

©Copyright 2023

Reina Kawai

# Evaluating Forecasts in Times of Financial Turmoil

Reina Kawai

A dissertation submitted in partial fulfillment of the  
requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Theo S. Eicher, Chair

Stephen J. Turnovsky

Fabio Ghironi

Hiro Ito

Program Authorized to Offer Degree:

Economics

University of Washington

**Abstract**

**Evaluating Forecasts in Times of Financial Turmoil**

Reina Kawai

Chair of the Supervisory Committee:

Theo S. Eicher

Department of Economics

This dissertation explores policy implications for the public sector, focusing on designing, assessing, and implementing macro-level economic performance improvement policies.

The first chapter, "**IMF Trade Forecasts for Crisis Countries: Bias, Inefficiency, and Their Origins**" (with Theo Eicher), investigates the accuracy of IMF external sector forecasts. IMF lends loans to countries facing crises as a lender of last resort. Loan disbursements are attached with the required policy reforms and quantitative performance criteria contingent on the IMF crisis assessment.

Findings in the paper have significant implications for improving IMF external sector forecasts in times of crisis that will bring an ultimate goal of restoring growth. We tested whether the forecast efficiently incorporated information feasible during the crisis and found that 11 covariates were not adequately integrated into IMF forecasts. Our findings confirm the informational values of IMF forecasts in times of crisis and guide how to further improve its efficiency to better design and assess the loan programs.

As a motivation for my dissertation's second chapter, Covid-19 has spurred governments worldwide to borrow more, leading to skyrocketing public debt to GDP ratio. As expected, a disproportionate share of the burden has fallen on poorer and developing nations. This has driven

the question of how an unanticipated increase in public debt will impact the economic growth rate. **"Public Debt and Real GDP: Revisiting the Impact"** (with Constance de Soyres and Mengxue Wang) examines how public debt to GDP shock will impact real GDP growth based on local projection and structural panel VAR.

Based on 178 countries' panel data from 1990-2020, we found a negative relationship between the debt and growth rate in our baseline results. Further, we address the responses based on various subsample analyses (i.e., initial debt level, debt trajectory over the five preceding years, income level). Interestingly, debt tends to be more detrimental to growth when the initial debt level is higher, raising the debt trajectory and high-income countries. We then discuss the effectiveness of the debt relief program (HIPC Initiative), the original sin hypothesis, and the debt overhang theory, where emerging countries with dollar-dominant debt will be sensitive to external shocks. Aligning with our impulse response results in income level classification, the HIPC Initiative, the debt relief program for the poorest of the poor countries, positively affected economic growth. Our findings highlight that macro-prudential policy targeting to reduce the default risk and boosting the capacity to repay is vital for economic growth over the long run.

The above second chapter's finding would extend to the following question; how does the usage of this additional government expenditure may impact economic growth differently? One may also wonder whether the government revenue originating from additional borrowings differs from the additional tax increase. Vast literature classifies government expenditure based on two sources; public investment and government consumption. We test how the usage of government expenditure may impact economic growth differently in the third chapter of the dissertation, **"Differential Growth Effects of Different Types of Government Expenditures"** (with Stephen Turnovsky).

Applying the endogenous growth model that we built based on Bruce and Turnovsky (1999) and the empirical approaches of Bayesian Model Averaging that dealt with model uncertainty as a contribution to existing literature, we found that when the government faces a trade-off in supporting current consumption or boosting infrastructure investment for the future, infrastructure investment remarkably enhances growth over the long run. Additionally, a tax increase seems to be less sensitive to economic growth compared to the case of an increase in additional borrowings.

# Table of Contents

	Page
List of Figures . . . . .	iii
List of Tables . . . . .	v
Chapter 1: IMF Trade Forecasts for Crisis Countries . . . . .	1
1.1 Introduction . . . . .	1
1.2 Data and Methodology . . . . .	6
1.3 Baseline Results: The Accuracy of IMF Trade Forecasts . . . . .	9
1.4 Forecast Accuracy by Country-Income, Crises-Types, Exchange Rate Regimes . . . . .	11
1.5 Sources of Forecast Inefficiency . . . . .	16
1.6 Did Forecast Bias and Efficiency Change Over Time? . . . . .	20
1.7 Do Forecast Horizons Affect Forecast Accuracy? . . . . .	21
1.8 Conclusion . . . . .	21
Tables . . . . .	24
Figures . . . . .	29
1.9 Data Correction . . . . .	30
1.10 Regressions for Figures 2 and 3 . . . . .	30

Chapter 2: Public Debt and Real GDP: Revisiting the Impact . . . . .	41
2.1 Introduction . . . . .	41
2.2 Data And Empirical Methodology . . . . .	45
2.3 Results . . . . .	48
2.4 Robustness Analysis . . . . .	58
2.5 Conclusion . . . . .	62
2.6 Statistics and Debt Shocks . . . . .	66
2.7 Other Results . . . . .	66
2.8 Data Cleaning . . . . .	67
Chapter 3: Usage of Government Expenditure and Growth: Empirical and Theoretical Standpoints . . . . .	74
3.1 Introduction . . . . .	74
3.2 Canonical Model of Fiscal Policy and Balanced Growth . . . . .	78
3.3 Equilibrium Changes in the Model . . . . .	82
3.4 Empirical Methodology . . . . .	83
3.5 Baseline Results of Government Borrowings on Growth . . . . .	86
3.6 Sensitivity Analysis . . . . .	95
3.7 Conclusion . . . . .	96

# List of Figures

Figure Number	Page
1.1 Prediction Realization Diagram, Full Sample . . . . .	29
1.2 Evolution of Forecast Bias and Efficiency Over Time . . . . .	30
1.3 Forecast Horizons and Forecast Accuracy, Full Sample . . . . .	30
2.1 Public Debt to GDP Ratio, 2018-21 . . . . .	43
2.2 Evolution of Public Debt Forecast Error, Median, 1995-2020 . . . . .	47
2.3 Response of Real GDP to Public Debt Shocks, Baseline . . . . .	49
2.4 Response of Real GDP to Public Debt Shocks, by Initial Debt Level Classification . . . . .	51
2.5 Response of Real GDP to Public Debt Shocks, by Public Debt Trajectory . . . . .	53
2.6 Response of Real GDP to Public Debt Shocks, by Income Classification . . . . .	56
2.7 Response of Real GDP to Public Debt Shocks, pre- vs. post-HIPC Initiative . . . . .	58
2.8 Response of Real GDP to Public Debt Shocks, pre- vs. post-HIPC Initiative . . . . .	60
2.9 Response of Real GDP to Public Debt Shocks, Baseline, Dropping Lags of Dependent Variable . . . . .	61
2.10 Response of Real GDP to Public Debt Shocks, Baseline, Using t+2 WEO Vintage for Actual Data . . . . .	63
2.11 Response of Real GDP to Public Debt Shocks, Baseline, Using t+3 WEO Vintage for Actual Data . . . . .	64

2.12	Evolution of Debt in the Dataset-Debt to GDP Ratio, All Countries,Median . . . . .	66
2.13	Public Debt Shocks-Public Debt to GDP Ratio, Deviation from Projection, Percent	67
2.14	Public External Debt Shocks-Public External Debt to GDP Ratio, Deviation from Projection, Percent . . . . .	68
2.15	Response of Real GDP to Public Debt Shocks-Middle-Income Countries . . . . .	69
2.16	Response of Real GDP to Public External Debt Shocks Baseline . . . . .	70
2.17	Evolution of Public Debt to GDP in Mongolia, Forecast and Actual Data . . . . .	71
2.18	Evolution of Public Debt to GDP in Angola, Forecast and Actual Data . . . . .	72
2.19	Evolution of Public Debt to GDP Ratio in Guinea-Bissau, Forecast and Actual Data	73
3.1	<i>IRFs</i> Characterizing the growth rate as a function of $g_p$ & $g_c$ . . . . .	83

# List of Tables

Table Number	Page
1.1 Forecast Bias and Inefficiency in IMF Forecasts . . . . .	24
1.2 Forecast Bias and Inefficiency by Subsamples . . . . .	24
1.3 Forecast Bias and Inefficiency [0.2cm] Sample Without Cancelled Programs . . . . .	27
1.4 Sources of Forecast Inefficiencies . . . . .	28
1.5 Survey of IMF Trade Forecast Accuracy Studies Sample Sizes, Data Sources, Variable Coverage . . . . .	31
1.6 Data Entered with Wrong Sign . . . . .	34
1.7 MONA Temporal Error . . . . .	34
1.8 Zeros Identify Missing Values Part I . . . . .	35
1.9 Zeros Identify Missing Values Part II . . . . .	36
1.10 Typos and Spelling Mistakes . . . . .	37
1.11 Wrong Line Item Entered . . . . .	37
1.12 Wrong Currency Unit Entered . . . . .	37
1.13 Rates vs. Levels Inconsistencies . . . . .	38
1.14 Unit Inconsistency: Base Years . . . . .	38
1.15 Missing Data Filled . . . . .	39

1.16	Corrected Outliers . . . . .	40
2.1	Regression Results, Baseline . . . . .	50
2.2	Regression Results, by Initial Debt Level Classification . . . . .	52
2.3	Regression Results, by Public Debt Trajectory . . . . .	54
2.4	Regression Results, by Income Classification . . . . .	57
2.5	Regression Results, pre- vs. post- HIPC Initiative . . . . .	59
2.6	Regression Results, Baseline, Adding Lags of Initial Debt Levels . . . . .	61
2.7	Regression Results, Baseline, Dropping Lags of Dependent Variable . . . . .	62
2.8	Regression Results, Baseline, Using $t + 2$ WEO Vintage for Actual Data . . . . .	63
2.9	Regression Results, Baseline, Using $t + 3$ WEO Vintage for Actual Data . . . . .	64
2.10	Data Cleaning, Rescaling . . . . .	71
2.11	Data Cleaning, Conversion to Missing Data . . . . .	72
2.12	Data Cleaning, Dropping Vintages . . . . .	73
3.1	OLS Bayesian Model Averaging: Joint Effect and Cuaresma (2011) Sensitivity Analysis	88
3.2	Fixed Effect Bayesian Model Averaging: Joint Effect and Cuaresma (2011) Sensitivity Analysis . . . . .	89
3.3	Fiscal Revenue to Infrastructure Investment (OLS: Growth Model on Bayesian Model Averaging) . . . . .	90
3.4	Fiscal Revenue to Infrastructure Investment (Fixed Effect: Growth Model on Bayesian Model Averaging) . . . . .	91
3.5	Fiscal Revenue to Government Consumption (OLS: Growth Model on Bayesian Model Averaging) . . . . .	93
3.6	Fiscal Revenue to Government Consumption (Fixed Effect: Growth Model on Bayesian Model Averaging) . . . . .	94
3.7	Further Sensitivity Analysis based on Cuaresma (2011) . . . . .	96

# Acknowledgments

I extend my deepest gratitude to my advisor, Professor Theo Eicher, whose unwavering support and insightful guidance have been invaluable throughout the years of my Ph.D. journey. His dedication and passion for research serve as a constant source of inspiration and learning for me. His timely and dedicated advice played a crucial role in shaping this dissertation, and his constructive feedback consistently pushed me to strive for excellence.

I am also extremely thankful to Professor Stephen Turnovsky for his mentorship in macroeconomics during our coauthored project. The knowledge and experience gained from this collaboration have been truly enriching, and I am excited for the outcomes of this and potential future projects. Additionally, I appreciate Professor Fabio Ghironi's guidance in improving my ability to present complex concepts to diverse audiences. His critical insights have been instrumental throughout my Ph.D. journey. I also thank Professor Hiro Ito for his unwavering advice and guidance, especially in refining my skills as a journal referee. His encouragement empowered me to take on this role early in my Ph.D., contributing to the development of my critical thinking abilities.

Aside from my committee advisors, I owe a debt of gratitude to my fellow macro graduate students. Particularly, I am thankful to my dear friends Jay, Kovid, Raj, and Ryan for their genuine kindness and encouragement, which have been a constant source of support throughout the ups and downs of the Ph.D. program.

My heartfelt appreciation also goes to all of my family, including my beloved husband, for believing in me and providing unwavering support throughout my Ph.D. journey. Their encouragements has been instrumental in helping me stay focused during challenging times.

# Dedication

My husband, Sefik Emre Eskimez

## Chapter 1

# IMF Trade Forecasts for Crisis Countries

### 1.1 Introduction

A core mission of IMF loan programs is *“to help countries restore macroeconomic stability by rebuilding their international reserves, stabilizing their currencies, and paying for imports - all necessary conditions for relaunching growth”* (IMF, 2020). IMF loans are accompanied by IMF conditionality related to policy reforms and economic performance targets. These targets are jointly determined by IMF’s crisis assessments and IMF recovery forecasts.<sup>1</sup> Reviews of IMF crisis loan performances suggest that a key indicator of the quality of program design is the absence of systematic bias in program forecasts (Baqir et al. (2005); and Mody and Rebucci (2006)). Since most IMF loan programs initiate during balance of payments crises, the viability of economic recoveries depends crucially on the quality of IMF external sector forecasts. Even if IMF programs do not originate with balance of payment crises, accurate external sector forecasts are crucial to predict countries’ financing needs to repay IMF loans. Hence our study focuses on the IMF’s external sector forecasts in times of crisis.

Only a single previous study of IMF trade forecasts for crisis countries includes import and

---

<sup>1</sup>Musso and Phillips (2002) note the term “forecast” is sedulously avoided by IMF documents. We label the “future values” included in IMF loan documents “forecasts,” since loan agreements represent formal contracts between the IMF and countries to implement policies and targets in exchange for funds.

export data. Most evaluations of IMF forecasts focused on the current account as a percent of GDP, conflating forecast errors that originate with imports, exports, and GDP.<sup>2</sup> Previous evaluations of IMF trade forecasts in crisis countries also employ surprisingly small and always different subsets of all available data, which may explain the surprising diversity of results. Forecast accuracy assessments range from exactly zero average forecast errors (Baqir et al. (2005)) to forecasts that cannot outperform random walks that predict future growth with past growth – a harrowingly low bar in times of crisis (see Artis (1988); Arora and Smyth (1990); and Musso and Phillips (2002)).<sup>3</sup> No previous forecast evaluation examines the accuracy of exchange rate forecasts, although exchange rate trajectories are arguably fundamental to all recovery trajectories in IMF programs, especially imports and exports. To address the lack of large sample forecast evaluations for imports, exports and exchange rates, we analyze nearly 30 years of import, export, and exchange rate forecasts in a dataset of over 600 IMF loan programs. Our dataset is over three times larger and more than a decade longer than the most comprehensive study to date. The size of our dataset allows for an unprecedented analysis of drivers of forecast bias and inefficiency by subsamples (e.g., country-income status (LICs/Non-LICs), crisis types (Hyperinflation/Non-hyperinflation/BOP), exchange rate regimes (fixed/flexible), or openness (financial/trade openness)).

The most recent paper to assess the accuracy of IMF trade forecasts in crisis countries also includes one of the largest datasets to date. It is also the only previous study that examines import and export forecasts (Eicher et al. (2019), EKPC from here on out). However, EKPC use estimated, not actual final data, which contaminates their assessments of forecast accuracy. EKPC are also transparent about their exclusion of a large number of observations for fear of errors in the IMF MONA loan database (IMF (2021a)). We audited the MONA database using original loan documents stored in the IMF archives, to verify extreme values and correct database errors (see Appendix C), which allowed us to add 10 additional years and 476 additional country/program observations to the EKPC dataset. We also use actual, not estimated final data.

The corrected/audited data, along with the additional data for additional years overturns the central findings of EKPC. First, EKPC suggested trade forecast errors were influenced by inefficient import growth forecasts. We find the opposite. Our expanded dataset shows IMF

---

<sup>2</sup>See Eicher et al. (2019) and Genberg and Martinez (2014) for surveys.

<sup>3</sup>For some financial data, especially if serially correlated, the inability to outperform the random walk is not an indictment. But for countries that careen into crisis, it is highly problematic when forecasts cannot beat the naïve random walk projection where past growth equals future growth.

import growth forecasts for crisis countries are remarkably accurate, without bias or inefficiency. We document that this result is largely independent of country-income status (LICs/Non-LICs), crisis types (Hyperinflation/Non-hyperinflation/BOP), or openness (financial/trade openness). The result does depend, however, on exchange rate regimes – IMF import forecasts for fixed exchange rate LICs feature substantial forecast inefficiency.

The second important finding of EKPC was that export growth forecasts were only inefficient but not biased. Our expanded dataset shows this EKPC result to be an artifact of the smaller dataset and with estimated rather than actual final data. For the corrected, full dataset, IMF export growth forecasts have, in fact, been systematically biased and inefficient across most subsamples: for Non-LICs, (Non)Hyperinflation countries, BOP programs, fixed/floating exchange rates, and for all degrees of trade and financial openness. Only LICs with floating exchange rate are unbiased and efficient. The stark difference in import vs. export forecast accuracies is noteworthy and surprising.<sup>4</sup> Being the first to examine exchange rate forecast accuracy, we can confirm that the divergence in forecast accuracies is not driven by inaccuracies in IMF exchange rate forecasts. Exchange rate forecasts for the full sample are unbiased and efficient, although LICs forecasts are spectacularly inaccurate. For LICs with floating exchange rates, IMF crisis forecasts cannot beat the naive random walk as they overestimate crisis depreciations by more than 100 percent on average.<sup>5</sup> Overall import, export, and exchange rate forecasts are shown to be systematically optimistic, on average, projecting larger than realized growth rates. As we disentangle average forecast errors, we can also show that IMF forecasts underestimate growth rates for slow recovering countries and overestimate growth rates for high-growth recoveries. This is an important insight since it suggests that the most vulnerable, slow recovering countries receive systematically pessimistic, excessively cautious external sector forecasts.

To identify the sources of IMF forecast inefficiencies, we examine whether IMF forecasters

---

<sup>4</sup>Boz et al. (2020) suggest a potential explanation for the differential bias and efficiency observed in IMF import and export forecasts: countries invoicing predominantly in foreign currency (dollars) may experience greater exchange rate pass-through to import prices and higher sensitivity of trade volumes to exchange rate fluctuations. Unfortunately, we cannot examine this hypothesis since the unique Boz et al. (2020) dataset covers only 40% of our full sample and 15% percent of our LICs sample.

<sup>5</sup>Meese and Rogoff (1983) and a voluminous subsequent literature point out that exchange rates are notoriously difficult to predict; out-of-sample forecasts seldom outperform random walks (Rossi (2013), surveys the literature). Itskhoki (2020) highlights that real exchange rate forecasts are difficult in part because they are virtually uncorrelated with most other macroeconomic variables, nominal or real, with the exception of the nominal exchange rate. In our findings it is noteworthy that only LIC real exchange rate forecasts struggle to outperform the random walk, but forecasts for the full sample are unbiased, efficient and outperform the random walk.

properly integrate all information known at the time of forecasts. We group this information into three areas. First, we confirm the findings of the previous literature that loan size is indeed a factor that affects IMF forecast accuracy for crisis countries (see [Beach and Schavey \(1999\)](#); [Dreher et al. \(2008\)](#); [Luna \(2014\)](#)). Second, we document the effects of conditionality on forecast inefficiency (as previously suggested by [Dreher et al. \(2008\)](#) for GDP forecasts). Specifically, debt conditionality is shown to be central to explaining forecast inefficiencies for exports, while a better account of arrears conditionality (domestic and external) could improve imports and exchange rate forecasts. Fiscal conditionality is not integrated properly into import and export forecasts, while inaccurate estimates of the effects of trade reforms played an important role in the unusual bias observed in LICs exchange rate forecasts. Third, we investigate whether geopolitical events are adequately integrated into forecasts, since [Aldenhoff \(2007\)](#) and [Dreher et al. \(2008\)](#) suggest geopolitical events drive errors in IMF GDP forecasts. We find that effects of elections and conflicts (internal/external known at the time of forecasts), were not properly accounted for in IMF crisis forecasts.

It is important to understand our results in the context of the previous IMF forecast evaluation literature that covers crisis countries. [Genberg and Martinez \(2014\)](#) survey 75 studies that evaluate IMF forecast accuracy since 1983, that focus only on GDP growth, inflation, or fiscal balances. IMF external sector forecasts are addressed in only 19 of these 75 studies, and only 7 of these include an analysis of import and export forecasts (see Table 5). Of these 7, only EKPC examine import and export data for crisis countries, while all other studies use IMF World Economic Outlook data ([IMF \(2021b\)](#)). WEO data does, however, not report forecasts for individual crisis countries, and before 2004 WEO data is instead aggregated into a single forecast for all developing countries per geographic region (Europe, Asia, Africa, Middle East, and Western Hemisphere).

The previous literature evaluating trade forecasts for IMF crisis countries is hampered by small and ever different datasets that produce an unusual array of distinctly different results. [Musso and Phillips \(2002\)](#) is the first study to evaluate IMF forecasts for crisis countries. They examine current account (percent of GDP) forecasts for 69 IMF loans over 5 years (1993-1997) to find IMF forecasts hold “very little predictive power beyond that of the random walk alternative.” Following the Asian Financial Crisis, the U.S. House of Representatives ([USGAO \(2003\)](#)) commissioned a report on the accuracy of IMF forecasts from 1990-2001. The report covered 87 countries, 57 of which were crisis countries that received IMF loans. Using Theil’s U statistic, the report finds that forecasts for non-crisis countries outperform forecasts for crisis countries, but neither sample can

beat the naïve random walk.

More recent studies of current account forecasts (as a percent of GDP) use 175 crisis countries for 8 years from 1993 to 2001 (Atoian et al. (2004)), or 183 crises countries for 17 years from 1993 to 2009 (Atoyán and Conway (2011)). Atoian et al. (2004) find forecasts biased and inefficient and Atoyán and Conway (2011) decompose the forecast error to identify an association of the forecast error with a single, aggregate conditionality indicator. We disaggregate conditionality into 18 different types of policy reforms and quantitative performance targets to highlight specific conditionalities associated with inefficient forecasts. No previous study correlated forecast inaccuracy with such detailed conditionality.

An unusual finding is reported by Baqir et al. (2005), who examined 94 IMF crisis loans over 13 years (1989-2002) to report a surprising zero average forecast error for current accounts (as a percent of GDP). Luna (2014) examines 103 program countries over 9 years (2002-2011) to find that current account forecasts (as a percent of GDP) consistently erred on the pessimistic side, but without statistically significant bias. The impressive range of forecast (in)accuracies throughout the previous literature thus ranges from zero information value to zero average errors, at times with and without bias and inefficiencies. Notice that none of these studies (other than EKPC) examine imports and exports, meaning that import, export, GDP, and inflation forecast errors were always conflated.

This diversity of results in the previous studies may not only be driven by the small and always different samples. Artis (1996) suggests “data for many of these [developing] countries are poor and tardy,” while Atoyán and Conway (2011) surmise the lack of forecast accuracy may have its origins not only in (i) poor data quality but also in (ii) “country-specific differences” in the forecast models, (iii) “incomplete policy implementations,” and (iv) “random errors.” Random errors should balance out in bias calculations, however. We rule out data quality issues by painstakingly auditing the MONA database using the original loan documents in the IMF Archives. We address the effects of incomplete policy implementation by examining whether canceled programs drive bias and inefficiency but find no evidence for that hypothesis. Timmermann (2007) and EKPC suggest IMF forecast accuracy is affected by outliers; we show, however, in Section II.2 that large observed forecast variances mirror large variations in the actual final data. While extreme values may exhibit high leverage, we examined results without extreme values and found results did not deviate

qualitatively from those reported below.<sup>6</sup>

Section II presents the data and methodology. Section III introduces baseline and robustness results. Section IV investigates if forecast accuracy varies by country-incomes, crisis-types, or exchange rate regimes. Section V explores sources of forecast inefficiencies, while Section VI documents the evolution of IMF forecast bias and inefficiency over time. Section VII investigates if the forecast horizon affects forecast accuracy, and Section VIII concludes.

## 1.2 Data and Methodology

### 1.2.1 Data

IMF forecasts were obtained from the IMF’s Monitoring of Fund Arrangements database [IMF \(2021a\)](#), which reports data from loan documents that are presented to the IMF’s executive board at the time of program approval.<sup>7</sup> We examine only current-year forecasts (in year  $t$  for year  $t$ ) that reflect the most recent information used in the loan program design. Missing observations rule out a meaningful analysis of longer forecast horizons. Our focus is on the growth of imports and exports of goods and services in US dollar and on real exchange rate depreciations vis-à-vis the dollar. As we prepared the paper, the MONA database reported data for 602 crisis countries programs in 123 countries over 29 years, from 1992 to 2020.<sup>8</sup>

Imports and exports in US dollar, together with real exchange rate depreciations avoid conflating forecast errors of these variables with local currency and inflation forecast errors. This is also why we do not examine the current account balance as a share of GDP, which conflates forecast errors for imports, exports, GDP, or inflation. [Kunze \(2020\)](#) suggests exchange rate forecast accuracy differs by exchange rate regimes. Hence, we also examine the quality of IMF forecasts in fixed and flexible exchange rate regimes. To identify exchange rate regimes as “fixed” or “flexible,”

---

<sup>6</sup>Results without extreme values are available upon request.

<sup>7</sup>The IMF’s WEO database provides forecasts in April and October of each year, but does not provide forecasts for countries in crisis, or forecasts set at the time program loans commence. WEO forecasts for individual developing countries are not available prior to 2004. We audited the MONA database; a list of errors/corrections is provided in Appendix C.

<sup>8</sup>26 programs provided only technical assistance through non-financing facilities (*Policy Support Instrument (PSI)*, *Policy Coordination Instrument (PCI)* programs, and *Flexible Credit Line (FCL)* programs without financial assistance); we include these programs for completeness; results do not change if these programs are excluded.

we follow [Ilzetzi et al. \(2019\)](#) and the exchange rate regime classifications of the IMF’s *Annual Report on Exchange Arrangements and Exchange Restrictions* ([IMF \(2021e\)](#)). When announced exchange rate bands cannot exceed  $\pm 2$  percent, regimes are labeled “fixed.”<sup>9</sup>

Actual final data was obtained from three IMF databases. Final inflation data was obtained from the IMF’s World Economic Outlook database ([IMF \(2021b\)](#)), final imports and exports from the IMF’s International Financial Statistics database ([IMF \(2021c\)](#)), and final exchange rate data from the IMF’s Balance of Payments and International Investment Position database ([IMF \(2021d\)](#)). Between missing observations in WEO, BOP, IFS, and MONA, we managed to gather forecasts and final outcome data for 576 programs. Our dataset is thus over three times larger and at least 10 years longer than the largest previous study of IMF external sector forecasts for program countries. Appendix Table A.1. provides a detailed data description.

## 1.2.2 Methodology to Evaluate Forecast Accuracy

One approach to measuring forecast accuracy is to produce forecast error statistics such as the mean absolute error (MAE) to compare different forecasts. Such statistics are useful only in the presence of two or more forecasts, but our case forecast evaluation involves only a single forecast, the IMF’s. This is not by choice; instead, the IMF is the sole institution with access to country-level data in times of crisis. This highlights the crucial importance of IMF forecasts: they are required for the lender of last resort’s loan documents, and the IMF is the only entity with access to country data in times of crisis.

One forecast statistic that can be employed in the presence of only a single forecast is Theil’s  $U_2$ , which is a relative accuracy measure. It compares IMF forecasts to a naive random walk forecast based on “no-change extrapolations,” where future values are predicted as past values.  $U_2$  is the root sum-of-squared forecast errors divided by the naive forecasting error where the forecast in  $t+1$  equals the actual in period  $t$ .  $U_2 = 0$  indicates perfect forecasts, while  $U_2 > 1$  indicates forecasts do not outperform the naïve random walk and have no informational value. In times of crisis, when

---

<sup>9</sup>We chose [Ilzetzi et al. \(2019\)](#) “coarse exchange rate regime classification” for fixed exchange rates that included i) no separate legal tender, ii) preannounced pegs, iii) currency boards, iv) de factor pegs, and v) preannounced horizontal bands  $\leq 2\%$ . If we use Ilzetzi’s “fine classification” for fixed exchange rate regimes, we can include additional countries with some form of exchange rate rigidities (including bands up to 5%, and “moving bands” which allow for appreciation and depreciations). Results (available upon request) are qualitatively similar.

programs are designed to reverse the economic trajectory of a country, the random walk is a low bar.

An assessment of forecast accuracy requires formal statistical tests. These tests are based on [Mincer and Zarnowitz \(1969\)](#), who extended the seminal work of [Theil \(1961\)](#). Theil’s “*Prediction-Realization Diagram*” (Figure 1) displays IMF forecasts for the current year,  $F_t$ , on the horizontal axis and official, actual final data for the current year,  $A_t$ , on the vertical axis. [Mincer and Zarnowitz \(1969\)](#) label the solid 45-degree line the “*Line of Perfect Forecasts*” as it represents coordinates where program forecasts equal actual final data.

Figure 1 visualizes that the extraordinary variation in forecasts is mirrored by a similarly unusual variance in the outcome data; perhaps this is not surprising for crisis countries that require IMF programs. [Timmermann \(2007\)](#) noted that outliers influence the evaluation of non-G7 countries’ forecasts, and EKPC eliminated 30 percent of their sample worrying about data errors or unusual extreme values. Figure 1 does highlight the existence of extreme values for import, export, and exchange rate forecasts. Such extreme values may or may not exert high leverage; we find they are not sufficiently influential to qualitatively change our results below. We see no reason to exclude them from the analysis.

[Mincer and Zarnowitz \(1969\)](#) suggest a formal test for unbiased and efficient forecasts, which was first employed by [Kenen and Schwartz \(1986\)](#) to evaluate IMF forecast accuracy. The technique has since been applied frequently in IMF forecast evaluations (albeit in much smaller samples), see [Artis \(1996\)](#), [Musso and Phillips \(2002\)](#), [Timmermann \(2007\)](#), [Genberg and Martinez \(2014\)](#), EKPC, and [Eicher and Rollinson \(2022\)](#):

$$A_t = \alpha + \beta F_t + \epsilon_t \tag{1.1}$$

Forecasts are chosen as the “independent” variable in (1) only because they are available before the actual final data is published. Forecasts are efficient when the forecast error is random and uncorrelated with forecasts. [Nordhaus \(1987\)](#) notes the similarity between forecast efficiency and stock market efficiency – both imply that all relevant and available information was considered, and all errors are white noise. In this case, the slope parameter,  $\beta$ , is unity and the intercept,  $\alpha$ , is zero. Since estimates of  $\alpha$  and  $\beta$  are generally correlated, individual T-Statistics are inappropriate, and the joint null hypothesis,  $\alpha = 0$  and  $\beta = 1$  is tested.

If the Mincer-Zarnowitz null ( $\alpha = 0$  &  $\beta = 1$ ) is rejected, forecasts are inefficient, but they may not be biased in the sense that  $E(A_t) \neq E(F_t)$ . Holden et al. (1987) demonstrate that  $\alpha = 0$  is sufficient but not necessary for unbiased forecasts. Holden et al. (1990) derived a necessary and sufficient condition for unbiased forecasts, which tests simply whether the regression line intersects the *Line of Perfect Forecasts* at  $E(A_t) = E(F_t)$ . When the Holden Peel test of  $A_t - F_t = \gamma + v_t$  rejects the null of  $\gamma = 0$ , forecasts are said to be biased. Here it is crucial to note that this metric of bias declares forecasts as “unbiased” when, for example, half are 40 percent higher, and half are 40 percent lower than the official final data. Bias is thus substantially less informative than efficiency.

### 1.3 Baseline Results: The Accuracy of IMF Trade Forecasts

Regressions (1a) - (3) in Table 1 present the results of Mincer-Zarnowitz regressions (eq. 1) associated with Figures 1a)-c). The regressions allow us to test for potential bias and inefficiency for IMF real exchange rate depreciations and import/export forecasts. To compare our results to the only previous analysis of IMF trade forecasts, regressions (1a) and (2a) restate EKPC results. Their sample uses 110 of the 602 available programs in nominal, local currency (EKPC did not examine exchange rate forecasts), which conflates price and currency forecasts. EKPC also used estimated final data and omitted programs when data exceeded four standard deviations from the mean or when program durations did not exceed 18 months. This led EKPC to reject efficiency for both imports and exports. Results change as indicated in (1b) and (2b), once omitted and excluded programs along with the entire available time series from 1992-2020 are included in the dataset. Import forecasts are shown to be efficient and unbiased, and export forecasts are not only inefficient but also biased.

Theil coefficients in Table 1 are substantially smaller for EKPC than those observed in our sample. This is because EKPC’s abundance of caution led them to exclude a sizable number of programs with data that represented large deviations from the mean. Many of these programs (46 for the EKPC time period, 2002-2015) were verified in our MONA audit and included in our dataset to produce substantially higher Theil coefficients. All Theil coefficients easily beat the random walk in the full sample.

Regression (3) indicates that the dichotomy in imports and exports forecast accuracy is

not necessarily driven by inaccurate exchange rate forecasts. Neither the Holden Peel test, nor the Mincer Zarnowitz test for unbiased and efficient forecasts can be rejected in the full sample (EKPC did not examine exchange rate forecasts). However, the beta coefficient for exchange rate forecasts is significantly below unity, indicating that, on average, IMF forecasts systematically overestimate real exchange rate depreciations for program countries. When the regression line intersects the line of perfect forecast from above, as in the case of exchange rate forecasts, it implies that IMF forecasts are systematically different for programs with small vs. large depreciations. For countries with small forecasted/realized exchange rate changes, the IMF underestimates the resulting depreciation, while programs with large depreciations receive excessively large forecasts in times of crisis. Given the slope coefficient, we know the excessive depreciation forecasts dominate our sample. The low slope coefficient along with Figure 1(c) also suggests that the forecast errors are substantial and the confidence interval seldom includes the line of perfect forecast, so that errors simply equal out to highlight the weakness of the Holden Peel test that we discussed above.

The excessively large depreciation forecasts (on average) may explain the IMF’s tendency to overestimate export growth (on average) in regression 2b, where the beta coefficient is also smaller than unity. On the other hand, imports are accurately forecasted, which thus presents a puzzle. Since the intercept for export growth in regression 2b is positive and significant, we know that exports for slow (fast) growth recoveries are under (over) estimated by IMF program forecasts.

Bias and inefficient export forecasts have important implications for IMF program countries. Forecasts that overestimated future exports may lead to underestimates of the required external capital inflows to jeopardize the recovery. Overly optimistic export forecasts also translate into optimistic program targets and performance criteria (e.g., trade balance, government revenues, unemployment rate, exchange rate), which thus become more challenging to meet. When forecasts are overly optimistic, the country review will produce seemingly below-par country program performance and suggest a lack of conditionality implementation, when in fact, the differential may only be due to inaccurate export and exchange rate forecasts.<sup>10</sup>

---

<sup>10</sup>A number of authors (Beach and Schavey (1999); Aldenhoff (2007); Luna (2014)) suggest IMF staff may face career incentives to forecast optimistic outcomes, while Musso and Phillips (2002) counter that the IMF Executive Board review process incentivizes staff to forecast initially pessimistic outcomes, to be able to characterize future results as “unexpectedly better” instead of “unexpectedly weak.” Beaudry and Willems (2022) show that the quality of IMF forecasts may be contingent on the optimism/pessimism of individual IMF Mission Chiefs. Independent of career incentives or personal proclivities, Genberg and Martinez (2014) document that IMF desk economists with longer work experiences produce more accurate forecasts.

## 1.4 Forecast Accuracy by Country-Income, Crises-Types, Exchange Rate Regimes

In a formal review of IMF program lending, [Ghosh et al. \(2006\)](#) suggest forecast accuracy may differ by programs-types or country-types since trade flows can recover at differential speeds depending on such country characteristics. Of course, the hope would be that forecasters appropriately incorporate country and crisis characteristics into their forecasts so that subsamples do not exhibit systematically different forecast performance. In this case, the external sector forecasts would not exhibit systematic differences by country or program-type. The unique size of our dataset allows, for the first time, an evaluation of country and program-specific subsample effects with sufficient statistical power of inference. We thus examine the forecast accuracy by (i) crisis-types (inflation/BOP<sup>11</sup>), (ii) country-types (LICs/NonLICs)<sup>12</sup>, (iii) exchange rate regimes (fixed/floating), and by (iv) openness (Financial/Trade).

### 1.4.1 Import Forecast Accuracy by Subsample

We examine import forecasts by subsample in Table 2a. Unbiased and efficient import forecasts are not universal across country-types. Specifically, fixed exchange rate regimes show a systematic pattern of inefficient import forecasts. This inefficiency is driven entirely by LICs. Once fixed exchange rate LICs are removed from the sample, IMF import forecasts are unambiguously unbiased and efficient. Our sample features 98 fixed exchange rate LICs that produce remarkably inefficient import forecasts. The vast majority of these loans (75) represent 14 African countries' programs that are members of the African Franc zones.<sup>13</sup> The low slope coefficient of 0.67 for fixed-exchange rate LICs implies forecasts overestimate import growth to the point where actual imports

---

<sup>11</sup>IMF BOP crisis loan programs include Extended Credit Facility (ECF), Extended Fund Facility (EFF), Exogenous Shock Facility (ESF), Flexible Credit Line (FCL), Stand-By Agreements (SBA), Standby Credit Facility (SCF), Precautionary Credit Line (PCL), Precautionary Liquidity Line (PLL). The remaining programs in our dataset focus on structural reforms following crises: Structural Adjustment Facility (SAF), Policy Reform Instrument (PSI), Poverty Reduction and Growth Trust (PRGT), Policy Coordination Instrument (PCI), Enhanced Structural Adjustment Facility (ESAF), see [IMF \(2021a\)](#).

<sup>12</sup>Strictly speaking, when we refer to “the LICs sample” below, we are referring to ‘Non-Hyperinflation-LICs.’ Bias and inefficiency results do not change if we use the full LICs sample instead of Non-Hyperinflation-LICs, we believe, however, the latter sample is more informative. Results for ‘all LICs’ are available upon request.

<sup>13</sup>Benin, Burkina Faso, Guinea-Bissau, Côte D’Ivoire, Mali, Niger, Senegal, and Togo (West African CFA franc zone), and Cameroon, Central African Republic, Chad, Republic of the Congo, Equatorial Guinea, and Gabon (Central African CFA franc zone).

are on average 50 percent smaller ( $1/0.67-1$ ) than predicted. Nevertheless, the forecast errors average out to indicate no systematic bias, highlighting that inefficiency is a more informative concept to assess forecast accuracy as we mentioned above.

The Theil U2 statistic confirms that import growth forecasts contain some informational value relative to naïve forecasts for all subsamples. We also include directional analysis introduced by [Theil \(1961\)](#) in the last row of Table 2. Directional analysis reports the percent of forecasts that do not match the direction of actual outcomes ( $F_t > 0$  &  $A_t < 0$ ,  $F_t < 0$  &  $A_t > 0$ ). [Henriksson and Merton \(1981\)](#) derived a test that pits this percentage against a coin flip to establish the null hypothesis that forecast-directions and outcome-directions are independent events.<sup>14</sup> Of course, this test sets a low bar, perhaps even lower than the threshold value of 1 in Theil’s U2, but directional accuracy is nevertheless insightful for a number of reasons. First, instead of squaring or averaging positive and negative errors, the directional error analysis provides insight into the forecasting fundamentals: to what extent do forecasters assume a continuation of a trend or not. Second, when the IMF provides loans to countries in crisis in exchange for difficult conditionality, it is crucial to understand if IMF forecasts of future recoveries coincide at least with the observed directional changes for imports, exports, and exchange rates. Third, our particular interest in forecasts in times of crisis stems from the fact that turning points in crisis countries are to be expected after they accept IMF loans and conditions. Hence one would expect IMF forecasters to do especially well against the random directional coin flips. Finally, we also examine the root mean squared error (RMSE), which turns out to correlate closely with Henriksson-Merton’s directional inaccuracy, as higher directional errors correlated with higher RMSE (0.88 correlation).

For imports, between 20 percent and 29 percent of forecasts from 1992-2019 were not only off the mark in terms of their quantitative levels but in terms of the direction in which imports changed during the recovery. Here it is of note that among the 4 subsamples (full, BOP-programs, floats, fixed), LICs have a substantially higher rate of directional errors than other subsamples, reaching at times 29 percent for LICs with balance of payments crises. Once LICs are purged from the sample, the directional error rate drops to about 20 percent for Non-LICs.

---

<sup>14</sup>Strictly speaking, the Merton-Henriksson test is based on [Merton \(1981\)](#) and [Henriksson and Merton \(1981\)](#), who examined whether market-timing forecasts of asset returns add informational value. [Schnader and Stekler \(1990\)](#) adapted Merton-Henriksson 2x2 contingency tables to assess directional accuracy and developed test statistics for the 2x2 case, see [Diebold and Lopez \(1996\)](#).

### 1.4.2 Export Forecast Accuracy by Subsample

Table 2b highlights that export and import forecast accuracies differ profoundly, especially at the subsample level. In contrast to import forecasts, all export forecasts other than LICs' are biased and inefficient, independent of exchange rate regimes or crisis-types. Among LICs, fixed exchange rate export forecasts are again inefficient, and the low slope coefficient of 0.75 implies that IMF forecasts substantially exceed actual outcomes by 33 percent on average ( $1/0.75-1$ ) as a one-unit increase in forecasts is associated with an increase in actual outcomes of 0.75 units only. The regression also suggests the mean actual export growth is overpredicted by 0.05 percent when forecasts are zero. Together with the statistically significant positive constant, this implies that IMF forecasts under (over) estimate exports for low (high) growth recoveries. The full sample is biased and inefficient throughout, independent of crisis type or exchange rate regimes.

The Theil index indicates that export forecasts do hold informational value over naïve forecasts, although the directional analysis indicates a substantial share of forecasts feature the wrong direction. For the full sample, about 23 percent of export forecasts are not just quantitatively but also directionally inaccurate. This figure increases to over 30 percent for LICs (LICs-Full, BOP-Program-LICs, and Float-LICs). Exports forecast RMSEs are even more closely related to export directional inaccuracy as for the case of import forecasts, as the correlation between directionally inaccuracy and RMSEs rises to 0.90 for IMF export forecasts.

### 1.4.3 Exchange Rate Forecast Accuracy by Subsample

Forecasts of real exchange rate depreciations in Table 2c exhibit a roughly similar pattern as imports, with some important exceptions. Similarly, once LICs are eliminated from the sample, exchange rate and import forecasts are unbiased and efficient for the remaining countries. For LICs, the inefficiency of exchange rate forecasts is driven by countries with floating exchange rates and countries with balance of payment crises. Of note is the staggeringly low slope estimate of 0.49 for LICs' floating exchange rate forecasts. It indicates that, on average, IMF forecasts predict 104 percent greater depreciation than realized ( $1/0.49-1$ ). The average depreciation forecasts can be parsed further, as the low slope coefficient also indicates that low (high) depreciation events are systematically under (over) forecasted by the IMF.

These results imply that the IMF is much more optimistic than justified regarding the stability of exchange rates for low depreciation countries while expecting much greater than realized depreciations for countries with large changes in exchange rates. These are concerning insights since IMF exchange rate forecasts provide crucial market guidance which has the potential to exacerbate crises if depreciations are much lower than forecast. At the same time, these results suggest the IMF imparts excessive confidence with regard to country performance for the most vulnerable crisis countries: LICs with larger than average depreciations. Such systematically optimistic real exchange rate depreciation forecasts have direct implications for the forecasted current account, reserve, and financing needs, rendering these program countries less likely to hit their actual quantitative performance targets.

LICs forecasts for floating exchange rate regimes are so inaccurate that the Theil coefficient indicates IMF forecasts hold zero information value as they cannot outperform the naïve, no-change random walk extrapolation. That sample also produces an unusually high RMSE. As noted in footnote 5, real exchange rate forecasts are notoriously difficult and a voluminous literature outlines the associated forecasting challenges. Hence perhaps the most noteworthy insight derived here is that only LICs forecasts feature substantial large forecasting challenges, while forecasts for the full sample are indeed unbiased, efficient, and outperform the random walk.

The directional analysis is a lower bar than the Theil coefficient, since it examines only if the sign of the forecast is accurate. The Henriksson-Merton test for directional accuracy for the floating LICs subsample cannot reject the null that the direction of forecasts and outcomes are independent, with directional inaccuracies exceeding 30 percent for just about all LICs subsamples. Once the sample is purged of LICs, directional accuracy improves dramatically, and Non-LICs exhibit only between 15 percent-17 percent of directional inaccuracies. Non-LICs, we thus find by far the lowest level of forecast inaccuracy among either of the three forecasts we examined (imports, exports, and exchange rates). Directional inaccuracies are less tightly correlated with the RMSE (correlation 0.38) than in the case of imports and exports, suggesting that directional errors are less of a driver of the forecast errors for real exchange rate depreciations.

#### 1.4.4 Disaggregation of Exchange Rate Forecast Accuracy

A remaining question is whether the source of real exchange rate forecast errors derive from forecast errors for nominal exchange rates or inflation. In Table 2d we show the contribution of inaccuracies in nominal exchange rate and inflation forecasts to real exchange rate forecasts. We find inaccuracies in the full sample are roughly equally explained by nominal exchange rates and inflation forecast errors. We cannot reject the null hypothesis that the two estimates ( $\beta_1 = \beta_2$ ) are equal for all but one subsample (floating exchange rate regimes). The inability to reject ( $\beta_1 = \beta_2$ ) is driven, however, entirely by hyperinflation countries. Once they are removed, nominal exchange rate and inflation contribute equally. There is thus a clear pattern that observed statistically significant deviations from perfect forecasts in Table 2c are driven equally by nominal exchange rate and inflation forecast errors.

Our findings regarding import, export and exchange rate accuracy reveal three important insights: First, import and exchange rate forecasts are unbiased and efficient once we eliminate LICs from the sample, while export forecasts are deeply flawed throughout all subsamples. Second, for import forecasts, fixed exchange rate LICs introduce noise, essentially produced by forecasts for African franc zones countries. Third, for exchange rate forecasts, flexible exchanges rate LICs introduce substantial noise to the point that IMF forecasts cannot beat the naïve random walk, and the forecasts and outcomes are so directionally challenged that we cannot reject that the two are independent. All three insights combined suggest that greater efficiency of LICs program forecasts should be a key policy goal for future IMF loan programs.

#### 1.4.5 Openness and Forecast Accuracy

The accuracy of external account forecasts may well be associated with trade and/or financial openness, if only because more open economies face greater exposure to contagion and idiosyncratic global shocks, whose advent and effects are difficult to forecast. In a panel of 29 OECD/BRIICS countries, [Lewis and Pain \(2015\)](#) document that OECD GDP growth forecast accuracy decreases with trade openness and with banking assets held by foreign banks. The authors surmise openness allows external imbalances and financial leverage to accumulate faster, to increase uncertainty and forecast inaccuracy. [Chatterjee \(2019\)](#) and [Chatterjee and Nowak \(2016\)](#) document

that financial and trade openness impact forecasts since openness increases the sensitivity of real variables (including trade) to uncertainty shocks.

In this section we examine whether forecast accuracy in IMF crisis countries differs systematically by GDP trade shares and financial openness (Chinn and Ito (2006)). We use the global sample mean for each measure to parse our sample into “more open” and “less open” trade/financial subsamples. Tables 2e)-g) report results suggesting forecast accuracy is remarkably stable across subsamples. Generally, both more and less open subsamples for both trade and capital flows show little difference in terms of bias and inefficiency when compared to the full sample for imports, exports and the exchange rate depreciations. This result holds even when we compare subsamples by income levels. For example, if NonLICs forecasts cannot reject the null that imports are unbiased and efficient, the same is true for NonLICs with more/less open trade/financial openness.

While the tables convey a strong overall sense that IMF forecast accuracy does not differ systematically for more/less trade/financial openness, there are two exceptions. For real exchange rate depreciations, the full sample is never biased for any country-income subsample, but countries with small trade shares are inefficient in all but the Non-LICs sample. Once hyperinflation countries are removed, financial openness is improving the accuracy of real exchange rate depreciation forecasts. For exports, all full samples for all income levels are biased and inefficient except for LICs, but countries with high degrees of financial openness buck the trend as they are unbiased in all subsamples, and efficient in all but the Non-LICs country-income subsample. Financial openness is again clearly improving the accuracy of IMF export forecasts.

## 1.5 Sources of Forecast Inefficiency

After establishing the extent to which specific subsamples suffer from systematic bias and inconsistency, we turn our attention to identifying possible sources of IMF forecast inaccuracies. We examine if all tangible information available at the time of the forecast was adequately integrated into the forecast. The unprecedented size of our dataset allows us to introduce a substantial number of covariates that represent possible sources of forecast errors for the samples with highly inaccurate forecasts.

### 1.5.1 Program Cancellations

Mussa and Savastano (1999) and Stiglitz (2011) note that MONA forecasts may not reflect actual forecasts but negotiated compromises between the IMF and country authorities. If this were the source of forecast errors, it would be difficult to explain why authorities systematically argue for excessively optimistic export growth or excessively large depreciations; both set up countries for debt, reserve, and overall program shortfalls.

Atoyan and Conway (2011), Luma (2014), and IMF (2019) point out that IMF forecasts are conditional on the assumption that conditionality is implemented; hence implementation failures may well explain overly optimistic IMF forecast bias. If this were the case, it is difficult to explain the asymmetric bias and inefficiency in IMF forecasts, where imports are unbiased and efficient, but exports and exchange rate forecasts show systematic forecast errors. Nevertheless, since this is an important line of reasoning, and we explore the effects of implementation failures by dropping programs from our dataset that experienced program cancellations.

Programs are canceled when IMF loan performance reviews do not produce sufficient evidence of conditionality implementation. The IMF monitors conditionality implementation in regular intervals as a part of its program surveillance (monthly, quarterly, or biannually depending on the program), and waivers may be obtained if conditionality or policy implementation is lagging. Lack of implementation may initially lead to “conditionality waivers,” but eventually, too many waivers may trigger program cancellation. In our dataset, cancellations occurred in 69 programs. In Table 3, we reproduce Table 1 without canceled programs to document that forecast inaccuracy was not driven by canceled programs. Indeed, results without canceled programs suffer, and forecast accuracy deteriorates. This is because most of the canceled programs (65 percent) are Non-LICs, which we now know to receive on average better forecasts than LICs. The absence of these Non-LICs thus weighs down the sample with a larger share of inefficiently forecasted LICs, so that overall forecast accuracy without these Non-LICs deteriorates. Limited conditionality implementations and cancellations thus cannot explain forecast inaccuracies.

### 1.5.2 Conditionality, Loan Size, and Geopolitics as Drivers of Inefficiency

In addition to program cancellation, the literature suggests three areas where information available at the time of the forecasts may not have been integrated properly into IMF forecasts. We examine all three areas below, which include IMF conditionality, loan size, and geopolitical events. A voluminous literature examines the effect of conditionality of forecast errors (see, for example, the survey contained in [Stubbs et al. \(2020\)](#)), but this literature is hampered by selection bias since conditionality is included only selectively. For example, [IMF \(2019\)](#) associates monetary conditionality with inflation forecast errors, while [Ismail et al. \(2020\)](#) examine whether the number of conditions affects GDP growth forecast errors. [Carriere-Swallow and Marzluf \(2021\)](#) examine whether credit and fiscal conditionality affects GDP growth forecasts. Only [Eicher and Rollinson \(2022\)](#) examine whether any of the entire palette of conditions affect forecast accuracy GDP growth and inflation. We follow their approach and, instead of selecting particular conditions, we test the entire range of conditionality as categorized into thematic topics by the IMF in the MONA database.<sup>15</sup>

Second, we investigate whether loan size affects IMF forecast accuracy since an important strand of the literature consistently links IMF forecast optimism to the size of the loan a program received (see, e.g, [Beach and Schavey \(1999\)](#); [Dreher et al. \(2008\)](#); [Luna \(2014\)](#)). Finally, we explore whether the economic effects of geopolitical events, known to IMF forecasters at the time of forecasts, were adequately integrated into IMF forecasts. In their reviews of IMF program design and conditionality, [IMF \(2019\)](#) note that forecast errors are impacted by political transitions, conflicts, and natural disasters (see also and [Przeworski and Vreeland \(2000\)](#); [Goldstein et al. \(2003\)](#); [Mody and Rebucci \(2006\)](#); [Kentikelenis et al. \(2016\)](#)). Hence, we consider variables related to elections (head of state/government and legislative) up to one year before the start of a program. For conflicts, we include indicators for (civil) wars that were initiated up to one year prior to program starts. Finally, we consider natural disasters that occurred up to one year prior to program starts (see Table A1 for election, conflict, and disaster data). Crucial is that all geopolitical events were known at the time of forecasts, so IMF forecasters were well aware of their potential effects on economic performance.

---

<sup>15</sup>The IMF MONA database provides a list of conditions for each program, grouped by IMF-specified programmatic objectives (see Table A1).

To investigate drivers of inefficiencies, we utilize the methodology of [Sinclair et al. \(2010\)](#), who extended Mincer-Zarnowitz regressions (1) to include additional covariates,  $X_t$ , that are thought to represent information available to forecasters at the time of the forecast:

$$A_t = \alpha + \beta F_t + \delta X_t + \epsilon_t, \quad (1.2)$$

If any entries in the vector  $\delta$  are non-zero, the information contained in the associated covariates can then, in part, explain forecasters' bias and inefficiencies. Significant covariates thus represent areas the IMF may consider with special caution in future IMF program forecasts. [Sinclair et al. \(2010\)](#) propose the joint null hypothesis of  $\beta = 1$  &  $\alpha = \delta = 0$  as a formal test of whether the information contained in the additional covariates was properly included in the forecast. If the null is rejected, [Sinclair et al. \(2010\)](#) note that the information contained in  $X_t$  was not fully integrated into the forecast, which then identifies possible sources of inefficiency. We apply the methodology to those three subsamples that were identified as drivers of forecast errors for imports, exports, and exchange rate forecasts (Table 4).

In Table 4, we are considering (i) fixed exchange rate LICs subsample for import forecasts, (ii) the full sample for export forecasts, and (iii) the floating exchange rate LICs subsample for exchange rate depreciation forecasts. The [Sinclair et al. \(2010\)](#) test indicates that the additional, significant regressors cannot be rejected as information known at the time of the forecast that was not properly integrated for import and export growth. For real exchange rate depreciations forecast, the [Sinclair et al. \(2010\)](#) test rejects the hypothesis that the additional regressors could have improved forecast accuracy; this finding is not surprising given the problems encountered by exchange rate forecasts in general (see, e.g., [Meese and Rogoff \(1983\)](#))

Given the sizable number of covariates, Table 4 reports only significant regressors in either subsample to economize on space. Significant coefficients are found for 11 covariates, whose effects were thus not properly integrated into IMF forecasts. Five of these regressors relate to quantitative conditionality: arrears ceilings (domestic & foreign), government finance (credit and balance limits), and short-term debt, while two relate to structural reform conditionality (trade and government). Geopolitical effects are also significant, specifically elections, wars, and civil wars.

Subsamples do not share statistically significant regressors, which is not surprising since the regressions cover not only different country-types but also different dependent variables (ex-

port/import/exchange rates) and different subsample characteristics (fixed/flex exchange rates). Nevertheless, several common themes present themselves as LICs subsamples share significant regressors in the arrears category (domestic arrears for imports and external arrears for exchange rates), and trade forecasts share significant regressors in the fiscal category (fiscal balance limits for imports and government credit limits for exports). Truly broad-based effects are established only for geopolitical events where either elections or (civil) wars are significant in all three subsamples. Finally, we also confirm the findings in the previous literature that loan size is indeed a factor that affects IMF forecast accuracy, but only for import forecasts in LICs with fixed exchange rates.

## 1.6 Did Forecast Bias and Efficiency Change Over Time?

Forecasts for IMF program countries are available dating back to 1992, which raises the question of whether the accuracy of IMF external sector forecasts has changed over time. Instead of reporting forecast accuracy for individual years, we report results for 5-year rolling windows to avoid large fluctuations in the number of observations across samples and retain sufficiently large samples for each time interval. Figures 2.1-2.3 provide visual summaries of the evolution of forecast accuracy over time, based on Mincer-Zarnowitz regressions reported in Appendix Table A.2. The solid black line in Figure 2.1 represents the values of  $\beta$  estimates in Table A.2, while the dotted lines represent 95 percent confidence intervals. Yellow dots indicate years with inefficient forecasts; bias years receive an orange dot.

Only export growth forecasts show a marked improvement in both bias and efficiency since 2013. Import growth forecasts and real exchange rate depreciations continue to exhibit substantial inaccuracies even in recent years. Real exchange rate forecasts were indeed much more accurate prior to the arrival of 2007 in the 5-year rolling window. It is fascinating to see that, even in windows when real exchange rate forecasts are accurate, import and export forecasts can exhibit substantial bias and/or inefficiency. The slope coefficients for real exchange rate depreciations spike in the 2000-2004 window to exceed unity while substantially declining thereafter.

## 1.7 Do Forecast Horizons Affect Forecast Accuracy?

Forecast accuracy is generally thought to decrease as forecast horizons increase ([Armstrong \(2001\)](#) and [USGAO \(2003\)](#)). One might suspect this to be particularly relevant for IMF program forecasts, as forecasters' information sets grow substantially in size and accuracy towards the end of the year as ever more data vintages are released. [Timmermann \(2007\)](#) previously found evidence that IMF forecast errors increase with time horizons (using WEO data, which does not include crisis countries). In this section, we examine whether forecast horizons drive bias and inefficiency for IMF trade forecasts. Figure 3a)-3c) provide a visual summary of forecast accuracy by the month in which programs were approved. The figure is a visual representation of the Mincer-Zarnowitz regressions in Appendix Table A.3.

Examining whether forecasts produced earlier in the program year exhibit a greater propensity towards bias and inefficiency than those formed later in the year does not produce the expected result. There is no clear pattern of improved forecast accuracy for programs that are approved later in the year. Import and export forecasts show inefficiencies as late as December and November, respectively. Exchange rate forecasts show inefficiencies as late as October. Biases and inefficiencies are distributed roughly evenly across the year without a clear pattern of either bias or efficiency improvements as forecast horizons change.

## 1.8 Conclusion

We analyze the accuracy of IMF's external sector forecasts for crisis countries. The previous literature on trade forecast evaluations was hampered by small-sample studies that cover ever different samples and time periods and often conflated import, export and exchange rate forecast errors. This may have been the reason for the vastly different results in the literature, ranging from exactly zero average forecast errors to zero informational values of IMF forecasts. We audited the IMF MONA database which contains external sector forecasts of crisis countries to overcome the small-sample issues; our sample is thus three times larger than any previous study and it contains the first evaluation of real exchange rate depreciation forecasts. The size of our dataset also allows us to drill down to different subsample levels (country-income levels, program-types, exchange rate regimes, financial openness, and trade shares) to understand the drivers of forecast inaccuracies.

Our results indicate import forecasts are broadly unbiased and efficient while export forecasts are broadly biased and inefficient. We can confirm that this surprising asymmetry in import and export forecast accuracy is not driven by exchange rate forecast inaccuracies. Once the sample is purged of LICs exchange rate, forecasts are unbiased and efficient in contrast to what one might expect given the [Meese and Rogoff \(1983\)](#) puzzle. Nevertheless, for the full sample and for floating exchange rate countries, unusually large forecast errors hamper exchange rate forecasts and at times cannot outperform the random walk. The exchange rate is a crucial variable and central to equilibrium conditions in both goods and asset markets of the macroeconomic model predicting crisis recovery. Hence, we suggest the explicit acknowledgment of uncertainties involved in IMF exchange rate forecasts and how these uncertainties translate to debt and reserve adequacy projections. This recommendation contrasts with the recent IMF's policy of not disclosing the exchange rate forecasts that underlie the entire crisis recovery program.

We also examine whether IMF conditionality is a driver of bias and inefficiency to find that IMF conditionality related to trade and government reforms, ceilings on domestic/foreign arrears, and limits on fiscal finance (government credit/fiscal balance) is not properly integrated into IMF forecasts. The policy implication being that IMF forecasts must be sensitive to the effects of IMF conditionality, especially as they relate to fiscal finance, arrears and reforms. We also find that geopolitical events that are known at the time of the forecast (elections and armed conflict), were not properly accounted for in IMF forecasts to cause systematic forecast inaccuracies. Future forecasts should thus be more sensitive to the economic impact of non-economic events in crisis countries' recoveries.

Contrary to conventional wisdom, we find the accuracy of IMF current-year forecasts does not improve as the time horizon shrinks, but our longest forecast horizon is only one year. While forecast accuracy did improve over the past 28 years for exports, the accuracy of imports and exchange rate forecasts remains a challenge even in the most recent forecasts. Clearly forecast accuracy is affected by exogenous shocks, such as the COVID crisis or the 2008 global financial crisis. However, if the forecast errors we observe were due to idiosyncratic shocks only, we would expect them to cancel instead of producing the systematic bias and inefficiency.

Our findings of systematically optimistic, biased and inefficient export and exchange rate forecasts, especially for Low Income Countries, point a direction towards further research as to

the origins of the forecast errors. One approach is to not publish exchange rate forecasts anymore (as evidenced in recent IMF loan documents); perhaps the better approach would be an upfront acknowledgment of the uncertainty surrounding exchange rate forecasts along with the inclusion of alternative exchange rate forecast scenarios for the associated macroeconomic implications.

## Tables and Figures:

Table 1.1: Forecast Bias and Inefficiency in IMF Forecasts

Dependent: Actual Final Data	Import Growth		Export Growth		Exchange Rate (Depreciation)
	EKPC*	Full	EKPC*	Full	Full
	1a	1b	2a	2b	3b
<b>Forecast (<math>\beta</math>)</b>	0.63**	0.91	0.72***	0.93	0.86**
<b>p-value (<math>\beta = 1</math>)</b>	0.03	0.54	0	0.43	0.04
<b>Intercept (<math>\alpha</math>)</b>	0.02	0	0.03	0.04***	0
<b>p-value (<math>\alpha = 0</math>)</b>	0.5	0.8	0.18	0	0.62
<b>Observations</b>	110	578	110	576	565
<b>Adj. <math>R^2</math></b>	0.47	0.35	0.55	0.24	0.73
<b>MZ p-value (<math>\alpha=0</math> &amp; <math>\beta=1</math>)</b>	0.08*	0.77	0.00***	0.00***	0.11
<b>HP p-value (<math>\gamma=0</math>)</b>	0.89	0.58	1	0.00***	0.4
<b>Theil U2</b>	0.65	0.76	0.55	0.79	0.54

*Note:* EKPC Table (1) regressions include regional dummies, which we do not include to produce proper Mincer Zarnowitz regressions to assess IMF forecast quality.

Table 1.2: Forecast Bias and Inefficiency by Subsamples

	(2a) Import Growth															
	All				Non-Hyperinflation				LICs				Non-LICs			
	Full	BOP	Float	Fixed	Full	BOP	Float	Fixed	Full	BOP	Float	Fixed	Full	BOP	Float	Fixed
Forecast ( $\beta$ )	0.912***	1.081***	0.963***	0.808***	0.925***	1.110***	1.003***	0.733***	0.940***	1.246***	1.047***	0.672***	0.898***	0.980***	0.942***	0.805***
p-value ( $\beta = 1$ )	0.54	0.61	0.85	0.05	0.64	0.53	0.99	0.01	0.83	0.32	0.87	0.01	0.31	0.86	0.61	0.33
Intercept ( $\alpha$ )	0	-0.01	-0.01	0.020*	0	-0.01	-0.01	0.029**	-0.01	-0.052*	-0.03	0.041**	0.01	0	0	0.02
p-value ( $\alpha=0$ )	0.8	0.52	0.72	0.07	0.97	0.3	0.39	0.01	0.77	0.09	0.33	0.02	0.39	0.93	0.99	0.16
Observations	578	308	371	207	512	263	324	188	269	63	171	98	243	200	153	90
Adj. $R^2$	0.35	0.53	0.35	0.34	0.38	0.57	0.43	0.26	0.37	0.71	0.43	0.22	0.37	0.41	0.42	0.26
p-value ( $\alpha=0$ & $\beta=1$ )	0.77	0.81	0.65	0.07*	0.7	0.58	0.29	0.01**	0.52	0.23	0.24	0.03**	0.5	0.98	0.87	0.31
HP p-value ( $\gamma=0$ )	0.58	0.85	0.36	0.66	0.41	0.62	0.16	0.56	0.27	0.45	0.15	0.83	0.88	0.98	0.72	0.5
Theil U2	0.76	0.66	0.77	0.75	0.74	0.63	0.72	0.78	0.73	0.53	0.72	0.76	0.75	0.74	0.72	0.81

<b>(2b) Export Growth</b>																
	All				Non-Hyperinflation				LICs				Non-LICs			
	Full	BOP	Float	Fixed	Full	BOP	Float	Fixed	Full	BOP	Float	Fixed	Full	BOP	Float	Fixed
Forecast ( $\beta$ )	0.928***	1.030***	0.995***	0.811***	0.912***	1.014***	0.983***	0.769***	0.946***	1.228***	1.043***	0.750***	0.880***	0.914***	0.902***	0.841***
p-value ( $\beta=1$ )	0.43	0.83	0.97	0.05	0.38	0.93	0.9	0.02	0.71	0.59	0.84	0.07	0.22	0.39	0.42	0.31
Intercept ( $\alpha$ )	0.036***	0.022**	0.029**	0.050***	0.036***	0.023*	0.025*	0.057***	0.03	-0.01	0.01	0.053**	0.045***	0.033***	0.037**	0.059***
p-value ( $\alpha=0$ )	0	0.04	0.03	0	0	0.06	0.07	0	0.13	0.79	0.58	0.03	0	0.01	0.02	0.01
Observations	576	307	369	207	511	263	323	188	268	63	170	98	243	200	153	90
Adj. $R^2$	0.24	0.34	0.24	0.25	0.23	0.32	0.25	0.2	0.25	0.32	0.26	0.21	0.2	0.31	0.22	0.16
p-value ( $\alpha=0$ & $\beta=1$ )	0***	0.03**	0.03**	0***	0***	0.04**	0.1*	0***	0.21	0.83	0.63	0.07*	0***	0.03**	0.07*	0.02**
HP p-value ( $\gamma=0$ )	0***	0.02**	0.01**	0.01**	0***	0.03**	0.05**	0.01***	0.14	0.57	0.35	0.18	0***	0.01***	0.05**	0.01***
Theil U2	0.79	0.74	0.8	0.76	0.79	0.75	0.79	0.78	0.77	0.75	0.78	0.74	0.81	0.74	0.8	0.82

<b>(2c) Exchange Rate Growth</b>																
	All				Non-Hyperinflation				LICs				Non-LICs			
	Full	BOP	Float	Fixed	Full	BOP	Float	Fixed	Full	BOP	Float	Fixed	Full	BOP	Float	Fixed
Forecast ( $\beta$ )	0.861***	0.814***	0.864***	0.795***	0.779***	0.947***	0.746***	0.888***	0.555***	0.884**	0.488*	0.746***	0.982***	0.975***	0.955***	1.107***
p-value ( $\beta=1$ )	0.044	0.059	0.077	0	0.049	0.597	0.061	0.371	0.029	0.782	0.046	0.104	0.761	0.673	0.481	0.428
Intercept ( $\alpha$ )	-0.004	-0.002	0.002	-0.005	0.005	0.016*	0.008	0.002	0.002	0.036	0.005	0.002	0.006	0.009	0.01	0.001
p-value ( $\alpha=0$ )	0.622	0.893	0.856	0.324	0.447	0.082	0.36	0.717	0.882	0.357	0.742	0.816	0.281	0.135	0.198	0.873
Observations	565	302	367	197	500	257	321	179	262	59	170	92	238	198	151	87
Adj. $R^2$	0.73	0.61	0.75	0.84	0.25	0.38	0.21	0.47	0.09	0.14	0.05	0.34	0.59	0.59	0.57	0.68
p-value ( $\alpha=0$ & $\beta=1$ )	0.11	0.17	0.21	0***	0.04**	0.03**	0.07*	0.45	0.03**	0.07*	0.08*	0.13	0.51	0.29	0.33	0.72
HP p-value ( $\gamma=0$ )	0.4	0.72	0.89	0.18	0.18	0.05**	0.29	0.27	0.35	0.2	0.58	0.14	0.26	0.13	0.19	0.77
Theil U2	0.54	0.65	0.52	0.46	0.88	0.79	0.9	0.68	0.98	0.93	1	0.79	0.64	0.64	0.66	0.53

<b>Table 2(d): Contributions of Nominal Exchange Rate and Inflation to Real Exchange Rate Depreciation Inaccuracies</b>																
	Full Sample				Non-Hyper Inflation				LICs				Non-LICs			
	Full	BOP	Float	Fixed	Full	BOP	Float	Fixed	Full	BOP	Float	Fixed	Full	BOP	Float	Fixed
<b>Nom. E Forecast (<math>\beta_1</math>)</b>	0.82***	0.75***	0.84***	0.80***	0.79***	0.96***	0.75***	0.94***	0.56***	0.89**	0.49*	0.80***	0.99***	0.98***	0.96***	1.17***
<b>p-val. (<math>\beta_1=0</math>)</b>	0	0	0	0	0	0	0	0	0.01	0.05	0.05	0	0	0	0	0
<b>Inflation Forecast (<math>\beta_2</math>)</b>	0.74***	0.68***	0.75***	1.36***	0.70***	0.86***	0.68***	0.79***	0.52*	1.01**	0.46	0.68**	0.85***	0.86***	0.84***	0.96***
<b>p-val. (<math>\beta_2=0</math>)</b>	0	0	0	0	0	0	0.01	0	0.09	0.04	0.29	0.01	0	0	0	0
<b>Constant (<math>\alpha</math>)</b>	-0.01	-0.01	-0.01	0.03	0	0.01	0	0	0	0.05	0	0	0	0	0	-0.01
<b>p-val. (<math>\alpha=0</math>)</b>	0.11	0.34	0.24	0.23	0.89	0.45	0.91	0.71	0.95	0.2	0.96	0.99	0.73	0.97	0.97	0.42
<b>Observations</b>	559	297	364	195	495	253	319	176	258	56	168	90	237	197	151	86
<b>Adjusted R-squared</b>	0.7	0.55	0.76	0.44	0.25	0.38	0.21	0.48	0.09	0.13	0.05	0.34	0.59	0.59	0.57	0.69
<b>p-value (<math>\beta_1=\beta_2</math>)</b>	0.16	0.33	0.05	0.19	0.56	0.56	0.71	0.35	0.87	0.77	0.9	0.62	0.42	0.49	0.61	0.32

Table 2(e): Import Forecast Bias and Inefficiency by Subsamples																				
	Full Sample					Non-Hyper Inflation					LICs					Non-LICs				
	Full		Fin. Open		Trade Share	Full		Fin. Open		Trade Share	Full		Fin. Open		Trade Share	Full		Fin. Open		Trade Share
	Full	Large	Small	Large	Small	Full	Large	Small	Large	Small	Full	Large	Full	Large	Small	Full	Large	Small	Large	Small
<b>forecast</b>	0.91	0.9	0.91	0.84	0.93	0.92	0.91	0.94	0.81	0.95	0.95	0.82	0.96	0.77	0.98	0.9	1	0.85	0.86	0.91
<b>p-value (<math>\beta = 1</math>)</b>	0.541	0.347	0.671	0.281	0.712	0.644	0.398	0.772	0.262	0.811	0.833	0.424	0.897	0.43	0.944	0.307	0.995	0.38	0.43	0.472
<b>Constant</b>	0.00	-0.00	0.00	0.01	0.00	0.00	-0.00	0.00	0.01	-0.00	-0.01	-0.01	-0.01	0.00	-0.01	0.01	0.00	0.01	0.01	0.00
<b>p-value (<math>\alpha = 0</math>)</b>	0.797	0.653	0.871	0.528	0.969	0.974	0.791	0.929	0.484	0.861	0.766	0.602	0.856	0.927	0.733	0.389	0.774	0.352	0.341	0.687
<b>Observations</b>	578	181	361	162	416	512	173	316	141	371	269	59	204	48	221	243	114	112	93	150
<b>Adjusted R-squared</b>	0.345	0.384	0.338	0.276	0.367	0.376	0.397	0.369	0.265	0.406	0.372	0.294	0.385	0.200	0.408	0.367	0.477	0.287	0.334	0.372
<b>F-test (<math>\alpha = 0</math> &amp; <math>\beta = 1</math>)</b>	0.267	0.870	0.151	0.589	0.167	0.355	0.606	0.107	0.636	0.286	0.657	1.456	0.240	0.464	0.472	0.687	0.0419	0.566	0.498	0.306
<b>p-value (<math>\alpha = 0</math> &amp; <math>\beta = 1</math>)</b>	0.765	0.421	0.86	0.556	0.847	0.701	0.547	0.898	0.531	0.752	0.519	0.242	0.787	0.631	0.624	0.504	0.959	0.569	0.609	0.737
<b>Holden-Peel T-Test (<math>\gamma = 0</math>)</b>	-0.559	-1.139	-0.464	-0.189	-0.540	-0.830	-0.898	-0.451	-0.358	-0.751	-1.112	-1.578	-0.667	-0.673	-0.897	0.147	0.269	0.275	0.259	0.00175
<b>p-value</b>	0.576	0.256	0.643	0.85	0.589	0.407	0.37	0.653	0.721	0.453	0.267	0.12	0.505	0.504	0.371	0.883	0.789	0.784	0.796	0.999
<b>RMSE</b>	0.175	0.131	0.188	0.176	0.175	0.164	0.130	0.178	0.158	0.166	0.185	0.153	0.194	0.210	0.180	0.135	0.114	0.145	0.123	0.143
<b>RMSPE</b>	28.68	44.87	10.49	15.13	32.46	27.98	45.68	11.21	16.21	31.31	36.76	76.89	8.495	11.22	40.22	12.39	10.37	14.94	18.27	6.456
<b>Theil U</b>	0.764	0.751	0.769	0.789	0.755	0.739	0.738	0.735	0.789	0.724	0.733	0.809	0.721	0.846	0.708	0.750	0.685	0.786	0.723	0.764

Table 2(f): Export Forecast Bias and Inefficiency by Subsamples																				
	Full Sample					Non-Hyper Inflation					LICs					Non-LICs				
	Full		Fin. Open		Trade Share	Full		Fin. Open		Trade Share	Full		Fin. Open		Trade Share	Full		Fin. Open		Trade Share
	Full	Large	Small	Large	Small	Full	Large	Small	Large	Small	Full	Large	Full	Large	Small	Full	Large	Small	Large	Small
<b>forecast</b>	0.93	0.82	0.91	1.02	0.85	0.91	0.83	0.91	1	0.84	0.95	0.94	0.95	1.02	0.89	0.88	0.79	0.87	1.01	0.79*
<b>p-value (<math>\beta = 1</math>)</b>	0.428	0.136	0.435	0.737	0.334	0.375	0.163	0.518	0.956	0.339	0.711	0.802	0.778	0.796	0.66	0.215	0.143	0.379	0.922	0.0979
<b>Constant</b>	0.04***	0.02*	0.05***	0.04**	0.04***	0.04***	0.02	0.05***	0.03**	0.04***	0.03	-0.02	0.03*	0.02	0.03	0.05***	0.03**	0.07***	0.04**	0.05***
<b>p-value (<math>\alpha = 0</math>)</b>	0.00017	0.0998	0.000678	0.0157	0.00281	0.000696	0.139	0.00165	0.0139	0.00878	0.126	0.495	0.08	0.532	0.203	0.000406	0.0171	0.00339	0.0169	0.00617
<b>Observations</b>	576	181	360	162	414	511	173	315	141	370	268	59	203	48	220	243	114	112	93	150
<b>Adjusted R-squared</b>	0.241	0.279	0.210	0.479	0.155	0.232	0.289	0.202	0.500	0.149	0.249	0.253	0.252	0.545	0.158	0.201	0.338	0.109	0.421	0.125
<b>F-test (<math>\alpha = 0</math> &amp; <math>\beta = 1</math>)</b>	8.334	1.709	6.676	4.640	5.409	6.788	1.441	6.288	3.777	4.233	1.573	0.494	2.279	0.321	1.392	6.440	2.982	4.487	4.987	4.424
<b>p-value (<math>\alpha = 0</math> &amp; <math>\beta = 1</math>)</b>	0.000270	0.184	0.00142	0.0110	0.00480	0.00123	0.239	0.00210	0.0253	0.0152	0.209	0.613	0.105	0.727	0.251	0.00189	0.0547	0.0134	0.00881	0.0136
<b>Holden-Peel T-Test (<math>\gamma = 0</math>)</b>	3.556	0.282	3.154	3.007	2.518	3.129	0.192	3.166	2.752	2.227	1.495	-0.967	1.874	0.782	1.296	3.117	1.341	2.729	3.157	1.947
<b>p-value</b>	0.000408	0.778	0.00174	0.00306	0.0122	0.00185	0.848	0.00169	0.00671	0.0266	0.136	0.338	0.0624	0.438	0.196	0.00205	0.183	0.00738	0.00216	0.0534
<b>RMSE</b>	0.203	0.136	0.225	0.162	0.217	0.202	0.137	0.232	0.151	0.219	0.216	0.172	0.225	0.188	0.222	0.185	0.113	0.243	0.129	0.213
<b>RMSPE</b>	116.2	6.815	146.8	5.812	137	15.96	6.843	19.69	6.103	18.38	21.80	10.73	24.37	9.160	23.68	3.388	3.381	3.636	3.629	3.230
<b>Theil U</b>	0.788	0.766	0.805	0.633	0.842	0.787	0.762	0.798	0.613	0.839	0.772	0.804	0.761	0.581	0.822	0.810	0.719	0.866	0.652	0.870

	Table 2(g): Exchange Rate Forecast Bias and Inefficiency by Subsamples																			
	Full Sample					Non-Hyper Inflation					LICs					Non-LICs				
	Full	Fin. Open	Trade Share	Full	Fin. Open	Trade Share	Full	Fin. Open	Trade Share	Full	Fin. Open	Trade Share	Full	Fin. Open	Trade Share	Full	Fin. Open	Trade Share	Full	Fin. Open
<b>forecast</b>	0.86**	0.38***	0.89	0.92	0.69***	0.78**	1.01	0.61***	0.85	0.74***	0.56**	1.10	0.43***	0.62	0.49***	0.98	0.95	0.92	1.01	0.96
<b>p-value (<math>\beta = 1</math>)</b>	0.0442	0.000198	0.385	0.254	5.52E-06	0.0486	0.968	0.00443	0.57	0.000949	0.0288	0.897	0.00236	0.432	0.00011	0.761	0.704	0.167	0.865	0.62
<b>Constant</b>	-0.00	0.01	-0.01	0.03	-0.01*	0.00	0.01	0.00	0.04**	-0.01	0.00	0.03	-0.00	0.09**	-0.02**	0.01	0.00	0.01	0.01	0.01
<b>p-value (<math>\alpha = 0</math>)</b>	0.622	0.508	0.293	0.107	0.0662	0.447	0.415	0.652	0.0452	0.261	0.882	0.428	0.673	0.0472	0.0274	0.281	0.746	0.298	0.407	0.449
<b>Observations</b>	565	178	350	158	407	500	170	308	135	365	262	58	198	48	214	238	112	110	87	151
<b>Adjusted R-squared</b>	0.731	0.0751	0.562	0.818	0.483	0.246	0.201	0.187	0.218	0.285	0.0885	0.0984	0.0817	0.0590	0.110	0.587	0.486	0.481	0.609	0.566
<b>F-test (<math>\alpha = 0</math> &amp; <math>\beta = 1</math>)</b>	2.225	7.963	1.381	1.807	14.50	3.321	1.072	4.390	3.050	5.553	3.632	0.791	4.839	3.527	7.989	0.677	0.298	1.505	0.353	0.467
<b>p-value (<math>\alpha = 0</math> &amp; <math>\beta = 1</math>)</b>	0.109	0.000489	0.253	0.168	8.27e-07	0.0369	0.345	0.0132	0.0507	0.00421	0.0278	0.458	0.00889	0.0376	0.000452	0.509	0.743	0.227	0.704	0.628
<b>Holden-Peel T-Test (<math>\gamma = 0</math>)</b>	-0.841	0.960	-1.537	1.583	-2.845	1.333	1.081	0.991	2.318	-0.530	0.940	0.947	0.610	2.220	-1.064	1.122	0.530	1.018	0.806	0.787
<b>p-value</b>	0.401	0.338	0.125	0.115	0.00467	0.183	0.281	0.323	0.0219	0.596	0.348	0.347	0.542	0.0313	0.289	0.263	0.597	0.311	0.422	0.432
<b>RMSE</b>	0.194	0.197	0.187	0.262	0.158	0.134	0.147	0.129	0.189	0.106	0.168	0.232	0.145	0.287	0.120	0.0842	0.0709	0.0931	0.0885	0.0819
<b>RMSPE</b>	381	24.66	483.8	2.022	448.9	14.94	25.23	3.304	2.170	17.43	13.99	28.87	3.854	1.391	15.46	15.91	23.12	1.959	2.498	19.89
<b>Theil U</b>	0.536	1.069	0.666	0.436	0.790	0.875	0.893	0.935	0.896	0.852	0.980	0.946	1.025	0.987	0.971	0.638	0.705	0.711	0.617	0.654

Table 1.3: Forecast Bias and Inefficiency  
Sample Without Cancelled Programs

Dependent: Actual Final Data	Imports	Exports	Exchange Rate
	No Cancelled	No Cancelled	No Cancelled
	1c	2c	3c
<b>Forecast (<math>\beta</math>)</b>	0.71***	0.85*	0.88*
<b>p-value (<math>\beta = 1</math>)</b>	0.00	0.08	0.07
<b>Intercept (<math>\alpha</math>)</b>	0.02*	0.05***	-0.00
<b>p-value (<math>\alpha = 0</math>)</b>	0.09	0	0.71
<b>Observations</b>	511	509	500
<b>Adj. <math>R^2</math></b>	0.20	0.18	0.74
<b>p-value (<math>\alpha = 0</math> &amp; <math>\beta = 1</math>)</b>	0.00***	0.00***	0.17
<b>HP p-value (<math>\gamma = 0</math>)</b>	0.34	0.00***	0.53
<b>Theil U2</b>	0.85	0.81	0.53

Table 1.4: Sources of Forecast Inefficiencies

Dependent Variable:		Import Growth (Fixed Exchange Rate LICs)	Export Growth (Full Sample)	Real Exchange Rate Depreciation (Float LICs)
Forecast	Forecast ( $\beta$ )	0.67**	0.90	0.58*
	p-value ( $\beta=1$ )	0.02	0.31	0.07
	Loan Size (Loan/Quota)	-7.81*** 0.00	-0.06 0.66	6.61 0.24
Conditionality: Quantitative Performance Criteria	Arrears: External Ceilings	0.02 0.74	0.02 0.27	-0.06* 0.07
	Arrears: Domestic Ceilings	0.09* 0.05	-0.03 0.15	-0.02 0.49
	Fiscal: Gov't/Pub. Credit	-0.00 0.92	-0.05** 0.04	0.04 0.45
	Fiscal: Gov't Balance Limits	-0.07* 0.09	-0.01 0.65	0.02 0.5
	Debt: Short-Term	0.02 0.73	0.04** 0.05	-0.03 0.25
	Reforms: General Gov't	0.20*** 0.01	0.00 0.91	0.08 0.35
	Reforms: Trade	-0.03 0.51	0.03 0.24	0.08** 0.02
	Geopolitics	Elections Wars: international Na Na	-0.01 0.81 Na Na	-0.03* 0.05 -0.06** 0.03
	Wars: intranational	0.07* 0.09	-0.03 0.25	0.03 0.54
	Constant ( $\alpha$ )	-0.14 0.13	0.025 0.03	0.03 0.19
	Observations	98	576	170
	Adjusted R-squared	0.31	0.24	0.18
	SJS F-test ( $\alpha=\delta=0$ & $\beta=1$ )	5.34***	1.46*	1.33
	p-value	0.00	0.07	0.16

Figure 1.1: Prediction Realization Diagram, Full Sample

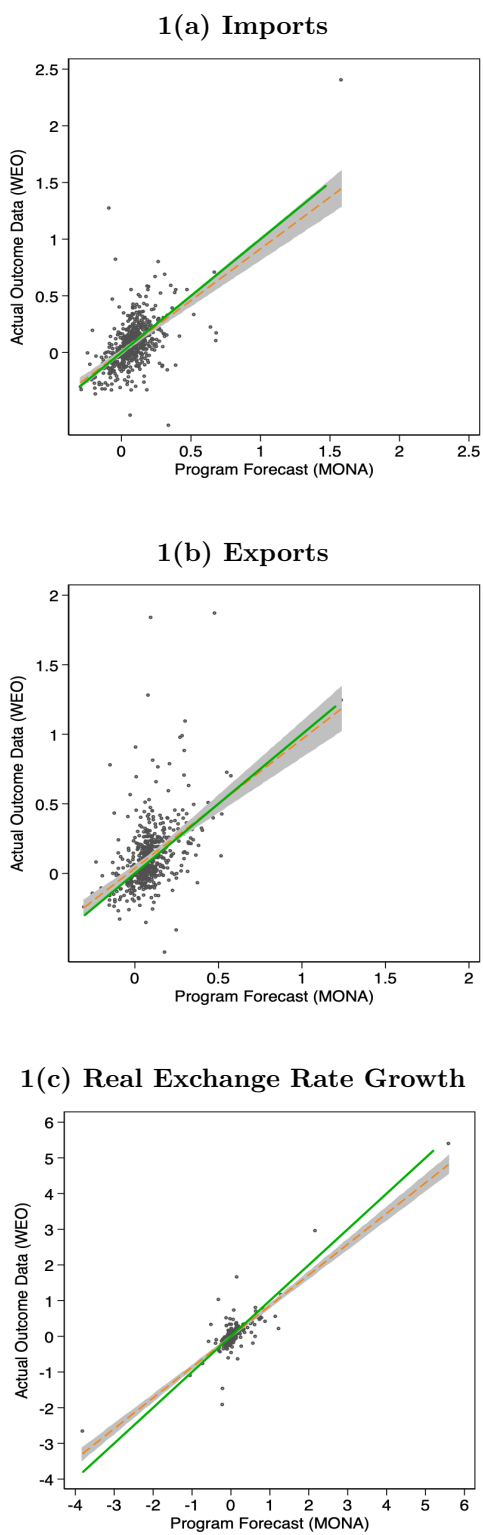


Figure 1.2: Evolution of Forecast Bias and Efficiency Over Time

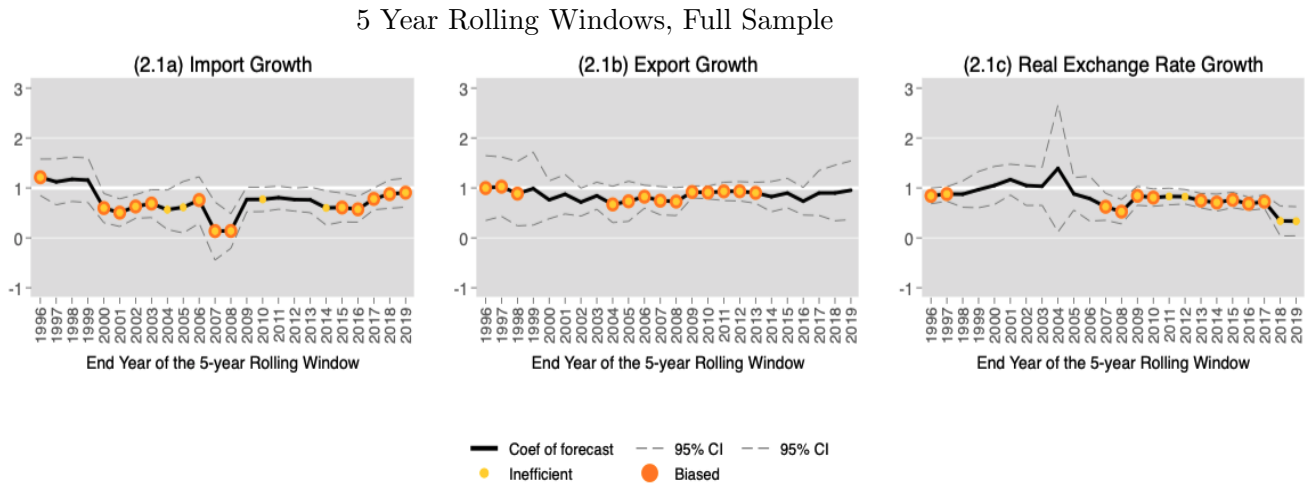
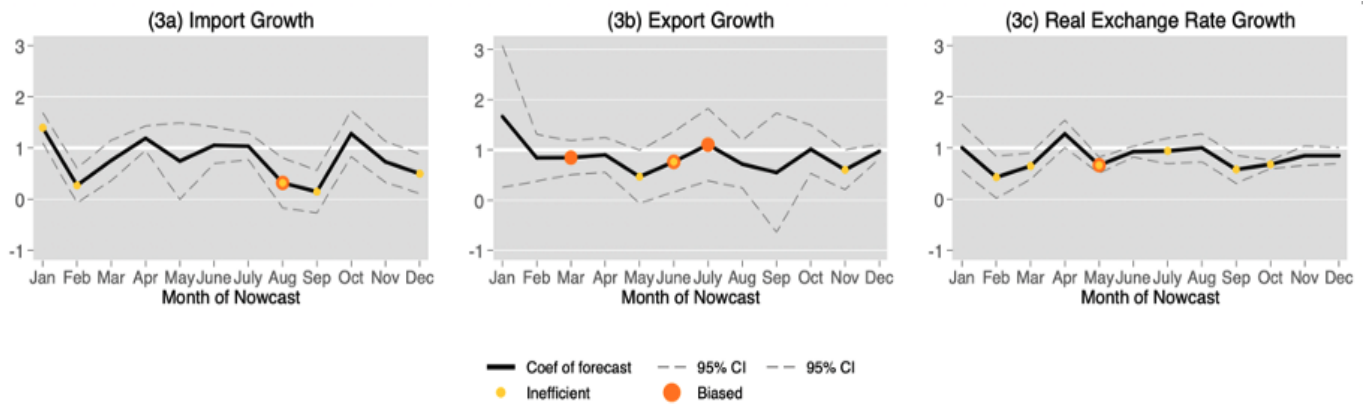


Figure 1.3: Forecast Horizons and Forecast Accuracy, Full Sample



## 1.9 Data Correction

## 1.10 Regressions for Figures 2 and 3

### 1.10.1 Regression Result for Figures 2

Table 1.5: Survey of IMF Trade Forecast Accuracy Studies Sample Sizes, Data Sources, Variable Coverage

Author/Year	Scope of Study	Years	Data Source	Trade Variables
Arora and Smyth (1990)	Developing countries aggregated into 1 forecast per continent	1981-1988	WEO	import and export, current account % GDP
Artis (1988)	7 industrial countries. Developing countries aggregated into 1 forecast per continent	1971-1986	WEO	import and export, current account % GDP
Artis (1996)	7 industrial countries. aggregated into 1 forecast per continent	1972-1994	WEO	import and export, current account % GDP
Artis and Zhang (1990)	G7 countries	1980-1987	WEO	current account % GDP
Atoyan and Conway (2011)	183 IMF loan programs	1993-2009	MONA	current account % GDP
Atoyan et al. (2004)	175 IMF loan programs	1993-2001	MONA	current account % GDP
Baqir et al. (2006)	94 IMF loan programs	1989-2002	MONA	current account % GDP
Beach et al. (1999)	14 countries	1971-1998	WEO	current account % GDP
EKPC (2019)	110 IMF loan programs	2002-2015	MONA	import and export
Fратиanni et al. (1991)	G7 countries	1980-1987	WEO	current account % GDP
Frenkel et al. (2013)	G7 countries	1989-2010	WEO	current account % GDP
IMF (2012)	148 IMF loan programs	2002-2011	MONA	current account % GDP
Kenen and Schwartz (1986)	7 industrialized countries. Non-oil LDCs are aggregated into 1 forecast per continent. Imports (exports) feature 6 (8) years of forecasts	1971-1985	WEO	import and export
Luna (2014b)	103 IMF loan programs	2002-2011	MONA	current account % GDP
Musso and Phillips (2002)	69 IMF loan programs	1993-1997	MONA	current account % GDP
Timmermann (2007)	29 industrialized countries, 149 developing countries aggregated into 1 forecast per continent	1990-2003	WEO	import and export, current account % GDP
US Government Accounting Office (2003)	87 emerging markets, of which 57 received loan programs, G7 countries	1990-2001	WEO	current account % GDP
Verbeek (1999)	23 countries	1991-1997	World Bank (Unified Survey Projections)	imports, exports, current account % GDP

### 1.10.2 Regression Result for Figures 3

Variable	Data Source	Description / Database Codenames
<b>Current-Year Forecasts</b>	IMF MONA database, (IMF 2021a). For the construction of our dataset and MONA error corrections, see Appendix C	"t-1" to "t" period growth rates for Real exchange rate (per \$) o MONA: ENDA - PCPIC o If ENDA is not available, MONA: ENDE o If PCPIC is not available, MONA: PCPIE Imports of goods and services in USD Pre 2002 MONA data o MONA: BMT+ BMS_O o If BMS_O are not available, MONA: BMG Post-2002 MONA data: o MONA: BMGS o If BMGS is not available, MONA: BMG +BMS o IF BMGS, BMG, BMS are not available, MONA: NM/ENDA Export of goods and services in USD Pre 2002 MONA data: o MONA: BXT+ BXS.O o IF BXS.O is not available, MONA: BXG Post-2002 MONA dataset: o MONA: BXGS, o If BXGS is not available, MONA: BXG + BXS If BXGS, BXG, BXS are not available, MONA: NX/ENDA
<b>Actual Final Data</b>	IMF IFS database, (IMF 2021c), IMF BOP database, (IMF 2021d), IMF WEO database, (IMF 2021b)	"t-1" to "t" period growth rates for Real exchange rate (per \$) o IFS: ENDA_XDC.USD_RATE – WEO: PCPIE o If ENDA_XDC.USD_RATE is not available IFS: ENDE_XDC.USD_RATE o If PCPIE is not available, WEO: PCPI Imports of goods and services in USD o IFS: BMGS_BP6.USD, o If BMGS_BP6.USD is not available, BOP: BM.GSR.GNFS.CD Export of goods and services in USD o IFS: BXGS_BP6.USD If BXGS_BP6.USD is not available, BOP: BX.GSR.GNFS.CD
<b>Elections Data</b>	Beck et al. (2001) pre 1998; IFES (2020) post1998	Election dummy covers head of state, government, and legislative election. Program received a "1" if election occurred up to 1 year prior to program start date. Details at <a href="http://www.electionguide.org">www.electionguide.org</a>
<b>Conflicts Data (UCDP/PRIO)</b>	Harbom et al. (2009)	Conflict dummy covers intra-state & inter-state conflicts. Program received a "1" if country experienced a conflict up to one year prior to program start date.
<b>Disasters Data</b>	EM-DAT (2020)	Disaster dummy covers natural disasters. Program received a "1" if a disaster occurred up to 1 year prior to the program start date.
<b>Conditionality: Quantitative Performance Criteria</b>	IMF MONA database, (IMF 2021a).	Dummy variables defined by MONA's Glossary (IMF 2021a) for quantitative performance targets: Domestic Credit Ceiling, Gov't/Public Sector Credit Ceilings, BOP Reserve Tests, Debt Ceilings (short, medium and long term), Arrears Ceilings (domestic and external), Fiscal Deficit Ceilings.
<b>Conditionality: Structural Performance Criteria</b>	IMF MONA database, (IMF 2021a).	Dummy variables defined based on the MONA Glossary (IMF 2021a) for structural performance criteria: General Gov't Reform, Central Bank Reforms, Civil Service Wage/Employment Reforms, Pension Reforms, Gov't Enterprise Pricing Reforms, Financial Sector Reforms, Current Capital Account Openness/Reforms, Tariff/Quota Reductions/Reforms, Labor Market Wage/Employment Reforms, Statistics Reforms, Legal/Market Reforms
<b>Trade Share</b>	World Bank (2022)	Trade (% of GDP), World Bank national accounts data, and OECD National Accounts data files
<b>Financial Openness</b>	Chinn Ito (2006)	Index measuring a country's degree of capital account openness, based on restrictions on cross-border financial transactions reported in the IMF's Annual Report on Exchange Arrangements and Exchange Restrictions (AREAER).

	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to
	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
<b>Imports</b>																								
<b>MONA Forecast, <math>\beta</math></b>	1.214***	1.121***	1.174***	1.155***	0.597***	0.505***	0.628***	0.687***	0.564***	0.613**	0.757***	0.136	0.139	0.769***	0.770***	0.804***	0.767***	0.761***	0.600***	0.608***	0.572***	0.780***	0.877***	0.906***
	(0.185)	(0.233)	(0.225)	(0.227)	(0.148)	(0.140)	(0.120)	(0.141)	(0.199)	(0.260)	(0.234)	(0.291)	(0.174)	(0.124)	(0.123)	(0.116)	(0.116)	(0.129)	(0.172)	(0.144)	(0.130)	(0.108)	(0.143)	(0.145)
<b>Constant, <math>\alpha</math></b>	0.026	0.007	-0.013	-0.019	-0.014	-0.019	-0.024*	-0.016	0.027	0.058**	0.078**	0.173***	0.185***	0.033	0.032*	0.018	0.009	0	0.035**	-0.011	-0.034**	-0.031**	-0.034**	-0.035**
	(0.024)	(0.024)	(0.020)	(0.019)	(0.013)	(0.013)	(0.013)	(0.014)	(0.023)	(0.027)	(0.031)	(0.035)	(0.027)	(0.021)	(0.018)	(0.016)	(0.016)	(0.014)	(0.018)	(0.017)	(0.015)	(0.015)	(0.015)	(0.013)
<b>Observations</b>	144	173	185	183	177	165	148	132	110	92	83	72	71	75	89	85	92	85	78	59	66	62	62	62
<b>Adjusted R-squared</b>	0.481	0.402	0.458	0.458	0.129	0.120	0.212	0.236	0.109	0.151	0.147	-0.00637	-3.72e-05	0.412	0.381	0.431	0.429	0.430	0.239	0.349	0.312	0.376	0.370	0.377
<b>p-value (<math>\alpha = 0</math>)</b>	0.289	0.754	0.522	0.318	0.281	0.142	0.0742	0.243	0.247	0.0336	0.0141	6.19e-06	2.51e-09	0.122	0.0748	0.288	0.581	0.979	0.0464	0.537	0.0218	0.0396	0.0246	0.0118
<b>p-value (<math>\beta = 1</math>)</b>	0.250	0.603	0.441	0.496	0.00704	0.000544	0.00225	0.0285	0.0307	0.140	0.302	0.00412	4.99e-06	0.0667	0.0655	0.0952	0.0482	0.0683	0.0227	0.00854	0.00157	0.0464	0.390	0.519
<b>MZ F-test (<math>\alpha=0, \beta=1</math>)</b>	3.727	0.632	0.304	0.502	8.241	13.82	13.82	6.353	2.775	2.481	5.902	17.18	24.04	2.117	2.555	1.677	2.011	1.707	3.080	5.861	11.76	8.078	5.382	5.156
<b>p-value (<math>\alpha=0, \beta=1</math>)</b>	0.0264	0.533	0.738	0.606	0.000380	2.85e-06	3.18e-06	0.00233	0.0668	0.0894	0.00405	8.50e-07	1.19e-08	0.128	0.0835	0.193	0.140	0.188	0.0517	0.00484	4.47e-05	0.000782	0.00708	0.00858
<b>HP T-test (<math>\gamma=0</math>)</b>	2.467	1.069	0.118	-0.469	-3.622	-4.532	-4.339	-3.178	-0.974	1.314	3.424	4.030	2.519	4.494	0.561	-0.141	-0.768	-0.843	-0.223	-2.117	-3.194	-2.686	-2.925	-3.027
<b>p-value (<math>\gamma=0</math>)</b>	0.0148	0.286	0.906	0.640	0.000382	1.12e-05	2.65e-05	0.00185	0.332	0.192	0.000967	0.000139	0.0141	0.623	0.576	0.888	0.445	0.402	0.824	0.0385	0.00216	0.00930	0.00484	0.00362

	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to
	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
<b>Exports</b>																								
<b>MONA Forecast, <math>\beta</math></b>	0.998***	1.026***	0.886***	0.986***	0.762***	0.874***	0.716***	0.843***	0.670***	0.733***	0.825***	0.745***	0.727***	0.919***	0.909***	0.931***	0.933***	0.902***	0.825***	0.898***	0.733***	0.898***	0.900***	0.954***
	(0.329)	(0.302)	(0.327)	(0.369)	(0.192)	(0.200)	(0.142)	(0.138)	(0.184)	(0.202)	(0.116)	(0.142)	(0.141)	(0.059)	(0.064)	(0.093)	(0.097)	(0.106)	(0.152)	(0.147)	(0.138)	(0.225)	(0.278)	(0.293)
<b>Constant, <math>\alpha</math></b>	0.068**	0.056**	0.047*	0.028	0.012	0.002	0.009	0.018	0.056**	0.075***	0.082***	0.120***	0.123***	0.053***	0.068***	0.063***	0.043**	0.037**	0.052*	0.009	-0.004	0.007	0.007	-0.006
	(0.028)	(0.027)	(0.026)	(0.030)	(0.017)	(0.017)	(0.016)	(0.017)	(0.024)	(0.027)	(0.023)	(0.024)	(0.028)	(0.016)	(0.017)	(0.018)	(0.018)	(0.017)	(0.028)	(0.028)	(0.021)	(0.023)	(0.024)	(0.015)
<b>Observations</b>	143	172	184	182	177	165	147	131	109	91	82	72	71	75	89	85	92	85	78	59	66	62	62	62
<b>Adjusted R-squared</b>	0.202	0.185	0.143	0.123	0.0927	0.106	0.0787	0.0974	0.113	0.150	0.419	0.439	0.419	0.697	0.542	0.472	0.444	0.420	0.176	0.255	0.154	0.202	0.165	0.333
<b>p-value (<math>\alpha = 0</math>)</b>	0.0157	0.0401	0.0772	0.346	0.482	0.910	0.573	0.276	0.0186	0.00692	0.000483	5.13e-06	4.89e-05	0.00149	0.000198	0.000941	0.0153	0.0321	0.0636	0.744	0.848	0.771	0.761	0.693
<b>p-value (<math>\beta = 1</math>)</b>	0.994	0.933	0.727	0.969	0.217	0.528	0.0476	0.257	0.0751	0.190	0.135	0.0777	0.0566	0.171	0.156	0.460	0.495	0.357	0.254	0.490	0.0582	0.654	0.720	0.875
<b>MZ F-test (<math>\alpha=0, \beta=1</math>)</b>	7.411	5.753	3.434	1.195	0.774	0.217	1.997	1.051	2.861	5.262	7.626	16.25	12.34	5.474	7.636	6.041	3.061	2.572	1.809	0.247	1.860	0.143	0.0925	0.133
<b>p-value (<math>\alpha=0, \beta=1</math>)</b>	0.000870	0.00382	0.0344	0.305	0.463	0.805	0.139	0.352	0.0616	0.00692	0.000931	1.59e-06	2.62e-05	0.00609	0.000880	0.00355	0.0517	0.0825	0.171	0.782	0.164	0.867	0.912	0.876
<b>HP T-test (<math>\gamma=0</math>)</b>	3.517	3.120	2.174	1.409	-0.426	-0.453	-0.604	0.424	1.810	3.156	3.755	5.219	4.393	3.120	3.487	3.274	2.199	1.729	1.373	0.0142	-0.700	0.0438	0.0517	-0.500
<b>p-value (<math>\gamma=0</math>)</b>	0.000587	0.00212	0.0310	0.160	0.670	0.651	0.547	0.672	0.0731	0.00217	0.000325	1.71e-06	3.88e-05	0.00258	0.000765	0.00154	0.0304	0.0874	0.174	0.989	0.487	0.965	0.959	0.619

	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to	to
	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
<b>Real Exchange Rates</b>																								
<b>MONA Forecast, <math>\beta</math></b>	0.842***	0.878***	0.875***	0.966***	1.050***	1.167***	1.049***	1.034***	1.392**	0.881**	0.786***	0.623***	0.525***	0.843***	0.807***	0.829***	0.825***	0.746***	0.708***	0.760***	0.682***	0.724***	0.338**	0.333**
	(0.080)	(0.073)	(0.134)	(0.185)	(0.193)	(0.156)	(0.202)	(0.193)	(0.641)	(0.164)	(0.225)	(0.135)	(0.123)	(0.095)	(0.088)	(0.084)	(0.073)	(0.085)	(0.077)	(0.065)	(0.068)	(0.149)	(0.145)	
<b>Constant, <math>\alpha</math></b>	-0.033	-0.027	-0.022	-0.002	0.015	0.013	0.022	0.008	-0.003	-0.006	-0.023	-0.060***	-0.079***	-0.033***	-0.020**	-0.013	-0.002	0.011**	0.006	0.024**	0.039***	0.032***	0.036***	0.034**
	(0.021)	(0.019)	(0.018)	(0.018)	(0.021)	(0.017)	(0.020)	(0.018)	(0.015)	(0.011)	(0.014)	(0.011)	(0.013)	(0.010)	(0.008)	(0.008)	(0.007)	(0.005)	(0.006)	(0.008)	(0.009)	(0.011)	(0.013)	(0.013)
<b>Observations</b>	145	175	186	185	177	166	146	129	107	90	79	71	71	76	90	86	92	84	75	55	61	55	54	53
<b>Adjusted R-squared</b>	0.833	0.830	0.557	0.490	0.443	0.574	0.312	0.386	0.217	0.268	0.178	0.377	0.300	0.581	0.559	0.616	0.639	0.707	0.634	0.696	0.535	0.460	0.302	0.300
<b>p-value (<math>\alpha = 0</math>)</b>	0.117	0.148	0.236	0.907	0.484	0.472	0.268	0.668	0.833	0.560	0.108	1.47e-06	2.85e-08	0.00170	0.0147	0.118	0.730	0.0349	0.322	0.00521	0.000119	0.00418	0.00604	0.0102
<b>p-value (<math>\beta = 1</math>)</b>	0.0483	0.0987	0.352	0.854	0.796	0.287	0.810	0.862	0.543	0.470	0.345	0.00684	0.000240	0.103	0.0310	0.0449	0.0189	0.000766	0.00102	0.00308	9.28e-06	0.000169	4.68e-05	2.82e-05
<b>MZ F-test (<math>\alpha=0, \beta=1</math>)</b>	4.123	3.049	1.499	0.0214	0.349	1.100	1.138	0.508	0.236	0.409	1.352	14.78	20.79	5.336	4.103	2.447	3.037	13.44	12.89	12.85	16.51	9.227	12.64	12.58
<b>p-value (<math>\alpha=0, \beta=1</math>)</b>	0.018	0.050	0.226	0.979	0.706	0.335	0.323	0.603	0.790	0.666	0.265	0.000	0.000	0.007	0.020	0.093	0.053	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<b>HP T-test (<math>\gamma=0</math>)</b>	-1.968	-1.883	-1.629	-0.135	0.790	1.128	1.339	0.540	-0.331	-0.270	-1.158	-4.94	-4.936	-2.864	-1.918	-0.914	0.189	2.350	2.395	3.004	3.436	2.581	0.324	0.155
<b>p-value (<math>\gamma=0</math>)</b>	0.051	0.061	0.105	0.893	0.430	0.261	0.183	0.590	0.741	0.788	0.250	0.000	0.000	0.005	0.058	0.363	0.851	0.021	0.019	0.004	0.001	0.013	0.747	0.878

## 1.10.3 Correction

We corrected 20194 observations with incorrect signs entered in MONA. To minimize measurement problems, we cleaned our dataset by eliminating negative signs in trade data. These errors are also not necessarily consistent within any given program, but it appears in new MONA with higher frequency than old MONA.

	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec
<b>IMPORTS</b>												
MONA Forecast, $\beta$	1.39***	0.27***	0.76	1.19	0.74	1.05	1.03	0.32***	0.15***	1.28	0.73	0.5**
p-value ( $\beta = 1$ )	0.01	0.00	0.22	0.12	0.49	0.77	0.80	0.01	0.00	0.21	0.18	0.01
Constant, ( $\alpha$ )	-0.04*	0.07*	-0.02	-0.01	0.05	0.02	-0.02	-0.01	0.09**	-0.04	0.02	0.03
p-value ( $\alpha = 0$ )	0.08	0.05	0.32	0.56	0.11	0.25	0.13	0.91	0.05	0.41	0.46	0.43
Observations	69	27	60	55	48	67	61	32	37	22	32	68
Adj. R-sq	0.70	0.04	0.27	0.52	0.17	0.37	0.51	-0.01	-0.02	0.55	0.34	0.12
MZ p-value ( $\alpha=0, \beta=1$ )	0.03**	0.00***	0.27	0.27	0.28	0.28	0.31	0.01***	0.00***	0.43	0.40	0.00***
HP p-value ( $\gamma=0$ )	0.97	0.90	0.15	0.96	0.39	0.12	0.21	0.09*	0.53	0.93	0.63	0.54
<b>EXPORTS</b>												
MONA Forecast, $\beta$	1.67	0.84	0.85	0.90	0.47**	0.76	1.10	0.72	0.55	1.01	0.60**	0.97
p-value ( $\beta = 1$ )	0.35	0.49	0.36	0.57	0.05	0.43	0.77	0.23	0.45	0.96	0.05	0.69
Constant, ( $\alpha$ )	-0.04	0.05	0.05**	0.01	0.07**	0.12**	0.04	0.01	0.09	0.01	0.03*	0.01
p-value ( $\alpha = 0$ )	0.47	0.17	0.04	0.72	0.04	0.02	0.11	0.79	0.14	0.87	0.07	0.46
Observations	69	27	60	55	48	65	61	32	37	22	32	68
Adj. R-sq	0.31	0.21	0.29	0.36	0.07	0.02	0.31	0.25	0.02	0.41	0.27	0.57
MZ p-value ( $\alpha=0, \beta=1$ )	0.62	0.39	0.1*	0.85	0.07*	0.05**	0.14	0.43	0.28	0.96	0.08*	0.73
HP p-value ( $\gamma=0$ )	0.46	0.27	0.07*	0.99	0.47	0.02**	0.09*	0.43	0.11	0.79	0.70	0.59
<b>REAL EXCHANGE RATE DEPRECIATION</b>												
MONA Forecast, $\beta$	1.00	0.43***	0.65***	1.27**	0.66***	0.93	0.94	1.00	0.58***	0.68***	0.85	0.85*
p-value ( $\beta = 1$ )	0.99	0.01	0.01	0.05	0.00	0.17	0.64	0.99	0.00	0.00	0.12	0.06
Constant, ( $\alpha$ )	0.04	0.01	0.04	-0.04	-0.06***	-0.02	-0.02	0.02	-0.03*	0.03	-0.01	0
p-value ( $\alpha = 0$ )	0.21	0.62	0.27	0.34	0.01	0.28	0.39	0.39	0.07	0.52	0.44	0.81
Observations	68	27	56	53	48	63	59	29	35	23	33	71
Adj. R-sq	0.27	0.17	0.49	0.73	0.55	0.93	0.58	0.49	0.49	0.86	0.61	0.64
MZ p-value ( $\alpha=0, \beta=1$ )	0.19	0.00	0.00	0.11	0.00	0.29	0.08	0.66	0.01	0.00	0.25	0.17
HP p-value ( $\gamma=0$ )	0.17	0.12	0.95	0.41	0.01	0.20	0.24	0.41	0.13	0.22	0.63	0.95

Robust p values. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 1.6: Data Entered with Wrong Sign

Total Revisions in Each Horizon:	t-3	t-2	t-1	t	t+1	t+2	t+3	t+4
20194	2495	2526	2536	2534	2537	2539	2529	2498

Table 1.7: MONA Temporal Error

#### 1.10.4 Correction

Count	Arrangement	Country Name	Prog. Year	Mnemonic	Review Type	Correction
1	7	Estonia	1993	All	Last	corrected data using IMF archives
2	15	El Salvador	1993	All	Last	corrected data using IMF archives
3	256	Indonesia	1998	PCPIC	All	corrected data using IMF archives
4	275	Indonesia	1999	ENDA, PCPIC	All	corrected data using IMF archives
5	552	Dominican Rep.	2005	All	All	corrected data using IMF archives
6	571	Modagscar	2006	All	Last	corrected data using IMF archives
7	579	Gabon	2007	All	R1-Last	corrected data using IMF archives

Table 1.8: Zeros Identify Missing Values Part I

## 1.10.5 Correction

Count	Arrangement	Country Name	Prog. Year	Mnemonic	Review Type
1	11	Honduras	1992	BMS_O	R0
2	16	Kyrgyz Republic	1993	BXS_O	R0
3	16	Kyrgyz Republic	1993	BMS_O	R0
4	25	Poland	1993	ENDA	R0
5	108	Kazakhstan	1994	BXS_O	R0
6	108	Kazakhstan	1994	ENDA	R0
7	114	Mozambique	1994	ENDA	R0
8	153	Kazakhstan	1995	ENDA	R0
9	153	Kazakhstan	1995	BXS_O	R0
10	159	Azerbaijan	1996	BMS_O	R0
11	163	Moldova	1996	BMS_O	R0
12	250	Mauritania	1997	PCPIC	R0
13	264	Central African Republic	1998	PCPIC	R0
14	281	Argentina	1998	PCPIC	R0
15	337	Indonesia	2000	PCPIC	R0
16	507	Albania	2002	PCPIE	R6
17	560	Benin	2005	PCPIE	R0
18	564	Iraq	2005	BXS	R0
19	572	Haiti	2006	ENDA	R0
20	590	Liberia	2008	PCPIE	R7
21	610	Sao Tome and Principe	2009	ENDA	R0
22	632	Malawi	2010	ENDA	R0
23	642	Tanzania	2010	ENDA	R0
24	649	Armenia	2010	ENDA	R0
25	651	Haiti	2010	ENDA	R0
26	656	Senegal	2010	PCPIE	R6-R7
27	658	North Macedonia	2011	ENDA	R0
28	678	Burundi	2012	ENDA	R0
29	679	Guinea	2012	ENDA	R0
30	681	Niger	2012	PCPIE	R2-R7
31	683	Georgia	2012	ENDA	R0
32	685	Gambia, The	2012	ENDA	R0
33	685	Gambia, The	2012	PCPIE	R0
34	686	Central African Republic	2012	ENDA	R0
35	687	Tanzania	2012	ENDA	R0
36	688	Sao Tome and Principe	2012	ENDA	R0
37	689	Malawi	2012	ENDA	R0
38	691	Morocco	2012	ENDA	R0
39	692	Bosnia and Herzegovina	2012	ENDA	R0
40	693	Liberia	2012	NX	R0
41	697	Jamaica	2013	ENDA	R0
42	698	Cyprus	2013	ENDA	R0
43	699	Tunisia	2013	ENDA	R0
44	701	Mozambique	2013	ENDA	R0
45	702	Uganda	2013	ENDA	R0
46	703	Pakistan	2013	ENDA	R0
47	705	Sierra Leone	2013	ENDA	R0
48	709	Albania	2014	PCPIE	R2-R3
49	710	Armenia	2014	ENDA	R0
50	712	Seychelles	2014	ENDA	R0
51	714	Tanzania	2014	ENDA	R0
52	714	Tanzania	2014	PCPIE	R0
53	717	Chad	2014	ENDA	R0

Table 1.9: Zeros Identify Missing Values Part II

Count	Arrangement	Country Name	Prog. Year	Mnemonic	Review Type
54	722	Kenya	2015	ENDA	R0
55	723	Serbia, Republic of	2015	ENDA	R0
56	724	Ukraine	2015	NX	R0
57	724	Ukraine	2015	NM	R0
58	724	Ukraine	2015	ENDA	R0
59	726	Kyrgyz Republic	2015	ENDA	R0
60	729	Senegal	2015	ENDA	R0
61	729	Senegal	2015	PCPIE	R3
62	730	Guinea-Bissau	2015	ENDA	R0
63	731	Sao Tome and Principe	2015	ENDA	R0
64	731	Sao Tome and Principe	2015	PCPIE	R0
65	735	Tunisia	2016	ENDA	R0
66	738	Sri Lanka	2016	ENDA	R0
67	739	Rwanda	2016	ENDA	R0
68	739	Rwanda	2016	PCPIE	R0
69	741	Iraq	2016	PCPIE	R0
70	742	Madagascar	2016	ENDA	R0
71	747	Bosnia and Herzegovina	2016	ENDA	R0
72	749	Moldova	2016	ENDA	R0
73	750	Cote d'Ivoire	2016	BXGS	R0
74	750	Cote d'Ivoire	2016	ENDA	R0
75	750	Cote d'Ivoire	2016	BMGS	R0
76	751	Niger	2017	ENDA	R0
77	753	Poland	2017	ENDA	R0
78	754	Benin	2017	ENDA	R0
79	755	Georgia	2017	ENDA	R0
80	756	Togo	2017	ENDA	R0
81	757	Mongolia	2017	ENDA	R0
82	757	Mongolia	2017	NX	R0
83	757	Mongolia	2017	NM	R0
84	758	Sierra Leone	2017	ENDA	R0
85	760	Cameroon	2017	ENDA	R0
86	761	Chad	2017	ENDA	R0
87	764	Mauritania	2018	ENDA	R0
88	765	Guinea	2017	ENDA	R0
89	766	Seychelles	2017	ENDA	R0
90	767	Burkina Faso	2018	ENDA	R0
91	768	Malawi	2018	ENDA	R0
92	769	Colombia	2018	ENDA	R0
93	771	Serbia, Republic of	2018	ENDA	R0
94	772	Barbados	2018	ENDA	R0
95	778	Armenia	2019	ENDA	R0
96	780	Rwanda	2019	ENDA	R0
97	781	Pakistan	2019	ENDA	R0
98	782	Honduras	2019	ENDA	R0
99	783	Cabo Verde	2019	ENDA	R0
100	784	Congo, Rep.	2019	ENDA	R0
101	785	Mali	2019	ENDA	R0
102	786	Sao Tome and Principe	2019	ENDA	R0
103	787	Mexico	2019	ENDA	R0
104	789	Liberia	2019	ENDA	R0
105	789	Liberia	2019	NX	R0
106	790	Central African Republic	2019	ENDA	R0

Table 1.10: Typos and Spelling Mistakes

**1.10.6 Correction**

Count	Arrangement	Country Name	Prog. Year	Mnemonic	Review Type	Correction
1	15	El Salvador	1993	programyear	R2	programyear corrected
2	18	Latvia	1993	programyear	R1-Last	programyear corrected
3	75	Turkey	1994	PCPIC	R0	data corrected using IMF archives
4	117	Albania	1994	countryname	All	wrong countryname corrected
5	117	Albania	1994	countryncode	All	wrong countryname corrected
6	132	Sierra Leone	1995	programyear	R0	programyear corrected
7	136	Haiti	1995	PCPIC	R0	data corrected using IMF archives
8	143	Pakistan	1996	programyear	R0-R1	programyear corrected
9	160	Russian Federation	1995	PCPIC	R0	data corrected using IMF archives
10	205	Vietnam	1996	boarddocno	R1	board document typo corrected
11	207	Ethiopia	1997	reviewtype	All	reviewtype lables corrected
12	212	Kyrgyz Republic	1997	ENDA	Last	data corrected using IMF archives
13	230	Burkina Faso	1996	programyear	R0	data corrected using IMF archives
14	250	Mauritania	1997	PCPIC	R0	data corrected using IMF archives
15	274	Ukraine	1998	PCPIC	R0	data corrected using IMF archives
16	274	Ukraine	1998	programyear	R5-R6	programyear corrected
17	402	Moldova	2000	BXS_O	R0	data corrected using IMF archives
18	510	Argentina	2003	PCPIE	R0	data corrected using IMF archives
19	521	Ghana	2003	PCPIE	R0	data corrected using IMF archives
20	527	Nicaragua	2002	programyear	R10	programyear corrected
21	535	Uruguay	2002	PCPIE	R0	data corrected using IMF archives
22	539	Domnican Rep.	2003	PCPIE	R0	data corrected using IMF archives
23	545	Peru	2004	PCPIE	R0	data corrected using IMF archives
24	560	Benin	2005	PCPIE	R0	data corrected using IMF archives
25	560	Benin	2005	boarddocno	R0	board document typo corrected
26	628	Kyrgyz Republic	2008	reviewtype	All	reviewtype lables corrected
27	681	Niger	2012	programyear	R8	programyear corrected
28	724	Ukraine	2015	reviewtype	All	reviewtype lables corrected
29	734	Kenya	2016	reviewtype	All	reviewtype lables corrected
30	764	Mauritania	2017	programyear	R0-R4	programyear corrected
31	All	All	All	initialenddate	All	spelling error corrected

Table 1.11: Wrong Line Item Entered

**1.10.7 Correction**

Count	Arrangement	Country Name	Prog Year	Mnemonic	Review Type	Correction
1	560	Benin	2005	PCPIE	R0	Corrected using IMF archives

Table 1.12: Wrong Currency Unit Entered

**1.10.8 Correction**

Count	Arrangement	Country Name	Prog. Year	Mnemonic	Review Type	Correction
1	70	Poland	1994	ENDA	R0	Divided by 1000
2	75	Turkey	1994	ENDA	Last	Divided by 1000
3	84	Algeria	1995	ENDA	R0	Divided by 1000
4	164	Russian Federation	1996	ENDA	R0	Divided by 1000
5	199	Croatia	1997	ENDA	Last	Divided by 1000
6	398	Bulgaria	2002	ENDA	R0	Divided by 1000
7	517	Croatia	2003	indicator currency of PCPIE	R0-R1	Changed to "Index Number"
8	580	Mozambique	2007	PCPIE	R0	Divided by 1000

Table 1.13: Rates vs. Levels Inconsistencies

**1.10.9 Correction**

Count	Arrangement	Country Name	Prog. Year	Mnemonic	Review Type	Correction
1	510	Argentina	2003	PCPIE	R0	corrected to rates
2	521	Ghana	2003	PCPIE	R0	corrected to rates
3	527	Nicaragua	2002	PCPIE	R0	corrected to rates
4	535	Uruguay	2002	PCPIE	R0	corrected to rates
5	539	Dominican Republic	2003	PCPIE	R0	corrected to rates
6	545	Peru	2004	PCPIE	R0	corrected to rates
7	556	Turkey	2005	PCPIE	All	corrected to rates
8	560	Benin	2005	PCPIE	R0	corrected to rates
9	562	North Macedonia, Rep	2005	PCPIE	R0-R1	corrected to rates
10	564	Iraq	2005	PCPIE	R0	corrected to rates
11	566	Grenada	2006	PCPIE	R0	corrected to rates
12	572	Haiti	2006	PCPIE	R0	corrected to rates
13	580	Mozambique	2007	PCPIE	R0	corrected to rates
14	588	Iraq	2007	PCPIE	R0	corrected to rates
15	591	Honduras	2008	PCPIE	R0	corrected to rates
16	685	Gambia, The	2012	PCPIE	R0	corrected to rates
17	709	Albania	2014	PCPIE	R0	corrected to rates
18	714	Tanzania	2014	PCPIE	R0	corrected to rates
19	718	Yemen	2014	PCPIE	R0	corrected to rates
20	731	Sao Tome and Principe	2015	PCPIE	R0	corrected to rates
21	739	Rwanda	2016	PCPIE	R0	corrected to rates
22	741	Iraq	2016	PCPIE	R0	corrected to rates

Table 1.14: Unit Inconsistency: Base Years

**1.10.10 Correction**

Count	Arrangement	Country Name	Prog Year	Mnemonic	Review Type	Correction
1	16	Kyrgyz Republic	1993	All	All	Unresolved, dropped
2	108	Kazakhstan	1994	All	All	Unresolved, dropped
3	532	Sierra Leone	2001	PCPIE	Last	Corrected to rates
4	533	Tanzania	2000	PCPIE	Last	Corrected to rates
5	538	Burundi	2004	PCPIE	Last	Corrected to rates
6	547	Zambia	2004	PCPIE	Last	Corrected to rates
7	549	Bulgaria	2004	PCPIE	Last	Corrected to rates
8	554	Kyrgyz Republic	2005	PCPIE	Last	Corrected to rates
9	561	Sao Tome and Principe	2005	PCPIE	Last	Corrected to rates
10	565	Albania	2006	PCPIE	Last	Corrected to rates
11	567	Moldova	2006	PCPIE	Last	Corrected to rates
12	568	Paraguay	2006	PCPIE	Last	Corrected to rates
13	596	Burundi	2008	PCPIE	Last	Corrected to rates
14	617	Romania	2009	PCPIE	Last	Corrected to rates
15	619	Ghana	2009	PCPIE	Last	Corrected to rates
16	620	Sri Lanka	2009	PCPIE	Last	Corrected to rates
17	623	Angola	2010	PCPIE	Last	Corrected to rates
18	625	Congo, Democ. Rep.	2010	PCPIE	Last	Corrected to rates
19	635	El Salvador	2010	PCPIE	Last	Corrected to rates
20	678	Burundi	2012	PCPIE	Last	Corrected to rates
21	697	Jamaica	2013	PCPIE	Last	Corrected to rates
22	704	Romania	2013	PCPIE	Last	Corrected to rates
23	712	Seychelles	2014	PCPIE	Last	Corrected to rates

Table 1.15: Missing Data Filled

## 1.10.11 Correction

Count	Arrangement	Country Name	Prog. Year	Mnemonic	Review Type	Correction
1	95	Ukraine	1995	ENDA	R0	entered data using IMF archives
2	153	Kazakhstan	1995	ENDA	R0	entered data using IMF archives
3	250	Mauritania	1997	RENDA	R0	entered data using IMF archives
4	261	Mozambique	1997	ENDA	R0	entered data using IMF archives
5	273	Ukraine	1997	ENDA	R0	entered data using IMF archives
6	377	Pakistan	2001	ENDA	R0	entered data using IMF archives
7	419	Congo, Dem. Rep.	2002	ENDA	R0	entered data using IMF archives
8	421	Rwanda	2001	ENDA	R0	entered data using IMF archives
9	502	Tajikistan	2002	NM	R0	entered data using IMF archives
10	502	Tajikistan	2002	NX	R0	entered data using IMF archives
11	506	Bosnia and Herzegovina	2002	NM	R0	entered data using IMF archives
12	506	Bosnia and Herzegovina	2002	NX	R0	entered data using IMF archives
13	537	Serbia	2002	NX	R0	entered data using IMF archives
14	537	Serbia	2002	NM	R0	entered data using IMF archives
15	572	Haiti	2006	ENDA	R0	entered data using IMF archives
16	610	Sao Tome and Principe	2009	ENDA	R0	entered data using IMF archives
17	624	Maldives	2009	ENDA	R0	entered data using IMF archives
18	648	Rwanda	2010	NM	R0	entered data using IMF archives
19	651	Haiti	2010	ENDA	R0	entered data using IMF archives
20	678	Burundi	2012	ENDA	R0	entered data using IMF archives
21	685	Gambia, The	2012	RENDA	R0	entered data using IMF archives
22	691	Morocco	2012	ENDA	R0	entered data using IMF archives
23	698	Cyprus	2013	ENDA	R0	entered data using IMF archives
24	699	Tunisia	2013	ENDA	R0	entered data using IMF archives
25	702	Uganda	2013	ENDA	R0	entered data using IMF archives
26	703	Pakistan	2013	ENDA	R0	entered data using IMF archives
27	705	Sierra Leone	2013	ENDA	R0	entered data using IMF archives
28	711	Ukraine	2014	NX	R0	entered data using IMF archives
29	711	Ukraine	2014	NM	R0	entered data using IMF archives
30	712	Seychelles	2014	ENDA	R0	entered data using IMF archives
31	713	Grenada	2014	NX	R0	entered data using IMF archives
32	714	Tanzania	2014	ENDA	R0	entered data using IMF archives
33	726	Kyrgyz Republic	2015	ENDA	R0	entered data using IMF archives
34	731	Sao Tome and Principe	2015	ENDA	R0	entered data using IMF archives
35	735	Tunisia	2016	ENDA	R0	entered data using IMF archives
36	739	Rwanda	2016	ENDA	R0	entered data using IMF archives
37	742	Madagascar	2016	ENDA	R0	entered data using IMF archives
38	749	Moldova	2016	ENDA	R0	entered data using IMF archives
39	750	Cote d'Ivoire	2016	RENDA	R0	entered data using IMF archives
40	756	Togo	2017	RENDA	R0	entered data using IMF archives
41	761	Chad	2017	RENDA	R0	entered data using IMF archives
42	764	Mauritania	2017	ENDA	R0	entered data using IMF archives
43	765	Guinea	2017	ENDA	R0	entered data using IMF archives
44	766	Seychelles	2017	ENDA	R0	entered data using IMF archives
45	768	Malawi	2018	ENDA	R0	entered data using IMF archives
46	769	Colombia	2018	ENDA	R0	entered data using IMF archives
47	770	Argentina	2018	ENDA	R0	entered data using IMF archives
48	772	Barbados	2018	ENDA	R0	entered data using IMF archives
49	778	Armenia	2019	ENDA	R0	entered data using IMF archives
50	781	Pakistan	2019	ENDA	R0	entered data using IMF archives
51	782	Honduras	2019	ENDA	R0	entered data using IMF archives
52	786	Sao Tome and Principe	2019	ENDA	R0	entered data using IMF archives
53	787	Mexico	2019	ENDA	R0	entered data using IMF archives
54	789	Liberia	2019	RENDA	R0	entered data using IMF archives

Table 1.16: Corrected Outliers

## 1.10.12 Correction

Count	Arrangement	Country Name	Prog. Year	Mnemonic	Correction
1	1	Albania	1993	NM	corrected data using IMF archives
2	14	Jamaica	1992	RENDA	corrected data using IMF archives
3	14	Jamaica	1992	ENDA	corrected data using IMF archives
4	16	Kyrgyz Republic	1993	PCPIC	Dropped due to currency reform.
5	17	Lao PDR	1993	NX	corrected data using IMF archives
6	19	Lithuania	1993	RENDA	corrected data using IMF archives
7	19	Lithuania	1993	ENDA	corrected data using IMF archives
8	29	Kenya	1993	NX	corrected data using IMF archives
9	29	Kenya	1993	NM	corrected data using IMF archives
10	75	Turkey	1994	PCPIC	Typos and spelling mistakes. Fixed
11	82	Moldova	1994	RENDA	corrected data using IMF archives
12	82	Moldova	1994	ENDA	corrected data using IMF archives
13	93	Cambodia	1994	NX	corrected data using IMF archives
14	108	Kazakhstan	1994	PCPIC	Dropped due to differences in base years
15	114	Mozambique	1994	RENDA	corrected data using IMF archives
16	114	Mozambique	1994	ENDA	corrected data using IMF archives
17	118	Congo, Republic Of	1994	PCPIC	Suspicious WEO data, MONA last review data used
18	129	Papua New Guinea	1995	RENDA	corrected data using IMF archives
19	129	Papua New Guinea	1995	ENDA	corrected data using IMF archives
20	132	Sierra Leone	1995	NX	corrected data using IMF archives
21	134	Belarus	1995	RENDA	corrected data using IMF archives
22	134	Belarus	1995	ENDA	corrected data using IMF archives
23	134	Belarus	1995	NX	corrected data using IMF archives
24	134	Belarus	1995	NM	corrected data using IMF archives
25	136	Haiti	1995	PCPIC	Error 5: Typos and spelling mistakes. Fixed
26	160	Russian Federation	1995	PCPIC	Error 5: Typos and spelling mistakes. Fixed
27	170	Cambodia	1995	NX	corrected data using IMF archives
28	202	Bulgaria	1997	RENDA	corrected data using IMF archives
29	202	Bulgaria	1997	ENDA	corrected data using IMF archives
30	218	Haiti	1997	NX	corrected data using IMF archives
31	228	Sierra Leone	1997	NM	corrected data using IMF archives
32	228	Sierra Leone	1997	NX	corrected data using IMF archives
33	242	Bosnia and Herzegovina	1998	NM	corrected data using IMF archives
34	256	Indonesia	1998	PCPIC	Error 1: Temporal issue. Fixed
35	275	Indonesia	1999	PCPIC	Error 1: Temporal issue. Fixed
36	307	Mozambique	1999	NM	corrected data using IMF archives
37	316	Albania	1999	NM	corrected data using IMF archives
38	316	Albania	1999	NX	corrected data using IMF archives
39	347	Kenya	2000	NM	corrected data using IMF archives
40	508	Argentina	2003	PCPIE	No data available from EBS document
41	510	Argentina	2003	PCPIE	Error 5: Typos and spelling mistakes. Fixed
42	521	Ghana	2003	PCPIE	Error 5: Typos and spelling mistakes. Fixed
43	532	Sierra Leone	2001	NM	corrected data using IMF archives
44	535	Uruguay	2002	RENDA	corrected data using IMF archives
45	535	Uruguay	2002	ENDA	corrected data using IMF archives
46	535	Uruguay	2002	PCPIE	Error 5: Typos and spelling mistakes. Fixed
47	538	Burundi	2004	NX	corrected data using IMF archives
48	539	Dominican Republic	2003	PCPIE	Error 5: Typos and spelling mistakes. Fixed
49	547	Zambia	2004	NX	corrected data using IMF archives
50	552	Dominican Republic	2005	PCPIE	Error 1: Temporal issue. Fixed
51	574	Mauritania	2006	NM	corrected data using IMF archives
52	579	Gabon	2007	NM	corrected data using IMF archives
53	603	Iceland	2008	NX	corrected data using IMF archives
54	625	Congo, Dem. Rep.	2010	NX	corrected data using IMF archives
55	645	Burkina Faso	2010	NX	corrected data using IMF archives
56	733	Mozambique	2016	ENDA	corrected data using IMF archives
57	765	Guinea	2017	NX	corrected data using IMF archives
58	791	Ethiopia	2019	NM	corrected data using IMF archives

## Chapter 2

# Public Debt and Real GDP: Revisiting the Impact

### 2.1 Introduction

Understanding the effects of debt to GDP ratio on macroeconomic variables (i.e., growth, consumption, etc.) is a crucial policy question (see, for example, [Domar \(1944\)](#); [Barro \(1980\)](#); [Bernheim \(1987\)](#)).<sup>1</sup> An increase in debt to GDP ratio can stem from a rise in government expenditures, an increase in public investment, a reduction in GDP, a decrease in tax revenues or other fiscal changes. While understanding the impact of the resulting increase in debt to GDP ratio on real GDP is essential to public debt sustainability assessments, there is little consensus on its direction and magnitude.

Public debt sustainability analyses are considered a key element in the IMF's work on member countries. They are the basis to IMF lending decisions – whether the primary balance needed to stabilize debt to GDP ratio under both the baseline and realistic shock scenarios is

---

<sup>1</sup>[Domar \(1944\)](#) studies debt and the issue of deficit financing and demonstrates that debt burden is primarily a problem of achieving a growing national income. The faster the income grows, the less debt burden the country holds. [Barro \(1980\)](#) studies the effect of public debt shocks at the federal level on taxation movements and finds that deficits can be considered economically efficient. [Bernheim \(1987\)](#) criticizes the Ricardian equivalence claiming that deficits merely postpone taxes as rational agents should be indifferent between either paying US\$1 today or paying US\$1 plus interest rate tomorrow. In particular, he finds that government deficits cause people to consume more.

economically and politically feasible, such that the level of debt to GDP ratio is consistent with an acceptably low rollover risk and with preserving potential growth at a satisfactory level. They also underpin the Fund's decisions in program contexts, particularly in helping to determine the timing and size of the financing, policy choices, and the member's capacity to repay. For those forward-looking assessments, it is crucial to understand how changes in public debt to GDP ratio are likely to impact real GDP over the short to medium-term.<sup>2</sup>

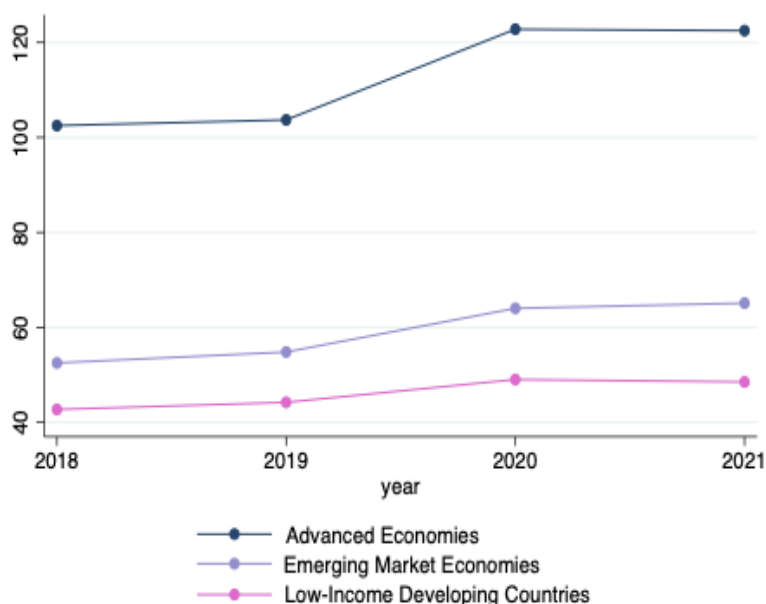
Understanding the impact of increases in public debt to GDP ratio on output is even more critical in the context of the COVID-19 pandemic. The pandemic led to a significant contraction in the world economy (Aizenman and Ito (2020)). Policy responses, including expansionary fiscal policies, have resulted in sharp increases in public debt to GDP ratio across the globe. The average public debt to GDP ratio across all countries is projected to reach 98.8 percent in 2021 compared to 83.7 percent in 2019, before the pandemic. Advanced Economies (AEs) average public debt to GDP ratio is projected to increase from 103.7 to 122.5 over the same period, while smaller increases are projected for Emerging Markets (EM) and Low-Income Developing Economies (LIDCs), from 54.8 to 65.1 and from 44.2 to 48.5 percentage points, respectively (Figure 1). While policy responses that resulted in an increase of public debt to GDP ratio may be effective in the short run and help boost growth, increased debt to GDP ratios may as well partly (or fully) negate the effects of the fiscal stimulus in the medium-term, which could slow down the recovery from the pandemic.

This paper revisits the relationship between public debt to GDP ratio and growth. It is mainly related to the significant strand of literature that theoretically and empirically assesses the impact of an unanticipated increase in public debt to GDP ratio on real GDP. On the theoretical side, existing literature demonstrates that if the country already has a high level of debt, the existing debt to GDP ratio prevents the government from smoothly issuing additional debt because of fear of default (Woo and Kumar (2015)) as the "debt overhang hypotheses" (a possible offsetting impact of external debt on real GDP growth beyond a certain level) illustrate. Despite knowing that issuing additional debt would generally benefit the economy, a high initial debt to GDP ratio hurts subsequent development (Myers (1977)). Overlapping generations models (Blanchard (1985); Diamond (1965); Modigliani (1961)) also illustrate that high debt to GDP levels negatively impact

---

<sup>2</sup>Public debt consists of all liabilities that require payment or payments of interest and/or principal by the debtor to the creditor at a date or dates in the future. This includes debt liabilities in the form of SDRs, currency and deposits, debt securities, loans, insurance, pensions and standardized guarantee schemes, and other accounts payable. Debt can be valued at current market, nominal, or face values.

Figure 2.1: Public Debt to GDP Ratio, 2018-21



Sources: World Economic Outlook, July 2021

growth, change expectations, or bring uncertainty (Cochrane (2011)), and change sovereign yield spreads (Codogno et al. (2003)) and real interest rate to affect lower private investment (Laubach (2009)). A surge in public debt uses a portion of national savings meant for future generations. The resulting reduction in savings pushes up the interest rate, which reduces incentives to invest—lower investment results in lower capital accumulation, which drags down economic growth. An increase in public debt to GDP ratio generally negatively affects long-run growth based on endogenous growth models (Barro (1990); Saint-Paul (1992)).

On the empirical side, zooming into the debt level rather than the debt to GDP ratio may give additional informative insights; Barro (1980) demonstrates that public debt shocks may impact output and unemployment. Still, the magnitude of the effect is milder than monetary policy shocks. Lo and Rogoff (2015) show that governments react to a rising public debt to GDP ratio by increasing the primary surplus or running smaller deficits. The seminal paper of Reinhart and Rogoff (2010) claims that the public debt to GDP ratio can drag down economic growth at high debt levels, roughly above 90 percent. Pattillo et al. (2011) use GMM to find a non-linear relationship between external debt to GDP ratio and growth. Using 93 developing countries, they find that doubling the external debt to GDP ratio reduces growth by a third to a half percentage

point. [Liaquat \(2019\)](#) analyzes domestic debt and output growth of 39 high-income countries using Panel VAR and finds a negative relationship based on 1980-2017 data. [Pescatori et al. \(2014\)](#) focus on the long-term relationship by looking at the impact of the debt to GDP ratio on future real GDP growth to get around the reverse causality and conclude that there is no simple threshold for debt to GDP ratios above which medium-term growth prospects are severely undermined. [Cecchetti et al. \(2011\)](#) use a similar growth model approach to conclude that the public debt to GDP ratio will drag down real GDP growth beyond a certain level. [Eberhardt and Presbitero \(2015\)](#) find a negative long-run relationship between public debt to GDP ratio and growth based on linear and nonlinear systems. [Panizza and Presbitero \(2013\)](#) use an IV approach to identify the public debt to GDP shocks and conclude that there is a negative correlation between public debt and real GDP growth for OECD countries, but that the link disappears when correcting for endogeneity.

Several other papers also zoom into different factors that may affect the impact of public debt to GDP ratio on real GDP. [Pescatori et al. \(2014\)](#) categorize the debt to GDP ratio trajectory based on the last period growth and previous 15 years of growth and find a slight negative impact on growth for the subsample with positive debt to GDP trajectories. [Henri \(2019\)](#) considers that there is little or no positive impact in the short- or medium-term growth of the HIPC Initiative using World Bank data of 1990-2015 and suggests that other factors such as improvement in governance or educational quality and trade agreements matter more. However, [Marcelino and Hakobyan \(2014\)](#) find a significantly positive response of real GDP growth following the HIPC Initiative, similar to [Hussain and Gunter \(2005\)](#) and [Pattillo et al. \(2011\)](#). [Clements et al. \(2003\)](#) also find that the reduction in external debt to GDP ratio may be associated with positive growth following the HIPC Initiative using GMM.

Although the literature on the relationship between the relative size of debt and GDP is ample, there is still significant uncertainty regarding the magnitude of the impact. One of the critical reasons for this uncertainty is the difficulty in identifying exogenous debt shocks, i.e., shocks that would be uncorrelated with contemporaneous macroeconomic shocks (see, for example, [Nakamura and Steinsson \(2018\)](#)).

This paper will revisit the debt to GDP ratio and growth relationship, with two main contributions to the literature. The first contribution is the construction of debt to GDP shocks using forecast errors on public debt to identify the causal impact of public debt to GDP shocks on

real GDP, following a similar approach as in [Furceri et al. \(2018\)](#) and [Blanchard and Perotti \(2002\)](#). Once public debt shocks are identified, we use the local projections approach of [Jordà \(2005\)](#) to trace out the short- and medium-run output responses for a panel of 178 countries over the period 1995-2020. The second contribution is the comprehensive subsample analysis, studying how the impact of public debt to GDP shock on GDP varies in different subsets of the data. The identified exogenous movements in debt to GDP ratio enable us to examine the role of several characteristics such as the initial level of debt to GDP ratio, debt trajectory, income level, and participation in the HIPC Initiative that can potentially shape this impact.

The paper’s main results can be summarized: an unanticipated increase in public debt to GDP ratio significantly reduces real GDP contemporaneously. This aggregate effect masks differences across countries. Based on subsample analysis, we find that the impulse response to an unanticipated increase in public debt to GDP ratio is even more damaging when the initial debt to GDP level is high and for countries with a positive debt trajectory over the five preceding years. However, the effect is positive for low-income countries and countries with debt relief from the Highly Indebted Poor Countries (HIPC) Initiative.

The rest of the paper is organized as follows. Section [2.2](#) presents the data and methodology. Section [2.3](#) discusses the main results. Section [2.4](#) presents some robustness analyses of our results. Section [2.5](#) concludes by summarizing the main findings and policy implications. More details on the dataset and other results are shown in Annexes [2.6-2.8](#).

## 2.2 Data And Empirical Methodology

### 2.2.1 Data

The analysis mainly uses the World Economic Outlook (WEO) dataset, covering all countries from 1995 to the present.<sup>3</sup> There are two WEO vintages published every year (in April and in October). Each WEO vintage contains actual data available at the publication and projections over a 5-year horizon. While baseline results are based on all countries from the dataset, the additional specifications focus on various subsets of countries, depending on the initial level of debt, the

---

<sup>3</sup>See Annex [2.6](#) for more details.

debt trajectory over preceding years, the income level, and participation in the HIPC Initiative. In particular, we use the common World Bank Low-Income Country (LIC) demarcation for income classification, a time-varying measure based on the World Bank's yearly data of GNI per capita.<sup>4</sup>

## 2.2.2 Methodology

To estimate the causal effect of public debt on real GDP, we identify exogenous shocks in public debt by using the forecast errors from various vintages of the IMF WEO publications. Debt shocks are computed in debt to GDP terms, applied to both public debt and GDP. More precisely, they are defined as the difference between the growth rate of the actual debt to GDP ratio ( $\Delta \ln debt_{i,t}^{actual}$ ) and the growth rate of the forecasted debt to GDP ratio by IMF analysts in October of the same year ( $\Delta \ln debt_{i,t}^{forecast}$ ).

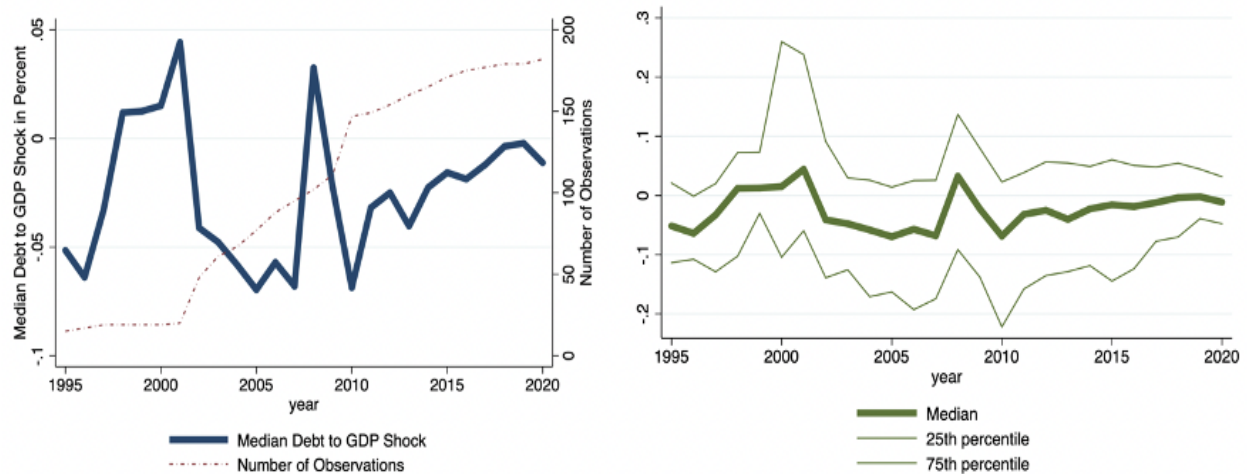
$$\begin{aligned} debt_{i,t}^{Shock} &= \left( \ln \frac{Debt_{i,t}^{actual}}{NGDP_{i,t}^{actual}} - \ln \frac{Debt_{i,t-1}^{actual}}{NGDP_{i,t-1}^{actual}} \right) - \left( \ln \frac{Debt_{i,t}^{forecast}}{NGDP_{i,t}^{forecast}} - \ln \frac{Debt_{i,t-1}^{forecast}}{NGDP_{i,t-1}^{forecast}} \right) \\ &= \Delta \ln debt_{i,t}^{actual} - \Delta \ln debt_{i,t}^{forecast} \end{aligned} \quad (2.1)$$

Constructing shocks using forecast errors solves the endogeneity problem between public debt and other macroeconomic variables, including real GDP. The most common difficulties encountered in the strand of literature attempting to assess the impact of an increase in debt to GDP ratio on output is that the debt level is always endogenous to a country's macroeconomic environment. There could be reverse causality and omitted variable bias issues in the regression of real GDP on public debt. In line with the literature (e.g., [Furceri et al. \(2018\)](#); [Blanchard and Perotti \(2002\)](#)), the forecast errors constructed as described above can be considered as exogenous changes. The projected debt level from the October WEO for the same year is assumed to have absorbed all the information up to the time of projection. The change between the projected debt level and the materialized debt level is thus identified as an exogenous shock to debt. There could be reverse causality of growth to these shocks. But, for this to be a concern, the adjustment of public debt needs to happen within the same quarter as the news about the state of the economy (from October to December). Since all the requisite information about public debt and economic performance until October are incorporated in the October forecasts, this is highly unlikely in the literature that uses

---

<sup>4</sup>See [World bank website](#).

Figure 2.2: Evolution of Public Debt Forecast Error, Median, 1995-2020



Sources: World Economic Outlook

Note: Time series plots of median debt to GDP ratio for all countries are given above. Data points are extremely limited before 2002 among Emerging and Developing Countries (EMDCs).

this approach. The time series of exogenous debt shocks is illustrated in Figure 2.2.<sup>5</sup> It shows the median of public debt to GDP shocks across economies by year using the entire dataset, starting from 1995. The most significant forecast error in terms of debt to GDP ratio deviation under a sufficient number of observations was 3 percent in 2008 during the GFC period.<sup>6</sup>

Once public debt shocks are identified, the local projections approach of Jordà (2005) can trace out the short- and medium-run output responses. Specifically, the baseline regression is

$$y_{i,t+k} - y_{i,t-1} = c_i^k + d_t^k + \beta^k \text{debt}_{i,t}^{\text{Shock}} + \theta^k Z_{i,t} + \epsilon_{it} \quad (2.2)$$

The left-hand side component  $y_{i,t+k} - y_{i,t-1}$  is the  $k$ -period ahead change of output level. In our analysis,  $k$  goes from zero to five to assess the short- and medium-term impulse responses to the exogenous public debt shock.  $\text{debt}_{i,t}^{\text{Shock}}$  is the exogenous shock on the debt to GDP ratio defined as the realized debt to GDP growth rate minus the forecasted debt to forecasted GDP growth rate.  $c_i^k$  and  $d_t^k$  are the country and time fixed effects, respectively. We cluster the standard error at the

<sup>5</sup>See also Annex 2.6 for more details.

<sup>6</sup>Given that 2020 data are not actual data yet, our dataset of debt shocks technically runs through 2018 and we don't have the data of debt shock under COVID-19 yet.

country level since debt responses are country-specific (Reinhart et al. (2003)).  $Z_{i,t}$  refers to the control variables, including two lags of real GDP growth and two lags of debt shocks.<sup>7</sup>

Based on the regression model, impulse responses are obtained from direct multistep regressions of empirical data. Each response is estimated by a single Ordinary Least Squares (OLS) regression without relying on a theoretical model, thereby preventing potential misspecification errors, as explained by Montiel Olea and Plagborg-Møller (2021). The robustness to misspecification errors also stems from a large number of observations to compute the impulse responses at each horizon, which helps to improve the precision and stability of the result.

This empirical model is first applied to all countries in the dataset, which serves as the baseline. We conduct further analysis by separating our sample based on other characteristics such as the initial level of debt, the debt trajectory over the preceding five years, the income level, and pre vs. post participation in the HIPC debt relief initiative.<sup>8</sup> Lastly, we perform robustness checks by (i) controlling for the initial level of debt, (ii) checking for a potential Nickell-bias, and (iii) using actual data from  $t + 2$  and  $t + 3$  WEO vintage.

## 2.3 Results

### 2.3.1 Baseline Result

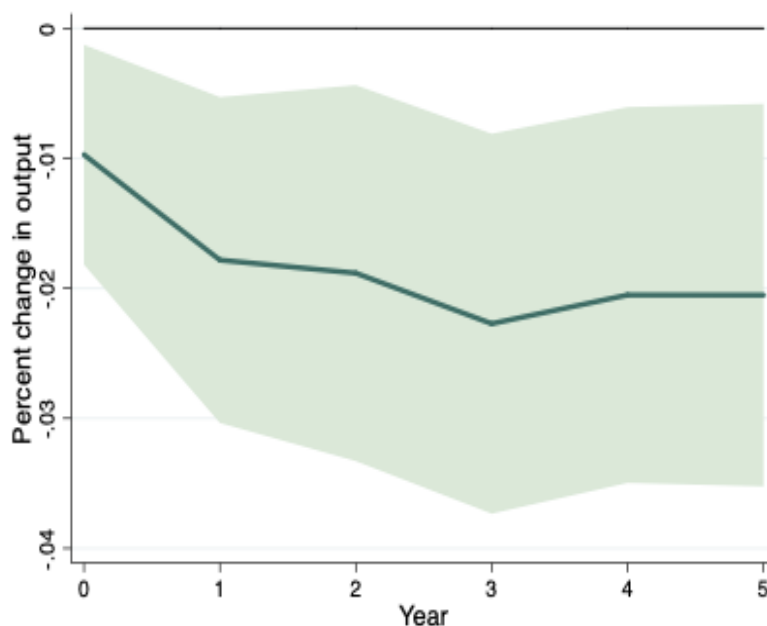
Our empirical results indicate that real GDP responds negatively to an unexpected increase in public debt on average. Figure 2.3 and Table 2.1 suggest that a 1 percent unanticipated increase in the debt to GDP ratio leads to a significant -0.02 percent decrease in the real GDP level immediately after the shock. In the short run, the impact on real GDP is negative, and the trends get more remarkable, suggesting that surprises on public debt take time to materialize. For example, applying the median shock under the Global Financial Crisis (GFC) to the median country in the sample, an unanticipated 1.38 percentage points increase in the public debt to GDP ratio reduces the output

---

<sup>7</sup>To test the potential concerns of Nickell Bias, we dropped two lags of dependent variables and found identical results. We also conducted other robustness checks (see Annex 2.7 for details).

<sup>8</sup>We define positive (negative) debt trajectories when countries had a positive (negative) debt to GDP ratio growth over the preceding five years.

Figure 2.3: Response of Real GDP to Public Debt Shocks, Baseline



Sources: World Economic Outlook and IMF staff calculation.

Note:  $t = 0$  is the year of shock. The Green shaded area denotes the 90 percent, confidence bands. The solid green line represents the output response to an unanticipated increase in public debt to GDP ratio. The median public debt to GDP ratio is 46.12 percent. Estimates based on equation (2.2).

level by up to -0.06 percent immediately after the shock.<sup>9</sup> Our empirical result is broadly in line with other results from the literature, such as [Reinhart and Rogoff \(2010\)](#) and [Woo and Kumar \(2015\)](#). Impulse responses also vary depending on other fundamental characteristics that we study in the following sections.

### 2.3.2 Initial Level of Public Debt

To analyze how the initial level of debt affects the response of real GDP to an unanticipated increase in public debt, we separate our sample into two groups: high and low initial level of public debt, using as a threshold the median of our selection (46.12 percent). The main result is that an unanticipated increase in public debt hurts growth for countries with a high initial level of debt ([Figure 2.4](#) and [Table 2.2](#)). In the high initial debt level subsample, we find that a 1 percent

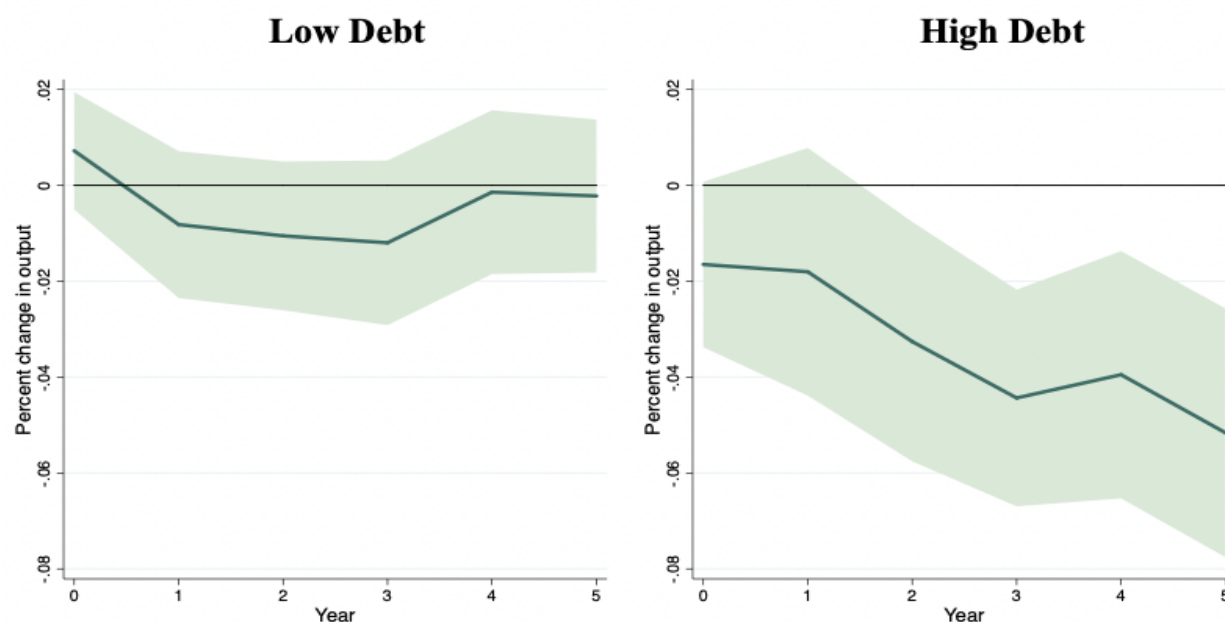
<sup>9</sup>The median shock during the GFC is 3 percent ([Figure 2.2](#)), and our sample's median public debt to GDP ratio is 46.12 percent. In this example, the median shock of a 3 percent increase under GFC is equivalent to a rise in the median debt to GDP ratio by 1.38 percentage points.

Table 2.1: Regression Results, Baseline

	<b>K=0</b>	<b>K=1</b>	<b>K=2</b>	<b>K=3</b>	<b>K=4</b>	<b>K=5</b>
<b>PUBLIC DEBT SHOCK(T)</b>	-0.00972*	-0.0178***	-0.0188**	-0.0227**	-0.0205**	-0.0205**
	0.00514	0.00761	0.00879	0.00888	0.00878	0.00894
<b>GROWTH(T-1)</b>	.084093	-.0189105	-.1449393	-.3038634***	-.4205633***	-.5503717***
	.0845462	.0891174	.1049426	.0961871	.0980341	.0957038
<b>GROWTH(T-2)</b>	-.1671624**	-.1756248**	-.1551317*	-.09265	-.0687622	-.0214411
	.0792387	.0754154	.0924446	.0917253	.1022521	.0975482
<b>PUBLIC DEBT SHOCK(T-1)</b>	-.0032329	-.002934	-.0089907	-.0083104	-.0103024	-.0040412
	.0041954	.0077448	.0085263	.0077699	.0087287	.0082071
<b>PUBLIC DEBT SHOCK(T-2)</b>	-.0001221	-.0031401	-.0016663	-.0036035	-.0017956	-.008136
	.0047932	.0055046	.0062853	.0074224	.0078455	.0077854
<b>NUMBER OF COUNTRIES</b>	179	178	176	173	168	163
	2229	2052	1876	1701	1530	1366
<b>R<sup>2</sup></b>	0.4047	0.3970	0.4182	0.4608	0.5034	0.5418

Note: Observations are at the country-vintage level. The dependent variable is the change in real GDP for up to five-year horizon(k).  $k = 0$  is the year of shock. Country and time fixed effects were incorporated. Standard errors are clustered at the country level. \*, \*\*, and \*\*\* indicate significance at the 10, 5, and 1 percent levels, respectively.

Figure 2.4: Response of Real GDP to Public Debt Shocks, by Initial Debt Level Classification



Sources: World Economic Outlook and IMF staff calculation.

Note:  $t = 0$  is the year of shock. The Green shaded area denotes the 90 percent, confidence bands. The solid green line represents the output response to an unanticipated increase in public debt to GDP ratio. The median public debt to GDP ratio is 46.12 percent. Estimates based on equation (2.2).

unanticipated increase in the debt to GDP ratio leads to a significant up to -0.05 percent decrease in real GDP 2-5 years after the shock.<sup>10</sup> On the contrary, responses in the low initial debt level subsample were not statistically significant.

Various papers are investigating whether there exists a debt to GDP ratio threshold beyond which macroeconomic outcomes would be negatively affected. Our result is broadly in line with the literature, which finds that high initial debt to GDP ratios tend to be associated with lower subsequent growth.

<sup>10</sup>To illustrate the result: the median shock during the GFC is 8 percent (Figure 2.2). The median public debt to GDP ratio in the high initial debt subsample is 73.98 percent. In this example, the median shock of an 8 percent increase in the GFC is equivalent to a rise in the median debt to GDP ratio by 5.92 percentage points. In this context, such an unanticipated increase in the debt to GDP ratio causes a -0.16 percent reduction in real GDP for countries with a high initial level of debt.

Table 2.2: Regression Results, by Initial Debt Level Classification

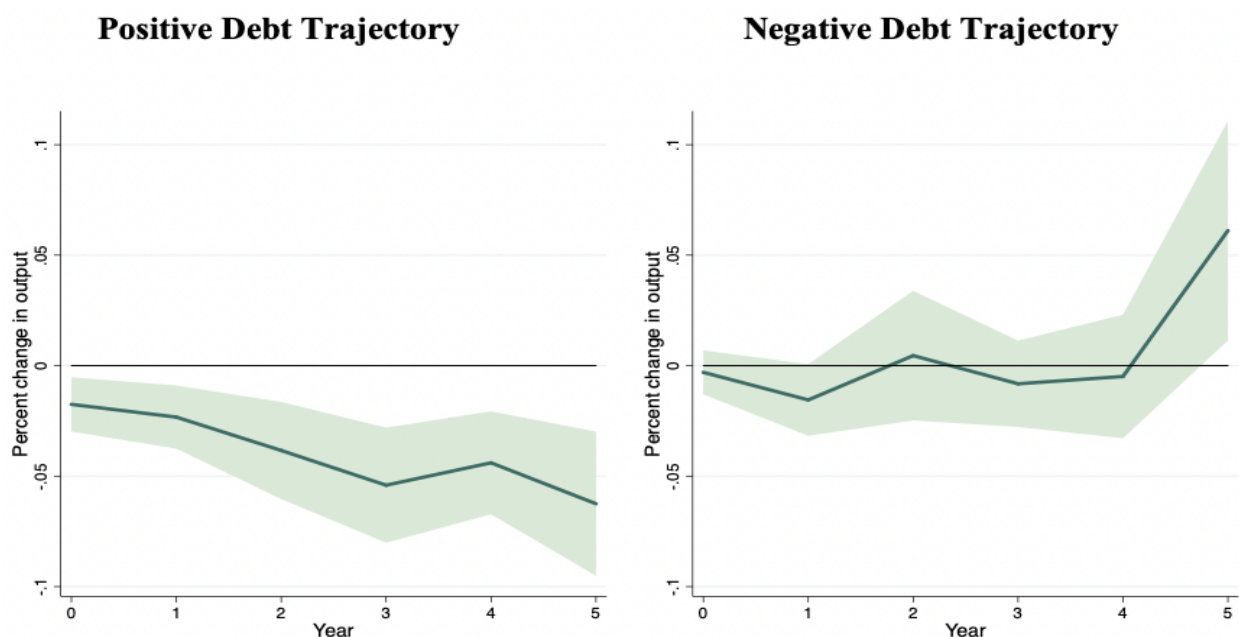
LOW DEBT	K=0	K=1	K=2	K=3	K=4	K=5
PUBLIC DEBT SHOCK(T)	0.00720	-0.00819	-0.0105	-0.0120	-0.00142	-0.00223
	0.00744	0.00930	0.00943	0.0104	0.0104	0.00970
NUMBER OF COUNTRIES	122	118	112	102	94	88
NUMBER OF OBSERVATIONS	1038	915	799	692	597	506
$R^2$	0.4146	0.4502	0.5083	0.5394	0.5777	0.6232

HIGH DEBT	K=0	K=1	K=2	K=3	K=4	K=5
PUBLIC DEBT SHOCK(T)	-0.0165	-0.0180	-0.0326**	-0.0444***	-0.0395***	-0.0516***
	0.0105	0.0157	0.0152	0.0137	0.0157	0.0158
NUMBER OF COUNTRIES	109	99	94	80	70	62
NUMBER OF OBSERVATIONS	956	837	735	640	561	494
$R^2$	0.4796	0.4585	0.4419	0.4785	0.5093	0.5412

Note: Observations are at the country-vintage level. The dependent variable is the change in real GDP for up to five-year horizon(k).  $k = 0$  is the year of shock. Country and time fixed effects were incorporated. Standard errors are clustered at the country level. \*, \*\*, and \*\*\* indicate significance at the 10, 5, and 1 percent levels, respectively.

Figure 2.5: Response of Real GDP to Public Debt Shocks, by Public Debt Trajectory



Sources: World Economic Outlook and IMF staff calculation.

Note:  $t = 0$  is the year of shock. The Green shaded area denotes the 90 percent, confidence bands. The solid green line represents the output response to an unanticipated increase in public debt to GDP ratio. The median public debt to GDP ratio is 46.12 percent. Estimates based on equation (2.2).

### 2.3.3 Public Debt Trajectory

To analyze how the public debt trajectory in previous years affects the response of real GDP to an unanticipated increase in public debt, we separate our sample into two groups based on the following definition: a debt trajectory is positive (negative) when countries had a positive (negative) growth in their debt to GDP ratio over the preceding five years. In [Figure 2.5](#) and [Table 2.3](#), the subsample analysis with positive public debt trajectories reveals that a 1 percent unanticipated increase in the debt to GDP ratio leads to a significant -0.05 percent decrease in real GDP starting immediately after the shock. On the contrary, results are not effective in the negative public debt trajectory subsample.

Our result can be related to [Pescatori et al. \(2014\)](#), who find that the debt trajectory can be an essential determinant in understanding future growth prospects.

Table 2.3: Regression Results, by Public Debt Trajectory

<b>POSITIVE DEBT TRAJECTORY</b>	<b>K=0</b>	<b>K=1</b>	<b>K=2</b>	<b>K=3</b>	<b>K=4</b>	<b>K=5</b>
<b>PUBLIC DEBT SHOCK(T)</b>	-0.0175**	-0.0233***	-0.0384***	-0.0541***	-0.0440***	-0.0625***
	0.00745	0.00868	0.0133	0.0158	0.0141	0.0198
<b>NUMBER OF COUNTRIES</b>	160	151	147	131	108	97
<b>NUMBER OF OBSERVATIONS</b>	1023	834	685	547	435	349
<b><math>R^2</math></b>	0.5298	0.4895	0.5066	0.5292	0.5454	0.6066

<b>NEGATIVE DEBT TRAJECTORY</b>	<b>K=0</b>	<b>K=1</b>	<b>K=2</b>	<b>K=3</b>	<b>K=4</b>	<b>K=5</b>
<b>PUBLIC DEBT SHOCK(T)</b>	-0.00304	-0.0155	0.00453	-0.00820	-0.00488	0.0611**
	0.00605	0.00984	0.0178	0.0119	0.0170	0.0302
<b>NUMBER OF COUNTRIES</b>	122	110	99	83	60	49
<b>NUMBER OF OBSERVATIONS</b>	628	489	378	283	207	157
<b><math>R^2</math></b>	0.3510	0.3959	0.4149	0.4883	0.4290	0.4998

Note: Observations are at the country-vintage level. The dependent variable is the change in real GDP for up to five-year horizon( $k$ ).  $k = 0$  is the year of shock. Country and time fixed effects were incorporated. Standard errors are clustered at the country level. \*, \*\*, and \*\*\* indicate significance at the 10, 5, and 1 percent levels, respectively.

### 2.3.4 Income Classification

To analyze how the income classification affects the response of real GDP to an unanticipated increase in public debt, we divide our sample between high-, middle-, and low-income countries following the World Bank income classification based on GNI per capita. The comparison between high-income countries and low-income countries in [Figure 2.6](#) and [Table 2.4](#) indicates that, in low-income countries, we didn't observe a remarkable response.<sup>11</sup> On the contrary, in high-income countries, the answer is negative with small magnitude; a 1 percent unanticipated increase in debt to GDP ratio leads to a significant -0.03 percent reduction in real GDP 4 years after the shock.

This result of high-income countries is in line with Rahman et al. (2019), who emphasize the importance of analyzing the causal implications of public debt shock to real GDP growth based on the income classification. The literature on low-income countries is limited because of relatively scarce data.

### 2.3.5 HIPC Initiative

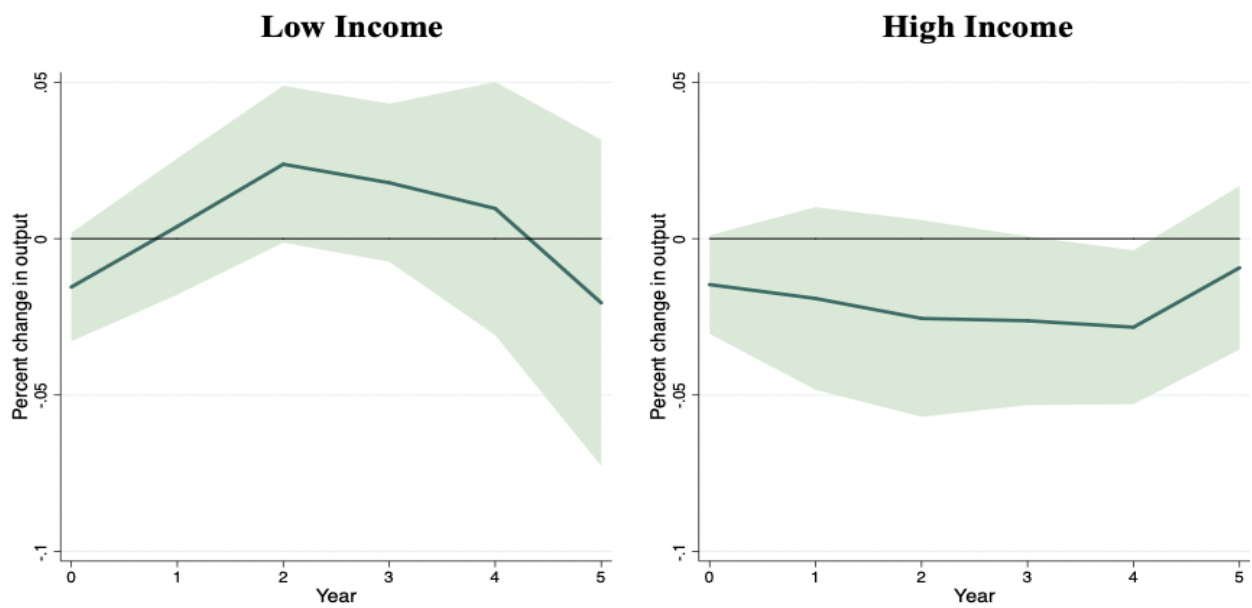
The HIPC Initiative was introduced in 1996 by the IMF and the World Bank to help reduce unmanageable debt burdens in emerging countries. It has provided US\$76 billion in debt relief so far. The debt relief was completed for 36 countries and is still in progress in 3 countries. To analyze how the HIPC debt relief initiative may have affected the response of real GDP to an unanticipated increase in public external debt, we divide our sample between data points before and after the completion of the HIPC debt relief.<sup>12</sup> In the HIPC process, there are two important milestones: (i) the decision point is the date when the country is assessed to qualify for the HIPC debt relief, which involves, among other requirements, establishing a credential track record of policy performance, attempting to clear the amount of arrears to the IMF and World Bank, preparing of Poverty Reduction Strategy Paper (PRSP), and agreeing on a floating completion point with the World Bank and the IMF, and (ii) the completion point when the country has completed the critical structural reforms agreed at the decision point. We divide the data between pre vs. post HIPC

---

<sup>11</sup>Results for middle-income countries can be found in [Annex 2.7](#).

<sup>12</sup>See [IMF Policy Paper, Heavily Indebted Poor Country \(HIPC\) Initiative and Multilateral Debt Relief Initiative \(MDRI\) – Statistical Update, August 6, 2019](#).

Figure 2.6: Response of Real GDP to Public Debt Shocks, by Income Classification



Sources: World Economic Outlook and IMF staff calculation.

Note:  $t = 0$  is the year of shock. The Green shaded area denotes the 90 percent, confidence bands. The solid green line represents the output response to an unanticipated increase in public debt to GDP ratio. The median public debt to GDP ratio is 46.12 percent. Estimates based on equation (2.2).

Table 2.4: Regression Results, by Income Classification

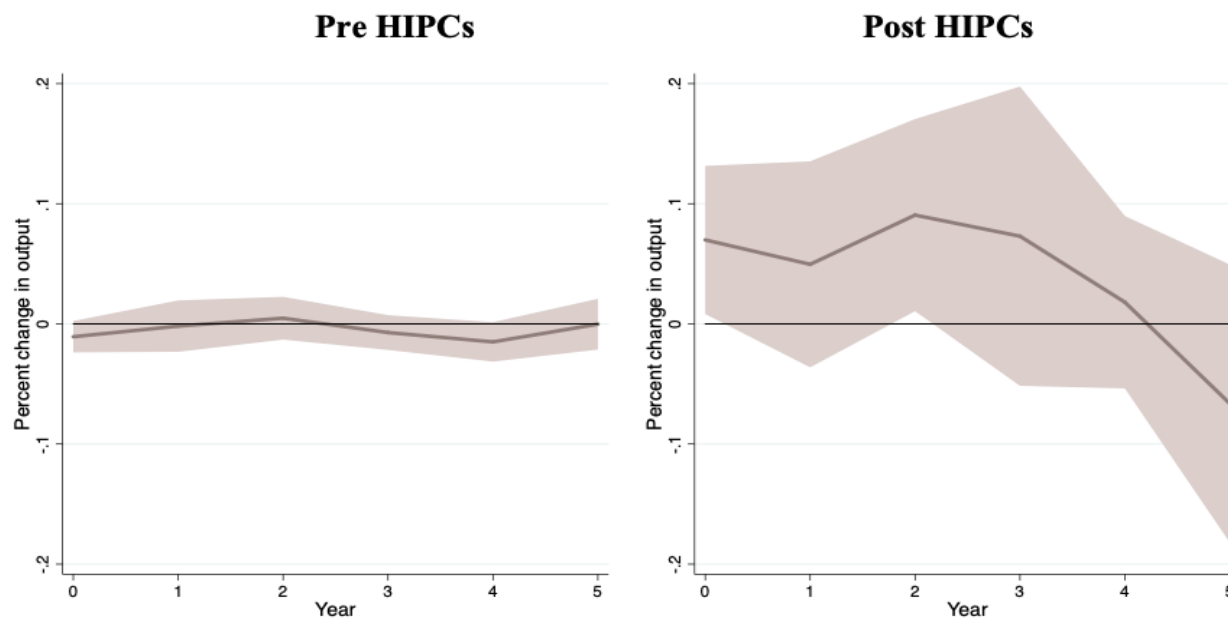
LOW-INCOME COUNTRIES	K=0	K=1	K=2	K=3	K=4	K=5
PUBLIC DEBT SHOCK(T)	-0.0155	0.00390	0.0238	0.0179	0.00965	-0.0205
	0.0105	0.0132	0.0152	0.0154	0.0246	0.0317
NUMBER OF COUNTRIES	38	37	32	30	27	27
NUMBER OF OBSERVATIONS	273	233	196	165	137	112
$R^2$	0.2693	0.4850	0.5596	0.6679	0.6951	0.7321

HIGH-INCOME COUNTRIES	K=0	K=1	K=2	K=3	K=4	K=5
PUBLIC DEBT SHOCK(T)	-0.0147	-0.0191	-0.0255	-0.0263	-0.0283*	-0.00930
	0.00957	0.0178	0.0191	0.0164	0.0149	0.0159
NUMBER OF COUNTRIES	55	53	52	50	50	49
NUMBER OF OBSERVATIONS	778	721	668	618	570	521
$R^2$	0.5066	0.5189	0.5453	0.5746	0.5965	0.6266

Note: Observations are at the country-vintage level. The dependent variable is the change in real GDP for up to five-year horizon(k).  $k = 0$  is the year of shock. Country and time fixed effects were incorporated. Standard errors are clustered at the country level. \*, \*\*, and \*\*\* indicate significance at the 10, 5, and 1 percent levels, respectively.

Figure 2.7: Response of Real GDP to Public Debt Shocks, pre- vs. post-HIPC Initiative



Sources: World Economic Outlook and IMF staff calculation.

Note:  $t = 0$  is the year of shock. The red shaded area denotes the 90 percent, confidence bands. The solid red line indicates the response to an unanticipated increase in public debt to GDP ratio. The median public debt to GDP ratio is 65.28 and 23.97 percent in pre- vs. post-HIPC Initiative subsamples. Estimates based on equation (2.2).

Initiative based on the completion point to assess the difference in impulse responses of real GDP to the unanticipated public debt increase. Given that the HIPC debt relief was focused on public external debt, the analysis is focused on public external debt.<sup>13</sup> Figure 2.7 and Table 2.5 indicate that, in the post-HIPC Initiative subsample, a 1 percent increase in the debt to GDP ratio leads to a significant 0.09 percent increase in real GDP 2 years after the shock.

## 2.4 Robustness Analysis

This section assesses several robustness checks to confirm whether our results are robust to different specifications. In our analysis, we add the initial debt levels as control variables, check for the Nickell-bias, and use actual data from  $t + 2$  and  $t + 3$  WEO vintages.

<sup>13</sup>The total liabilities of a country with foreign creditors on resident general government and monetary authorities to all foreign (non-resident) sectors. Creditors often determine all the terms of the debt contracts, which are normally subject to the jurisdiction of the foreign creditors or, for multilateral credits, to international law.

Table 2.5: Regression Results, pre- vs. post- HIPC Initiative

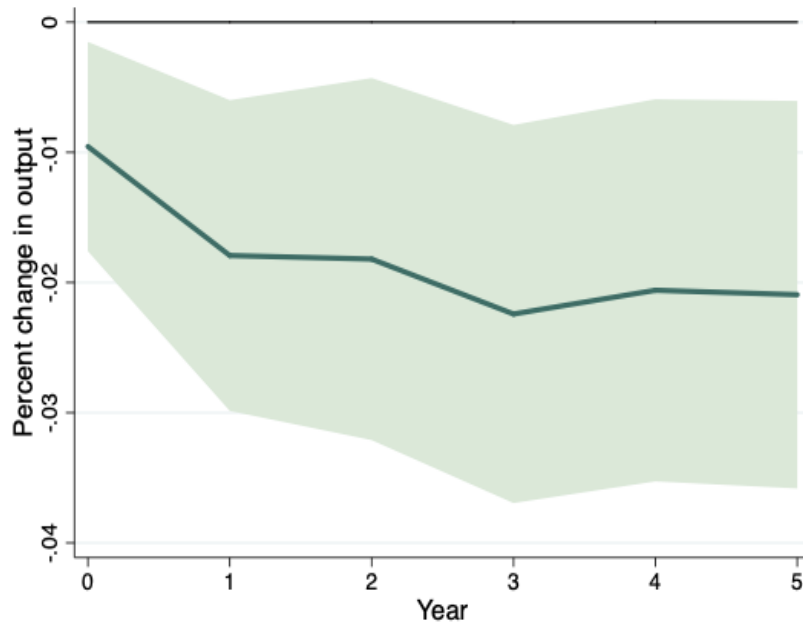
<b>PRE HIPC</b>	<b>K=0</b>	<b>K=1</b>	<b>K=2</b>	<b>K=3</b>	<b>K=4</b>	<b>K=5</b>
<b>PUBLIC DEBT SHOCK(T)</b>	-0.0107	-0.00189	0.00471	-0.00724	-0.0150	-0.000272
	0.00794	0.0130	0.0108	0.00879	0.00997	0.0128
<b>NUMBER OF COUNTRIES</b>	32	32	32	32	32	31
<b>NUMBER OF OBSERVATIONS</b>	224	223	221	219	217	215
<b><math>R^2</math></b>	0.2003	0.3036	0.3897	0.5053	0.6218	0.6165

<b>POST HIPC</b>	<b>K=0</b>	<b>K=1</b>	<b>K=2</b>	<b>K=3</b>	<b>K=4</b>	<b>K=5</b>
<b>PUBLIC DEBT SHOCK(T)</b>	0.0698	0.0495	0.0905*	0.0730	0.0180	-0.0663
	0.0374	0.0521	0.0486	0.0757	0.0436	0.0703
<b>NUMBER OF COUNTRIES</b>	35	35	35	35	34	33
<b>NUMBER OF OBSERVATIONS</b>	418	382	346	310	277	245
<b><math>R^2</math></b>	0.1857	0.2456	0.3959	0.4846	0.5807	0.6546

Note: Observations are at the country-vintage level. The dependent variable is the change in real GDP for up to five-year horizon(k).  $k = 0$  is the year of shock. Country and time fixed effects were incorporated. Standard errors are clustered at the country level. \*, \*\*, and \*\*\* indicate significance at the 10, 5, and 1 percent levels, respectively.

Figure 2.8: Response of Real GDP to Public Debt Shocks, pre- vs. post-HIPC Initiative



Sources: World Economic Outlook and IMF staff calculation.

Note:  $t = 0$  is the year of shock. The Green shaded area denotes the 90 percent, confidence bands. The solid green line represents the response to an unanticipated increase in public debt to GDP ratio. The median public debt to GDP ratio is 46.12 percent. Estimates based on equation (2.2).

### 2.4.1 Adding Lags of Initial Debt Level as Controls

To check the robustness of our results, we add two lags of the initial debt to GDP Ratio as controls. Potential endogeneity issues could arise despite the approach of shock construction based on forecast errors, defined as the gap between the materialized outcome and October WEO projection of the same year. However, results remain essentially identical (Figure 2.8 and Table 2.6).

### 2.4.2 Dropping Two Lags of Dependent Variable

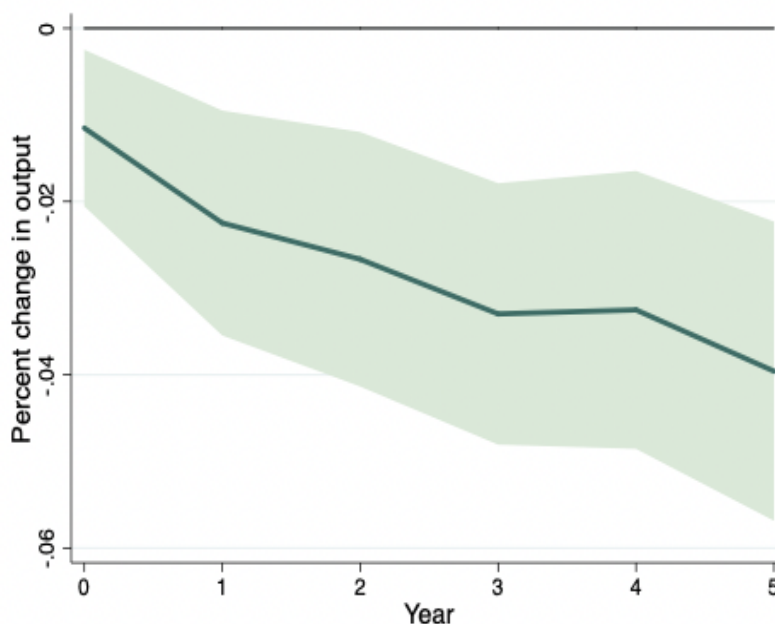
Nickell bias is caused by adding two lags of dependent variables as controls, which may cause a potential endogeneity issue. To prevent the potential Nickell-bias, Furceri, and Li (2017) drop two lags of the dependent variable. Our results remain broadly identical (Figure 2.9 and Table 2.7).

Table 2.6: Regression Results, Baseline, Adding Lags of Initial Debt Levels

ADDING INITIAL DEBT LEVELS	K=0	K=1	K=2	K=3	K=4	K=5
PUBLIC DEBT SHOCK(T)	-0.00956*	-0.0179**	-0.0182**	-0.0224**	-0.0206**	-0.0209**
	0.00549	0.00789	0.00893	0.00917	0.00975	0.0105
NUMBER OF COUNTRIES	179	178	176	173	168	163
NUMBER OF OBSERVATIONS	2229	2052	1876	1701	1530	1366
$R^2$	0.3597	0.2943	0.2578	0.2419	0.5118	0.5513

Note: Observations are at the country-vintage level. The dependent variable is the change in real GDP for up to five-year horizon(k).  $k = 0$  is the year of shock. Country and time fixed effects were incorporated. Standard errors are clustered at the country level. \*, \*\*, and \*\*\* indicate significance at the 10, 5, and 1 percent levels, respectively.

Figure 2.9: Response of Real GDP to Public Debt Shocks, Baseline, Dropping Lags of Dependent Variable



Sources: World Economic Outlook and IMF staff calculation.

Note:  $t = 0$  is the year of shock. The Green shaded area denotes the 90 percent, confidence bands. The solid green line represents the response to an unanticipated increase in public debt to GDP ratio. The median public debt to GDP ratio is 46.12 percent. Estimates based on equation (2.2).

Table 2.7: Regression Results, Baseline, Dropping Lags of Dependent Variable

NICKELL-BIAS	K=0	K=1	K=2	K=3	K=4	K=5
<b>PUBLIC DEBT SHOCK(T)</b>	-0.0115**	-0.0225***	-0.0267***	-0.0330***	-0.0325***	-0.0396***
	0.00549	0.00789	0.00893	0.00917	0.00975	0.0105
<b>NUMBER OF COUNTRIES</b>	179	178	176	173	168	163
<b>NUMBER OF OBSERVATIONS</b>	2229	2052	1876	1701	1530	1366
$R^2$	0.3597	0.2943	0.2578	0.2419	0.2281	0.2240

Note: Observations are at the country-vintage level. The dependent variable is the change in real GDP for up to five-year horizon(k).  $k = 0$  is the year of shock. Country and time fixed effects were incorporated. Standard errors are clustered at the country level. \*, \*\*, and \*\*\* indicate significance at the 10, 5, and 1 percent levels, respectively.

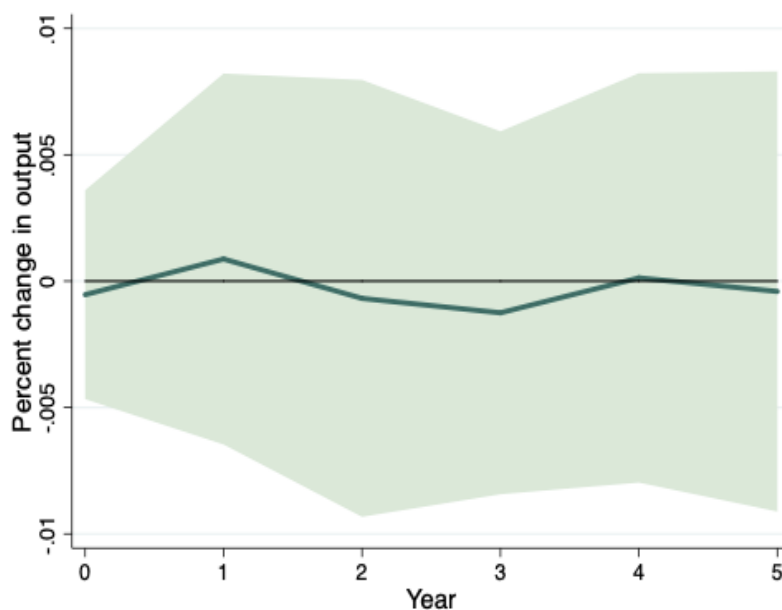
### 2.4.3 Alternative Constructions of the Forecast Error

Since the beginning of our dataset, WEO data may have encountered some revisions, making our use of actual data from the April 2021 WEO to construct forecast errors potentially less accurate. To tackle this potential concern, we use realized actual data published two and three years after the forecasted October WEO instead of actual data available in the latest WEO vintage to calculate unanticipated shocks. Our results remain identical (Figure 2.10, Figure 2.11, Table 2.8, and Table 2.9).

## 2.5 Conclusion

Using a novel way of constructing exogenous shocks in public debt, this paper provides new empirical evidence on the impact of an unanticipated change in the public debt to GDP ratio on real GDP. Overall, real GDP responds negatively to unanticipated debt-to- GDP ratio increases. The response also varies depending on several fundamental characteristics. We find that an unanticipated increase in the public debt to GDP ratio hurts real GDP for countries with (i) a high initial debt level, (ii) a positive debt trajectory over the five preceding years, and (iii) a high-income level. On the contrary, it boosts real GDP for countries (iv) that have completed the

Figure 2.10: Response of Real GDP to Public Debt Shocks, Baseline, Using  $t+2$  WEO Vintage for Actual Data



Sources: World Economic Outlook and IMF staff calculation.

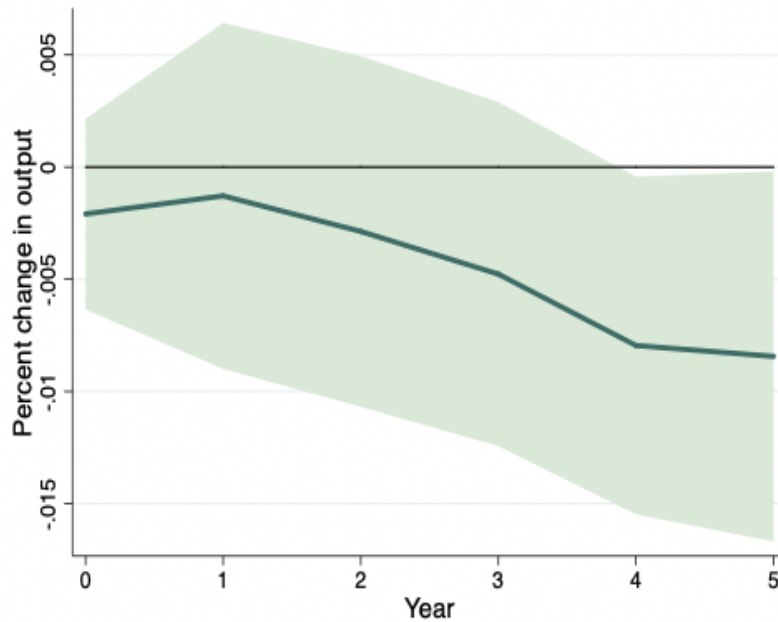
Note:  $t = 0$  is the year of shock. The Green shaded area denotes the 90 percent, confidence bands. The solid green line represents the response to an unanticipated increase in public debt to GDP ratio. The median public debt to GDP ratio is 35.64 percent. Estimates based on equation (2.2).

Table 2.8: Regression Results, Baseline, Using  $t + 2$  WEO Vintage for Actual Data

USING T+2 ACTUAL DATA	K=0	K=1	K=2	K=3	K=4	K=5
PUBLIC DEBT SHOCK(T)	-0.000536	0.000874	-0.000683	-0.00125	0.000127	-0.000409
	0.00251	0.00446	0.00525	0.00436	0.00492	0.00529
NUMBER OF COUNTRIES	177	177	177	174	169	163
NUMBER OF OBSERVATIONS	1890	1890	1890	1715	1544	1380
$R^2$	0.2498	0.3187	0.4078	0.4511	0.4930	0.4845

Note: Observations are at the country-vintage level. The dependent variable is the change in real GDP for up to five-year horizon( $k$ ).  $k = 0$  is the year of shock. Country and time fixed effects were incorporated. Standard errors are clustered at the country level. \*, \*\*, and \*\*\* indicate significance at the 10, 5, and 1 percent levels, respectively.

Figure 2.11: Response of Real GDP to Public Debt Shocks, Baseline, Using  $t+3$  WEO Vintage for Actual Data



Sources: World Economic Outlook and IMF staff calculation.

Note:  $t = 0$  is the year of shock. The Green shaded area denotes the 90 percent, confidence bands. The solid green line represents the response to an unanticipated increase in public debt to GDP ratio. The median public debt to GDP ratio is 35.10 percent. Estimates based on equation (2.2).

Table 2.9: Regression Results, Baseline, Using  $t + 3$  WEO Vintage for Actual Data

USING T+3 ACTUAL DATA	K=0	K=1	K=2	K=3	K=4	K=5
PUBLIC DEBT SHOCK(T)	-0.00209	-0.00128	-0.00287	-0.00477	-0.00795*	-0.00844*
	0.00258	0.00469	0.00475	0.00466	0.00457	0.00501
NUMBER OF COUNTRIES	173	173	173	173	168	162
NUMBER OF OBSERVATIONS	1710	1710	1710	1710	1539	1375
$R^2$	0.2542	0.3311	0.3719	0.4510	0.4957	0.4896

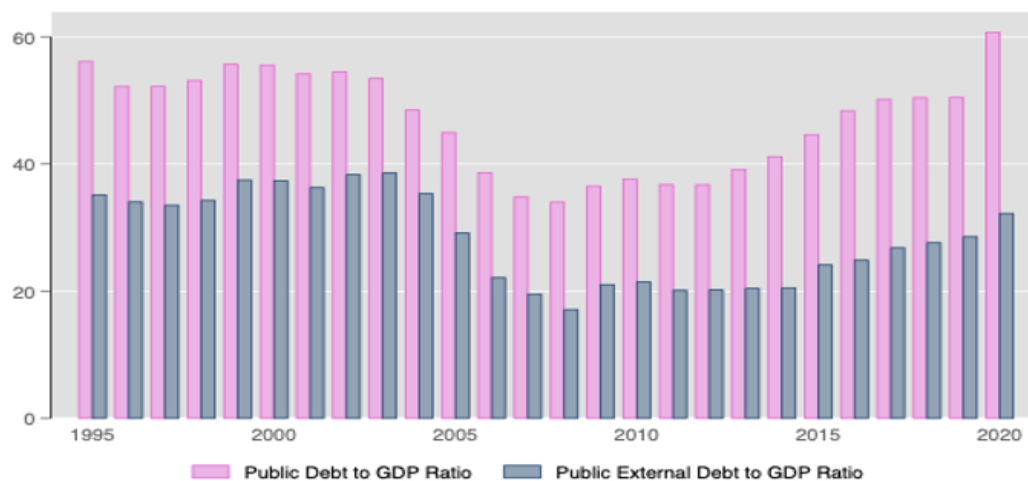
Note: Observations are at the country-vintage level. The dependent variable is the change in real GDP for up to five-year horizon( $k$ ).  $k = 0$  is the year of shock. Country and time fixed effects were incorporated. Standard errors are clustered at the country level. \*, \*\*, and \*\*\* indicate significance at the 10, 5, and 1 percent levels, respectively.

HIPC debt-service relief initiative. To conclude, country-specific characteristics need to be taken into account to fully assess the impact of increasing public debt on real GDP.

Future research could look into the channels through which the unanticipated increase in public debt impacts real GDP, particularly for each subsample with different macroeconomic characteristics. If the rise in public debt is used for public investment, it is likely to increase subsequent growth (Furceri et al. (2018)). However, the increase in public debt may be channeled to other uses, such as tax cuts or additional fiscal spending. Those channels can also be impacted by other factors such as governance or corruption. In those cases, the impact of an increase in public debt on real GDP may not be the same. Several papers in the literature have shed light on those impacts. Using a difference-in-differences approach, Kim et al. (2017) find that corruption hurts the effect of debt on growth in countries with high corruption.

Our results prescribe various policy implications. First, high-income countries can negatively impact an increase in the debt to GDP ratio. Second, lowering the initial debt levels or keeping a downward debt trajectory increases countries' potential benefit from additional borrowings. Third, participation in the HIPC Initiative increases countries' potential to gain from extra borrowings, which can be a positive experience to be considered for other debt relief initiatives.

Figure 2.12: Evolution of Debt in the Dataset-Debt to GDP Ratio, All Countries,Median



Sources: World Economic Outlook and IMF staff calculation

## ANNEXES

### 2.6 Statistics and Debt Shocks

This annex illustrates our dataset and shows the distribution of the constructed debt shocks, both for public debt shocks and public external debt shocks.

### 2.7 Other Results

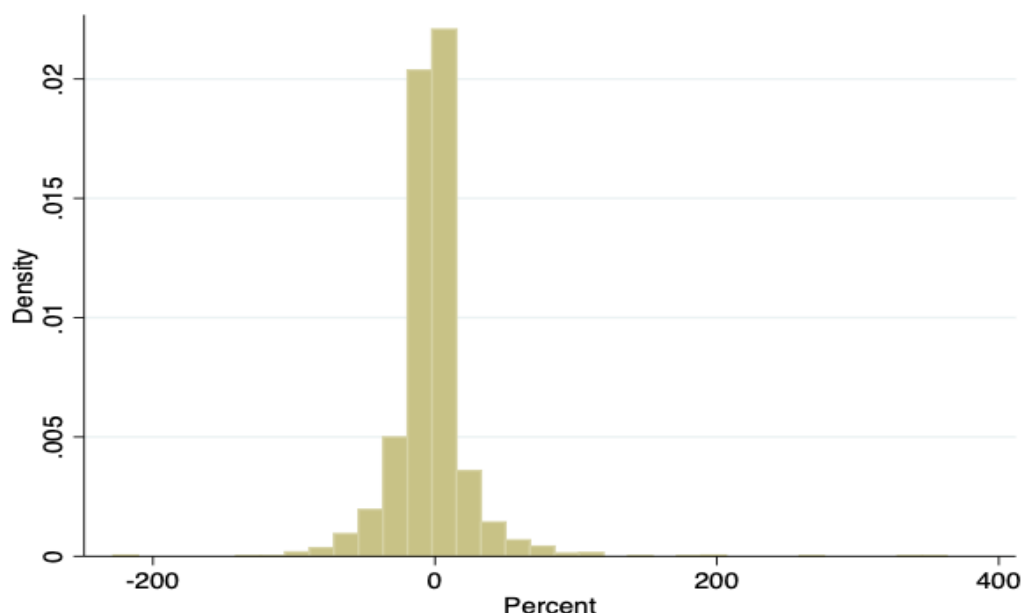
This annex provides results for two other income classifications and public external debt.

#### Additional results for income classification

As mentioned in section 2.32.3.4, here we provide the results for the other two income classifications that are not shown in section 2.32.3.4. The results suggest that there is a small positive impact of exogenous debt shocks for lower middle income countries, while a small and negative impact for higher middle-income countries.

#### Public external debt

Figure 2.13: Public Debt Shocks-Public Debt to GDP Ratio, Deviation from Projection, Percent



Sources: World Economic Outlook and IMF staff calculation

Note: The figure reports the shock distribution of public debt shock(percent). The Median is 1.23 percent. 94.23 percent of data lies within 50% range of deviation, and 55.72% of data lies within 10% range of deviation.

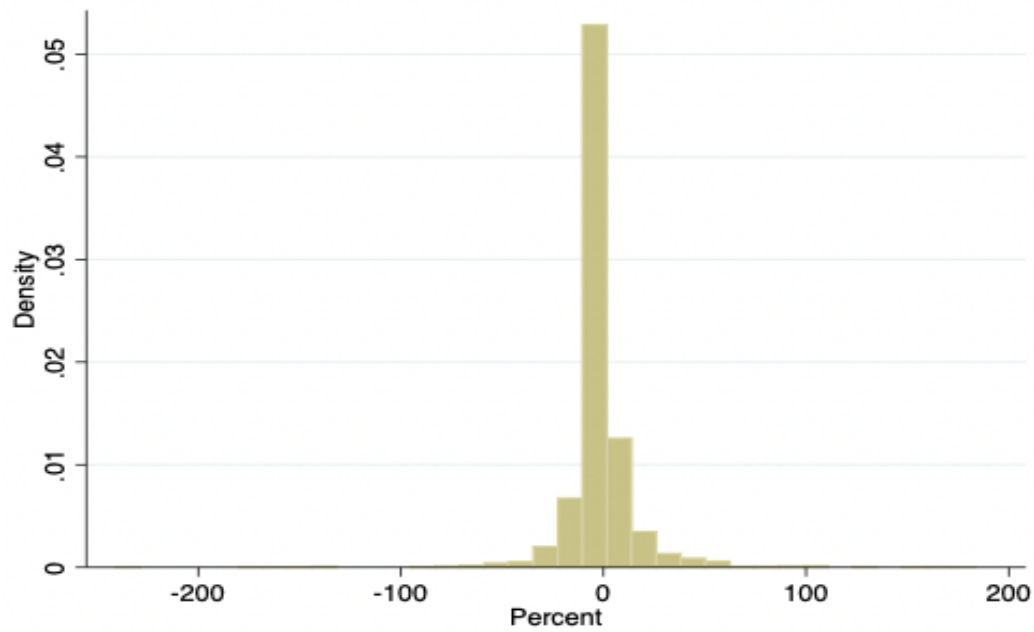
We perform additional analysis on public external debt shocks, and the baseline results are shown in [Figure 2.12](#). The results suggest that real GDP overall responds negatively to public external debt shocks and the response is contemporaneous.

## 2.8 Data Cleaning

This annex explains how we cleaned our data to prevent possible data errors from influencing the results. The data cleaning process that we used closely follows the methodology of the paper “Worse Than You Think: Public Debt Forecast Errors in Advanced and Developing Economies” (Flores et al., 2020). Looking at each country individually, they rescale, convert, or eliminate data points when forecast errors are unusually large. We use their strategy when applicable to our context.

### Rescaling

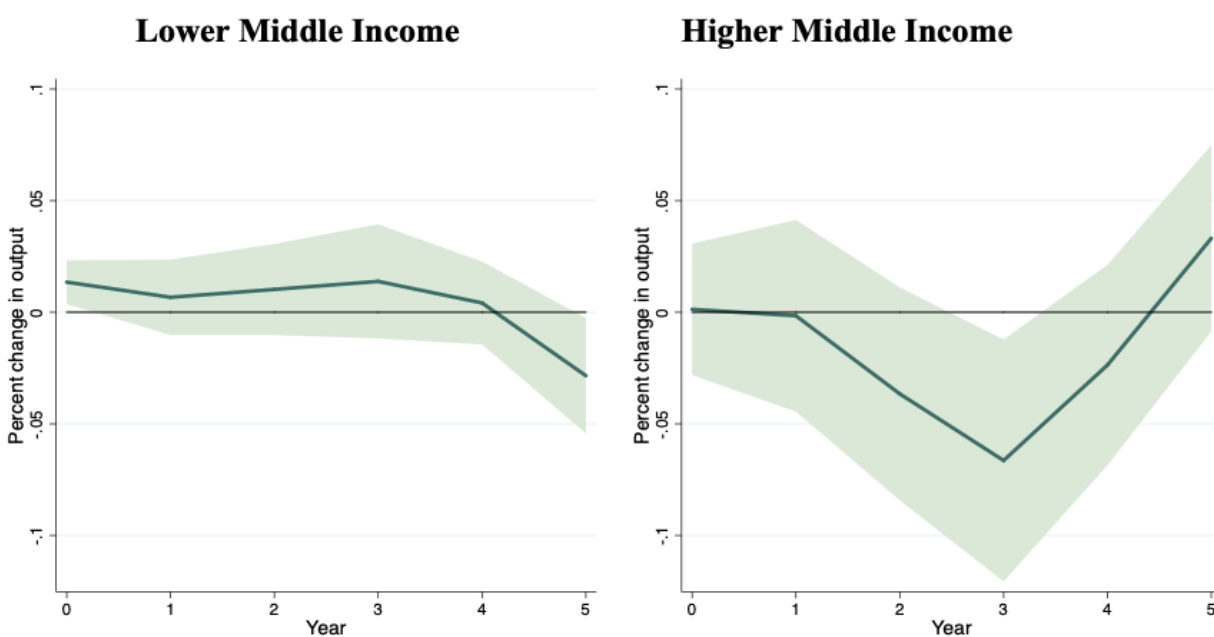
Figure 2.14: Public External Debt Shocks-Public External Debt to GDP Ratio, Deviation from Projection, Percent



Sources: World Economic Outlook and IMF staff calculation

Note: The figure reports the shock distribution of public debt shock(percent). The Median is -0.88 percent. 97.66 percent of data lies within 50% range of deviation, and 76.40% of data lies within 10% range of deviation.

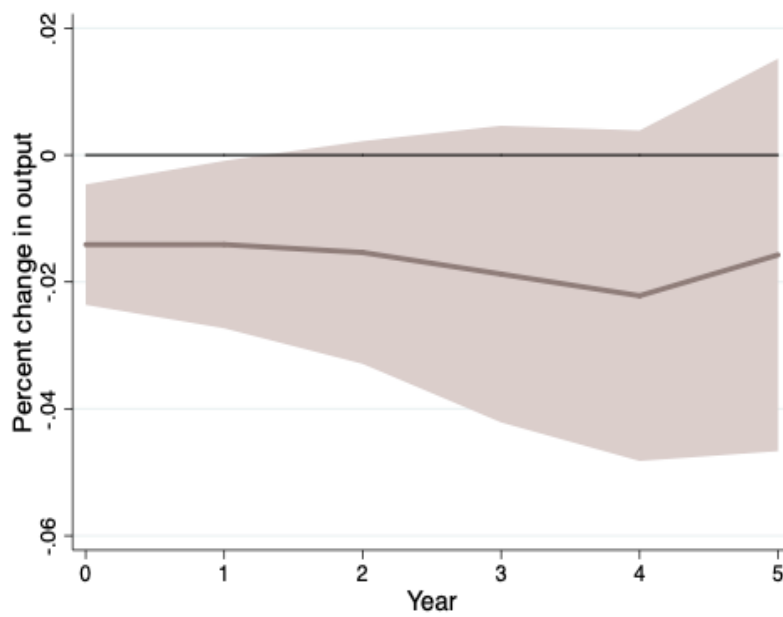
Figure 2.15: Response of Real GDP to Public Debt Shocks-Middle-Income Countries



Sources: World Economic Outlook and IMF staff calculation

Note:  $t = 0$  is the year of shock. The Green shaded area denotes the 90 percent confidence bands. The solid green line denotes the response to an unanticipated increase in public debt to GDP ratio. Median public debt to GDP ratio is 40.36 and 42.18 percent in the lower and upper middle-income countries subsamples, respectively. Estimates based on equation (2.2).

Figure 2.16: Response of Real GDP to Public External Debt Shocks Baseline



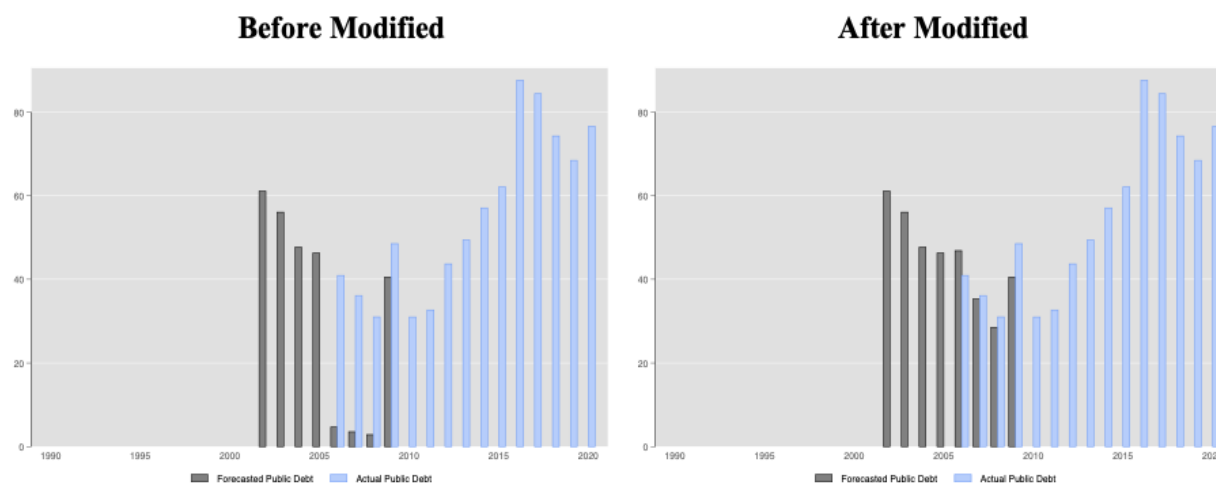
Sources: World Economic Outlook and IMF staff calculation

Note:  $t = 0$  is the year of shock. Red shaded area denotes the 90 percent confidence bands. The solid red line denotes the response to an unanticipated increase in public debt to GDP ratio. Median public external debt to GDP ratio is 27.5 percent. Estimates based on equation (2.2).

Table 2.10: Data Cleaning, Rescaling

Country Name	Correction
Mongolia	Gross debt series multiplied by 10 for vintages 2005-2008

Figure 2.17: Evolution of Public Debt to GDP in Mongolia, Forecast and Actual Data



Sources: World Economic Outlook and IMF staff calculation.

Following [Estefania Flores et al. \(2021\)](#), we rescaled the time series when there was an important shift in the public debt forecast time series. [Table 2.6](#) lists which data points were rescaled and [Figure 2.13](#) illustrates how the adjustment changed the time series before and after the rescaling in the case of Mongolia.

### Converting some Vintages to Missing

Following [Estefania Flores et al. \(2021\)](#), we converted data points to missing data when there was an important shift in the public debt forecast time series and rescaling was not feasible. [Table 2.7](#) lists which data points were converted to missing and [Figure 2.14](#) illustrates how the adjustment changed the time series before and after the conversion to missing data in the case of Angola.

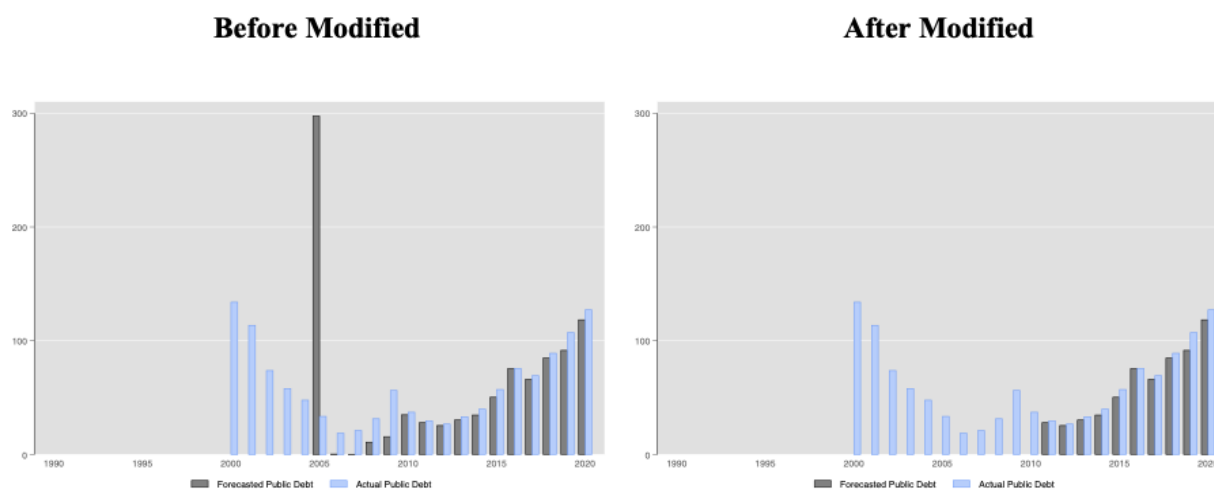
### Dropped All Vintages

Following [Estefania Flores et al. \(2021\)](#), we dropped all vintages when there was an important shift in the public debt forecast time series and rescaling or conversion to missing data points was not feasible. [Table 2.8](#) lists which data points were dropped and [Figure 2.15](#) illustrates how the

Table 2.11: Data Cleaning, Conversion to Missing Data

Country Name	Correction
Angola	All gross debt series converted to missing prior to 2010 vintage
Albania	All gross debt series converted to missing when its equal to 1.000e-13
Hong Kong	All gross debt series converted to missing between the 2011 to 2014
Kazakhstan	All gross debt series converted to missing in 2005
Kyrgyz Republic	All gross debt series converted to missing for in 2014
Niger	All gross debt series converted to missing for all vintages before 2014
Nicaragua	All gross debt series converted to missing for all vintages before 2013
Pakistan	All gross debt series converted to missing for all vintages before 2003
Sweden	All gross debt series converted to missing for all vintages before 2006
Seychelles	All gross debt series converted to missing for all vintages 2005-2020

Figure 2.18: Evolution of Public Debt to GDP in Angola, Forecast and Actual Data

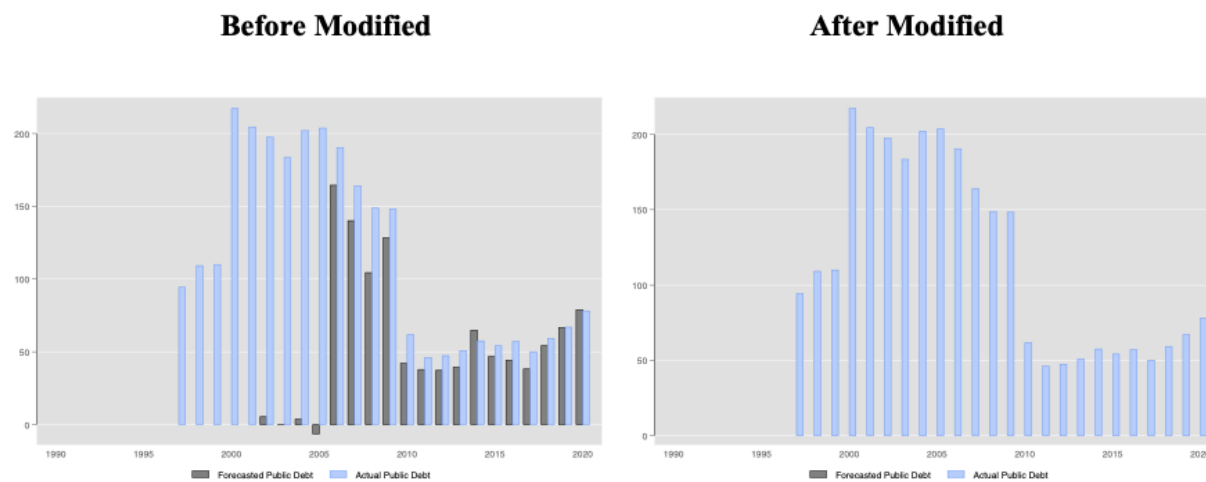


Sources: World Economic Outlook and IMF staff calculation.

Table 2.12: Data Cleaning, Dropping Vintages

Country Name	Correction
Guinea-Bissau	All gross debt series converted to missing.
Mauritania	All gross debt series converted to missing.
Uruguay	All gross debt series converted to missing.

Figure 2.19: Evolution of Public Debt to GDP Ratio in Guinea-Bissau, Forecast and Actual Data



Sources: World Economic Outlook and IMF staff calculation.

adjustment changed the time series before and after the change in the case of Guinea-Bissau.

## Chapter 3

# Usage of Government Expenditure and Growth: Empirical and Theoretical Standpoints

### 3.1 Introduction

$$\frac{\partial \phi}{\partial g_c} = -\frac{\theta}{(\alpha g_p^\beta)^{\frac{1}{\beta-1}} (\theta - \gamma (\eta + \theta + 1) + 1)} < 0$$

$$\frac{\partial \phi}{\partial g_p} = -\frac{\beta \tau - \beta - \beta \theta + g_p \theta - \beta^2 \tau + \beta^2 + \beta g_c \theta}{g_p (\beta - 1) (\alpha g_p^\beta)^{\frac{1}{\beta-1}} (\gamma - \theta + \eta \gamma + \gamma \theta - 1)} \leq 0$$

Larger fiscal deficits have historically concurred with inefficient government expenditures ([Afonso et al. \(2005\)](#); [Afonso and Jalles \(2014\)](#)). Fiscal deficits arise when government expenditures exceed its tax revenue in a persistent manner. They also translate into a surge in public debt that typically occurs during periods of economic turmoil. Allocation of fiscal spending is a more crucial decision under recession to mitigate the economic downturn, therefore understanding the impact of usage of government revenue, primarily stemming from tax or additional public debt, on economic growth is of particular interest among researchers. To address this issue in a general framework,

we construct an endogenous growth model based on [Bruce and Turnovsky \(1999\)](#). We then follow up by analyzing the resulting empirical nuances by applying Bayesian Model Averaging, a novel technique in this context. The approach of incorporating the model uncertainties in the empirical analysis is generally superior to relying on any single model ([Raftery \(1995\)](#); [Draper \(1995\)](#); [Fernandez et al. \(2001\)](#); [Eicher and García-Peñalosa \(2008\)](#); [Eicher et al. \(2011\)](#); [Asatryan and Feld \(2015\)](#)). Therefore, this step is considered innovative since erstwhile results regarding the usage of government revenues for growth are contradictory, stemming primarily from the diverse focus of country sets, lack of model uncertainty incorporation, and improper assessment of disaggregated government revenue/expenditure.

Literature in government expenditure on growth typically selects a single model based on different criteria. The inferences based on the selected model thereby disregards the empirical uncertainty in the model selection approach. The different approaches to modeling have led to diverse implications for the relationship between government expenditure and economic growth. [Chen et al. \(2017\)](#) is one of recent literature that dealt with the topic of government expenditure on growth. They analyze a panel smoothing transitional regression (PSTR) model with flexible parameters to incorporate the heterogeneity among countries across time. They found a possible positive relationship between debt and infrastructure investment on growth as long as the government holds below the threshold; 59.72% debt to GDP ratio and 20.04% public investment to GDP ratio. They show that once public investment hits this threshold or exceeds it, there will be a subsequent negative response to growth (crowding-out effect). [Agell et al. \(2006\)](#) controlled for simultaneity and sample-selection bias before demonstrating that the relationships between government expenditure and tax to GDP ratio on growth are not robust when the initial level of GDP and demographic factors are controlled. To accommodate possible intercountry heterogeneities in structural forms of nature, [Dar and AmirKhalkhali \(2002\)](#) deploy a random coefficient model. They proceed to demonstrate that total government expenditure significantly negatively impacted growth across the year, 1971-1999. [Afonso and Furceri \(2010\)](#) consider IV-estimates to deal with potential endogeneity and reverse causality in the model to claim that the share of government revenue and spendings are detrimental to growth. There is considerable work that accounts for GMM and/or fixed effects to rule out the potential omitted variable bias. They have largely found a negative relationship (occasionally non-significant) between government expenditure and growth ([Miller and Russek \(1997\)](#); [Bellettini and Ceroni \(2000\)](#); [Afonso and Jalles \(2013\)](#)). Even though their selection

of IV is transparent, lagged IV approach may potentially bring bias by increasing RMSE, thereby skyrocketing the type I error when a lagged explanatory variable has some form of a causal relationship with the dependent variable (Wang and Bellemare (2019)). Various stylized results have emphasized that if the government expenditure is used specifically for public investment, there will be a significant positive response in the medium to long run (Romero-Avila and Strauch (2008); Furceri and Li (2017)). On the aspect of government consumption, a vast literature conclude that government consumption negatively affects growth (Cameron (1982); Landau (1986); Hansson and Henrekson (1994); Arin et al. (2019)). These findings highlight the importance of decomposing government expenditure and monitor them closely at the same time.

On the topic of taxation policies, Romer and Romer (2010) found that a tax increase of 1 percent of GDP reduces output by nearly three percent over the subsequent next three years based on their novel measure of fiscal shocks, while Mendoza et al. (1997) found insignificant quantitative response of tax policy on growth based on their theoretical and empirical approaches thereby suggesting its ineffectiveness. On the contrary in the sphere of public debt, Reinhart and Rogoff (2010) claim that there exists potential negative relationship between debt and growth post 90% debt to GDP ratio, while Panizza and Presbitero (2014) use an IV approach to identify public debt shocks and conclude that there is a negative correlation between public debt and real GDP growth for OECD countries, but that the link disappears once endogeneity issues are corrected. Vast literature so far show debt is generally detrimental to economic growth (Myers (1977); Cecchetti et al. (2011); Soyres et al. (2022)). While literature that assesses tax and debt individually reveal a certain indicator for growth, we think that tax and public debt need to be analyzed simultaneously since these are interlinked in the context of the fiscal policy perspective. This is not only feasible but also important because whenever governments want to increase expenditure, they face trade offs in deciding whether to increase tax or issue additional borrowings.

As we see from previous discussion, existing literature so far implicitly assume not only the existence of “true model”, but also rule out different potential model candidates that have been considered in the literature. To deal with the model uncertainties that underpin the diverse findings, we take an initiative to address this issue by using Bayesian Model Averaging. One of the earlier empirical works that incorporated model uncertainty in growth literature is Sala-i Martin (1997) who analyzed growth determinants using Bayesian Averaging over Classical Estimates (BACE) and showed 67 determinants that may influence economic growth. Bergh and Karlsson (2010) then

comprehensively assess the growth effects of fiscal policy. They address the relationship based on OECD panel data using Bayesian Averaging over Classical Estimates (BACE) analysis to deal with model uncertainty. They have found a negative correlation between the size of government and the economy's growth rate. While their findings are remarkable in highlighting an important remedy for negative growth stemming from increase in government expenditure through further economic openness and sound policies, they do not account for the disaggregation of government expenditure that may have opposing implications for growth as we see in [Chen et al. \(2017\)](#). Further, [Arin et al. \(2019\)](#) address model uncertainty using the Bayesian Model Averaging (BMA). Based on OECD countries from 1990-2013, they analyzed how various kinds of government spending in conjunction with an overall budget surplus/deficit affect economic growth. Their remarkable finding of fiscal effects becoming robust with a lag of two years is to be noted. To obtain the medium-term response, we follow their empirical findings by generating every two-year GDP growth rate. They also found that the income tax effect is insignificant, and budget surplus positively impacts growth. However, they did not account for the specific mechanisms by which the disaggregation of usage of fiscal revenues can be directly used as an expenditure that we address in this paper. We believe their choice of diagnosis in budget surplus may mask the effects of public debt, tax, and government expenditure. Therefore, we take an initiative to assess decomposition of fiscal balance to identify which fiscal component(s) will impact growth by following the literature by [Chen et al. \(2017\)](#). While [Chen et al. \(2017\)](#) have comprehensively analyzed the optimal threshold of infrastructure investment and public debt, they accounted less for tax components in their empirical analysis. We took an important step to provide comprehensive reviews of fiscal analysis. Further, the optimal threshold they provided may not properly account for a time-varying threshold that exists under changing circumstances of already high public debt, high inflation rate, and high interest rates; a situation that resonates with current economic conditions. In this paper, we are not trying to assess the optimal threshold, but rather assess the directional growth responses stemming explicitly from the four fiscal components (government consumption, infrastructure investment, tax, and public debt).

Our contribution to the literature can be broadly summarized as follows: we assessed how economic growth may differ by usage of government expenditure (government consumption vs. infrastructure investment) that can arise from fiscal revenues (primarily from either additional tax or additional public debt). To account for model uncertainty that brought the diverse interpretation

in literature, we use Bayesian Model Averaging. Further, we analyzed 180 countries from 1990-2019 to address the diverse findings in literature stemming from different country subsets and periods. We found that regardless of revenue sources, government consumption is detrimental to growth. On the contrary, infrastructure investment arising from additional borrowing has a positive relationship with economic growth. This finding is consistent after accounting for the potential bias in interaction terms presented by [Cuaresma \(2011\)](#). Our theoretical analogy that we will discuss in the next section further clarify vital mechanisms to these directions. These analyses are crucial steps towards measuring the sound policy implication for sustainable growth. Addressing the empirical assessment in incorporating the model uncertainty will provide an important role to fill in the missing pieces that will account for a total of  $2^k$  (assuming  $k$  being the number of regressors) kinds of model selection for a single simulation.

The rest of the paper is organized as follows. Canonical model summary and equilibrium changes are provided in [Section 3.2](#) and [Section 3.3](#). [Section 3.4](#) describes the data and methodology, and [3.5](#) discuss our findings. [Section 3.6](#) deals with additional sensitivity analysis. Finally, [section 2.5](#) concludes. The Appendix presents a detailed description of the data and robustness checks.

## 3.2 Canonical Model of Fiscal Policy and Balanced Growth

### 3.2.1 Fiscal Policy

In this section, we present a small model, the purpose of which is to identify the channels whereby the different fiscal instruments impact the growth rate. The model is adapted from [Bruce and Turnovsky \(1999\)](#), which itself is a modification of the basic [Barro \(1990\)](#) model of endogenous growth.

We consider an economy in which output,  $Y$ , is determined according to a simple Cobb-Douglas technology of the form

$$Y = \alpha G_p^\beta K^{1-\beta} \equiv \alpha (G_p/K)^\beta K \quad 0 \leq \beta \leq 1 \quad (3.1)$$

where  $K$  denotes the private capital stock, assumed to be infinitely durable, and  $G_p$  denotes the flow of services from government spending on the economy's infrastructure. Although the firm

faces diminishing returns to the accumulation of its private capital stock, the overall technology is linearly homogeneous in the two factors of production taken together. This permits an equilibrium in which the economy follows a path of ongoing endogenous growth.

The economy is populated by an infinitely-lived representative agent who maximizes the intertemporal isoelastic utility function

$$U \equiv \int_0^{\infty} \frac{1}{\gamma} \left( C G_c^\eta K^\theta \right)^\gamma e^{-\rho t} dt \quad -\infty < \gamma < 1, \eta > 0, \theta > 0, 1 > \gamma(1 + \eta + \theta) \quad (3.2)$$

where  $C$  denotes private consumption,  $G_c$  denotes the consumption services of a government-provided public good, and the parameter  $\eta$  measures the impact of public consumption on the welfare of the private agent. We assume that both private and public consumption yield positive marginal utility, so that  $\eta > 0$ . In addition, we assume that the household derives utility from its wealth,  $K$ . This term was first introduced into the utility function by [Kurz \(1968\)](#) and subsequently by [Zou \(1994\)](#) and others in their analysis of the “spirit of capitalism”. Our reason for doing so is to demonstrate that it is theoretically the only channel whereby government consumption impacts the equilibrium growth rate. The parameter  $\gamma$  is related to the intertemporal elasticity of substitution (IES),  $s$  say, by  $s = 1/(1 - \gamma)$ . The remaining constraints on the coefficients appearing in (3.2) are required to ensure that the utility function remains concave in the quantities  $C, G_c, K$ .

At any point in time the household faces the flow budget constraint

$$\dot{K} + \dot{B} = (1 - \tau)(rB + Y) - C - T \quad (3.3)$$

where  $B$  denotes the value of government bonds held by the household and  $r$  denotes the real rate on government bonds. The government taxes income at the rate  $\tau$ . We also allow for a lump-sum tax (transfer if negative) denoted by  $T$ . This tax plays the role of a “balancing item” in our analysis, and we do not relate it to any existing part of the government tax structure. All the variables, except for the two taxes, are time dependent, although for notational convenience, the time notation is suppressed. Finally, a dot indicates a time derivative.

The household chooses the time path for consumption, and the accumulation of capital and government bonds, so as to maximize its intertemporal utility function, (3.2), subject to the production function, (3.1), and its flow budget constraint, (3.3). In performing this optimization

of all government fiscal variables, tax rates as well as expenditure levels, are taken as given.

The agent's optimality conditions can be expressed as

$$C^{\gamma-1} G_c^{\eta\gamma} K^{\theta\gamma} = \lambda \quad (3.4a)$$

$$r(1 - \tau) = \theta(C/K) + \alpha(1 - \beta)(1 - \tau)(G_p/K)^\beta = \rho - \dot{\lambda}/\lambda \quad (3.4b)$$

where  $\lambda$  is the marginal utility of wealth. The first condition equates the marginal utility of consumption to the tax-adjusted marginal utility of wealth. The second condition equates the after-tax rate of return on savings, either in the form of bonds or physical capital, to the rate of return on consumption, in accordance with the Keynes-Ramsey consumption rule.<sup>1</sup> In addition, the following transversality conditions must hold

$$\lim_{t \rightarrow \infty} \lambda B e^{-\rho t} = \lim_{t \rightarrow \infty} \lambda K e^{-\rho t} = 0 \quad (3.4c)$$

Taking the time derivative of (3.4a) and combining with (3.4b), leads to the growth relationship

$$(\gamma - 1) \frac{\dot{C}}{C} + \eta\gamma \frac{\dot{G}_c}{G_c} + \theta\gamma \frac{\dot{K}}{K} = \rho - r(1 - \tau) \quad (3.5a)$$

while taking the first equality in (3.4b) implies that the interest rate is determined by

$$r = \frac{\theta(C/K)}{(1 - \tau)} + \alpha(1 - \beta)(G_p/K)^\beta \quad (3.5b)$$

Thus, as the ratio of government infrastructure per unit of capital increases rendering capital more productive, the equilibrium rate of interest must rise. Also, to the extent that capital enhances utility, the interest rate must adjust to reflect that.

---

<sup>1</sup>Ramsey states the Keynes-Ramsey rule as: "rate of saving multiplied by marginal utility of consumption should always equal bliss minus actual rate of utility enjoyed."

### 3.2.2 Balanced Growth

To see the role of the various fiscal instruments, we restrict our attention to fiscal policies that are consistent with a balanced growth equilibrium. Thus, we assume that the income tax rate is constant over time, and that government consumption and infrastructure, though taken as given at any point in time by the household, grow at the same constant rate as output. Specifically, the government is assumed to set its expenditures as constant spending ratios,  $G_c = g_c Y$  and  $G_p = g_p Y$ , where  $0 < g_c, g_p < 1$  are constants through time, so

$$\frac{\dot{Y}}{Y} = \frac{\dot{K}}{K} = \frac{\dot{B}}{B} = \frac{\dot{C}}{C} = \frac{\dot{G}_c}{G_c} = \frac{\dot{G}_p}{G_p} \equiv \phi \quad (3.6)$$

denotes the balanced growth rate at which all real quantities grow. Combining this relationship with (3.5a) immediately yields the following expression for the equilibrium growth rate

$$\phi = \frac{r(1 - \tau) - \rho}{1 - \gamma(1 + \eta + \theta)} \quad (3.7)$$

Also, we can use the aggregate production function to express  $G_p/K = (\alpha g_p)^{1/(1-\beta)}$  and substitute in (3.1) and (3.5b) to express output and the interest rate in the form

$$\frac{Y}{K} = \left( \alpha g_p^\beta \right)^{1/(1-\beta)} \quad (1')$$

$$r(1 - \tau) = \theta(C/K) + (1 - \beta)(1 - \tau) \left( \alpha g_p^\beta \right)^{1/(1-\beta)} \quad (5b')$$

which, with the spending share constant over time, also remains constant. Finally, dividing the economy-wide resource constraint,  $Y = C + G_c + G_p + \dot{K}$ , by  $K$ , we can solve for the constant ratio of consumption to capital, in the form

$$\frac{C}{K} = (1 - g_c - g_p) \left( \alpha g_p^\beta \right)^{1/(1-\beta)} - \phi \quad (3.8)$$

Substituting into (5b'),

$$r(1 - \tau) = \theta \left( (1 - g_c - g_p) \left( \alpha g_p^\beta \right)^{1/(1-\beta)} - \phi \right) + (1 - \beta)(1 - \tau) \left( \alpha g_p^\beta \right)^{1/(1-\beta)} \quad (3.9)$$

and hence equilibrium growth rate is

$$\phi = \frac{[\theta(1 - g_c - g_p) + (1 - \beta)(1 - \tau)] (\alpha g_p^\beta)^{1/(1-\beta)} - \rho}{1 - \gamma(1 + \eta + \theta) + \theta} \quad (3.10)$$

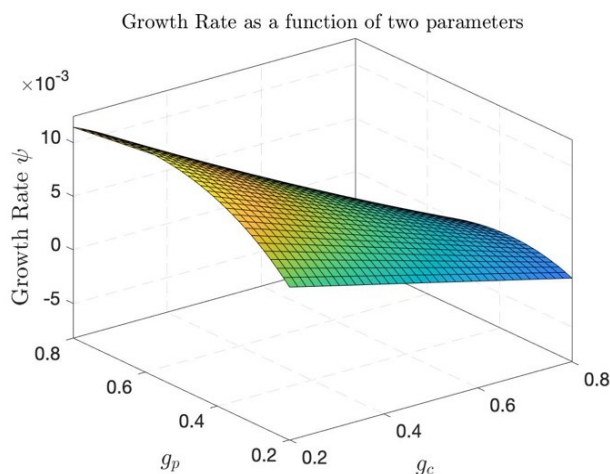
### 3.3 Equilibrium Changes in the Model

The following conclusions regarding fiscal policy can be drawn immediately. These are partial effects and are based on the assumption that the fiscal shock is financed either by lump-sum taxation or by issuing more debt.

- (i) An increase in the income tax rate raises the private consumption ratio and reduces the growth rate.
- (ii) If  $\theta > 0$ , an increase in the government consumption ratio reduces the growth rate. In most endogenous growth models,  $\theta = 0$  in which case government consumption has no effect on the growth rate (see [Eaton \(1981\)](#)). On the other hand, some growth models introduce elastic labor supply, in which case they find that an increase in government consumption increases the growth rate.
- (iii) An increase in the fraction of output spent by the government on infrastructure raises both the real interest rate and the growth rate unless it is too large ([Barro \(1990\)](#)).

The model is more general than much of this literature in that we extend the range of fiscal instruments available to the policy maker thus permitting a more general treatment of fiscal issues. This accounts for differences in some of the results from those obtained in the literature. For example, by assuming that, [Barro \(1990\)](#) identifies two offsetting effects of an increase in government infrastructure expenditure on the growth rate. With our more general fiscal specification, these are now decomposed into the two components, the negative income tax effect in (i) and the positive productivity effect in (iii). On the other hand, the invariance of the growth rate with respect to a bond-financed increase in government consumption expenditure is familiar from early work by [Eaton \(1981\)](#).

Before proceeding further with the analysis, two limitations of the model should be noted.

Figure 3.1: *IRFs* Characterizing the growth rate as a function of  $g_p$  &  $g_c$ 

Note - The remaining parameters were selected following the literature.

First, like [Barro \(1990\)](#), [Rebelo \(1991\)](#), [Ireland \(1994\)](#) we assume that labor is supplied inelastically. This assumption is important in that it implies the neutrality of the consumption tax, noted in (iii) above. An important reason for assuming an inelastic labor supply is to eliminate the static source of supply side effects of tax cuts and focus purely on the dynamic growth effects. Some implications of endogenizing labor supply will be noted in our conclusions. A second assumption of the model is that it assumes complete compliance with the tax code. This ignores the fact that one margin along which agents respond to changes in tax rates is in the resources they devote to tax avoidance activities; see e.g. [Cremer and Gahvari \(1993\)](#), [Fullerton and Karayannis \(1994\)](#). Both of these effects could modify our results and should be kept in mind when considering them.

## 3.4 Empirical Methodology

### 3.4.1 Methodology (Bayesian Model Averaging)

In this section, we briefly discuss an empirical growth model and Bayesian Model Averaging to illustrate the methodology we follow. The stylized relationship between the usage of government revenues on growth in literature is mixed based on its specifications of government expenditure and assumption of models. The vastly different results in the literature on government expenditure

on growth range from a positive relationship to growth (Chen et al. (2017)), insignificant or no remarkable responses (Agell et al. (1997)), to a negative relationship between the two (Fölster and Henrekson (2001); Bergh and Karlsson (2010); Afonso and Jalles (2014)). Primarily, infrastructure investment seems to be a driver of positive contribution to growth (Romero-Avila and Strauch (2008); Furceri and Li (2017)). These variety of interpretations are stemming from different focuses on country sets, government expenditure components (i.e., infrastructure investment, government consumption), source of government revenues (i.e., tax or public debt), and lack of model uncertainty incorporation (Fernandez et al. (2001); Barro and Sala-i Martin (2004); Bergh and Karlsson (2010); Arin et al. (2019)). Therefore, we incorporate them to have a comprehensive assessment regarding how the usage of government revenue will impact growth.

Rather than generating regressions relying on a single model to test the robustness of findings, we use Bayesian Model Averaging. Our interest is in the unconditional posterior distribution of a parameter vector given data with respect to specific models. Suppose we have  $j$  candidate of models  $M_j$  where  $j \in (1, J)$  with corresponding model parameters,  $\theta_j$ . For each model of the prior distribution  $p(M_j)$ , we have a prior distribution for each parameter  $p(\theta_j|M_j)$ . Then the posterior probability of any one of  $j$  models can be expressed as

$$p(M_j|D) = \frac{p(D|M_j)p(M_j)}{\sum_{h=1}^J p(D|M_h)p(M_h)}$$

where  $p(D|M_j)$  is the conditional probability generated by the model, which in terms of integrated likelihood is equivalent to  $\int p(D|\theta_j, M_j)p(\theta_j|M_j)d\theta_j$ , and  $p(M_j)$  is the prior probability of  $J$  being the true model. This setup of Bayesian Model Averaging allows the selection of weights to favor good models over weak ones, while results incorporate the essence of data, as each weight was selected based on all possible model and the goodness of fit (or adjusted  $R^2$ ) performance of each individual model (Raftery (1995); Barro and Sala-i Martin (2004); Bergh and Karlsson (2010); Arin et al. (2019)). Further, we average parameter estimates obtained using fixed effect. Among the  $H$  possible regressors, there is a set of  $J = 2^H$  models  $M_j$  where  $j \in (1, J)$ , for which each is a particular linear combination of growth determinants (see Raftery (1995); Draper (1995); Fernandez et al. (2001); Eicher and García-Peñalosa (2008) for details).

### 3.4.2 Our Approaches in Empirical Model Averaging

As an illustration, consider the following equation

$$\Delta_2 y_{i,t} = a + bx + cD + dInt(D, x) + \theta' X_{i,t-s} + e_{i,t} \quad (3.11)$$

where  $x$  is a continuous variable (i.e., tax and public debt in our model), and  $D$  is a dummy variable (i.e.,  $G_p$  (infrastructure investment to GDP ratio) and  $G_c$  (government consumption to GDP ratio) in our model), where dummy is “1” when the two-year growth rate of  $G_p$  or  $G_c$  is greater than the median of its growth rate, otherwise zero, and  $Int(\cdot)$  are interaction between the two terms.

Then, if  $x$  increases by one-unit, main effects can be calculated by

$$\begin{cases} d\Delta_2 y_{i,t} = b + c & \text{if } D = 1 \\ d\Delta_2 y_t = b & \text{if } D = 0 \end{cases}$$

Further, if  $x$  increases by one-unit, total effects of an interaction term can be calculated by

$$\begin{cases} \Delta_2 y_{i,t} = \hat{b} + \hat{c} + \hat{d} & \text{if } D = 1 \\ \Delta_2 y_t = \hat{b} & \text{if } D = 0 \end{cases}$$

The growth rate of country  $I = 1, \dots, N$  in year  $t = 0, \dots, T$  is given by

$$\Delta_2 y_{i,t} = a + bx + cD + dInt(D, x) + \theta' x_{i,t-s} + \nu_{i,t} \quad \text{Where } \Delta_2 y_{i,t} = y_{i,t} - y_{i,t-2}$$

We constructed the growth rate in every two years following the recent findings of most fiscal measures identified as robust growth determinants within two years (Arin et al. (2019)). The error term consists of:  $\nu_{i,t} = \mu_i + \varepsilon_{i,t}$ , where  $\mu_i$  is a country specific effects, and  $\varepsilon_{i,t}$  is idiosyncratic errors.  $a + bx + cD + dInt(D, x)$  is the key term that we are interested in testing jointly and individually.  $x_{i,t-s}$  are control variables, which includes the log of GDP per capita at  $t - 1$  to control for initial condition. We closely followed the selection of control variables based on the most recent literature Arin et al. (2019) that selected growth determinants following the existing growth literature (Kneller et al. (1999); and Lee and Gordon (2005); Durlauf et al. (2008)).

To identify the flow components of infrastructure investment and government consumption overtime, we generated the dummy variable based on the 2-year growth rate of its variable; if the 2-year growth rate is greater than the median of its own, then we have set it as 1, otherwise 0. Then we have interacted it with variables like 2-year lags of government debt to GDP ratio and 2-year lags of tax to GDP ratio. Literature addresses the way to carefully apply interaction terms in Bayesian Model Averaging framework. [Masanjala and Papageorgiou \(2008\)](#) compared growth determinants of Africa vs. the rest of the world by adding dummy and interaction terms in a single generation based on unit information prior (UIP), then found that relevant growth variables for Africa are quite different, supporting the views by [Brock and Durlauf \(2001\)](#). While there should not be statistical problems, [Cuaresma \(2011\)](#) concerns with potential unobservable that the dummy variable may capture and suggest testing BMA without interactions or models with strong heredity prior. For completeness and transparency, we have addressed this as a sensitivity analysis and carefully diagnosed the results in our interpretation. Our results did not largely change by these additional sensitivity checks.

### 3.4.3 Data Description

The analysis mainly uses the International Monetary Fund (IMF) and World Bank (WB) datasets for fiscal variables, growth determinants, and GDP per capita data, covering 180 countries from 1990 to 2019. For the income tax data, we used the data from the Organization for Economic Cooperation and Development (OECD). Other data sources of control variables are primarily obtained by closely following the literature on other growth determinants ([Barro \(1990\)](#); [Sala-i Martin \(1997\)](#); [Barro and Sala-i Martin \(2004\)](#); [Arin et al. \(2019\)](#)). Detailed descriptions of the origins of other control variables we used in this analysis are listed in the Appendix of [Arin et al. \(2019\)](#).

## 3.5 Baseline Results of Government Borrowings on Growth

As [Arin et al. \(2019\)](#) illustrates, most effects of government expenditure occur with a lag of two years under their analysis of Bayesian Model Averaging, therefore, we looked at growth rate for every two years. [Table 3.1](#) and [Table 3.2](#) eliminate all control variables that literature indicates essential to first isolate the effects of our variable of interest. Before moving on to the growth model,

testing the individual effects are essential in BMA analysis based on existing results from previous literature (Cuaresma (2011) and Masanjala and Papageorgiou (2008)). Further, we address how the joint selection of two instruments viz.  $G_p$ ,  $G_c$  or tax, debt may lead to the total effect on economic growth through the interaction terms in the Bayesian Model Averaging framework.

Looking at individual effects of interaction terms, we see that both debt and income tax are detrimental to growth in general. However, tracing the usage of this additional fiscal spending, when this additional debt is used for infrastructure investment, total effect indicate there is a possible positive impact on growth rate regardless of fiscal revenue sources. On the contrary, if this additional debt is used for government consumption side, the impact is consistently negative on growth. Our data for 180 countries reveal a consistent trend among the different specifications.

In the sphere of usage of government tax, we found that effect can be mixed as we observe a possible opposing directional response by an increase in tax spent on government consumption based on incorporation of country fixed effect. Interestingly, we observe a largely similar trend as the case of additional borrowings; if the additional tax were to be used for infrastructure investment, then response on growth is positive, while if it were to be used for government consumption, response can be possibly negative over two years. This illustrates government consumption will generally hurt growth in two years. Robustness can be indicated in bold whenever posterior inclusion probability (PIP) being greater than 0.5. Posterior inclusion probability (PIP) indicates the ranking measure in importance of inclusion of a variable in the regression. Aligning with the approach by Arin et al. (2019), we use a hierarchical binomial beta prior that are found to be less sensitive to prior choice according to Ley and Steel (2009). Our findings are largely consistent to Arin et al. (2019), while our results are generally more robust by including over 150 additional countries. Furthermore, to expand our horizon beyond the scope of their findings, we have assessed the usage of fiscal revenues on growth using interaction terms in the subsequent sections.

Once we analyzed how the interaction terms indicate the possible relationship between the usage of government revenues on economic growth rate, in Table 3.3 and Table 3.4, we incorporated them into the framework of economic growth model that is widely used in the literature on growth determinants (Barro (1990); Barro and Sala-i Martin (2004); Arin et al. (2019)). We closely followed the selection of control variables listed by Arin et al. (2019). Aligning with our findings in Table 3.1 and Table 3.2, we found that increase in infrastructure investment is strictly positively impact on

Table 3.1: OLS Bayesian Model Averaging: Joint Effect and Cuaresma (2011) Sensitivity Analysis

<b>Pooled OLS</b>			<b>Pooled OLS: w/o interaction</b>		
<b>Debt X Gp</b>			<b>Debt X Gp</b>		
	Mean	PIP		Mean	PIP
Gp.dummy	8.94E-04	0.03397242	Gp.dummy	1.04E-03	0.04562214
debt X Gp.dummy	-3.19E-06	0.01856692	lag2_Debt_GDP	-5.22E-08	0.02176037
lag2_Debt_GDP	5.02E-07	0.01545838			
<b>Total effect:</b>	<b>8.91E-04</b>		<b>Total effect:</b>	<b>1.04E-03</b>	
<b>Tax X Gp</b>			<b>Tax X Gp</b>		
	Mean	PIP		Mean	PIP
<b>lag2_Tax_GDP</b>	<b>-0.002764</b>	<b>0.6729317</b>	<b>lag2_Tax_GDP</b>	<b>-0.0035951</b>	<b>0.8524742</b>
tax X Gp.dummy	-0.0010418	0.2184748	Gp.dummy	0.00245493	0.1440291
Gp.dummy	0.00981338	0.1873919			
<b>Total effect:</b>	<b>6.01E-03</b>		<b>Total effect:</b>	<b>-1.14E-03</b>	
<b>Debt X Gc</b>			<b>Debt X Gc</b>		
	Mean	PIP		Mean	PIP
debt X Gc.dummy	-7.17E-05	0.15541131	Gc.dummy	-8.56E-04	0.04118831
Gc.dummy	-4.29E-04	0.02828265	lag2_Debt_GDP	-4.02E-06	0.03202469
lag2_Debt_GDP	-1.83E-06	0.02309064			
<b>Total effect:</b>	<b>-5.02E-04</b>		<b>Total effect:</b>	<b>-8.60E-04</b>	
<b>Tax X Gc</b>			<b>Tax X Gc</b>		
	Mean	PIP		Mean	PIP
<b>lag2_Tax_GDP</b>	<b>-0.0034753</b>	<b>0.80236519</b>	<b>lag2_Tax_GDP</b>	<b>-0.003799</b>	<b>0.8776336</b>
tax X Gc.dummy	-0.0001264	0.09187128	Gc.dummy	-0.0009409	0.1167175
Gc.dummy	-0.0004409	0.07229917			
<b>Total effect:</b>	<b>-4.04E-03</b>		<b>Total effect:</b>	<b>-4.74E-03</b>	

Note - Dependent variable: every two years of GDP growth rate. To capture the flow components of fiscal expenditure, we generated the dummy variable of infrastructure investment ( $G_p$ ) and government consumption ( $G_c$ ) based on the 2-year growth rate; if the growth rate is greater than the median of sample, then we have set it as 1, otherwise 0. Further, to address the long run implication without temporal noise, we calculate the five-year average in sample year before running the regression following the methodology by Arin et al. (2019).

Table 3.2: Fixed Effect Bayesian Model Averaging: Joint Effect and [Cuaresma \(2011\)](#) Sensitivity Analysis

<p><b>Country Fixed Effect</b></p> <p><b>Debt X Gp</b></p> <table border="1"> <thead> <tr> <th></th> <th>Mean</th> <th>PIP</th> </tr> </thead> <tbody> <tr> <td><b>Gp_dummy</b></td> <td><b>0.08052851</b></td> <td><b>1</b></td> </tr> <tr> <td>debt X Gp_dummy</td> <td>-0.0002398</td> <td>0.4223042</td> </tr> <tr> <td>lag2_Debt_GDP</td> <td>0.00014098</td> <td>0.3634615</td> </tr> <tr> <td><b>Total effect:</b></td> <td><b>8.03E-02</b></td> <td></td> </tr> </tbody> </table> <p><b>Tax X Gp</b></p> <table border="1"> <thead> <tr> <th></th> <th>Mean</th> <th>PIP</th> </tr> </thead> <tbody> <tr> <td><b>Gp_dummy</b></td> <td><b>0.14033589</b></td> <td><b>1</b></td> </tr> <tr> <td>tax X Gp_dummy</td> <td>-0.0104303</td> <td>0.99993453</td> </tr> <tr> <td>lag2_Tax_GDP</td> <td>0.00014474</td> <td>0.09629178</td> </tr> <tr> <td><b>Total effect:</b></td> <td><b>1.30E-01</b></td> <td></td> </tr> </tbody> </table> <p><b>Debt X Gc</b></p> <table border="1"> <thead> <tr> <th></th> <th>Mean</th> <th>PIP</th> </tr> </thead> <tbody> <tr> <td><b>Gc_dummy</b></td> <td><b>-1.71E-02</b></td> <td><b>0.6444767</b></td> </tr> <tr> <td>debt X Gc_dummy</td> <td>-5.80E-06</td> <td>0.03645746</td> </tr> <tr> <td>lag2_Debt_GDP</td> <td>5.38E-07</td> <td>0.01469132</td> </tr> <tr> <td><b>Total effect:</b></td> <td><b>-1.71E-02</b></td> <td></td> </tr> </tbody> </table> <p><b>Tax X Gc</b></p> <table border="1"> <thead> <tr> <th></th> <th>Mean</th> <th>PIP</th> </tr> </thead> <tbody> <tr> <td>tax X Gc_dummy</td> <td>1.63E-05</td> <td>0.01699274</td> </tr> <tr> <td>Gc_dummy</td> <td>8.86E-05</td> <td>0.01411139</td> </tr> <tr> <td>lag2_Tax_GDP</td> <td>-1.93E-05</td> <td>0.0125085</td> </tr> <tr> <td><b>Total effect:</b></td> <td><b>1.05E-04</b></td> <td></td> </tr> </tbody> </table>		Mean	PIP	<b>Gp_dummy</b>	<b>0.08052851</b>	<b>1</b>	debt X Gp_dummy	-0.0002398	0.4223042	lag2_Debt_GDP	0.00014098	0.3634615	<b>Total effect:</b>	<b>8.03E-02</b>			Mean	PIP	<b>Gp_dummy</b>	<b>0.14033589</b>	<b>1</b>	tax X Gp_dummy	-0.0104303	0.99993453	lag2_Tax_GDP	0.00014474	0.09629178	<b>Total effect:</b>	<b>1.30E-01</b>			Mean	PIP	<b>Gc_dummy</b>	<b>-1.71E-02</b>	<b>0.6444767</b>	debt X Gc_dummy	-5.80E-06	0.03645746	lag2_Debt_GDP	5.38E-07	0.01469132	<b>Total effect:</b>	<b>-1.71E-02</b>			Mean	PIP	tax X Gc_dummy	1.63E-05	0.01699274	Gc_dummy	8.86E-05	0.01411139	lag2_Tax_GDP	-1.93E-05	0.0125085	<b>Total effect:</b>	<b>1.05E-04</b>		<p><b>w/o interaction</b></p> <p><b>Debt X Gp</b></p> <table border="1"> <thead> <tr> <th></th> <th>Mean</th> <th>PIP</th> </tr> </thead> <tbody> <tr> <td><b>Gp_dummy</b></td> <td><b>6.71E-02</b></td> <td><b>1</b></td> </tr> <tr> <td>lag2_Debt_GDP</td> <td>6.43E-06</td> <td>0.05899015</td> </tr> <tr> <td><b>Total effect:</b></td> <td><b>6.71E-02</b></td> <td></td> </tr> </tbody> </table> <p><b>Tax X Gp</b></p> <table border="1"> <thead> <tr> <th></th> <th>Mean</th> <th>PIP</th> </tr> </thead> <tbody> <tr> <td><b>Gp_dummy</b></td> <td><b>0.05717113</b></td> <td><b>0.99999432</b></td> </tr> <tr> <td>lag2_Tax_GDP</td> <td>-0.0001647</td> <td>0.07535043</td> </tr> <tr> <td><b>Total effect:</b></td> <td><b>5.70E-02</b></td> <td></td> </tr> </tbody> </table> <p><b>Debt X Gc</b></p> <table border="1"> <thead> <tr> <th></th> <th>Mean</th> <th>PIP</th> </tr> </thead> <tbody> <tr> <td><b>Gc_dummy</b></td> <td><b>-1.97E-02</b></td> <td><b>0.74411764</b></td> </tr> <tr> <td>lag2_Debt_GDP</td> <td>6.79E-07</td> <td>0.02731722</td> </tr> <tr> <td><b>Total effect:</b></td> <td><b>-1.97E-02</b></td> <td></td> </tr> </tbody> </table> <p><b>Tax X Gc</b></p> <table border="1"> <thead> <tr> <th></th> <th>Mean</th> <th>PIP</th> </tr> </thead> <tbody> <tr> <td>Gc_dummy</td> <td>1.40E-04</td> <td>0.02107061</td> </tr> <tr> <td>lag2_Tax_GDP</td> <td>-2.83E-05</td> <td>0.01855332</td> </tr> <tr> <td><b>Total effect:</b></td> <td><b>1.12E-04</b></td> <td></td> </tr> </tbody> </table>		Mean	PIP	<b>Gp_dummy</b>	<b>6.71E-02</b>	<b>1</b>	lag2_Debt_GDP	6.43E-06	0.05899015	<b>Total effect:</b>	<b>6.71E-02</b>			Mean	PIP	<b>Gp_dummy</b>	<b>0.05717113</b>	<b>0.99999432</b>	lag2_Tax_GDP	-0.0001647	0.07535043	<b>Total effect:</b>	<b>5.70E-02</b>			Mean	PIP	<b>Gc_dummy</b>	<b>-1.97E-02</b>	<b>0.74411764</b>	lag2_Debt_GDP	6.79E-07	0.02731722	<b>Total effect:</b>	<b>-1.97E-02</b>			Mean	PIP	Gc_dummy	1.40E-04	0.02107061	lag2_Tax_GDP	-2.83E-05	0.01855332	<b>Total effect:</b>	<b>1.12E-04</b>	
	Mean	PIP																																																																																																											
<b>Gp_dummy</b>	<b>0.08052851</b>	<b>1</b>																																																																																																											
debt X Gp_dummy	-0.0002398	0.4223042																																																																																																											
lag2_Debt_GDP	0.00014098	0.3634615																																																																																																											
<b>Total effect:</b>	<b>8.03E-02</b>																																																																																																												
	Mean	PIP																																																																																																											
<b>Gp_dummy</b>	<b>0.14033589</b>	<b>1</b>																																																																																																											
tax X Gp_dummy	-0.0104303	0.99993453																																																																																																											
lag2_Tax_GDP	0.00014474	0.09629178																																																																																																											
<b>Total effect:</b>	<b>1.30E-01</b>																																																																																																												
	Mean	PIP																																																																																																											
<b>Gc_dummy</b>	<b>-1.71E-02</b>	<b>0.6444767</b>																																																																																																											
debt X Gc_dummy	-5.80E-06	0.03645746																																																																																																											
lag2_Debt_GDP	5.38E-07	0.01469132																																																																																																											
<b>Total effect:</b>	<b>-1.71E-02</b>																																																																																																												
	Mean	PIP																																																																																																											
tax X Gc_dummy	1.63E-05	0.01699274																																																																																																											
Gc_dummy	8.86E-05	0.01411139																																																																																																											
lag2_Tax_GDP	-1.93E-05	0.0125085																																																																																																											
<b>Total effect:</b>	<b>1.05E-04</b>																																																																																																												
	Mean	PIP																																																																																																											
<b>Gp_dummy</b>	<b>6.71E-02</b>	<b>1</b>																																																																																																											
lag2_Debt_GDP	6.43E-06	0.05899015																																																																																																											
<b>Total effect:</b>	<b>6.71E-02</b>																																																																																																												
	Mean	PIP																																																																																																											
<b>Gp_dummy</b>	<b>0.05717113</b>	<b>0.99999432</b>																																																																																																											
lag2_Tax_GDP	-0.0001647	0.07535043																																																																																																											
<b>Total effect:</b>	<b>5.70E-02</b>																																																																																																												
	Mean	PIP																																																																																																											
<b>Gc_dummy</b>	<b>-1.97E-02</b>	<b>0.74411764</b>																																																																																																											
lag2_Debt_GDP	6.79E-07	0.02731722																																																																																																											
<b>Total effect:</b>	<b>-1.97E-02</b>																																																																																																												
	Mean	PIP																																																																																																											
Gc_dummy	1.40E-04	0.02107061																																																																																																											
lag2_Tax_GDP	-2.83E-05	0.01855332																																																																																																											
<b>Total effect:</b>	<b>1.12E-04</b>																																																																																																												

Note - Dependent variable: every two years of GDP growth rate. To capture the flow components of fiscal expenditure, we generated the dummy variable of infrastructure investment ( $G_p$ ) and government consumption ( $G_c$ ) based on the 2-year growth rate; if the growth rate is greater than the median of sample, then we have set it as 1, otherwise 0. Further, to address the long run implication without temporal noise, we calculate the five-year average in sample year before running the regression following the methodology by [Arin et al. \(2019\)](#).

Table 3.3: Fiscal Revenue to Infrastructure Investment (OLS: Growth Model on Bayesian Model Averaging)

<b>Pooled OLS</b>			<b>Pooled OLS: w/o interaction</b>		
<b>Debt X Gp</b>			<b>Debt X Gp</b>		
	Mean	PIP		Mean	PIP
<b>laglny1</b>	<b>-1.25E-02</b>	<b>0.60467462</b>	<b>laglny1</b>	<b>-1.25E-02</b>	<b>0.60585792</b>
lag_Lifeexp	8.30E+00	0.42467542	lag_Lifeexp	8.94E+00	0.45226923
lag_TFR	-5.48E-03	0.24218901	lag_TFR	-6.37E-03	0.27917969
lag_Corruption	-2.78E-03	0.17910066	lag_Corruption	-2.88E-03	0.19041866
Gp_dummy	7.16E-03	0.15190674	Gp_dummy	7.60E-03	0.16802701
lag_Inflation	2.52E-04	0.13597056	lag_Trade	4.42E-05	0.1559567
lag_Trade	3.78E-05	0.13390534	lag_Inflation	2.73E-04	0.14953983
lag_Yrschool	-1.51E-05	0.03511894	lag_Yrschool	-1.72E-05	0.04066233
lag_population	1.12E-12	0.02869985	lag_population	1.30E-12	0.03382209
debt X Gp_dummy	-4.19E-06	0.02390315	lag2_Debt_GDP	-3.78E-06	0.0265651
lag2_Debt_GDP	-3.29E-06	0.02289766			
<b>Total effect:</b>	<b>7.15E-03</b>		<b>Total effect:</b>	<b>7.59E-03</b>	
<b>Tax X Gp</b>			<b>Tax X Gp</b>		
	Mean	PIP		Mean	PIP
<b>lag_Lifeexp</b>	<b>5.78E+01</b>	<b>0.92618377</b>	<b>lag_Lifeexp</b>	<b>5.76E+01</b>	<b>0.92336473</b>
lag_Trade	3.61E-05	0.1299829	lag_Trade	3.98E-05	0.1433578
laglny1	-3.55E-03	0.10990672	laglny1	-3.77E-03	0.11820669
lag_Corruption	-7.83E-04	0.08186637	lag_Corruption	-8.81E-04	0.09201122
lag2_Tax_GDP	-1.90E-04	0.07944192	lag2_Tax_GDP	-2.13E-04	0.08935106
lag_Inflation	-1.10E-04	0.0752698	lag_Inflation	-1.23E-04	0.08396768
lag_population	-7.99E-12	0.0499653	lag_population	-8.88E-12	0.05582195
Gp_dummy	1.40E-03	0.04448608	Gp_dummy	1.10E-03	0.04383789
tax X Gp_dummy	-7.07E-05	0.03551023	lag_Yrschool	-1.35E-05	0.03472384
lag_Yrschool	-1.15E-05	0.0302227	lag_TFR	-3.34E-04	0.02951718
lag_TFR	-3.00E-04	0.02617914			
<b>Total effect:</b>	<b>1.14E-03</b>		<b>Total effect:</b>	<b>8.83E-04</b>	

Note - Dependent variable: every two years of GDP growth rate. To capture the flow components of fiscal expenditure, we generated the dummy variable of infrastructure investment ( $G_p$ ) and government consumption ( $G_c$ ) based on the 2-year growth rate; if the growth rate is greater than the median of sample, then we have set it as 1, otherwise 0. Further, to address the long run implication without temporal noise, we calculate the five-year average in sample year before running the regression following the methodology by [Arin et al. \(2019\)](#).

Table 3.4: Fiscal Revenue to Infrastructure Investment (Fixed Effect: Growth Model on Bayesian Model Averaging)

<b>Country Fixed Effect</b>			<b>w/o interaction</b>		
<b>Debt X Gp</b>			<b>Debt X Gp</b>		
	Mean	PIP		Mean	PIP
lag_Lifeexp	7.46E+01	1	lag_Lifeexp	7.56E+01	1
lag_TFR	-7.30E-02	1	Gp_dummy	6.82E-02	1
Gp_dummy	9.85E-02	1	lag_TFR	-7.60E-02	1
lag_Corruption	-3.22E-02	0.99999988	lag_Corruption	-3.30E-02	0.9999999
lag_Trade	1.09E-03	0.99999972	lag_Yrschool	-2.00E-03	0.9999998
lag_Yrschool	-1.98E-03	0.9999997	lag_Trade	1.10E-03	0.9999996
debt X Gp_dummy	-5.72E-04	0.92852215	laglny1	-2.30E-02	0.8695442
laglny1	-2.48E-02	0.91195823	lag_population	3.99E-10	0.8536138
lag_population	4.38E-10	0.89302631	lag2_Debt_GDP	-2.62E-04	0.7376265
lag2_Debt_GDP	-2.15E-05	0.14117426	lag_Inflation	1.04E-05	0.1418442
lag_Inflation	9.73E-07	0.09831248			
<b>Total effect:</b>	<b>9.79E-02</b>		<b>Total effect:</b>	<b>6.79E-02</b>	
<b>Tax X Gp</b>			<b>Tax X Gp</b>		
	Mean	PIP		Mean	PIP
Gp_dummy	1.51E-01	1	Gp_dummy	7.05E-02	1
lag_Corruption	-3.37E-02	1	lag_Corruption	-3.36E-02	1
lag_Lifeexp	1.12E+02	0.99999986	lag_Lifeexp	1.21E+02	0.99999999
tax X Gp_dummy	-1.02E-02	0.99997421	lag_TFR	-9.14E-02	0.99972651
lag_TFR	-9.26E-02	0.99901451	lag2_Tax_GDP	-1.29E-02	0.9976022
lag2_Tax_GDP	-7.83E-03	0.84098896	lag_Yrschool	-3.81E-04	0.45996151
lag_Yrschool	-7.55E-04	0.76680809	lag_Trade	1.78E-04	0.40460654
lag_Trade	2.02E-04	0.46211445	laglny1	-5.46E-03	0.22614341
laglny1	-8.01E-03	0.31047463	lag_Inflation	-5.57E-05	0.10377459
lag_Inflation	-4.95E-05	0.11491954	lag_population	5.96E-12	0.06636804
lag_population	3.20E-11	0.09677815			
<b>Total effect:</b>	<b>1.33E-01</b>		<b>Total effect:</b>	<b>5.77E-02</b>	

Note - Dependent variable: every two years of GDP growth rate. To capture the flow components of fiscal expenditure, we generated the dummy variable of infrastructure investment ( $G_p$ ) and government consumption ( $G_c$ ) based on the 2-year growth rate; if the growth rate is greater than the median of sample, then we have set it as 1, otherwise 0. Further, to address the long run implication without temporal noise, we calculate the five-year average in sample year before running the regression following the methodology by [Arin et al. \(2019\)](#).

growth regardless of fiscal sources. Once we accounted the country fixed effect, this response becomes robust. This finding is persistent when we assess the sensitivity analysis introduced by [Cuaresma \(2011\)](#). We incorporated both OLS and fixed effect results because if there is correlation between unobservable country characteristics and control variables, then coefficients of the variable of interest will be biased unless the model properly controls for the country characteristics. On the contrary, fixed-effects approach also eliminates all cross-country variation and some of this variation may contain valuable information about the relationship between the growth rate and our variable of interests. We therefore explore both of these approaches for transparency and comprehensiveness in line with the existing literature.

Interestingly, we observe a mild negative association between the years of education and growth rate. [Barro \(2001\)](#) illustrates the insignificant relationship between growth rate and the years of school attainment of females at the secondary and higher levels. This may contribute to our results of near zero responses by an increase in years of education. Further, the total fertility rate and higher corruption index are signaled negatively affecting the growth rate aligning with literature ([Mauro \(1995\)](#); [Barro \(1991\)](#); [Barro and Becker \(1989\)](#)).

On the aspect of the usage of government revenue on government consumption in [Table 3.5](#) and [Table 3.6](#), we also found a remarkable consistent trend following the [Table 3.1](#) and [Table 3.2](#) in terms of both public debt and income tax. If additional borrowings/income were to be spent on government consumption, then it will directly hurt growth. This result is consistent to our theoretical findings in the section [3.2](#) and [3.3](#). These findings are remarkable in terms of policy perspectives. Our findings largely align to the existing literature on this same issue ([Cameron \(1982\)](#); [Landau \(1986\)](#); [Hansson and Henrekson \(1994\)](#); [Arin et al. \(2019\)](#)).

Besides fiscal components, reciprocal life expectancy is significantly associated with economic growth. Once a country fixed effect is incorporated, this trend becomes more robust persistently. These findings can also be confirmed in [Arin et al. \(2019\)](#). Robust results in various fiscal components may indicate that changes in fiscal expenditure generally have a sound effect on economic growth, strengthening the policy implication.

These findings highlight the importance of focusing on infrastructure investment for future economic development. Further, a generally greater magnitude of posterior mean in our results indicates that additional government borrowing spending on government consumption tends to

Table 3.5: Fiscal Revenue to Government Consumption (OLS: Growth Model on Bayesian Model Averaging)

<b>Pooled OLS</b>			<b>Pooled OLS</b>		
<b>Debt X Gc</b>			<b>Debt X Gc</b>		
	Mean	PIP		Mean	PIP
<b>laglny1</b>	<b>-1.06E-02</b>	<b>0.5574394</b>	<b>laglny1</b>	<b>-1.06E-02</b>	<b>0.55718769</b>
lag_Corruption	-7.49E-03	0.40771023	lag_Corruption	-7.49E-03	0.41357948
lag_TFR	-5.08E-03	0.23007566	lag_TFR	-5.92E-03	0.26411101
lag_Lifeexp	3.80E+00	0.21100448	lag_Lifeexp	4.37E+00	0.2403437
lag_Trade	1.21E-05	0.05582819	lag_Trade	1.40E-05	0.06472569
lag_Inflation	7.76E-05	0.05412905	lag_Inflation	8.54E-05	0.06049619
Gc_dummy	-1.69E-03	0.0503727	<b>Gc_dummy</b>	<b>-2.01E-03</b>	<b>0.0597477</b>
debt X Gc_dummy	-8.79E-06	0.02710534	lag_Yrschool	-9.74E-06	0.03042626
lag_Yrschool	-8.18E-06	0.02638928	lag_population	1.09E-12	0.02748353
lag_population	9.66E-13	0.02390169	lag2_Debt_GDP	-2.60E-06	0.02029294
lag2_Debt_GDP	-2.28E-06	0.0174595			
<b>Total effect:</b>	<b>-1.70E-03</b>		<b>Total effect:</b>	<b>-2.01E-03</b>	
<b>Tax X Gc</b>			<b>Tax X Gc</b>		
	Mean	PIP		Mean	PIP
<b>lag_Lifeexp</b>	<b>6.01E+01</b>	<b>0.93952624</b>	<b>lag_Lifeexp</b>	<b>5.99E+01</b>	<b>0.93763112</b>
lag_Trade	3.53E-05	0.12901109	lag_Trade	3.89E-05	0.14242251
lag_Corruption	-1.07E-03	0.10453628	lag_Corruption	-1.19E-03	0.11686202
laglny1	-3.10E-03	0.09746518	laglny1	-3.28E-03	0.10493405
lag_Inflation	-1.42E-04	0.09176984	lag_Inflation	-1.57E-04	0.10224833
lag2_Tax_GDP	-1.80E-04	0.07649151	lag2_Tax_GDP	-2.02E-04	0.08602854
lag_population	-7.84E-12	0.04941996	lag_population	-8.78E-12	0.0555963
tax X Gc_dummy	-5.46E-05	0.03193893	lag_Yrschool	-1.22E-05	0.0333124
lag_TFR	-3.87E-04	0.02911928	lag_TFR	-4.30E-04	0.03283563
lag_Yrschool	-1.05E-05	0.02895372	Gc_dummy	-2.96E-05	0.02202349
Gc_dummy	-3.24E-06	0.01958487			
<b>Total effect:</b>	<b>-2.38E-04</b>		<b>Total effect:</b>	<b>-2.31E-04</b>	

Note - Dependent variable: every two years of GDP growth rate. To capture the flow components of fiscal expenditure, we generated the dummy variable of infrastructure investment ( $G_p$ ) and government consumption ( $G_c$ ) based on the 2-year growth rate; if the growth rate is greater than the median of sample, then we have set it as 1, otherwise 0. Further, to address the long run implication without temporal noise, we calculate the five-year average in sample year before running the regression following the methodology by [Arin et al. \(2019\)](#).

Table 3.6: Fiscal Revenue to Government Consumption (Fixed Effect: Growth Model on Bayesian Model Averaging)

Country Fixed Effect			Country Fixed Effect		
Debt X Gc			Debt X Gc		
	Mean	PIP		Mean	PIP
lag_Lifeexp	7.68E+01	1	lag_Lifeexp	7.66E+01	1
lag_Yrschool	-2.46E-03	1	lag_Yrschool	-2.45E-03	1
lag_Corruption	-3.34E-02	0.99999999	lag_Corruption	-3.37E-02	0.99999999
lag_TFR	-6.95E-02	0.99999998	lag_TFR	-6.88E-02	0.99999998
lag_Trade	1.13E-03	0.99999997	lag_Trade	1.13E-03	0.99999998
Gc_dummy	-4.84E-02	0.99700142	Gc_dummy	-4.77E-02	0.99998426
lag_population	2.41E-10	0.56652521	lag_population	2.78E-10	0.64665413
lag2_Debt_GDP	-1.40E-04	0.43236803	lag2_Debt_GDP	-1.58E-04	0.50573695
laglnyl	-1.61E-03	0.12881101	laglnyl	-2.43E-03	0.18782101
debt X Gc_dummy	1.93E-05	0.0844005	lag_Inflation	9.19E-06	0.07621187
lag_Inflation	6.40E-06	0.05053828			
<b>Total effect:</b>	<b>-4.85E-02</b>		<b>Total effect:</b>	<b>-4.78E-02</b>	
Tax X Gc			Tax X Gc		
	Mean	PIP		Mean	PIP
lag_Lifeexp	1.11E+02	1	lag_Lifeexp	1.11E+02	1
lag_Corruption	-3.54E-02	1	lag_Corruption	-3.53E-02	1
lag_TFR	-7.12E-02	0.95925989	lag_TFR	-7.29E-02	0.96836828
lag2_Tax_GDP	-8.67E-03	0.87252351	lag2_Tax_GDP	-8.85E-03	0.89076011
lag_Yrschool	-4.56E-04	0.46956178	lag_Yrschool	-4.93E-04	0.50895112
lag_Trade	3.26E-05	0.09900949	lag_Trade	3.96E-05	0.11949151
Gc_dummy	-5.09E-04	0.05117568	Gc_dummy	-6.65E-04	0.06495431
tax X Gc_dummy	-4.47E-05	0.0467116	lag_population	-2.34E-11	0.04695025
lag_Inflation	-1.61E-05	0.03853033	lag_Inflation	-1.86E-05	0.04607432
lag_population	-1.91E-11	0.03848836	laglnyl	-4.57E-05	0.03602974
laglnyl	-3.63E-05	0.02960706			
<b>Total effect:</b>	<b>-9.23E-03</b>		<b>Total effect:</b>	<b>-9.52E-03</b>	

Note - Dependent variable: every two years of GDP growth rate. To capture the flow components of fiscal expenditure, we generated the dummy variable of infrastructure investment ( $G_p$ ) and government consumption ( $G_c$ ) based on the 2-year growth rate; if the growth rate is greater than the median of sample, then we have set it as 1, otherwise 0. Further, to address the long run implication without temporal noise, we calculate the five-year average in sample year before running the regression following the methodology by [Arin et al. \(2019\)](#).

hurt the economic growth rate more than the increase in government consumption financed by the income tax rate. However, we need to note that this does not necessarily imply that higher income tax is always recommended; theoretically speaking, an extreme rise in income tax will raise the private consumption ratio, directly negating economic growth. An excessive income tax increase is not expected to be desirable for economic growth.

Based on our empirical findings, one may conclude that more infrastructure is always better for economic growth. As discussed in sections 3.2 and 3.3, we claim that since countries so far have not yet reached the threshold level of infrastructure investment to GDP ratio, concerning our baseline results, more infrastructure is expected to bring higher economic growth returns in this state of the economy. Excessive infrastructure investment may be detrimental to economic growth, as it will lead to a crowding-out effect, and balancing it with government consumption expenditure is recommended.

### 3.6 Sensitivity Analysis

Cuaresma (2011) brings concerns of using interaction terms within the Bayesian Model Averaging framework. Even though statistically speaking, the approach does not have an issue, there could be a potential unobservable that the dummy variable may capture. They suggest testing BMA without interactions or models with a strong heredity prior. Therefore, we eliminated the interaction term to see if there is a difference in growth implications. We found that exclusion of the interaction term did not change as in Table 3.7.

Table 3.7: Further Sensitivity Analysis based on [Cuaresma \(2011\)](#)

<b>Pooled OLS w/o all interests</b>			<b>Country Fixed Effect w/o all interests</b>		
	Mean	PIP		Mean	PIP
<b>laglny1</b>	<b>-1.43E-02</b>	<b>0.6984252</b>	<b>lag_Lifeexp</b>	<b>7.09E+01</b>	<b>1</b>
lag_TFR	-9.29E-03	0.410844	<b>lag_TFR</b>	<b>-6.92E-02</b>	<b>1</b>
lag_Corruption	-5.55E-03	0.3343172	<b>lag_Trade</b>	<b>1.16E-03</b>	<b>1</b>
lag_Lifeexp	5.48E+00	0.30692934	<b>lag_Corruption</b>	<b>-3.39E-02</b>	<b>0.99999999</b>
lag_Trade	2.83E-05	0.12233817	<b>lag_Yrschool</b>	<b>-2.15E-03</b>	<b>0.99999999</b>
lag_Inflation	6.63E-05	0.06522087	<b>lag_population</b>	<b>2.68E-10</b>	<b>0.64220369</b>
lag_Yrschool	-1.57E-05	0.04529272	<b>laglny1</b>	<b>-1.29E-02</b>	<b>0.63057435</b>
lag_population	1.28E-12	0.03723678	lag_Inflation	-5.30E-06	0.09789323

Note - Dependent variable: every two years of GDP growth rate. To capture the flow components of fiscal expenditure, we generated the dummy variable of infrastructure investment ( $G_p$ ) and government consumption ( $G_c$ ) based on the 2-year growth rate; if the growth rate is greater than the median of sample, then we have set it as 1, otherwise 0. Further, to address the long run implication without temporal noise, we calculate the five-year average in sample year before running the regression following the methodology by [Arin et al. \(2019\)](#).

### 3.7 Conclusion

We analyze the mechanisms of how the usage of government expenditures impact on economic growth. The existing literature so far on this field was puzzling because of different focus of country sets, the lack in addressing the model uncertainty and the insufficient assessment in the decomposition of government revenue/expenditure, which may have been the reason for the mixed findings in the previous literature. We address this issue by first building the endogenous growth model based on [Bruce and Turnovsky \(1999\)](#), which is a modification of the basic [Barro \(1990\)](#) model of endogenous growth. To assess the empirical nuances with sensibly addressing the diverse findings in literature, we use comprehensive 180 countries data, and accounted for model uncertainty by applying the Bayesian Model Averaging. IMF, World Bank, and OECD databases enable us to assess based on primal four compositions of fiscal expenditures. Namely, how eco-

conomic growth may differ by government expenditure (government consumption vs. infrastructure investment) that can arise from fiscal revenues primarily from either additional tax or additional public debt.

Our results indicate that increase in debt and income tax are persistently detrimental to growth. This confirms debt overhang hypothesis (Myers (1977)) that claim having additional debt may hurt the country rather than benefit as it potentially increase default risk. We also confirm that if the government use this additional fiscal revenue for increase in infrastructure investment, it will positively impact on growth regardless of fiscal revenues of taxes or debt (additional borrowings). A negative response arises from them being used to increasing government consumption. This mechanism was also supported by our theoretical model. These effects we observed in the Bayesian Model Averaging were consistent after we address the possible bias introduced by Cuaresma (2011). Our findings reveal that when governments face a trade-off in supporting current consumption or boosting infrastructure investment, as policy implication they are encouraged to spend more for the latter since it is forecasted to provide robust increases in economic growth. In theory, if there is an increase in infrastructure investment for efficient transportation, this will provide more convenient and instant labor mobility and easy market accessibility to the society that will enable the higher productivity.

Our findings on the importance of focusing on infrastructure investment for future economic growth do not necessarily mean that the higher the infrastructure investment, the better. Since countries so far have not yet reached the threshold level of infrastructure investment to GDP ratio concerning our baseline results, more infrastructure is expected to bring higher economic growth returns in this state of the economy. Excessive infrastructure investment may be detrimental to economic growth, and we need to note that balancing it with government consumption expenditure is recommended.

We acknowledge that while our findings are largely aligning with our theoretical grounds, precise magnitude of how much growth rate is feasible is still not concrete in this literature. This step is challenging to address under changing economic circumstances of recent years surge in public debt, high inflation rate, and high-interest rates; a situation that resonates in our recent macroeconomic trend. As a future assessment on this topic could also improve by more distinct fiscal spending data (corporate tax, income tax, consumption tax, property tax, etc).

# Bibliography

- AFONSO, ANTONIO AND DAVIDE FURCERI (2010) “Government size, composition, volatility and economic growth,” *European Journal of Political Economy*, 26 (4), 517–532.
- AFONSO, ANTÓNIO AND JOÃO TOVAR JALLES (2013) “Growth and productivity: The role of government debt,” *International Review of Economics & Finance*, 25, 384–407.
- AFONSO, ANTÓNIO AND JOAO TOVAR JALLES (2014) “Fiscal composition and long-term growth,” *Applied Economics*, 46 (3), 349–358.
- AFONSO, ANTONIO, LUDGER SCHUKNECHT, AND VITO TANZI (2005) “Public sector efficiency: an international comparison,” *Public choice*, 123 (3), 321–347.
- AGELL, JONAS, THOMAS LINDH, AND HENRY OHLSSON (1997) “Growth and the public sector: A critical review essay,” *European Journal of Political Economy*, 13 (1), 33–52.
- AGELL, JONAS, HENRY OHLSSON, AND PETER SKOGMAN THOURSIE (2006) “Growth effects of government expenditure and taxation in rich countries: A comment,” *European Economic Review*, 50 (1), 211–218.
- AIZENMAN, JOSHUA AND HIRO ITO (2020) “US macro policies and global economic challenges,” Technical report, National Bureau of Economic Research.
- ALDENHOFF, FRANK-OLIVER (2007) “Are economic forecasts of the International Monetary Fund politically biased? A public choice analysis,” *The Review of International Organizations*, 2 (3), 239–260.
- ARIN, K PEREN, ELIAS BRAUNFELS, AND GERNOT DOPPELHOFFER (2019) “Revisiting the growth effects of fiscal policy: A Bayesian model averaging approach,” *Journal of Macroeconomics*, 62, 103158.

- ARMSTRONG, J SCOTT (2001) "Selecting forecasting methods," in *Principles of forecasting*, 365–386: Springer.
- ARORA, HARJIT K AND DAVID J SMYTH (1990) "Forecasting the developing world: An accuracy analysis of the IMF's forecasts," *International Journal of Forecasting*, 6 (3), 393–400.
- ARTIS, MICHAEL J (1988) "I How Accurate Is the World Economic Outlook? A Post Mortem on Short-Term Forecasting at the International Monetary Fund," in *World Economic and Financial Surveys, 1988*: International Monetary Fund.
- ARTIS, MR MICHAEL J (1996) *How accurate are the IMF's short-term forecasts? Another examination of the World Economic Outlook*: International Monetary Fund.
- ASATRYAN, ZAREH AND LARS P FELD (2015) "Revisiting the link between growth and federalism: A Bayesian model averaging approach," *Journal of Comparative Economics*, 43 (3), 772–781.
- ATOIAN, ROUBEN, PATRICK CONWAY, AND MARCELO SELOWSKY (2004) "Macroeconomic adjustment in IMF-supported programs: projections and reality," in *Globalization and the Nation State*, 345–370: Routledge.
- ATOYAN, RUBEN AND PATRICK CONWAY (2011) "Projecting macroeconomic outcomes: Evidence from the IMF," *The Review of International Organizations*, 6 (3-4), 415.
- BAQIR, REZA, RODNEY RAMCHARAN, AND RATNA SAHAY (2005) "IMF programs and growth: Is optimism defensible?" *IMF Staff Papers*, 52 (2), 260–286.
- BARRO, R AND XAVIER SALA-I MARTIN (2004) "Economic growth second edition."
- BARRO, ROBERT J (1980) "Federal deficit policy and the effects of public debt shocks," Technical report, National Bureau of Economic Research.
- (1990) "Government spending in a simple model of endogenous growth," *Journal of political economy*, 98 (5, Part 2), S103–S125.
- (1991) "Economic growth in a cross section of countries," *The quarterly journal of economics*, 106 (2), 407–443.
- (2001) "Education and economic growth," *The contribution of human and social capital to sustained economic growth and well-being*, 79, 13–41.

- BARRO, ROBERT J AND GARY S BECKER (1989) "Fertility choice in a model of economic growth," *Econometrica: journal of the Econometric Society*, 481–501.
- BEACH, WILLIAM AND AARON SCHAVEY (1999) "How Reliable Are IMF Economic Forecasts."
- BEAUDRY, PAUL AND TIM WILLEMS (2022) "On the macroeconomic consequences of over-optimism," *American Economic Journal: Macroeconomics*, 14 (1), 38–59.
- BELLETTINI, GIORGIO AND CARLOTTA BERTI CERONI (2000) "Social security expenditure and economic growth: an empirical assessment," *Research in economics*, 54 (3), 249–275.
- BERGH, ANDREAS AND MARTIN KARLSSON (2010) "Government size and growth: Accounting for economic freedom and globalization," *Public Choice*, 142 (1), 195–213.
- BERNHEIM, B DOUGLAS (1987) "Ricardian equivalence: An evaluation of theory and evidence," *NBER macroeconomics annual*, 2, 263–304.
- BLANCHARD, OLIVIER J (1985) "Debt, deficits, and finite horizons," *Journal of political economy*, 93 (2), 223–247.
- BLANCHARD, OLIVIER AND ROBERTO PEROTTI (2002) "An empirical characterization of the dynamic effects of changes in government spending and taxes on output," *the Quarterly Journal of Economics*, 117 (4), 1329–1368.
- BOZ, EMINE, CAMILA CASAS, GEORGIOS GEORGIADIS, GITA GOPINATH, HELENA LE MEZO, ARNAUD MEHL, AND TRA NGUYEN (2020) "Patterns in invoicing currency in global trade."
- BROCK, WILLIAM A AND STEVEN N DURLAUF (2001) "What have we learned from a decade of empirical research on growth? Growth empirics and reality," *the world bank economic review*, 15 (2), 229–272.
- BRUCE, NEIL AND STEPHEN J TURNOVSKY (1999) "Budget balance, welfare, and the growth rate:" dynamic scoring" of the long-run government budget," *Journal of Money, Credit, and Banking*, 162–186.
- CAMERON, DAVID R (1982) "On the limits of the public economy," *The Annals of the American Academy of Political and Social Science*, 459 (1), 46–62.

- CARRIERE-SWALLOW, MR YAN AND JOSÉ MARZLUF (2021) *Macrofinancial Causes of Optimism in Growth Forecasts*: International Monetary Fund.
- CECCHETTI, STEPHEN G, MADHUSUDAN S MOHANTY, AND FABRIZIO ZAMPOLLI (2011) “The real effects of debt.”
- CHATTERJEE, PRATITI (2019) “Asymmetric impact of uncertainty in recessions: are emerging countries more vulnerable?” *Studies in Nonlinear Dynamics & Econometrics*, 23 (2).
- CHATTERJEE, PRATITI AND SYLWIA NOWAK (2016) *Forecast errors and uncertainty shocks*: International Monetary Fund.
- CHEN, CHUANGLIAN, SHUJIE YAO, PEIWEI HU, AND YUTING LIN (2017) “Optimal government investment and public debt in an economic growth model,” *China Economic Review*, 45, 257–278.
- CHINN, MENZIE D AND HIRO ITO (2006) “What matters for financial development? Capital controls, institutions, and interactions,” *Journal of development economics*, 81 (1), 163–192.
- CLEMENTS, BENEDICT, RINA BHATTACHARYA, AND TOAN QUOC NGUYEN (2003) “External debt, public investment, and growth in low-income countries.”
- COCHRANE, JOHN H (2011) “Presidential address: Discount rates,” *The Journal of finance*, 66 (4), 1047–1108.
- CODOGNO, LORENZO, CARLO FAVERO, AND ALESSANDRO MISSALE (2003) “Yield spreads on EMU government bonds,” *Economic Policy*, 18 (37), 503–532.
- CREMER, HELMUTH AND FIROUZ GAHVARI (1993) “Tax evasion and optimal commodity taxation,” *Journal of Public Economics*, 50 (2), 261–275.
- CUARESMA, JESUS CRESPO (2011) “How different is Africa? A comment on Masanjala and Papa-georgiou,” *Journal of Applied Econometrics*, 26 (6), 1041–1047.
- DAR, ATUL A AND SAL AMIRKHALKHALI (2002) “Government size, factor accumulation, and economic growth: evidence from OECD countries,” *Journal of policy modeling*, 24 (7-8), 679–692.
- DIAMOND, PETER A (1965) “National debt in a neoclassical growth model,” *The American Economic Review*, 55 (5), 1126–1150.

- DIEBOLD, FRANCIS X AND JOSE A LOPEZ (1996) “8 Forecast evaluation and combination,” *Handbook of statistics*, 14, 241–268.
- DOMAR, EVSEY D (1944) “The” burden of the debt” and the national income,” *The American Economic Review*, 34 (4), 798–827.
- DRAPER, DAVID (1995) “Assessment and propagation of model uncertainty,” *Journal of the Royal Statistical Society: Series B (Methodological)*, 57 (1), 45–70.
- DREHER, AXEL, NOEL GASTON, AND PIM MARTENS (2008) “Consequences of globalization re-considered: Applying the KOF index,” in *Measuring Globalisation*, 75–171: Springer.
- DURLAUF, STEVEN N, ANDROS KOURTELLOS, AND CHIH MING TAN (2008) “Are any growth theories robust?” *The Economic Journal*, 118 (527), 329–346.
- EATON, JONATHAN (1981) “Fiscal policy, inflation and the accumulation of risky capital,” *The Review of Economic Studies*, 48 (3), 435–445.
- EBERHARDT, MARKUS AND ANDREA F PRESBITERO (2015) “Public debt and growth: Heterogeneity and non-linearity,” *Journal of International Economics*, 97 (1), 45–58.
- EICHER, THEO AND CECILIA GARCÍA-PEÑALOSA (2008) “Endogenous strength of intellectual property rights: Implications for economic development and growth,” *European Economic Review*, 52 (2), 237–258.
- EICHER, THEO S, DAVID J KUENZEL, CHRIS PAPAGEORGIOU, AND CHARIS CHRISTOFIDES (2019) “Forecasts in times of crises,” *International Journal of Forecasting*, 35 (3), 1143–1159.
- EICHER, THEO S, CHRIS PAPAGEORGIOU, AND ADRIAN E RAFTERY (2011) “Default priors and predictive performance in Bayesian model averaging, with application to growth determinants,” *Journal of Applied Econometrics*, 26 (1), 30–55.
- EICHER, THEO S AND YUAN GAO ROLLINSON (2022) “The accuracy of IMF crises nowcasts,” *International Journal of Forecasting*.
- ESTEFANIA FLORES, JULIA, DAVIDE FURCERI, SIDDHARTH KOTHARI, AND JONATHAN D OSTRY (2021) “Worse Than You Think: Public Debt Forecast Errors in Advanced and Developing Economies.”

- FERNANDEZ, CARMEN, EDUARDO LEY, AND MARK FJ STEEL (2001) “Model uncertainty in cross-country growth regressions,” *Journal of applied Econometrics*, 16 (5), 563–576.
- FÖLSTER, STEFAN AND MAGNUS HENREKSON (2001) “Growth effects of government expenditure and taxation in rich countries,” *European economic review*, 45 (8), 1501–1520.
- FULLERTON, DON AND MARIOS KARAYANNIS (1994) “Tax evasion and the allocation of capital,” *Journal of Public Economics*, 55 (2), 257–278.
- FURCERI, DAVIDE AND BIN GRACE LI (2017) *The macroeconomic (and distributional) effects of public investment in developing economies*: International Monetary Fund.
- FURCERI, DAVIDE, MR GIOVANNI MELINA, JUN GE, MR PRAKASH LOUNGANI ET AL. (2018) “The Distributional Effects of Government Spending Shocks in Developing Economies,” Technical report, International Monetary Fund.
- GENBERG, HANS AND ANDREW MARTINEZ (2014) “On the accuracy and efficiency of IMF forecasts: A survey and some extensions,” *IEO Background Paper No. BP/14/04 (Washington: Independent Evaluation Office of the IMF)*.
- GHOSH, ATISH, TIMOTHY LANE, ALUN THOMAS, AND JUAN ZALDUENDO (2006) “Program Objectives and Program Success,” *PART I. DESIGN*.
- GOLDSTEIN, MORRIS, TIMOTHY F GEITHNER, PAUL KEATING, AND YUNG CHUL PARK (2003) “IMF structural programs,” in *Economic and financial crises in emerging market economies*, 363–458: University of Chicago Press.
- HANSSON, PÄR AND MAGNUS HENREKSON (1994) “A new framework for testing the effect of government spending on growth and productivity,” *Public choice*, 81 (3), 381–401.
- HENRI, ATANGANA ONDOA (2019) “Heavily indebted poor countries initiative (HIPC), debt relief, economic stability and economic growth in Africa,” *Economic Change and Restructuring*, 52 (2), 89–121.
- HENRIKSSON, ROY D AND ROBERT C MERTON (1981) “On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills,” *Journal of business*, 513–533.
- HOLDEN, KAREN, DAVID PEEL, AND B SANDHU (1987) “The accuracy of OECD forecasts,” *Empirical Economics*, 12 (3), 175–186.

- HOLDEN, KENNETH, DAVID A PEEL ET AL. (1990) “On testing for unbiasedness and efficiency of forecasts,” *The Manchester School of Economic & Social Studies*, 58 (2), 120–127.
- HUSSAIN, M NURELDIN AND BERNHARD G GUNTER (2005) “External shocks and the HIPC Initiative: Impacts on growth and poverty in Africa,” *African Development Review*, 17 (3), 461–492.
- ILZETZKI, ETHAN, CARMEN M REINHART, AND KENNETH S ROGOFF (2019) “Exchange arrangements entering the twenty-first century: Which anchor will hold?” *The Quarterly Journal of Economics*, 134 (2), 599–646.
- IMF (2019) “Review of Program Design and Conditionality.,” *Washington DC: International Monetary Fund*.
- (2021a) “*Monitoring of Fund Arrangements*, various issues.,” Data retrieved from International Monetary Fund, <https://www.imf.org/external/np/pdr/mona/index.aspx>.
- (2021b) “*World Economic Outlook*, various issues.,” Data retrieved from International Monetary Fund, <https://www.imf.org/en/Publications/SPROLLS/world-economic-outlook-databases>.
- (2021c) “*International Financial Statistics*, various issues.,” Data retrieved from International Monetary Fund, <https://data.imf.org/?sk=4c514d48-b6ba-49ed-8ab9-52b0c1a0179b>.
- (2021d) “*Balance of Payments Statistics Yearbook (Balance of Payments and International Investment Position)*, various issues.,” Data retrieved from International Monetary Fund, <https://data.imf.org/?sk=7A51304B-6426-40C0-83DD-CA473CA1FD52>.
- (2021e) “*Annual Report on Exchange Arrangements and Exchange Restrictions*, various issues.,” Data retrieved from International Monetary Fund, <https://www.imf.org/en/Publications/Annual-Report-on-Exchange-Arrangements-and-Exchange-Restrictions/Issues/2021/08/25/Annual-Report-on-Exchange-Arrangements-and-Exchange-Restrictions-2020-49738>.
- IRELAND, PETER N (1994) “Supply-side economics and endogenous growth,” *Journal of Monetary Economics*, 33 (3), 559–571.

- ISMAIL, KAREEM, ROBERTO PERRELLI, AND JESSIE YANG (2020) “Optimism Bias in Growth Forecasts-The Role of Planned Policy Adjustments,” *Available at SSRN 3758062*.
- ITSKHOKI, OLEG (2020) “The story of the real exchange rate,” Technical report, National Bureau of Economic Research.
- JORDÀ, ÒSCAR (2005) “Estimation and inference of impulse responses by local projections,” *American economic review*, 95 (1), 161–182.
- KENEN, PETER B AND STEPHEN B SCHWARTZ (1986) *An assessment of macroeconomic forecasts in the International Monetary Fund’s World Economic Outlook*: International Finance Section, Department of Economics, Princeton University.
- KENTIKELÉNIS, ALEXANDER E, THOMAS H STUBBS, AND LAWRENCE P KING (2016) “IMF conditionality and development policy space, 1985–2014,” *Review of International Political Economy*, 23 (4), 543–582.
- KIM, EUNJI, YOONHEE HA, AND SANGHEON KIM (2017) “Public debt, corruption and sustainable economic growth,” *Sustainability*, 9 (3), 433.
- KNELLER, RICHARD, MICHAEL F BLEANEY, AND NORMAN GEMMELL (1999) “Fiscal policy and growth: evidence from OECD countries,” *Journal of public economics*, 74 (2), 171–190.
- KUNZE, FREDERIK (2020) “Predicting exchange rates in Asia: New insights on the accuracy of survey forecasts,” *Journal of Forecasting*, 39 (2), 313–333.
- KURZ, MORDECAI (1968) “Optimal economic growth and wealth effects,” *International Economic Review*, 9 (3), 348–357.
- LANDAU, DANIEL (1986) “Government and economic growth in the less developed countries: an empirical study for 1960-1980,” *Economic Development and Cultural Change*, 35 (1), 35–75.
- LAUBACH, THOMAS (2009) “New evidence on the interest rate effects of budget deficits and debt,” *Journal of the European economic Association*, 7 (4), 858–885.
- LEE, YOUNG AND ROGER H GORDON (2005) “Tax structure and economic growth,” *Journal of public economics*, 89 (5-6), 1027–1043.

- LEWIS, CHRISTINE AND NIGEL PAIN (2015) "Lessons from OECD forecasts during and after the financial crisis," *OECD Journal: Economic Studies*, 2014 (1), 9–39.
- LEY, EDUARDO AND MARK FJ STEEL (2009) "On the effect of prior assumptions in Bayesian model averaging with applications to growth regression," *Journal of applied econometrics*, 24 (4), 651–674.
- LIAQAT, ZARA (2019) "Does government debt crowd out capital formation? A dynamic approach using panel VAR," *Economics letters*, 178, 86–90.
- LO, STEPHANIE AND KENNETH ROGOFF (2015) "Secular stagnation, debt overhang and other rationales for sluggish growth, six years on."
- LUNA, FRANCESCO (2014) "IMF Forecasts in the Context of Program Countries," *IMF Independent*.
- MARCELINO, MS SANDRA AND MS IVETTA HAKOBYAN (2014) *Does lower debt buy higher growth? The impact of debt relief initiatives on growth*: International Monetary Fund.
- SALA-I MARTIN, XAVIER (1997) "I just ran four million regressions."
- MASANJALA, WINFORD H AND CHRIS PAPAGEORGIOU (2008) "Rough and lonely road to prosperity: a reexamination of the sources of growth in Africa using Bayesian model averaging," *Journal of applied Econometrics*, 23 (5), 671–682.
- MAURO, PAOLO (1995) "Corruption and growth," *The quarterly journal of Economics*, 110 (3), 681–712.
- MEESE, RICHARD A AND KENNETH ROGOFF (1983) "Empirical exchange rate models of the seventies: Do they fit out of sample?" *Journal of international economics*, 14 (1-2), 3–24.
- MENDOZA, ENRIQUE G, GIAN MARIA MILESI-FERRETTI, AND PATRICK ASEA (1997) "On the ineffectiveness of tax policy in altering long-run growth: Harberger's superneutrality conjecture," *Journal of public economics*, 66 (1), 99–126.
- MERTON, ROBERT C (1981) "On market timing and investment performance. I. An equilibrium theory of value for market forecasts," *Journal of business*, 363–406.

- MILLER, STEPHEN M AND FRANK S RUSSEK (1997) “Fiscal structures and economic growth: international evidence,” *Economic Inquiry*, 35 (3), 603–613.
- MINCER, JACOB A AND VICTOR ZARNOWITZ (1969) “The evaluation of economic forecasts,” in *Economic forecasts and expectations: Analysis of forecasting behavior and performance*, 3–46: NBER.
- MODIGLIANI, FRANCO (1961) “Long-run implications of alternative fiscal policies and the burden of the national debt,” *The Economic Journal*, 71 (284), 730–755.
- MODY, MR ASHOKA AND MR ALESSANDRO REBUCCI (2006) *IMF-supported programs: recent staff research*: International Monetary Fund.
- MONTIEL OLEA, JOSÉ LUIS AND MIKKEL PLAGBORG-MØLLER (2021) “Local projection inference is simpler and more robust than you think,” *Econometrica*, 89 (4), 1789–1823.
- MUSSA, MICHAEL AND MIGUEL SAVASTANO (1999) “The IMF approach to economic stabilization,” *NBER macroeconomics annual*, 14, 79–122.
- MUSSO, ALBERTO AND STEVEN PHILLIPS (2002) “Comparing projections and outcomes of IMF-supported programs,” *IMF Staff Papers*, 49 (1), 22–48.
- MYERS, STEWART C (1977) “Determinants of corporate borrowing,” *Journal of financial economics*, 5 (2), 147–175.
- NAKAMURA, EMI AND JÓN STEINSSON (2018) “Identification in macroeconomics,” *Journal of Economic Perspectives*, 32 (3), 59–86.
- NORDHAUS, WILLIAM D (1987) “Forecasting efficiency: concepts and applications,” *The Review of Economics and Statistics*, 667–674.
- PANIZZA, UGO AND ANDREA F PRESBITERO (2013) “Public debt and economic growth in advanced economies: A survey,” *Swiss Journal of Economics and Statistics*, 149 (2), 175–204.
- (2014) “Public debt and economic growth: is there a causal effect?” *Journal of Macroeconomics*, 41, 21–41.
- PATTILLO, CATHERINE, LUCA ANTONIO RICCI ET AL. (2011) “External debt and growth,” *Review of economics and institutions*, 2 (3), 30.

- PESCATORI, MR ANDREA, MR DAMIANO SANDRI, AND JOHN SIMON (2014) *Debt and growth: is there a magic threshold?:* International Monetary Fund.
- PRZEWORSKI, ADAM AND JAMES RAYMOND VREELAND (2000) “The effect of IMF programs on economic growth,” *Journal of Development Economics*, 62 (2), 385–421.
- RAFTERY, ADRIAN E (1995) “Bayesian model selection in social research,” *Sociological methodology*, 111–163.
- REBELO, SERGIO (1991) “Long-run policy analysis and long-run growth,” *Journal of political Economy*, 99 (3), 500–521.
- REINHART, CARMEN M AND KENNETH S ROGOFF (2010) “Growth in a Time of Debt,” *American economic review*, 100 (2), 573–78.
- REINHART, CARMEN M, KENNETH S ROGOFF, AND MIGUEL SAVASTANO (2003) “Debt intolerance.”
- ROMER, CHRISTINA D AND DAVID H ROMER (2010) “The macroeconomic effects of tax changes: estimates based on a new measure of fiscal shocks,” *American Economic Review*, 100 (3), 763–801.
- ROMERO-AVILA, DIEGO AND ROLF STRAUCH (2008) “Public finances and long-term growth in Europe: Evidence from a panel data analysis,” *European Journal of Political Economy*, 24 (1), 172–191.
- ROSSI, BARBARA (2013) “Exchange rate predictability,” *Journal of economic literature*, 51 (4), 1063–1119.
- SAINT-PAUL, GILLES (1992) “Fiscal policy in an endogenous growth model,” *The Quarterly Journal of Economics*, 107 (4), 1243–1259.
- SCHNADER, MARJORIE H AND HERMAN O STEKLER (1990) “Evaluating predictions of change,” *Journal of Business*, 99–107.
- SINCLAIR, TARA M, FRED JOUTZ, AND HERMAN O STEKLER (2010) “Can the Fed predict the state of the economy?” *Economics Letters*, 108 (1), 28–32.
- DE SOYRES, CONSTANACE, REINA KAWAI, AND MENGXUE WANG (2022) “Public Debt and Real GDP: Revisiting the Impact.”

- STIGLITZ, JOSEPH E (2011) “A short note on surveillance and how reforms in surveillance can help the IMF to promote global financial stability.”
- STUBBS, THOMAS, BERNHARD REINSBERG, ALEXANDER KENTIKELLENIS, AND LAWRENCE KING (2020) “How to evaluate the effects of IMF conditionality,” *The Review of International Organizations*, 15 (1), 29–73.
- THEIL, HENRI (1961) “Economic forecasts and policy.”
- TIMMERMANN, ALLAN (2007) “An evaluation of the World Economic Outlook forecasts,” *IMF Staff Papers*, 54 (1), 1–33.
- USGAO (2003) “International financial crises: challenges remain in IMF’s ability to anticipate, prevent, and resolve financial crises,” *GAO-03-734, Washington: Government Accountability Office*.
- WANG, YU AND MARC F BELLEMARE (2019) “Lagged variables as instruments,” Technical report, Working Paper, Department of Applied Economics, University of Minnesota . . . .
- WOO, JAEJOON AND MANMOHAN S KUMAR (2015) “Public debt and growth,” *Economica*, 82 (328), 705–739.
- ZOU, HENG-FU (1994) “‘The spirit of capitalism’ and long-run growth,” *European Journal of Political Economy*, 10 (2), 279–293.