2.1]	>*	.7 [-.3, 1.9]	
PCE	Target ≥ 0	All	41	1.4 [.7, 2.3]	>***	-.2 [-2.3, .7]	1.6***
Wage Bill	Target ≥ 0	All	45	.5 [.3, .8]	>***	.0 [-.5, .6]	.5***
PB	Target ≥ 0	All	42	1.9 [1.1, 3.7]	>***	.1 [-2.2, 1.4]	1.8***
PB XOXG	Target ≥ 0	F	22	2.3 [.6, 5.0]	≈	2.3 [-3.9, 5.8]	.0 ≈ 2.3***
		NF	27	2.8 [1.4, 5.0]	>***	.6 [-1.4, 3.0]	
Debt	Target ≥ 0	All	36	8.3 [3.4, 13]	>***	-5.5 [-12, 4.8]	14***
CAB	Target ≥ 0	All	31	2.4 [1.2, 5.2]	>**	.5 [-1.5, 3.8]	1.9**
Reserves	Target ≥ 0	All	43	.5 [.2, 1.0]	>*	.4 [-.3, .9]	.1*

To compare Targets with Outcomes, we used the Wilcoxon signed-rank test. This non-parametric test checks whether the median of the difference between paired samples is different from zero. To compare optimism for F and NF, we used the Mann-Whitney U Test. This test, which is also non-parametric, checks whether medians of two unmatched, independent samples is different from zero.

Table 3. Optimism—On-Track Only

Medians [Inter-quartile Range]

Variable		#Obs (Paired)	Target	1-sided	Outcome	Optimism
Growth		39	5.5 [4.7, 6.7]	>***	4.5 [3.7, 5.6]	1.0***
Inflation		39	4.8 [2.1, 6.7]	≈	3.6 [1.1, 7.1]	1.2
Revenue XOXG	Target ≥ 0	32	1.4 [.8, 2.1]	>*	1.1 [-.4, 2.0]	.2
PCE	Target ≥ 0	26	1.5 [.9, 2.8]	>***	-.4 [-2.3, .6]	1.9***
Wage Bill	Target ≥ 0	29	.6 [.3, .8]	>***	.1 [-.4, .6]	.5***
PB	Target ≥ 0	26	2.1 [1.4, 4.5]	>***	-.3 [-.2.6, 1.3]	2.3***
PB XOXG	Target ≥ 0	30	2.3 [.9, 5.0]	>***	1.4 [-2.5, 2.7]	.9***
Debt	Target ≥ 0	21	6.3 [3.6, 12]	>**	-.1 [-8.9, 7.1]	6.5**
CAB	Target ≥ 0	20	2.8 [1.2, 6.2]	>**	1.1 [-1.9, 3.8]	1.7*
Reserves	Target ≥ 0	28	.5 [.2, 1.0]	≈	.5 [-.3, 1.2]	.0

Appendix C. Excluding COVID-Affected Programs

Table 1. Optimism in Pre-2017 Programs

Arithmetic Averages (Standard Deviations)

Variable		Country Type	#Obs (Paired)	Target	1-sided	Outcome	Optimism 2-sided
Growth		F	23	6.4 (2.8)	>**	4.0 (3.8)	2.4** (38%) ≈ .9** (15%)
		NF	27	5.7 (2.0)	>**	4.8 (1.8)	
Inflation		F	23	5.8 (3.3)	≈	6.0 (3.8)	.2 (4%) ≈ .7 (14%)
		NF	27	5.2 (2.2)	≈	5.9 (5.6)	
Revenue XOXG	Target ≥ 0	All	41	1.6 (1.1)	>**	1.0 (2.0)	.7** (41%)
PCE	Target ≥ 0	All	32	1.7 (1.3)	>***	-.5 (2.4)	2.2*** (127%)
Wage Bill	Target ≥ 0	All	35	.8 (.8)	>***	-.2 (1.5)	1.0*** (127%)
PB	Target ≥ 0	All	34	2.6 (2.6)	>***	.06 (3.0)	2.5*** (98%)
PB XOXG	Target ≥ 0	All	39	3.8 (3.9)	>**	2.4 (4.3)	1.3** (36%)
Debt	Target ≥ 0	All	27	8.6 (6.0)	>***	-2.4 (15)	11*** (128%)
CAB	Target ≥ 0	All	21	4.0 (4.2)	>**	1.1 (4.8)	2.9** (73%)
Reserves	Target ≥ 0	All	33	.8 (1.0)	>**	.3 (1.4)	.5** (64%)

To compare Targets with Outcomes, we tested for a difference in means using a paired t-test. To compare optimism for F versus NF, we tested for a difference in average optimism using an independent t-test

Table 2. Optimism in Pre-2017 Programs

Medians [Inter-quartile Range]

Variable		**Country Type**	**#Obs (Paired)**	**Target**	1-sided	**Outcome**	**Optimism** 2-sided
Growth		F	23	5.8 [4.5, 7.6]	>**	3.8 [2.7, 5.0]	2.0** ≈ 1.1**
		NF	27	5.8 [4.4, 6.7]	>**	4.7 [4.0, 5.9]	
Inflation		F	23	5.5 [3.0, 8.1]	≈	6.2 [2.6, 8.4]	.7 ≈ -.0
		NF	27	5.3 [2.9, 7.1]	≈	5.3 [1.8, 6.8]	
Revenue XOXG	Target ≥ 0	All	41	1.5 [.7, 2.1]	>**	1.1 [-.7, 2.0]	.4**
PCE	Target ≥ 0	All	32	1.4 [.7, 2.7]	>***	-.1 [-2.0, 1.1]	1.5***
Wage Bill	Target ≥ 0	All	35	.6 [.3, .9]	>***	.0 [-.5, .6]	.6***
PB	Target ≥ 0	All	34	1.7 [1.0, 3.1]	>***	.2 [-1.7, 1.4]	1.5***
PB XOXG	Target ≥ 0	All	39	2.4 [1.4, 5.0]	>**	2.0 [-.8, 4.4]	.4**
Debt	Target ≥ 0	All	27	8.3 [3.6, 12]	>***	-3.3 [-12, 7.1]	12***
CAB	Target ≥ 0	All	21	3.0 [1.4, 5.2]	>**	1.7 [.2, 3.8]	1.3**
Reserves	Target ≥ 0	All	33	.5 [.3, .8]	≈	.4 [-.3, .7]	.1

To compare Targets with Outcomes, we used the Wilcoxon signed-rank test. This non-parametric test checks whether the median of the difference between paired samples is different from zero. To compare optimism for F and NF, we used the Mann-Whitney U Test. This test, which is also non-parametric, checks whether medians of two unmatched, independent samples is different from zero.

Table 3. Optimism in Pre-2017 Programs—On-Track Only

Arithmetic Averages (Standard Deviations)

Variable		#Obs (Paired)	Target	1-sided	Outcome	Optimism
Growth		32	5.8 (2.2)	>**	4.5 2.3)	1.2** [21%]
Inflation		32	5.1 (2.4)	≈	5.3 (4.8)	.2 [4%]
Revenue XOXG	Target ≥ 0	26	1.7 (1.2)	≈	1.3 (2.0)	.3 [21%]
PCE	Target ≥ 0	21	2.0 (1.4)	>***	-.8 (2.7)	2.8*** [137%]
Wage Bill	Target ≥ 0	24	.8 (.7)	>***	.1 (1.0)	.7*** [86%]
PB	Target ≥ 0	20	3.1 (3.0)	>***	-.6 (3.2)	3.6*** [118%]
PB XOXG	Target ≥ 0	24	3.3 (3.1)	>**	1.4 (3.9)	1.9** [57%]
Debt	Target ≥ 0	17	7.1 (4.5)	>**	.8 (13)	6.3** [89%]
CAB	Target ≥ 0	13	4.9 (5.0)	>**	1.3 (5.3)	3.6** [74%]
Reserves	Target ≥ 0	22	.8 (.9)	≈	.4 (1.3)	.4 [46%]

Table 4. Optimism in Pre-2017 Programs—On-Track Only

Medians [Inter-quartile Range]

Variable		#Obs (Paired)	Target	1-sided	Outcome	Optimism
Growth		32	5.5 [4.4, 6.8]	>**	4.5 [3.8, 5.2]	1.0**
Inflation		32	5.1 [2.7, 6.9]	≈	4.9 [2.0, 7.2	-.1
Revenue XOXG	Target ≥ 0	26	1.4 [.7, 2.1]	≈	1.4 [-.3, 2.6]	-.0
PCE	Target ≥ 0	21	1.6 [.9, 2.8]	>***	-.2 [-2.3, .6]	1.8***
Wage Bill	Target ≥ 0	24	.6 [.3, 1.1]	>***	.1 [-.4, .7]	.5***
PB	Target ≥ 0	20	2.0 [1.3, 4.0]	>***	-.4 [-2.4, 1.3]	2.4***
PB XOXG	Target ≥ 0	24	2.2 [.7, 5.1]	>**	2.0 [-1.8, 3.0]	.2**
Debt	Target ≥ 0	17	6.3 [3.6, 11]	>*	1.1 [-7.8, 7.1]	5.2*
CAB	Target ≥ 0	13	3.4 [1.4, 6.1]	>*	1.7 [-1.5, 3.8]	1.7*
Reserves	Target ≥ 0	22	.6 [.3, 1.0]	≈	.5 [-.3, .9]	.1

Table 5. Correlations in Pre-2017 Programs

Variable		Country Type	#Obs (Paired)	Slope	Pearson Corr.	Spear-man ρ	Ken-dall τ	R^2	Theil U_2	
									Mean	Median
Growth		F	23	-.04	-.03	.25	.19	-1.05	1.12	1.40
		NF	27	.44	**.51*** **	**.36***	**.27***	-.41	.55	1.11
Inflation		F	23	**.9*** **	**.79*** **	**.77*** **	**.60*** **	.61	.40	.62
		NF	27	**1.7*** **	**.71*** **	**.77*** **	**.61*** **	.39	.62	.78
Rev. XOXG	Target ≥ 0	All	41	.25	.14	.11	.08	-.27	1.08	1.09
PCE	Target ≥ 0	All	32	.34	.19	.13	.08	-.96	1.37	1.37
Wage Bill	Target ≥ 0	All	36	.25	.13	.17	.11	-1.01	1.23	1.23
PB	Target ≥ 0	All	34	.29	.25	.21	.14	-.99	1.39	1.40
PB XOXG	Target ≥ 0	All	39	**.51*** **	**.46*** **	**.36** **	**.25** **	-.08	.98	1.02
Debt	Target ≥ 0	All	27	-.34	-.15	-.23	-.15	-.93	1.36	1.37
CAB	Target ≥ 0	All	21	.26	.23	.18	.07	-.74	1.29	1.25
Reserves	Target ≥ 0	All	33	.01	.01	.17	.11	-.58	1.23	1.24

Table 6. Correlations in Pre-2017 Programs—On Track Only

Variable		Type	#Obs (Paired)	Slope	Pearson Corr.	Spear-man ρ	Ken-dall τ	R^2	Theil U_2	
									Mean	Median
Growth		All	32	.05	.05	.23	.27	-1,14	1.42	1.45
Inflation		All	32	**1.47*** **	**.74*** **	**.79*** **	**.62*** **	.48	.70	.71
Revenue XOXG	Target ≥ 0	All	26	.26	.16	.08	.09	-.20	1.05	1.09
PCE	Target ≥ 0	All	21	.60	.32	.29	.21	-1.09	1.38	1.38
Wage Bill	Target ≥ 0	All	24	.43	.30	.34	.21	-.60	1.21	1.23
PB	Target ≥ 0	All	20	.37	.35	.35	.24	-1.62	1.54	1.48
PB XOXG	Target ≥ 0	All	24	**.50***	**.40***	.31	.10	-.23	1.06	1.09
Debt	Target ≥ 0	All	17	-.27	-.09	-.13	-.12	-.44	1.13	1.17
CAB	Target ≥ 0	All	13	.24	.24	.14	.05	-.93	1.28	1.31
Reserves	Target ≥ 0	All	22	-.20	-.14	.08	.06	-.80	1.28	1.32

Appendix D. Including Negative Adjustments

Table 1. Optimism with Negative Adjustment Targets

Arithmetic Averages (Standard Deviations)

Variable	Country Type	#Obs (Paired)	Target	1-sided	Outcome	Optimism 2-sided
Revenue XOXG	F	30	1.3 (1.8)	>**	.5 (2.2)	1.1** (61%) ≈ .2 (15%)
	NF	32	1.1 (1.5)	≈	.9 (1.9)	
PCE	F	27	.6 (1.8)	>**	-.6 (2.5)	1.2** (211%) ≈ 2.3*** (181%)
	NF	31	1.2 (1.7)	>***	-1.0 (3.1)	
Wage Bill	F	28	.3 (.6)	>***	-.3 (1.0)	.6*** (189%) ≈ 1.0*** (178%)
	NF	32	.6 (.9)	>***	-.4 (1.6)	
PB	F	30	.6 (2.8)	>*	-.4 (3.4)	1.0* (175%) <* 2.6*** (130%)
	NF	32	2.0 (2.7)	>***	-.6 (2.8)	
PB XOXG	F	30	2.2 (5.0)	≈	1.1 (6.0)	1.1 (51%) ≈ 2.3*** (86%)
	NF	32	2.7 (3.2)	>***	.4 (3.1)	
Debt	F	28	4.8 (8.9)	>***	-3.3 (12)	8.1*** (168%) ≈ 9.7*** (410%)
	NF	32	2.4 (7.6)	>***	-7.3 (12)	
CAB	F	23	-.2 (5.7)	≈	-.4 (5.7)	.2 (-78%) <* 3.0*** (124%)
	NF	26	2.4 (3.4)	>***	-.6 (4.6)	
Reserves	F	30	.3 (1.1)	≈	.1 (5.7)	.2 (63%) ≈ .4 (129%)
	NF	32	.3 (1.1)	≈	-.1 (1.7)	

Table 2. Optimism with Negative Adjustment Targets—On Track Only

Arithmetic Averages (Standard Deviations)

Variable		#Obs (Paired)	Target	1-sided	Outcome	Optimism
Growth		39	5.7 (2.0)	>***	4.5 (2.1)	1.2*** [21%]
Inflation		39	4.7 (2.5)	≈	4.8 (4.6)	.1 [2%]
Revenue XOXG	Target ≥ 0	38	1.1 (1.7)	≈	1.0 (2.1)	.1 [13%]
PCE	Target ≥ 0	37	1.0 (1.9)	>***	-1.1 (3.2)	2.1*** [217%]
Wage Bill	Target ≥ 0	39	.5 (.7)	>***	-.1 (1.0)	.6*** [128%]
PB	Target ≥ 0	39	1.6 (3.1)	>***	-1.1 (3.1)	2.7*** [165%]
PB XOXG	Target ≥ 0	39	2.0 (3.6)	>***	-.2 (4.3)	2.3*** [112%]
Debt	Target ≥ 0	39	2.7 (7.1)	>***	-3.2 (11)	5.8*** [218%]
CAB	Target ≥ 0	34	1.3 (5.2)	>**	-1.1 (5.3)	2.4** [190%]
Reserves	Target ≥ 0	39	.3 (1.1)	≈	.1 (1.6)	.2 [57%]

Table 3. Correlation with Negative Adjustment Targets

Variable	Type	#Obs (Paired)	Slope	Pearson Corr.	Spearman ρ	Kendall τ	R^2	Theil U_2	
								Mean	Median
Revenue XOXG	All	62	.21	.17	.13	.09	-.44	1.18	1.19
PCE	All	58	.43**	.27**	.14	.09	-.45	1.19	1.18
Wage Bill	All	59	.47***	.36***	.38**	.27**	-51	1.21	1.21
PB	All	62	.22	.20	.20	.13	-.82	1.33	1.32
PB XOXG	All	62	.55***	.49***	.39**	.27**	-.06	1.01	1.00
Debt	All	60	.08	.06	.03	.03	-1.00	1.39	1.40
CAB	All	49	.29*	.27*	.28*	.20**	-.46	1.18	1.19
Reserves	All	62	.22	.16	.18	.12	-.32	1.13	1.09

Table 4. Correlation with Negative Adjustment Targets—On Track Only

Variable	Type	#Obs (Paired)	Slope	Pearson Corr.	Spear-man ρ	Ken-dall τ	R^2	Theil U_2	
								Mean	Median
Revenue XOXG	All	39	.28	.23	.10	.08	-.29	1.11	1.07
PCE	All	38	.55	.32**	.19	.14	-.42	1.16	1.15
Wage Bill	All	37	.63	.46***	.48	.33	-.26	1.09	1.09
PB	All	39	.23	.23	.26	.17	-1.30	1.48	1.48
PB XOXG	All	39	.62	.51***	.39**	26**	-.12	1.03	1.04
Debt	All	39	.26	.17	.10	.06	-.50	1.19	1.19
CAB	All	34	.36	.36**	.32*	.22*	-.48	1.18	1.13
Reserves	All	39	.26	.17	.20	.14	-.25	1.09	1.09

Appendix E. Initial vs Repeat Programs

One may wonder whether there is evidence of learning in our data, in the sense of more realistic program designs for "repeat" (R) programs. Here, repeat programs are programs with countries that have a preceding program in our sample. All other programs are referred to as "initial" or "non-repeat" (NR) programs. Of course, this labeling is imperfect, since some initial programs may have had predecessors before the start of our sample period. Nonetheless, it may be interesting to see if we can find differences between the two groups.

Tables 1 to 5, below, reveal very little material differences between NR versus R programs in terms of ambitiousness, optimism, or correlation. Our sample contains 84 programs across 43 countries for the ambitiousness assessment and, maximally, 62 programs across 37 countries for optimism and correlation. This implies that we have 43 and 37 NR programs versus 41 and 25 R programs, respectively.

Table 1: Absolute Ambitiousness

Arithmetic Averages (Standard Deviations)

Variable	Initial Condition				Target			
	#Obs (NR, R)	NR	2-sided	R	#Obs (NR, R)	NR	2-sided	R
Growth	(43, 41)	4.0 (2.8)	≈	3.4 (2.9)	(43, 41)	**6.1** (2.5)	>**	**5.0** (1.8)
Inflation	(42, 41)	6.8 (5.1)	≈	5.8 (4.9)	(42, 41)	5.3 (3.1)	≈	5.3 (3.1)
Revenue XOXG	(43, 41)	18 (8.3)	≈	16 (7.0)	(43, 41)	1.5 (1.9)	≈	1.3 1.2
PCE	(41, 41)	19 (8.9)	≈	18 (8.8)	(40, 40)	.7 (2.1)	≈	.9 (1.3)
Wage Bill	(42, 41)	7.6 (3.6)	≈	7.5 (2.5)	(42, 40)	.3 (.9)	≈	.4 (.6)
PB	(43, 41)	-2.2 (3.4)	≈	-1.4 (3.2)	(43, 41)	1.3 (2.8)	≈	1.2 (2.6)
PB XOXG	(43, 41)	-11 (7.6)	≈	-9.0 (6.0)	(43, 41)	2.6 (4.6)	≈	2.0 (2.4)
Debt	(43, 41)	61 (45)	≈	50 (22)	(41, 41)	4.3 (8.8)	≈	3.0 (7.9)
CAB	(31, 40)	-10 (9.9)	≈	-7.6 (9.1)	(31, 40)	1.7 (5.3)	≈	.5 (4.6)
Reserves	(42, 41)	4.0 (1.9)	≈	4.1 (3.3)	(42, 41)	.2 (1.0)	≈	.4 (1.2)

To compare F with NF, we used a Mann-Whitney U Test, i.e., a nonparametric 2-sample test for the equality of medians of unmatched data.

Table 2: Absolute Ambitiousness

Medians [Inter-quartile Range]

Variable	Initial Condition				Target			
	#Obs (NR, R)	NR		R	#Obs (NR, R)	NR		R
Growth	(43, 41)	3.8 [2.3, 5.8]	≈	3.6 [1.9, 5.0]	(43, 41)	5.8 [4.8, 6.9]	>**	4.9 [3.8, 6.2]
Inflation	(42, 41)	5.1 [3.1, 9.1]	≈	4.6 [1.9, 7.5]	(42, 41)	5.0 [2.4, 7.1]	≈	4.7 [2.7, 6.7]
Revenue XOXG	(43, 31)	15 [12, 24]	≈	14 [12, 20]	(43, 41)	1.5 [.6, 2.2]	≈	1.4 [.6, 2.1]
PCE	(41, 41)	16 [13, 24]	≈	15 [12, 21]	(40, 40)	.3 [-.3, 1.8]	≈	.8 [-.0, 1.7]
Wage Bill	(42, 41)	6.6 [5.4, 11]	≈	7.1 [5.3, 8.8]	(42, 40)	.1 [-.1, .7]	≈	.3 [-.1, .7]
PB	(43, 41)	-2.3 [-3.2, -.7]	≈	-1.8 [-3.4, .1]	(43, 41)	1.1 [-.2, 2.3]	≈	1.2 [-.3, 2.1]
PB XOXG	(43, 41)	-8.1 [-12, -5.6]	≈	-7.6 [-10, 5.0]	(43, 41)	2.1 [.4, 4.9]	≈	1.5 [.4, 3.5]
Debt	(43, 41)	49 [33, 73]	≈	48 [38, 58]	(41, 41)	3.4 [-2.5, 11]	≈	1.6 [-1.6, 6.5]
CAB	(31, 40)	-8.9 [-12, -5.3]	≈	-7.4 [-12, -5.3]	(31, 40)	1.4 [-1.3, 3.1]	≈	.9 [-.9, 2.9]
Reserves	(42, 41)	3.8 [2.8, 5.3]	≈	3.5 [2.3, 4.6]	(42, 41)	.3 [-.3, .7]	≈	.4 [-.1, 1.0]

To compare F with NF, we used a Mann-Whitney U Test, i.e., a nonparametric 2-sample test for the equality of medians of unmatched data.

Table 3. Optimism

Arithmetic Averages (Standard Deviations)

Variable		Country Type	#Obs (Paired)	Target	1-sided	Outcome	Optimism 2-sided
Growth		NR	36	6.2 (2.6)	>***	4.5 (3.5)	1.7*** ≈ 1.5***
		R	26	5.3 (1.7)	>***	3.8 (2.1)	
Inflation		NR	36	5.4 (3.0)	≈	5.9 (4.0)	.5 ≈ .5
		R	26	4.8 (2.5)	≈	5.3 (3.0)	
Revenue XOXG	Target ≥ 0	NR	29	2.0 (1.3)	>**	.9 (2.2)	1.1** ≈ .7**
		R	22	1.4 (.8)	>**	.7 (1.5)	
PCE	Target ≥ 0	All	41	1.7 (1.4)	>***	-.6 (2.2)	2.3***
Wage Bill	Target ≥ 0	All	44	.7 (.7)	>***	-.0 (.9)	.7***
PB	Target ≥ 0	All	42	2.6 (2.4)	>***	-.2 (3.2)	2.8***
PB XOXG	Target ≥ 0	NR	29	4.2 (4.3)	>**	2.0 (5.1)	2.2** ≈ 2.5***
		R	20	2.9 (2.3)	>***	.8 (3.6)	
Debt	Target ≥ 0	All	36	9.6 (8.4)	>***	-3.2 (14)	13***
CAB	Target ≥ 0	All	31	3.6 (3.7)	>***	.9 (4.8)	2.7***
Reserves	Target ≥ 0	NR	23	.7 (.9)	≈	.4 (1.5)	.3 ≈ .8**
		R	20	.9 (1.0)	>**	.1 (1.4)	

The table compares targets to outcomes in terms of arithmetic averages and standard deviations. Optimism is defined as the difference between the two. "Target ≥ 0" indicates that we limit attention to programs targeting a positive adjustment. If the number of paired observations, #Obs (paired), is less than 20 for R or NR, then programs are pooled ("All"). Otherwise, in the last column, average optimism is compared across Repeat and non-Repeat programs. Numbers in square brackets express optimism as a percentage of target. To compare Targets with Outcomes, we tested for a difference in means using a paired t-test. To compare optimism for R versus NR, we tested for a difference in average optimism using an independent t-test.

Table 4. Optimism

Medians [Inter-quartile Range]

Variable		Country Type	#Obs (Paired)	Target		Outcome	Optimism
Growth		NR	36	6.0 [4.7, 6.9]	>***	4.2 [3.2, 5.8]	1.8*** ≈ 1.2***
		R	26	5.4 [4.0, 6.3]	>***	4.2 [2.7, 4.9]	
Inflation		NR	36	5.2 [2.7, 7.4]	≈	5.7 [2.5, 7.2]	.5 ≈ -1.5
		R	26	4.5 [2.4, 6.7]	≈	3.1 [1.0, 7.4]	
Revenue XOXG	Target ≥ 0	NR	29	1.5 [1.2, 2.2]	>**	.6 [-.7, 1.5]	.9** ≈ .7**
		R	22	1.3 [.6, 2.1]	>**	.6 [-.3, 2.0]	
PCE	Target ≥ 0	All	41	1.4 [.7, 2.3]	>***	-.2 [-2.3, .7]	1.6***
Wage Bill	Target ≥ 0	All	44	.5 [.3, .8]	>***	.0 [-.5, .6]	.5***
PB	Target ≥ 0	All	42	1.9 [1.1, 3.7]	>***	.1 [-2.2, 1.4]	1.8***
PB XOXG	Target ≥ 0	NR	29	2.8 [1.7, 5.0]	>**	1.5 [-1.2, 4.4]	1.3** ≈ .7**
		R	20	2.3 [.8, 4.8]	>**	1.6 [-2.6, 2.8]	
Debt	Target ≥ 0	All	36	8.3 [3.4, 13]	>***	-5.5 [-12, 4.8]	14***
CAB	Target ≥ 0	All	31	2.4 [1.2, 5.2]	>**	.5 [-1.5, 3.8]	1.9**
Reserves	Target ≥ 0	NR	23	.5 [.2, .8]	≈	.5 [-.0, 1.0]	-.0** ≈ .4
		R	20	.6 [.2, 13]	>**	.2 [-.4, .9]	

To compare Targets with Outcomes, we used the Wilcoxon signed-rank test. This non-parametric test checks whether the median of the difference between paired samples is different from zero. To compare optimism for R and NR, we used the Mann-Whitney U Test. This test, which is also non-parametric, checks whether medians of two unmatched, independent samples is different from zero.

Table 5. Correlations

Variable		Type	#Obs (Paired)	Slope	Pearson Corr.	Spear-man ρ	Ken-dall τ	R^2	Theil U_2	
									Mean	Median
Growth		NR	36	**.32***	.27	**.30***	**.22***	-.53	.94	1.34
		R	26	-.20	-.16	**.40****	**.34****	-1.47	.96	1.54
Inflation		NR	36	**.95*****	**.72*****	**.80*****	**60*****	.51	.53	.69
		R	26	**1.9*****	**.83*****	**.85*****	**.69*****	.53	.57	.68
Revenue XOXG	Target ≥ 0	NR	29	.02	.02	-.02	-.01	-.62	1.21	1.26
		R	22	.71	**.40***	.4	.31	-.09	.97	1.04
PCE	Target ≥ 0	All	41	.38	.24	.19	.11	-1.14	1.43	1.42
Wage Bill	Target ≥ 0	All	44	.23	.12	.14	.08	-.91	1.35	1.38
PB	Target ≥ 0	All	42	.20	.15	.13	.09	-1.11	1.42	1.43
PB XOXG	Target ≥ 0	NR	29	**.79***	**.46***	.25	.17	-.31	1.10	1.12
		R	20	**.42****	**.51****	**.44***	**.31***	-.12	1.03	1.01
Debt	Target ≥ 0	All	35	-.15	-.07	-.12	-.07	-1.21	1.44	1.47
CAB	Target ≥ 0	All	31	.13	.10	.01	-.01	-.75	1.28	1.26
Reserves	Target ≥ 0	NR	23	.06	.03	.00	-.04	-.32	1.11	1.14
		R	20	-.07	-.05	.11	.05	-.94	1.37	1.34

The table reports on the strength of the monotone relationship between targets and outcomes, as well as on how well the former forecasts the latter. "Slope" refers to the regression coefficient of a simple linear regression between outcomes and targets, while "Pearson Corr." represents the associated correlation coefficient. "Spearman ρ" and "Kendall τ" test for statistical independence between targets and outcomes. "Theil U_2" measures how well outcomes are forecasted by the mean and median outcomes of other programs compared to using program targets. A number greater than 1 means that the naive forecasts do better on average. "Handicap" only uses those programs whose outcomes were known at the time a program was initiated. "R^2" gives the proportion of the variance in outcomes that is forecasted by program targets. A negative value means that targets tend to deviate more from outcomes than the uniform average.

Appendix F. How Much Does Growth Optimism Explain?

Here we ask to what extent optimism, observed in all dimensions other than inflation, could have a single, common cause, namely, optimism about growth. We first consider this question conceptually. Let Y denote nominal GDP, while X denotes another (nominal) variable, such as revenue or expenditure. Deflating X and Y by price level is P yields real values x and y, while the ratio of X over Y is given *r*. Let $\eta_{X,Y}$ and $\eta_{x,y}$ denote the elasticities of X with respect to Y and x with respect to y, respectively. It can be easily verified that $\eta_{X,Y} = \eta_{x,y}$. Furthermore,

$$\frac{dr}{dY/Y} \begin{matrix} < \\ = \\ > \end{matrix} 0 \Leftrightarrow \eta_{X,Y} \begin{matrix} < \\ = \\ > \end{matrix} 1 \,. \tag{1}$$

Thus, a negative surprise for growth helps to explain optimism about variable *r* iff $\eta_{X,Y} > 1$. This is intuitive, because $\eta_{X,Y} > 1$ means that a one percent drop in growth, Y, is associated with a greater than one percent drop in X.

Tables 1 and 2 report on the correlations between forecast errors for growth and those for the other variables, as well the R^2 of regressing the latter on the former. Here, forecast errors are defined as outcome minus target. Linking correlations to precise values for $\eta_{X,Y}$ is not straightforward. However, for a variable, *r*, expressed as a percentage of GDP, equation (1) tells us that we should expect this correlation to take on the sign of $Log[\eta_{X,Y}]$. That is, forecast errors for growth and those for *r* are positively (negatively) correlated iff $\eta_{X,Y}$ is greater (smaller) than 1.

Table 1. Correlations of Forecast Errors for Growth and Other Variables

Growth	Inflation	Revenue XOXG	PCE	Wage Bill	PB	PB XOXG	Debt	CAB	Reserves
R^2	.04	.02	.14	.00	.00	.00	.38	.01	.00
Pearson Corr.	-.20	.15	**.38****	.05	-.04	.01	**.62*****	-.08	.02
Spearman ρ	**-.36*****	.20	**.36****	.07	-.04	.05	**.55*****	-.01	.03
Kendall τ	**-.26*****	.13	**.26****	.05	-.01	.03	**.40*****	-.00	.3

The table reports on the strength of the monotone relationship between forecast errors for Growth and those for other variables, using Targets as our forecasting variable. "R^2"refers to the coefficient of determination of a simple linear regression between the forecast errors, while "Pearson Corr." refers the associated correlation coefficient. "Spearman ρ" and "Kendall τ" test for statistical independence between targets and outcomes.

Table 2. Correlations of Forecast Errors for Growth and Other Variables—On Track Only

Growth	Inflation	Revenue XOXG	PCE	Wage Bill	PB	PB XOXG	Debt	CAB	Reserves
R^2	.01	.00	.06	.00	.05	.02	.44	.01	.04
Pearson Corr.	-.11	.07	.25	.05	.22	.14	**.67****	-.11	-.20
Spearman ρ	-.18	.16	.27	.16	.19	.10	**.54****	.09	-.13
Kendall τ	-.11	.10	.20	.11	.15	.07	**.39****	.07	-.07

First, consider Revenue XOXG. Estimates for the elasticity of government revenue with respect to shocks to GDP vary across countries, time periods, and the types of revenues considered (see, e.g., Easterly and Rebello, 1993). On average, estimates suggest that the elasticity is typically slightly greater than one. This implies that, in the short run, a 1 percent increase in GDP tends to lead to a slightly more than 1 percent increase in government revenue. This finding aligns with the notion that economic expansions both broaden the tax base and increase tax revenues from progressive tax sources, such as income taxes, at a faster rate than GDP growth. However, LICs have weak administrative capacity and rely heavily on international trade taxes, while income taxes are only important in developed economies. Thus, LICs often have lower elasticities, whereas advanced economies with stronger institutions and broader tax bases tend to have higher elasticities. This suggests that, to a first order, an elasticity of 1 is an appropriate estimate, which implies that revenues as a percentage of GDP would not have been materially different if growth had been as predicted. Thus, we would expect growth optimism not to be a good explanation for shortfalls in revenues. And indeed, Tables 1 and 2 show that for Revenue XOXG, the point estimate of the correlation is small and not statistically significant. This suggests an elasticity of around 1, which is consistent with our prediction.

Next, consider government current expenditure. The elasticity of expenditure with respect to a GDP shock is even more variable than that for revenue and depends, of course, heavily on government policy choices and economic conditions. The overall elasticity is typically somewhat less than one, meaning that government spending tends to grow slower than GDP, at least in the short run. Obviously, this generalization varies significantly depending on the fiscal policy stance. With few automatic stabilizers and considerable liquidity constraints an elasticity that is positive by decidedly less than 1 would be a reasonable guess. Such an elasticity implies a negative relationship between unexpectedly higher growth and expenditure as a percentage of GDP. In turn, it yields a *positive* relationship between realized growth and expenditure *reduction.* In the full sample (Table 1), this prediction is also borne out by the data, at least for primary current expenditure, PCE: at .38, the correlation is significantly positive, both economically and statistically. For on-track programs (Table 2), the correlation is somewhat lower (.25) and not statistically significant, perhaps due to the smaller number of observations.

The essentially zero correlations for PB (XOXG) in the full sample may seem surprising, considering the positive correlations for PCE and the positive but insignificant one for Revenue XOXG. The missing element is, of course, public investment. If public investment spending is fixed in the short run, then the correlation between the forecast errors for growth and those for public investments as a percentage of GDP is negative. This could, in principle, explain the zero correlation for PB (XOXG), despite the positive correlation for PCE and

the non-significant one for Revenue XOXG. For on-track programs, the correlations are positive but not statistically significant.

Finally, consider debt, inflation, CAB, and reserves. The strong positive correlations between forecast errors for growth and those for debt (reduction) are consistent with the “snowball effect” of a negative growth shock on the debt-to-GDP ratio, especially at high initial debt levels. The negative correlations between forecast errors for growth and inflation suggest that, in our sample, growth shocks came predominantly from the supply side, rather than the demand side. Finally, we do not have strong priors about the correlations for CAB and Reserves and simply note that they are not statistically significant.

Overall, the results in Table 1 and 2 suggest that growth optimism is an important contributor to optimism about Debt reduction, and to a lesser degree for PCE.

Appendix G. GDP Re-benchmarking

The analysis conducted involved comparing the actual economic developments with the initial projections/targets set under the IMF-supported program. For each program, we utilized two vintages of WEO data: for the initial conditions and program targets, we used the first WEO vintage published after the program was adopted; and for outcomes, we used the October 2023 WEO. At times, we noted substantial differences between the two vintages regarding historical GDP data. This reflected the fact that the GDP series of some countries had been re-benchmarked during the period between the two WEO vintages. To allow us to compare the initial conditions/targets, which were expressed on the basis of the initial GDP series, to the outcomes, which were expressed on the basis of the re-benchmarked GDP, we had to conduct an "un-re-benchmarking" process as follows:

Calculation of Ratios: For the pre-sample period years 2003 to 2007, we determined for each country the ratios between the nominal GDP estimates from the 2023 WEO and those from the first WEO after program approval. This involved dividing the GDP of the 2023 vintage by that of the original vintage for each of the five years, generating five distinct (but usually similar) ratios.

Geometric Mean: For each country, we calculated the geometric means of these ratios, taking the fifth root of the product of the five ratios.

Un-re-benchmarking: The geometric means of the ratios served as an adjustment factor. Dividing the nominal GDP for *T*+3 as reported in the October 2023 WEO by this mean provided us with an estimate of what nominal GDP would have been under the original benchmark. This enabled a comparison between outcomes and targets.

In our study of 84 programs, 26 had a ratio equal to 1, indicating that no re-benchmarking had taken place. An additional 40 programs had ratios within the [0.95,1.05] range, indicating limited re-benchmarking. The remaining 18 had undergone more sizable re-benchmarking. We applied the same methodology to Non-oil GDP, where 29 programs had ratios equal to 1 and 42 programs featured geometric ratios within the [0.95,1.05] range.

Appendix H. Country Classifications

Table 1. FCS Status 2008-2023

Country	ifs	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Afghanista	512	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Albania	914	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Angola	614	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0
Bosnia and	963	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0
Burundi	618	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Cambodia	522	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Cameroon	622	1	1	1	0	0	0	0	0	0	1	1	0	0	0	0	1
Central Afr	626	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Chad	628	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Comoros	632	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Congo, Re	634	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Côte d'Ivoi	662	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
Democrati	636	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Djibouti	611	1	1	1	0	0	0	0	0	0	1	1	1	1	1	1	0
Eritrea	643	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Gambia, T	648	1	1	1	0	0	0	0	0	0	1	1	1	1	1	1	0
Georgia	915	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0
Guinea	656	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
Guinea-Bis	654	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Haiti	263	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Iraq	433	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Kiribati	826	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
Kosovo	967	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Lao P.D.R	544	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Lebanon	446	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1
Liberia	668	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
Libya	672	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
Madagasc	674	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0
Malawi	676	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0
Maldives	556	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0
Mali	678	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
Marshall Is	867	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
Micronesia	868	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
Myanmar	518	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Nepal	558	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0
Niger	692	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1
Nigeria	694	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1
Papua Nev	853	1	1	1	0	0	0	0	0	0	1	1	1	1	1	1	1
Rwanda	714	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
São Tomé	716	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
Sierra Leo	724	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
Solomon Is	813	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Somalia	726	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
South Sud	733	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
Sudan	732	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Syria	463	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
Tajikistan	923	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0
Timor-Lest	537	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Togo	742	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
Tonga	866	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Tuvalu	869	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	1
Uzbekistan	927	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
West Bank	487	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1
Western S	#N/A	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0
Yemen	474	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1
Zimbabwe	698	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Ukraine	926	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Mozambiq	688	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Burkina Fa	748	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Source: IMF. Between 2008 and 2022, the IMF had its own FCS classification. Since 2023, it has been harmonized with the World Bank's FCS classification.

Table 2. PRGT-Eligibility 2008-2023

Country	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Afghanistan	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Albania	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Angola	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Armenia	1	1	1	1	1	1	0	0	0	0	0	0	0	0
Azerbaijan, Rep. of	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Bangladesh	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Benin	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Bhutan	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Bolivia	1	1	1	1	1	1	1	1	0	0	0	0	0	0
Burkina Faso	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Burundi	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Cabo Verde	0	0	0	0	0	0	1	1	1	1	1	1	1	1
Cambodia	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Cameroon	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Cape Verde	1	1	1	1	1	1	0	0	0	0	0	0	0	0
Central African Rep.	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Chad	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Comoros	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Congo, Dem. Rep.	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Congo, Rep.	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Côte d'Ivoire	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Djibouti	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Dominica	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Eritrea	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Ethiopia	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Gambia, The	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Georgia	1	1	1	1	1	1	0	0	0	0	0	0	0	0
Ghana	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Grenada	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Guinea	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Guinea-Bissau	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Guyana	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Haiti	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Honduras	1	1	1	1	1	1	1	1	1	1	1	1	1	1
India	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Kenya	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Kiribati	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Kyrgyz Republic	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Lao P.D.R.	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Lesotho	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Liberia	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Madagascar	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Malawi	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Maldives	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Mali	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Marshall Islands	0	0	0	0	0	0	1	1	1	1	1	1	1	1
Mauritania	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Micronesia	0	0	0	0	0	0	1	1	1	1	1	1	1	1
Moldova	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Mongolia	1	1	1	1	1	1	1	1	0	0	0	0	0	0
Mozambique	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Myanmar	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Nepal	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Nicaragua	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Niger	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Nigeria	1	1	1	1	1	1	1	1	0	0	0	0	0	0
Pakistan	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Papua New Guinea	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Rwanda	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Samoa	1	1	1	1	1	1	1	1	1	1	1	1	1	1
São Tomé & Príncipe	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Senegal	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Sierra Leone	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Solomon Islands	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Somalia	1	1	1	1	1	1	1	1	0	0	0	1	1	1
South Sudan	0	0	0	0	1	1	1	1	1	1	1	1	1	1
Sri Lanka	1	0	0	0	0	0	0	0	0	0	0	0	0	0
St. Lucia	1	1	1	1	1	1	1	1	1	1	1	1	1	1
St. Vincent & Grenadines	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Sudan	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Tajikistan	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Tanzania	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Timor-Leste	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Togo	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Tonga	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Tuvalu	0	0	0	0	0	0	1	1	1	1	1	1	1	1
Uganda	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Uzbekistan	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Vanuatu	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Vietnam	1	1	1	1	1	1	1	1	0	0	0	0	0	0
Yemen	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Zambia	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Zimbabwe	0	0	0	0	0	0	0	0	1	1	1	1	1	1

Source: IMF.

PUBLICATIONS