



HAL
open science

The Interplay between Macroeconomic Forecasting, Financial Crises and Complexity

Umberto Collodel

► **To cite this version:**

Umberto Collodel. The Interplay between Macroeconomic Forecasting, Financial Crises and Complexity. Economics and Finance. Université Panthéon-Sorbonne - Paris I, 2022. English. NNT : 2022PA01E003 . tel-04360051

HAL Id: tel-04360051

<https://theses.hal.science/tel-04360051>

Submitted on 21 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE PARIS I PANTHÉON SORBONNE
UFR de Sciences Économiques

THÈSE

Pour l'obtention du titre de Docteur en Sciences Économiques
Présentée et soutenue publiquement
le 02 Mars 2022 par
Umberto COLLODEL

**The Interplay between Macroeconomic Forecasting,
Financial Crises and Complexity**

**Sous la direction de Mme Agnès BÉNASSY-QUÉRÉ, PU, Université
Paris 1 Panthéon-Sorbonne**

Membre du Jury

**Mme. Nicoletta BATINI - Professeure associée - Ecole supérieure
Sainte-Anne de Pise - Suffragant**

**M. Jean Bernard CHATELAIN - PU, Université Paris 1 Panthéon-
Sorbonne - Président de Jury**

**M. Pierre-Olivier GOURINCHAS - PU, Université de Californie
(Berkeley) - Rapporteur**

M. Ugo Panizza - PU, IHEID - Rapporteur

Résumé:

Cette thèse explore différents sujets liés à la prévision macroéconomique. Elle commence par tirer de la littérature sur la finance internationale quelques faits stylisés sur la façon dont l'économie mondiale est devenue de plus en plus complexe après l'effondrement du système de Bretton Woods; cette sophistication croissante, à son tour, génère un besoin de plus de recherche en matière de prévision pour étayer les décisions politiques. Le chapitre 3 montre que les modèles de prévision des crises financières (systèmes d'alerte précoce) sont relativement performants en termes de prévision. Il examine et compare le rôle des indicateurs mondiaux et nationaux sur la matérialisation d'une crise externe et se penche sur la différence entre les crises manquées et prévues en termes de pertes de production relatives. Le chapitre 4 évalue la performance des prévisions à court terme de l'activité économique produites par les principales institutions économiques et le secteur privé. Il étudie la corrélation entre les erreurs de prévision et les différents états du cycle économique, la comparaison entre les prévisions produites par différentes institutions et aborde les questions liées à la dimension politique des erreurs de prévision. Enfin, le chapitre 5 fournit un nouvel ensemble de données provenant des documents d'archives du FMI; par le biais d'un dictionnaire compilé manuellement et d'une approche de fréquence des termes, il mesure la profondeur de la discussion sur 20 crises économiques et non économiques différentes pour l'ensemble des membres du Fonds. Il exploite cette riche source de données pour analyser l'évolution de la complexité des prévisions macroéconomiques au cours des dernières décennies.

Mots-clés: crises financières, prévision, complexité

Summary:

This thesis explores different topics related to macroeconomic forecasting. It starts drawing from the literature on international finance some stylised facts on how the global economy became increasingly complex after the demise of the Bretton Woods system; this growing sophistication, in turn, generates a need for more forecasting research to buttress policy-making decisions. Chapter 3 shows that models for financial crises prediction (Early Warning Systems) fair relatively well in terms of forecasting performance. It discusses and compares the role of global and domestic indicators on the materialisation of an external crisis and delves into the difference between missed and predicted crises in terms of relative output losses. Chapter 4 evaluates the performance of short-term forecasts of economic activity produced by the main economic institutions and private sector. It studies how forecast errors correlate to different states of the business cycle, how forecasts produced by different institutions compare to each other and touches on questions related to the political dimension of forecast errors. Finally, Chapter 5 provides a novel dataset of IMF archival documents; through a manually compiled dictionary and a term-frequency approach, it measures the depth of discussion about 20 different economic and non-economic crises for the whole Fund membership. It harnesses this rich data source to analyse how the complexity of macroeconomic forecasting has changed over the last decades.

Keywords: financial crises, forecasting, complexity

PARIS SCHOOL OF ECONOMICS
UNIVERSITÉ PARIS 1 PANTHÉON-SORBONNE

The Interplay between Macroeconomic Forecasting, Financial Crises and
Complexity

BY

UMBERTO COLLODEL
March 2022

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in International Economics

Paris, France

Defense Committee

Thesis Director: Agnès BÉNASSY-QUÉRÉ
President of Jury: Jean-Bernard CHATELAIN
Jury Member: Nicoletta BATINI
Rapporteur: Pierre-Olivier GOURINCHAS
Rapporteur: Ugo PANIZZA

Copyright by
Umberto Collodel
2022

七転び八起き

A Guido, Andrea e ai miei genitori

Graxie per avermi fatto crescere e

diventare quello che sono

ACKNOWLEDGMENTS

To start with, my greatest gratitude goes to my supervisor Agnès Bénassy-Quéré. It was following her course on International Economics, now already four years ago, that I fell in love with the topics I am here discussing and that I started to think about a research career. Without her never ending encouragement and suggestions, it would not have been possible to finish this work. I am also deeply indebted to Jean-Bernard Chatelain whose guidance and advices helped me going through difficult moments. He always offered his comments and critiques in a constructive way and never refrained from compliments and kind words.

I am extremely grateful to Prakash Loungani and Nicoletta Batini for selecting me for the Ph.D. Summer Internship at the Independent Evaluation Office (IEO) of the IMF and with which I am currently working. The time I spent and I am spending at the IEO as well as the conversations with them have profoundly shaped my academic work and more in general, made me look at modern macroeconomic problems with anew eyes. I would also like to thank the whole IEO department for the warmth demonstrated me since my first day of work there.

I am immensely grateful to Éric Monnet. Our inspiring talks about boom-bust cycles, Bretton Woods and the development of the International Monetary System laid the early foundations for this work. I also wish to thank him for strongly endorsing and supporting me in the application process for this PhD. I am also profoundly grateful to Katheline Schubert who introduced

ACKNOWLEDGMENTS (Continued)

me to the joys (and the pains!) of teaching.

A great part of this thesis has been written in collaboration. I cannot find the words to express how indebted I am to Manuel Béтин. Not only he is a great friend whose positive and humble attitude toward life helped me in moments of great personal turmoil, he is also an excellent co-author with as many original research ideas as smiles per day. I will always remember with fondness our walks with a tea in the PSE campus. I am also profoundly grateful to Prakash Loungani (again!) and Zidong An for introducing me to the forecasts evaluation literature and for being such knowledgeable and supporting co-authors.

Now, it is time to move to the many friends that accompanied me closely throughout this journey. First of all, I would like to thank all the friends of the Macroeconomics department at PSE: Jayme Leiva, Brendan Vannier-Harnoys, Alessandro De Sanctis, Emanuele Franceschi, Andrea Mencarelli, Cyril Verluise, Pauline Wibaux and Louis-Marie Harpedanne de Belleville. Each one of you gave his/her contribution to the chapters in this thesis and kept me going through this stage of life with the tale of their own hurdles and successes. I am proud to have been part of this group of brilliant people. Similarly, I would like to express my gratitude to all friends that I had the luck to meet in the PSE and Paris 1 campus and ever since, became an important part of my life. Amit Dekel and Nina Rapoport, apart from being great friends, your intellectual curiosity and attitude toward life has been source of great inspiration for me. I will never forget the sense of peace and calmness I experienced at your place sharing a simple meal.

ACKNOWLEDGMENTS (Continued)

I wish you all the best for your future as a family. Elisa Baku, we walked this difficult path together. From you I learned what tenacity and dedications are: thank you for all your kind words, conversations and coffees. I am going to turn at your door once you become millionaire (I am sure you will be!). For the moment, just worry less and enjoy more: you have already come a long way señorita. João Henrique Avilla Dimas, thank you for teaching me how to not put a bound on my dreams. I would not be where I am without you and the afternoons spent together in Maison du Brésil study room (and of course thank you also for your witty jokes, or maybe not...). Alexia Lochmann, thank you for listening my rants and together complain about the crazy PhD life. Thank you also for carrying forward and showing me that there was light at the end of the tunnel. I would also like to thank all the people with whom I shared an office at PSE and I have not previously mentioned: Sandro Provenzano, Adrien Fabre, Stephane Zouri, Kentaro Asai, Karin Gourdon, Gabriela Pilay; without you the days spent working would have been certainly much more dull.

And now it is the turn of personal friends outside the academic sphere. Andrea Zago, you are part of my family. It would be far too long to list, but thank you for everything you have done for me in these four years. I hope I manage to demonstrate how an important piece of my life you are. Elisa Nardo, your smile, your will to improve yourself and to help others are a constant source of inspiration for me. Thank you for teaching me how to be a better person and not to care about the “ill” thoughts that try to disrupt our daily stability. Francesco Garatti, you constantly spur me to follow my inner voice and to not be afraid of external judge-

ACKNOWLEDGMENTS (Continued)

ment. Without you, much of this writing would have been more sad and lonely than it has been. Gracias hermano! Alexander Grosser, not only you are one of the best listeners I have ever met, but your kindness is unparalleled. Thank you for all the conversations, the support and, of course, the New Year's wishes (you will understand!). Mathieu Gard, your postcards and unexpected visits buttressed me in the most difficult moments over the past few years. I will always treasure your precious words in a little alcove of my heart. Giacomo Labbri, your curiosity toward economics prodded me toward explaining difficult concepts in simple words. Toward the end of this path, we faced similar challenges: our conversation helped me maintain the focus and gave me the strength necessary to finish this work. Shalmali Ghaisas, you had a glimpse at the darkest part of my soul and witnessed my worst moment. I never said it loud enough, but this thank you goes beyond words. Giada Borsoi, your company diverted me from the worries and anxieties that constellated this long path and allowed to take the much needed breaks while being at home (I probably also still have to thank you for those early and "unpaid" math lessons, not sure I would have coped without those...). Alberto Venturin, thank you for all the career advices and the long calls. I am proud of what we achieved together starting from a small village in the north-east of Italy and following a similar path. I have only one thing to say to you: *ad maiora!* Thank you also to Alice Dominici for her support and suggestions and for sharing her *locus amoenus* in the quaint Spoleto. Margot Falcoz-Vigne, your smiles, serenity and kind hospitality brought me temporary relief during challenging times. Thank you, I really hope you will get the best out of this life. Family Krishnamony, thank you for treating me like a king in your own home and exposing me to new ways of seeing life. Thank

ACKNOWLEDGMENTS (Continued)

you Krishna for sharing your wisdom while I was navigating in dire straits and Darshana for supporting me in the road preceding this path. Eugen Akimov, your courage in taking tough decisions and the steps taken toward your self-development and well-being have been source of unwavering inspiration for me. You helped me greatly in reaching the stability necessary to finish this work. Pravali Vangeti, since we both love winter walks, I will use an appropriate simile to describe how thankful I am to you: your words and presence have been like a stove for me, keeping me warm during the coldest and most rigid months of self doubt and loneliness.¹

Finally, since this dissertation has been written while being in different parts of the world (or at least Europe), thank you to all the people I shared a roof with and met on the road. To quote a famous Italian song “Buon viaggio hermano querido e buon cammino ovunque tu vada, forse un giorno potremmo incontrarci, di nuovo, lungo la strada!”

UC

¹A little game in all this seriousness: who said what? “Umbe, portami al mare (o a Westleveteren)!” , “Another life does not make sense”, “Helloo youu”, “I sacchi della spazzatura Umbe”, “Mi sto per schiantare”, “Chesstile”, “My frienddd, don’t break my heart!”, “Solo 50 dirham vec!”, “Cooool”, “Oddio sto mmale”, “How many kids Umberto?”.

TABLE OF CONTENTS

<u>CHAPTER</u>	<u>PAGE</u>
1 INTRODUCTION	1
1.1 Context and Motivation	1
1.2 Chapter 2: Are Financial Crises Predictable?	3
1.3 Chapter 3: Stylized Facts on Economic Forecasting and Crises	6
1.4 Chapter 4: Understanding Complex Crises	8
1.5 Summary: Research Questions	12
1.6 Summary: Contributions	13
2 INTRODUCTION (EN FRANÇAIS)	15
2.1 Contexte et Motivation	15
2.2 Chapitre 3: Les crises financières sont-elles prévisibles?	17
2.3 Chapitre 4: Faits stylisés sur les prévisions économiques et les crises	19
2.4 Chapitre 5: Comprendre les crises complexes	21
2.5 Résumé: Questions de recherche	24
2.6 Résumé : Contributions	25
3 FINDING A NEEDLE IN A HAYSTACK	
DO EARLY WARNING SYSTEM FOR SUDDEN STOPS WORK?	28
3.1 Introduction	29
3.2 Related Literature	32
3.3 Data	35
3.3.1 Sudden stops	35
3.3.2 Explanatory Variables	38
3.3.3 Data Transformation	41
3.4 Methodology	43
3.4.1 Dependent Variable	43
3.4.2 Estimation Strategy	44
3.4.3 Evaluation	46
3.4.4 Forecasting Procedure	48
3.5 Results	50
3.5.1 Determinants of Sudden Stops	50
3.5.2 Out-of-Sample Performance and Forecast Horizon	56
3.5.3 Timing	58
3.5.4 Sudden Stop Impact and Fitted Probabilities	62
3.6 Conclusion	66

TABLE OF CONTENTS (Continued)

<u>CHAPTER</u>		<u>PAGE</u>
4	WHEN (WHERE AND WHY) FORECASTERS GET IT WRONG?	
	30 YEARS OF GROWTH FORECASTS	68
4.1	Introduction	70
4.2	Data	73
4.3	Surveillance Forecasts	77
4.3.1	Surveillance Forecasts and Optimism	77
4.3.2	Optimism, Recessions and Financial Crises	88
4.3.3	Comparison between Forecasters	94
4.4	Program Forecasts	102
4.5	Conclusion	109
5	A (MODERN) NARRATIVE ON 70 YEARS OF MACROECONOMIC CRISES	
	A DATABASE FOR COMPLEXITY	111
5.1	Introduction	113
5.2	From Qualitative Judgment to Quantitative Measures	116
5.2.1	The Corpus of IMF Documents	120
5.2.2	The Lexicon of Crises	130
5.2.3	The Extraction Method	136
5.3	A Database on Crises Discussion	140
5.3.1	Comparison with Benchmarks: Stylized Facts	143
5.3.2	Comparison with Benchmarks: Econometric Estimation	148
5.3.3	Zeroing in on Non-Economic Indicators	154
5.4	Looking Back: A Rising Complexity	160
5.5	Conclusion	166
6	CONCLUSION	168
	Bibliography	171
	APPENDICES	184
	Appendix A	185
A.1	Data	185
A.1.1	Sudden Stops	189
A.1.2	Explanatory Variables and Data Transformation	199
A.2	Results	199
A.2.1	Determinants of Sudden Stops	199
A.2.2	Out-of-Sample Performance and Forecast Horizon	205
A.2.3	Sudden Stops Impact and Fitted Probabilities	210
	Appendix B	211
B.1	Data	211
B.2	Surveillance Forecasts	212
B.2.1	Surveillance Forecasts and Optimism	212

TABLE OF CONTENTS (Continued)

<u>CHAPTER</u>		<u>PAGE</u>
B.2.2	Optimism, Recessions and Financial Crises	220
B.2.3	A Comparison Between Forecasters	221
B.3	Program Forecasts	241
	Appendix C	251
C.1	From Qualitative Judgements to Quantitative Measures	251
C.1.1	The Corpus of IMF Documents	251
C.1.1.1	Scraping and Semantic Analysis	251
C.1.1.2	First Page Analysis	252
C.1.1.3	PDF Extraction	253
C.1.1.4	Final Corpus	254
C.1.2	The Lexicon of Crises	265
C.2	A Database On Crises Discussion	303
C.2.1	Comparison with Benchmarks	303
C.2.2	Zeroing In Non-Economic Indicator	305

LIST OF TABLES

<u>TABLE</u>	<u>PAGE</u>
I List of variables tested	42
II Example of Confusion Matrix	46
III Full Sample Logit	51
IV In-sample Performance	52
V Full Sample Logit - Benchmark Model With Interactions	55
VI Out-of-sample Performance	57
VII Different forecast horizons out-of-sample performance	58
VIII Predicted Probability and Sudden Stop Incidence	65
IX Forecasts Availability	76
X Magnitude of Short-Term Biases	80
XI Magnitude of Short-Term Biases - Region	83
XII Regional Specialization - World Bank and IMF	97
XIII Regional Specialization - European Commission and IMF	99
XIV Regional Specialization - Consensus and IMF	101
XV Descriptive Statistics - Amount Programs	104
XVI Amount Program and Bias	106
XVII Start Date by Income Group	130
XVIII Lexicon summary	133
XIX Typology of Crises	142
XX Confusion Matrix: Sovereign discussion and Default	147
XXI Confusion Matrix: Sovereign Discussion and Default Over Time	148
XXII Comparison with Benchmark: Severe Recession	149
XXIII Comparison with Benchmark: Sovereign	153
XXIV Average Shortest Path	164
XXV Countries List	185
XXVI List of Sudden Stops	194
XXVII Raw Data - Description and Sources	199
XXVIII Robustness Analysis	202
XXIX $H=0$ & $H=1$	212
XXX List of countries - IMF/World Bank	221
XXXI Full RMSE comparison - IMF/World Bank	224
XXXII List of countries - IMF/EC	228
XXXIII Full RMSE comparison - IMF/EC	230
XXXIV List of countries - IMF/Consensus	232
XXXV Full RMSE comparison - IMF/Consensus	236
XXXVI List of IMF programs - 2002-2018	241
XXXVII List of Keywords for Problematic Documents	252
XXXVIII Description of IMF Documents	254

LIST OF TABLES (Continued)

<u>TABLE</u>		<u>PAGE</u>
XXXIX	Country Coverage	256
XL	Full Lexicon	265
XLI	Comparison with Benchmark: Soft Recession	304

LIST OF FIGURES

<u>FIGURE</u>		<u>PAGE</u>
1	Sudden Stop Probabilities and Ex-Post Growth	5
2	Asymmetries in Growth Forecast Errors - Financial Crises, Recessions and Non-Recessions	7
3	A Network of Crises - 70 Years Later	10
4	Countries Sample	35
5	Examples of Sudden Stops Identification	37
6	Number of Sudden Stops over time	38
7	Dependent Variable Definition	44
8	Marginal Effects Covariates	53
9	GFC Sudden Stops	59
10	Post GFC Sudden Stops	61
11	Fitted Probabilities and Ex-Post Growth	63
12	Share of Countries with Short-Term Biases	79
13	Share of Countries with Short-Term Biases - Region	82
14	Share of Countries with Short-Term Biases - Consensus and World Bank	85
15	Simulated Debt-to-GDP: with and without bias	87
16	Distribution of Actual Real GDP Growth	89
17	Major Forecasters and Recessions	91
18	IMF Forecasts, Recessions and Financial Crises	93
19	Comparison between World Bank and IMF Forecasts	95
20	Comparison between European Commission and IMF Forecasts	98
21	Comparison between Consensus and IMF Forecasts	100
22	Comparison with Consensus: Anecdotal Evidence	107
23	Comparison with Consensus: Statistical Relationship	108
24	Converting Documents to Text Data	125
25	Size of the Corpus	127
26	Country coverage	129
27	Sketch of the Data Generating Process	137
28	Difference with Benchmark - Sovereign and Banking	145
29	Contemporaneous, Backward or Forward looking?	151
30	Examples of Violence Indicator	155
31	Non-economic crises: Breadth of Discussion	157
32	Non-economic crises: Depth of Discussion	159
33	A Network of Economic Complexity	162
34	Centrality of Expectations Channel	165
35	Sudden Stops Identification	189
36	Rolling Coefficients	203
37	Out-of-sample Estimated Probabilities	205

LIST OF FIGURES (Continued)

<u>FIGURE</u>		<u>PAGE</u>
38	Fitted Probabilities and Ex-Post Growth - Robustness	210
39	Country availability	211
40	Largest Optimistic Biases by Region	219
220figure.caption.80		
42	Number of IMF Documents per Year	255
43	Locating Turning Points in Economic Activity: United States	303
44	Examples of Natural Disasters Indicator	305

LIST OF ABBREVIATIONS

EM	Emerging Market
EWS	Early Warning System
GEP	Global Economic Prospects
GFC	Global Financial Crisis
IMF	International Monetary Fund
WEO	World Economic Outlook

CHAPTER 1

INTRODUCTION

1.1 Context and Motivation

The global economy has, in recent decades, become increasingly complex and interconnected, more than it has ever been in the past. The magnitude of cross-border financial positions and the number of links in the financial system mesh have increased swiftly from the mid-1990s onwards (Haldane, 2009; Lane and Milesi-Ferretti, 2007). Simultaneously, supply chains have markedly moved from a domestic or, at most, regional nature to a global one (Johnson and Noguera (2017)), and migration flows came to occupy an ever-larger share of the world population. Technological advances fostered this globalization process speeding up distance transactions, reducing their relative cost, and simplifying the overall movement of goods and people. At the same time, however, the development and spread of modern communication technologies have amplified hysteria and herd-like behaviours, facilitating contagion across markets and amplifying macroeconomic fluctuations. In an environment that currently exhibits multiple sources of both up and down-side risk, non-linearities, and tipping points, macroeconomic predictability has become a challenging endeavour for policymakers.

Despite these trends, the future is, however, born by anticipation: that is why forecasts are pervasive in all scientific fields and ultimately, in human life. Macroeconomic projections, in par-

ticular, direct the actions and decisions of individual governments, companies, and international organizations alike. On the one hand, reliable forecasts improve decision-making and planning, leading to the minimization of risks, the maximization of utilities, and eventually, higher general prosperity. Inaccurate forecasts, on the other hand, can lead economies fully astray (see for instance [Beaudry et al. \(2021\)](#)). Notwithstanding its importance, economic forecasting has only picked up recently in academic economic research, with the surge of machine learning related studies. In a more complex, interconnected, and nonlinear world, if anything, there is a need for more forecasting research, not less.

With this research, I contribute to the understanding of macroeconomic crises and forecasting biases, seeking to improve predictability in an increasingly complex world economy. Enhanced predictability, in turn, will allow governments and international organizations to prescribe and enforce more effective economic policies. In the rest of the dissertation, the main questions I will address are: can we accurately predict financial crises? To what extent are forecasters biased in their predictions? What are the drivers? Can we disentangle the multiple facets of macroeconomic crises?

1.2 Chapter 2: Are Financial Crises Predictable?

Financial crises are usually associated with large external adjustment, plummeting currencies, and ultimately, sharp increases in unemployment rates and rising poverty. Timely identification of the vulnerabilities giving rise to the aforementioned can avert the painful consequences associated with them; thus, it constitutes one of the key challenges for today's policymakers.

While previous literature has mostly focused on exchange-rate crises, in Chapter 3 I concentrate on a more general event, sudden stops, where external investors financing suddenly and abruptly comes to a halt, and ask whether these episodes constitute a completely random and unforeseeable shock or are the result of an endogenous build-up of observable vulnerabilities. Also for the summary of your contribution at the end. Concretely, I use a discrete choice model to predict the materialization of sudden stops within six quarters of advance and understand its determinants for a sample of 30 Emerging Markets over the period 1995q4-2017q1. At the same time, I discuss different "conceptual" issues related to the present relevance of this class of models for central banks and international organizations.

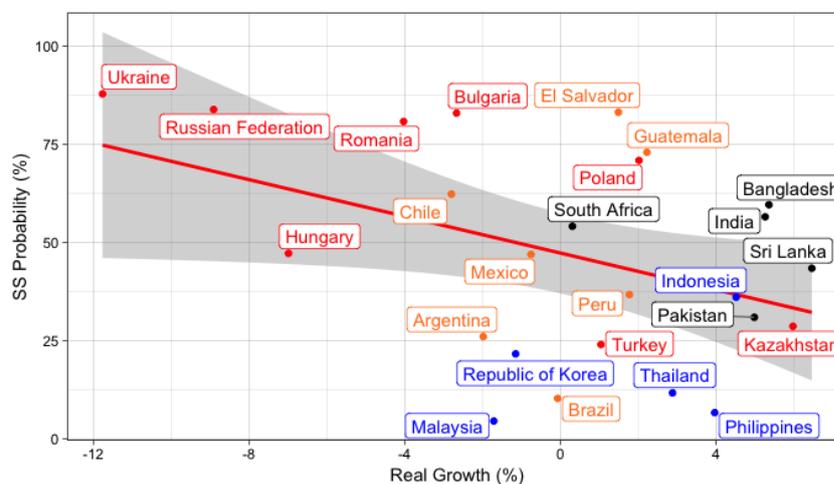
My contribution to the literature is manifold. First, I test a large pool of both domestic and global indicators of macro-financial vulnerabilities and relative transformations. I successfully overcome the large sample heterogeneity through country-specific percentile standardization of the indicators and find that domestic idiosyncrasies correspond to half of the explanation for sudden stops in EMs. Since domestic factors play a substantial role as leading indicators, the

reception of a signal by policymakers can trigger a pronounced correction of the fundamentals responsible for the rise in probability. This response would be, instead, highly constrained in case of a Global Cycle dominance.

Second, I set up an algorithmic framework to analyze the performance of the model recursively, mirroring accurately the information set available to forecasters at each step in time and appraise the model performance based on it. The model exhibits good sensitivity i.e. number of crises correctly called (47%) and very high specificity i.e. tranquil times correctly called (85%), largely improving over the chosen alternative.

Finally, I bridge two parts of the literature that have so far been kept separated: one is the classic Early Warnings literature that tries to predict in advance the occurrence of a crisis (e.g. [Bussière and Fratzscher \(2006\)](#), [Duca and Peltonen \(2013\)](#)), while the second, instead, tries to predict its incidence (for example, [Blanchard et al. \(2010\)](#); [Bussière \(2014\)](#); [Frankel and Saravelos \(2012\)](#)). I show that the estimated ex-ante probabilities of a sudden stop are highly correlated with the output impact of the ensuing event ([Figure 1](#)).

Figure 1: Sudden Stop Probabilities and Ex-Post Growth



Note: The figure shows the relationship between the median out-of-sample probability in the pre-crisis period for the GFC related sudden stops - obtained through the estimation of a discrete choice model for 30 Emerging Markets over the period 1995q4-2017q1 (see chapter 2) - and the median growth during the associated sudden stop. Red line is regression line with 95% confidence intervals. Red, orange, blue and gray labels indicate geographical group: respectively, Emerging Europe and Central Asia, Latin America, South-East Asia and an undefined category.

Taken altogether, these results assert the importance of EWSs as a tool for crises prevention, challenging skepticism on their use.

Moving forward, in the next chapter, I and co-authors review the performance of professional forecasters over the last 30 years, devoting particular attention to asymmetries in errors between

crises and no-crises periods.

1.3 Chapter 3: Stylized Facts on Economic Forecasting and Crises

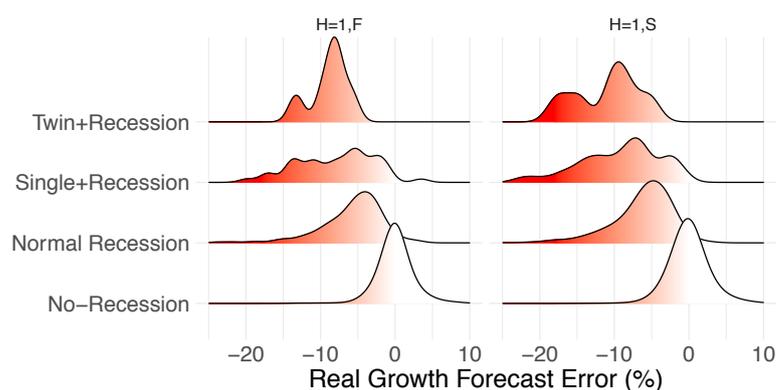
In Chapter 4, Zidong An, Prakash Loungani, and I study and evaluate short-term real GDP growth projections formulated by the main international economic institutions and private sector, starting from 1990 until the outburst of the Covid-19 crisis. Forecasts of economic activity are a critical component of international economics: indeed, they play a crucial role in shaping governments' policies, constituting the building stone of other pivotal institutional products such as fiscal and debt forecasts, and guiding private investors' allocation of capital.

In particular, in this chapter, we focus on the optimism usually ascribed by specialized news agencies, governments, and scholars alike to growth forecasts, relying on standard methodologies found in the forecasts evaluation literature (Clark and McCracken, 2013; Diebold and Lopez, 1996). Building on previous work (An et al., 2018; Loungani, 2001; Timmermann, 2007), we extend the analysis to the widest range of forecasters and longest time span at present: this completeness allows us, in turn, to draw general conclusions about the performance of the economic forecasting profession in the last decades. Are forecasts of economic activity too optimistic? To what extent? Is optimism shared across forecasters? What are its underlying causes?

We find evidence of widespread optimism in growth forecasts: optimism is much more prevalent than pessimism and ampler in magnitude. Optimistic biases stem from a sheer inability to

forecast accurately during downturns: errors during expansion are, on average, balanced and small, while errors during recessions are, instead, optimistic and sizeable. This result holds not only for the Fund, as put forward by previous literature, but for the whole sample of forecasters. In addition, we expand the aforementioned literature showing that the most optimistic forecast errors correspond to complex, multifaceted episodes of crises (Figure 2).

Figure 2: Asymmetries in Growth Forecast Errors - Financial Crises, Recessions and Non-Recessions



Source: World Economic Outlook

Note: Distribution of real GDP growth WEO forecast errors for episodes of no-recession, only recession and recession accompanied by a single or twin financial crises. Actual real GDP growth rate from October WEO issue of following year. Recessions are periods of negative growth. Financial crises correspond to currency, banking and sovereign debt crises. The corresponding dummy is from [Laeven and Valencia \(2018\)](#). H=1 indicates year-ahead forecast errors. F and S indicate, respectively, Fall and Spring issue forecast errors. Brighter red indicates more optimistic errors.

We complement these novel stylized facts with evidence on the high degree of cross-pollination between different forecasters and analyzing the specific case of IMF program forecasts.

Observing the dismal performance of the economic forecasting profession in the last 30 years and the close link between large negative forecast errors and multifaceted crises lays the foundations for the next chapter in the dissertation. In Chapter 5, as a matter of fact, Manuel Betín and I dig deeper into complex crises, seeking to go beyond the traditional “twin/triple crises” narrative.

1.4 Chapter 4: Understanding Complex Crises

Macroeconomic crisis analysis necessarily suffers from substantial data limitations that often limit the general understanding of these highly chaotic and painful episodes. While the narrative literature has tried to capture the entirety of events at play in a qualitative way through lengthy descriptions (Diaz-Alejandro, 1985; Kindleberger, 1975), the empirical crises literature has devoted particular attention to a restricted subset of crises, focusing on their co-occurrence (e.g. Kaminsky and Reinhart (1999); Laeven and Valencia (2018); Reinhart and Rogoff (2009)). In this final chapter, tapping from the IMF archives and the text mining literature, we aim at providing researchers with new quantitative material to fill this gap.

We focus on the International Monetary Fund as preferred text source for different reasons related to the mandate and expertise of the Fund as well as the characteristics of the documents produced over time:

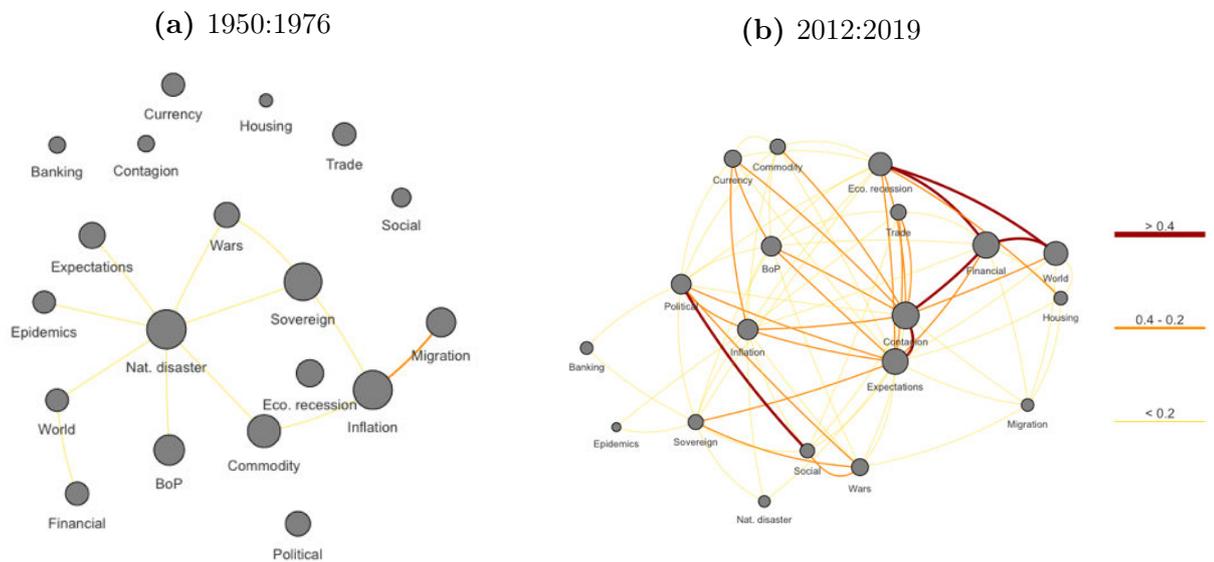
1. **Explicit surveillance mandate:** the Fund regularly monitors and evaluates the situation of economic and financial systems to identify contingent sources of risk. This guarantees a thorough coverage of volatility episodes.
2. **Credibility:** the Fund is the leading institution for matters related to financial stability. The close interaction of staff with leading scholars, national economic actors, and politicians ensures the quality and resiliency to structural breaks of the reports.
3. **Long and detailed:** the documents produced depict meticulously the real-time outlook of the country under scrutiny.
4. **Extremely cautious:** each word is discussed, weighed, and negotiated and follows a strict protocol of revision and publishing. In contrast to other textual sources (e.g. press or tweets), this painstaking revision process leads to a situation of homogeneous linguistic.

Exploiting Optical Character Recognition (OCR) and auto-correction techniques to overcome the accuracy hurdle that hindered previous work (Mihalyi and Mate, 2019), we provide accessibility to a raw text database of roughly 23,000 documents, country reports and program-related, for a grand total of over 20 million pages, covering in an unbalanced way the whole IMF membership throughout the period 1950-2019. Manually reading a large number of documents, we then compile a crises vocabulary and use a flexible and computationally feasible

term-frequency approach to capture Fund discussions about 20 different types of economic and non-economic events, ranging from standard recessions to more unconventional events such as violence crises and natural disasters.

We validate these new narrative indicators against the benchmarks established in the literature and examine the evolution of the “crisis system” (crises co-occurrence) from the Bretton Woods period until the present.

Figure 3: A Network of Crises - 70 Years Later



Note: Adjacency matrix built from pairwise correlations between term-frequencies: minimum correlation to display edge equal to 0.1. Size of nodes proportional to their eigencentality. Legend indicates correlations between categories. Visualisation of the network through the ForceAtlas2 algorithm (Jacomy et al., 2014).

We document a clear shift from a system dominated by real and non-economic crises, characterized by fewer links, to a financially dominated network with the characteristics of a “small world”, where crisis spill over one another with ease ([Figure 3](#)).

Closing the loop I opened with the first chapter, this result highlights the intrinsic difficulty in forecasting: while the current economic system can remain stable for long periods, a small perturbation can spread quickly across sectors and breed into complex outcomes. This conclusion opens up the discussion on new ways of tailoring economic policies and forecasting practices.

1.5 Summary: Research Questions

On Financial Crises Prediction

- Can we identify the build-up of vulnerabilities in the external sector of EMs that precedes a sudden stop of capital flows?
- All considered, are Early Warning Systems still relevant for policymakers?

On Forecasts Evaluation

- How good have economic forecasters predicted real economic activity over the last 30 years?
- What are the determinants of their performance? How do they differ between themselves?

On Narrative Economics and Crises Complexity

- Can we provide accessibility to researchers and extract relevant crises information from IMF documents? How?
- What can narrative elements teach us on the evolution of the “crisis” system in the last 70 years?

1.6 Summary: Contributions

On Financial Crises Prediction

- I provide detailed literature review and debate the “conceptual” issues behind the use of Early Warnings Systems as a tool of financial crises prevention
- I test a large set of macro-financial variables (domestic and global) and relative transformations on the predictability of sudden stops for 30 Emerging Markets over the period 1995-2017. I show that domestic and global factors contribute in a similar way to the reversal in capital flows
- I develop a recursive algorithm to test in a realistic way the performance of an Early Warning System (i.e. reflecting the information set of the policymaker at each point in time)
- I show that ex-ante estimated probabilities and output cost of the sudden stop are negatively correlated

On Forecasts Evaluation

- We provide detailed literature review on different aspects of forecast evaluation
- We gather a large panel of short-term growth projections that includes forecasts from all major international institutions (IMF, World Bank and European Commission) and the private sector (at the aggregate and individual level) over the period 1990-2019

- We provide evidence of a startling general inability of the economic profession to forecast turning points in economic activity
- We link large forecast errors to the the occurrence of multifaceted crises

On Narrative Economics and Crises Complexity

- We provide detailed literature review on the use of text analysis in macroeconomics
- We develop a rigorous and automated procedure to scrape, select relevant documents and extract text data in a statistical-friendly format from the IMF archives. We furnish a new database of roughly 23,000 IMF country documents covering in an unbalanced way the whole IMF membership over the period 1950-2019
- We develop a term-frequency approach composed by a supervised vocabulary of almost 1000 keywords to capture Fund discussions about 20 different categories of economic and non economic events. We furnish a new database of IMF crises discussion for the whole sample of IMF documents
- We show that crises co-occurrence has risen considerably in the last 70 years

CHAPTER 2

INTRODUCTION (EN FRANÇAIS)

2.1 Contexte et Motivation

Au cours des dernières décennies, l'économie mondiale est devenue de plus en plus complexe et interconnectée, plus qu'elle ne l'a jamais été par le passé. L'ampleur des positions financières transfrontalières et le nombre de liens dans le système financier se sont rapidement multipliés à partir du milieu des années 1990 (Haldane (2009), Lane and Milesi-Ferretti (2007)). Simultanément, les chaînes d'approvisionnement ont nettement passées d'une nature domestique ou, tout au plus, régionale, à une nature mondiale (Johnson and Noguera (2017)), et les flux migratoires en sont venus à occuper une part toujours plus grande de la population mondiale. Les progrès technologiques ont favorisé ce processus de mondialisation en accélérant les transactions à distance, réduisant leur coût relatif et simplifiant la circulation globale des biens et des personnes. Les progrès technologiques ont favorisé ce processus de mondialisation en accélérant les transactions à distance, en réduisant leur coût relatif et en simplifiant la circulation générale des biens et des personnes. Dans le même temps, le développement et la diffusion des technologies de communication modernes ont amplifié l'hystérie et les comportements grégaires, facilitant la contagion entre les marchés et amplifiant les fluctuations macroéconomiques. Dans un environnement qui présente actuellement des sources multiples de sources de risques à la hausse et à la baisse, de non-linéarités et de points de basculement, la prévisibilité macroéconomique est

devenue un défi pour les responsables politiques.

Malgré ces tendances, l'avenir naît pourtant de l'anticipation : c'est pourquoi les prévisions sont dans tous les domaines scientifiques et, en fin de compte, dans la vie humaine. Les projections macroéconomiques, en particulier, orientent les actions et les décisions des gouvernements, des entreprises et des organisations internationales. D'une part, des prévisions fiables améliorent la prise de décision et la planification, ce qui permet de minimiser les risques, de maximiser les bénéfices et, finalement, d'accroître la prospérité générale. D'autre part, des prévisions inexactes peuvent entraîner les économies dans une dérive totale (voir par exemple [Beaudry et al. \(2021\)](#)). En dépit de son importance, la prévision économique tend à être négligée par la recherche économique académique au profit d'autres sujets. Dans un monde plus complexe, interconnecté et non linéaire, il est nécessaire de développer la recherche sur les prévisions, et non de la réduire.

Avec cette recherche, je contribue à la compréhension des crises macroéconomiques et des biais de prévision, en cherchant à améliorer la prévisibilité dans une économie mondiale de plus en plus complexe. Une meilleure prévisibilité, à son tour, permettra aux gouvernements et aux organisations internationales de prescrire et d'appliquer des politiques économiques plus efficaces. Dans le reste de la thèse, les principales questions que j'aborderai sont les suivantes : peut-on prédire avec précision les crises financières? Dans quelle mesure les prévisionnistes sont-ils biaisés dans leurs prédictions? Quels en sont les moteurs? Pouvons-nous démêler les multiples facettes des crises macroéconomiques?

2.2 Chapitre 3: Les crises financières sont-elles prévisibles?

Les crises financières sont généralement associées à d'importants ajustements externes, à l'effondrement des monnaies et, en fin de compte, à une forte augmentation des taux de chômage et de la pauvreté. L'identification en temps utile des vulnérabilités à l'origine de ces phénomènes peut permettre d'éviter les conséquences douloureuses qui y sont associées. Il s'agit donc de l'un des principaux défis que doivent relever les décideurs politiques d'aujourd'hui. Dans le Chapitre 3, je me concentre sur un type particulier de crise financière, les arrêts soudains, où le financement des investisseurs extérieurs s'arrête soudainement et brutalement, et je me demande si ces événements constituent un choc totalement aléatoire et imprévisible ou sont le résultat d'une accumulation endogène de vulnérabilités observables. Concrètement, j'utilise un modèle de choix discret pour prédire la matérialisation des arrêts soudains dans les six trimestres à venir et comprendre ses déterminants pour un échantillon de 30 marchés émergents sur la période 1995q4-2017q1. Dans le même temps, je discute différentes questions "conceptuelles" liées à la pertinence actuelle de cette classe de modèles pour les banques centrales et les organisations internationales.

Ma contribution à la littérature est multiple. Tout d'abord, je teste un large éventail d'indicateurs de vulnérabilité indicateurs nationaux et internationaux de vulnérabilités macro-financières et de transformations relatives. Je réussis à surmonter la grande hétérogénéité de l'échantillon grâce à la normalisation du percentile des indicateurs par pays, et je trouve que les idiosyncrasies domestiques correspondent à la moitié de l'explication des arrêts soudains

dans les pays émergents. Comme les facteurs domestiques jouent un rôle important en tant qu'indicateurs avancés, la réception d'un signal par les décideurs politiques peut déclencher une correction prononcée des fondamentaux responsables de la hausse de la probabilité. Cette réponse serait, en revanche, très limitée en cas de dominance du cycle global.

Deuxièmement, j'ai mis en place un cadre algorithmique pour analyser la performance du modèle de manière récursive, en reflétant avec précision l'ensemble des informations disponibles pour les prévisionnistes à chaque étape du temps et en évaluant la performance du modèle sur cette base. Le modèle présente une bonne sensibilité, c'est-à-dire le nombre de crises correctement annoncées (47%), et une spécificité très élevée, c'est-à-dire le nombre de périodes de tranquillité correctement annoncées (85%), ce qui constitue une nette amélioration par rapport à l'alternative choisie.

Enfin, je relie deux parties de la littérature qui sont restées séparées jusqu'à présent : l'une est la littérature classique sur les alertes précoces qui tente de prédire à l'avance l'apparition d'une crise (par exemple, [Bussière and Fratzscher \(2006\)](#), [Duca and Peltonen \(2013\)](#)), tandis que la seconde, au contraire, tente de prédire son incidence (par exemple, [Blanchard et al. \(2010\)](#); [Bussière \(2014\)](#); [Frankel and Saravelos \(2012\)](#)). Je montre que les probabilités ex ante estimées d'un arrêt brutal sont fortement corrélées avec l'impact de l'arrêt sur la production de l'événement qui s'ensuit ([Figure 1](#)).

Dans l'ensemble, ces résultats affirment l'importance des EWS en tant qu'outil de prévention des crises, remettant en cause le scepticisme quant à leur utilisation.

Dans le chapitre suivant, mes coauteurs et moi-même passons en revue les performances des prévisionnistes professionnels au cours des 30 dernières années, en accordant une attention particulière aux asymétries des erreurs entre les périodes de crise et les périodes sans crise.

2.3 Chapitre 4: Faits stylisés sur les prévisions économiques et les crises

Dans le Chapitre 4, Zidong An, Prakash Loungani et moi-même étudions et évaluons les prévisions de croissance à court terme du PIB réel formulées par les principales institutions économiques internationales et le secteur privé, depuis 1990 jusqu'à l'éclatement de la crise Covid-19. Les prévisions de l'activité économique sont une composante essentielle de l'économie internationale : en effet, elles jouent un rôle crucial dans l'élaboration des politiques des gouvernements, constituent la pierre angulaire d'autres produits institutionnels essentiels tels que les prévisions budgétaires et de la dette, et guident l'allocation des capitaux par les investisseurs privés.

En particulier, dans ce chapitre, nous nous concentrons sur l'optimisme généralement attribué par les agences de presse spécialisées, les gouvernements et les universitaires aux prévisions de croissance, en nous appuyant sur les méthodologies standard que l'on trouve dans la littérature d'évaluation des prévisions (Clark and McCracken (2013); Diebold and Lopez (1996)). En

nous appuyant sur des travaux antérieurs (An et al. (2018) ; Loungani (2001) ; Timmermann (2007)), nous étendons l'analyse au plus large éventail de prévisionnistes et à la plus longue période actuelle : cette exhaustivité nous permet, à son tour, de tirer des conclusions générales sur la performance de la profession de prévisionniste économique au cours des dernières décennies. Les prévisions de l'activité économique sont-elles trop optimistes? Dans quelle mesure? L'optimisme est-il partagé par tous les prévisionnistes? Quelles sont ses causes sous-jacentes?

Nous trouvons des preuves d'un optimisme généralisé dans les prévisions de croissance : l'optimisme est beaucoup plus répandu que le pessimisme et son ampleur est plus grande. Les biais optimistes découlent d'une incapacité pure et simple à prévoir avec précision pendant les ralentissements : les erreurs pendant l'expansion sont, en moyenne, équilibrées et faibles, tandis que les erreurs pendant les récessions sont, au contraire, optimistes et importantes. Ce résultat est valable non seulement pour le Fonds, comme l'indique la littérature antérieure, mais aussi pour l'ensemble de l'échantillon de prévisionnistes. En outre, nous développons la littérature susmentionnée qui montre que les erreurs de prévision les plus optimistes correspondent à des situations complexes et multiformes (Figure 2).

Nous complétons ces nouveaux faits stylisés par des preuves du haut degré de pollinisation croisée entre les différents prévisionnistes et en analysant le cas spécifique des prévisions des programmes du FMI.

L’observation des performances déplorables de la profession de prévisionniste économique au cours des 30 dernières années et du lien étroit entre les erreurs de prévision négatives importantes et les crises multiformes jette les bases du chapitre suivant de la thèse. En effet, dans le Chapitre 5, Manuel Betín et moi-même approfondissons les crises complexes, en cherchant à aller au-delà du récit traditionnel des “crises jumelles/triples”.

2.4 Chapitre 5: Comprendre les crises complexes

L’analyse des crises macroéconomiques souffre nécessairement d’importantes limitations des données qui limitent souvent la compréhension générale de ces épisodes hautement chaotiques et douloureux. Alors que la littérature narrative a tenté de saisir l’intégralité des événements en jeu de manière qualitative par le biais de longues descriptions ([Diaz-Alejandro \(1984\)](#), [Kindleberger \(1975\)](#)), la littérature empirique sur les crises a accordé une attention particulière à un sous-ensemble restreint de crises, en se concentrant sur leur cooccurrence (par exemple, [Kaminsky \(1998\)](#); [Laeven and Valencia \(2018\)](#); [Reinhart and Rogoff \(2009\)](#)). Dans ce dernier chapitre, en puisant dans les archives du FMI et dans la littérature sur l’exploration de textes, nous visons à fournir aux chercheurs de nouveaux éléments quantitatifs pour combler cette lacune.

Nous nous concentrons sur le Fonds monétaire international comme source de texte privilégiée pour différentes raisons liées au mandat et à l’expertise du Fonds ainsi qu’aux caractéristiques des documents produits au fil du temps:

1. Mandat de surveillance explicite : le Fonds surveille et évalue régulièrement la situation des systèmes économiques et financiers afin d'identifier les sources de risque éventuelles. Cela garantit une couverture approfondie des épisodes de volatilité.
2. Crédibilité : le Fonds est l'institution de référence pour les questions liées à la stabilité financière. L'interaction étroite du personnel avec des universitaires de renom, des acteurs économiques nationaux et des politiciens garantit la qualité et la résistance aux ruptures structurelles des rapports.
3. Longs et détaillés : les documents produits décrivent méticuleusement les perspectives en temps réel du pays examiné.
4. Extrêmement prudents : chaque mot est discuté, pesé et négocié et suit un protocole strict de révision et de publication. Contrairement à d'autres sources textuelles (par exemple la presse ou les tweets), ce processus de révision minutieux conduit à une situation d'homogénéité linguistique.

En exploitant les techniques de reconnaissance optique de caractères (ROC) et d'autocorrection pour surmonter l'obstacle de la précision qui a entravé les travaux précédents (Mihalyi and Mate, 2019), nous donnons accès à une base de données de texte brut d'environ 23 000 documents, rapports nationaux et programmes connexes, pour un grand total de plus de 20 millions de pages, couvrant de manière déséquilibrée l'ensemble des membres du FMI sur la période 1950-2019. En lisant manuellement un grand nombre de documents, nous compilons ensuite un vocabulaire de crise et utilisons une approche flexible et calculable de la fréquence des termes pour capturer les discussions du FMI sur 20 types différents d'événements économiques et non économiques,

allant des récessions standard aux événements plus non conventionnels tels que les crises de violence et les catastrophes naturelles.

Nous validons ces nouveaux indicateurs narratifs par rapport aux repères établis dans la littérature et examinons l'évolution du "système de crise" (cooccurrence des crises) depuis la période de Bretton Woods jusqu'à aujourd'hui.

Nous constatons un passage clair d'un système dominé par des crises réelles et non économiques, caractérisé par un nombre réduit de liens, à un réseau dominé par la finance et présentant les caractéristiques d'un "petit monde", où les crises se déversent facilement les unes sur les autres (Figure 3).

Fermant la boucle que j'ai ouverte avec le premier chapitre, ce résultat met en évidence la difficulté intrinsèque de la prévision : si le système économique actuel peut rester stable pendant de longues périodes, une petite perturbation peut se propager rapidement dans tous les secteurs et se traduire par des résultats complexes. Cette conclusion ouvre la discussion sur de nouvelles façons d'adapter les politiques économiques et les pratiques de prévision.

2.5 Résumé: Questions de recherche

Sur la prédiction des crises financières

- Pouvons-nous identifier l'accumulation de vulnérabilités dans le secteur extérieur des économies émergentes qui précède un arrêt soudain des flux de capitaux?
- Tout bien considéré, les systèmes d'alerte précoce sont-ils toujours pertinents pour les décideurs politiques?

Sur l'évaluation des prévisions

- Dans quelle mesure les prévisionnistes économiques ont-ils prédit l'activité économique réelle au cours des 30 dernières années?
- Quels sont les déterminants de leur performance? Comment diffèrent-ils entre eux?

Sur l'économie narrative et la complexité des crises

- Pouvons-nous rendre accessible aux chercheurs et extraire des informations pertinentes sur les crises à partir des documents du FMI? Comment?
- Que peuvent nous apprendre les éléments narratifs sur l'évolution du système de "crise" au cours des 70 dernières années?

2.6 Résumé : Contributions

Sur la prédiction des crises financières

- Je fournis une revue détaillée de la littérature et débat des questions "conceptuelles" derrière l'utilisation des systèmes d'alerte précoce comme outil de prévention des crises financières.
- Je teste un large ensemble de variables macro-financières (domestiques et globales) et leurs transformations relatives sur la prévisibilité des arrêts soudains pour 30 marchés émergents sur la période 1995-2017. Je montre que les facteurs domestiques et globaux contribuent de manière similaire à l'inversion des flux de capitaux.
- Je développe un algorithme récursif pour tester de manière réaliste la performance d'un système d'alerte précoce (c'est-à-dire reflétant l'ensemble des informations du décideur politique à chaque moment).
- Je montre que les probabilités estimées ex ante et le coût de production de l'arrêt brutal sont négativement corrélés.

Sur l'évaluation des prévisions

- Nous fournissons une revue détaillée de la littérature sur les différents aspects de l'évaluation des prévisions.
- Nous rassemblons un large panel de projections de croissance à court terme qui inclut les prévisions de toutes les grandes institutions internationales (FMI, Banque

mondiale et Commission européenne) et du secteur privé (au niveau agrégé et individuel) sur la période 1990-2019.

- Nous apportons la preuve d'une incapacité générale étonnante de la profession économique à prévoir les points de retournement de l'activité économique
- Nous établissons un lien entre les erreurs de prévision importantes et l'apparition de crises à multiples facettes.

Sur l'économie narrative et la complexité des crises

- Nous fournissons une revue détaillée de la littérature sur l'utilisation de l'analyse de texte en macroéconomie.
- Nous développons une procédure rigoureuse et automatisée pour gratter, sélectionner les documents pertinents et extraire des données textuelles dans un format convivial pour les statistiques à partir des archives du FMI. Nous fournissons une nouvelle base de données d'environ 23 000 documents nationaux du FMI couvrant de manière déséquilibrée l'ensemble des membres du FMI sur la période 1950-2019.
- Nous développons une approche de fréquence des termes composée d'un vocabulaire supervisé de près de 1000 mots-clés pour capturer les discussions du FMI sur 20 catégories différentes d'événements économiques et non économiques. Nous fournissons une nouvelle base de données des crises du FMI discussion pour l'ensemble de l'échantillon de documents du FMI

- Nous montrons que la cooccurrence des crises a considérablement augmenté au cours des 70 dernières années.

CHAPTER 3

FINDING A NEEDLE IN A HAYSTACK

DO EARLY WARNING SYSTEM FOR SUDDEN STOPS WORK?

Abstract: The chapter develops an Early Warning System (EWS) to identify the build up of vulnerabilities in the external sector of 31 Emerging Markets (EMs) across the period 1995-2017 and avoid the painful sudden reversal of capital flows associated to them. It contributes to the literature on the prediction of financial discontinuities in three ways. First, it uses a discrete choice model to calculate and compare the marginal effect of different domestic and global factors on the probability of a sudden stop materializing. Second, it analyzes the performance of the model with a recursive framework that reflects accurately the information set available to policymakers at the time of the prediction. Third, it investigates the relationship between ex-ante probability of a sudden stop and the ensuing output loss. We find that domestic and global factors contribute to the reversal of capital flows in a comparable way. Our model calls half of the pre-crisis periods, exhibiting a high specificity and a proper timing. Moreover, we find a positive link between the ex-ante probability of a sudden stop and the associated ex-post loss. These results call for an active use of Early Warnings in the policy-making sphere.

3.1 Introduction

This time may seem different, but all too often a deeper look shows it is not. Encouragingly, history does point to warning signs that policy makers can look at to assess risk—if only they do not become too drunk with their credit bubble-fueled success and say, as their predecessors have for centuries, “This time is different”.

Reinhart and Rogoff (2009)

The global retrenchment of capital flows during the Great Recession and the upheavals concomitant with the raise of US target rate at the end of 2015 have once again raised the issue of inflows-fueled booms and following busts in Emerging Markets (EMs). Fickle capital markets put under severe strain these countries: plummeting currencies, external adjustment and in some cases, defaults, resulted in dramatic output losses and rising poverty. The recent Covid-19 crisis, although differs markedly from the usual boom-bust cycle, is exacerbating existing domestic vulnerabilities in the aforementioned countries: while capital flows appear rather stable, investors could soon decide to withdraw from their most risky positions, hence jeopardising financial stability (Kalemli-Ozcan (2020)). A timely identification of the vulnerabilities giving rise to sudden stops can prevent the painful consequences associated to them.

For this reason, interest toward Early Warning System (EWSs) has re-kindled in policy institutions such as the IMF and national central banks (e.g. Basu et al. (2019), Suss and Treitel (2019), Beutel et al. (2018), Duca and Peltonen (2013)). Nevertheless, many scholars remain doubtful about the predictive power of EWSs: while in-sample they generally work well,

out-of-sample they “consistently fail to predict the upcoming wave of external crises” (Rose and Spiegel, 2010, 2011).

In this chapter, we exploit a discrete choice model to predict the materialization of sudden stops within six quarters of advance and understand its determinants. We expand the literature in different directions. First, we test a large pool of both domestic and global indicators of macro-financial vulnerabilities and relative transformations. We uncover a near equivalence between the marginal effect of domestic and global factors on the probability of a sudden stop: this result stands in stark contrast with the proposition that EMs are solely at the mercy of a Global Financial Cycle (Rey, 2015). Since domestic factors play a more substantial role than previously maintained, the usefulness of Early Warnings for financial stability purposes increases significantly: indeed, the reception of a signal by policymakers can trigger a pronounced correction of the fundamentals responsible for the rise in probability. Policy options are, instead, far more constrained in the case of a Global Cycle dominance.

Second, we offer a framework to analyze the performance of the model recursively, mirroring accurately the information set available to forecasters at the time of the prediction and appraise our model based on it. The model exhibits good sensitivity i.e. number of crises correctly called (47%) and very high specificity i.e. tranquil times correctly called (85%), largely improving over the chosen alternative, a naive-decision benchmark. Moreover, we show that the sensitivity of a classifier can be a misleading evaluation metric when dealing with problems such as financial crises prediction. We find that the estimated ex-ante probabilities of a sudden stop are highly correlated with the output impact of the ensuing event: episodes with a catastrophic impact on

real activity are predicted with virtual certainty by the model, while those that entail only mild slowdowns are either missed or called with low probability. This result is robust to the choice of different measures of output impact.

The paper is structured as follows. Section 3.2 summarizes the relevant literature. Section 3.3 explains our identification of sudden stops and introduces the indicators tested. Section 3.4 delves into our methodology. Section 3.5 presents all our main results. Lastly, section 3.6 concludes.

3.2 Related Literature

This paper relates to different strands of the international finance literature. First, it links to the growing literature on the determinants of capital flows cyclical behaviour in emerging markets and the disruptive events associated to their reversal. In particular, the recent debate revolves around the relative importance of global (“push”) and domestic (“pull”) factors. While initially these works focused on net capital flows (Calvo et al. (2004), Levchenko and Mauro (2007)), after the pioneering work of Forbes and Warnock (2012) the attention has gradually shifted to monitoring gross flows. For a large sample of emerging and advanced economies alike spanning the period 1980-2009, Forbes and Warnock (2012) find that sudden stops in gross inflows are mainly caused by global factors: surge in risk aversion, proxied by the VIX, and slowdown in global economic activity. Local factors, including different capital controls measures, are, instead, not relevant. Fratzscher (2012) studies the high-frequency dynamics of portfolio flows for 50 emerging and advanced economies in the period around the 2007-2008 crisis. He finds that common shocks, spike in risk aversion and liquidity risk, drove the reallocation of capital flows from emerging markets to advanced economies amidst the GFC. Nevertheless, sensitivity to these common factors is largely explained by country-specific characteristics. Moreover, in the immediate recovery period, there seems to have been a re-balancing between push and pull factors. Eichengreen and Gupta (2016) analyze sudden stops in gross inflows for a sample of 34 emerging markets. The authors compare the magnitude and significance of different correlates between the years 1980-2003 and 2003-2013. They find that risk aversion played a key role in the more recent waves of sudden stops, while the impact of local factors have been mostly

insignificant. The opposite holds for the earlier period. [Cerutti et al. \(2017\)](#) claim that the importance of global factors has been overstated by the literature. Working on a panel of 63 advanced and emerging economies, they show that the goodness-of-fit of push regressions is always extremely low. [Eichengreen et al. \(2018\)](#) employ data on capital flows disaggregated by type and instrument for the same sample as in [Eichengreen and Gupta \(2016\)](#). They ask whether different flows react to the same set of covariates. Their result suggest that FDI inflows react more to pull factors, while portfolio debt and equity inflows to push factors. Other inflows, that represent the greatest share of total inflows to emerging markets and are mainly composed by banking flows, respond to both similarly.

Second, it clearly relates to the EWS literature on financial crises. This strand aims at the construction of a model able to forecast in advance the occurrence of different types of rare and disruptive events. Timely and reliable signals, in turn, can allow the intervention of policymakers and avoid the devastating macroeconomic consequences usually ensuing. EWSs have been developed and tailored specifically for different type of crises: mainly currency crises ([Frankel and Rose \(1996\)](#), [Reinhart et al. \(1998\)](#), [Berg and Patillo \(1999\)](#), [Kaminsky \(2003\)](#), [Bussière and Fratzscher \(2006\)](#), [Bussière \(2007\)](#), [Bussière \(2013\)](#)) and banking crises ([Alessi and Detken \(2009\)](#), [Babecký et al. \(2012\)](#), [Duca and Peltonen \(2013\)](#), [Alessi and Detken \(2018\)](#), [Aldasoro et al. \(2018\)](#), [Beutel et al. \(2018\)](#)), but also sovereign debt crises ([Manasse et al. \(2003\)](#), [Manasse and Roubini \(2009\)](#)), external crises ([Catão and Milesi-Ferretti, 2014](#)), IMF interventions ([Frankel and Saravelos, 2012](#)) and more recently, also sudden stops ([Basu et al., 2019](#)). The techniques used range from the discrete-dependent variable approach (logit and

probit models) to the leading indicators approach, in which single indicators send a signal when crossing a critical threshold and those signals are then aggregated and weighed and more modern machine learning (ML) techniques.

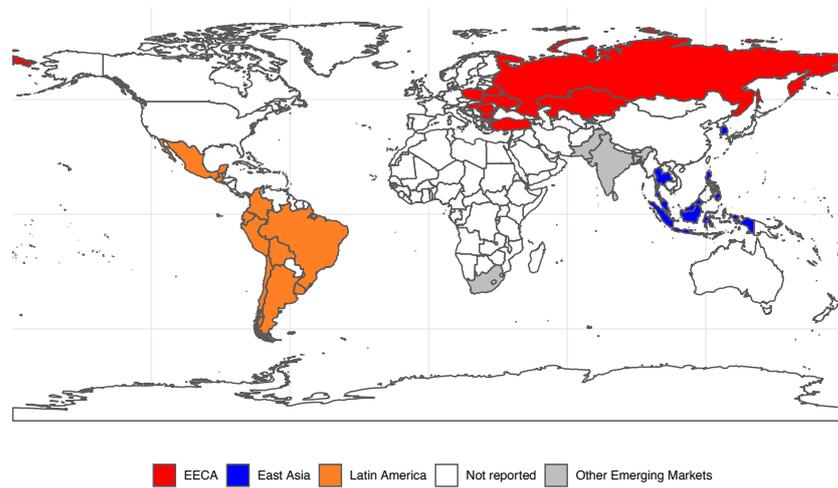
Third, it is connected to the literature on the 2007-2008 financial crisis and the heterogeneous cross-country incidence it had on real activity. [Gourinchas and Obstfeld \(2012\)](#) analyze the behaviour of different indicators around crisis events for advanced and emerging economies during the second part of the 20th century and compare it to the pre-GFC period. They conclude that countries that avoided large appreciations of their currencies, credit booms and hoarded international liquidity during the 2000s also were most likely to avoid the worst effects of the twenty-first century first global crisis. [Frankel and Saravelos \(2012\)](#) review the most consistent early warnings indicators found in the literature and ask whether these were able to predict the incidence of the GFC across countries. Their results comply with the findings of [Gourinchas and Obstfeld \(2012\)](#). On the other hand, earlier papers such as [Blanchard et al. \(2010\)](#) and [Rose and Spiegel \(2010\)](#) do not find any relationship between the causes of the crisis and its severity.

Compared to previous literature, in this paper we test a large pool of global and local factors and explicitly quantify their marginal effect on the probability of sudden stops. In addition, from the methodology standpoint, we introduce a recursive framework to appraise realistically the past performance of an Early Warning and debate the use of sensitivity as a correct evaluation metric.

3.3 Data

We collect quarterly frequency data for 31 emerging markets over the period 1995Q4-2017Q1. In [Figure 4](#) we display the regional composition of our sample.¹

Figure 4: Countries Sample



3.3.1 Sudden stops

Our definition of sudden stop follows step-by-step the algorithm developed by [Forbes and Warnock \(2012\)](#) and matches our theoretical understanding of sudden stops: a large drop in

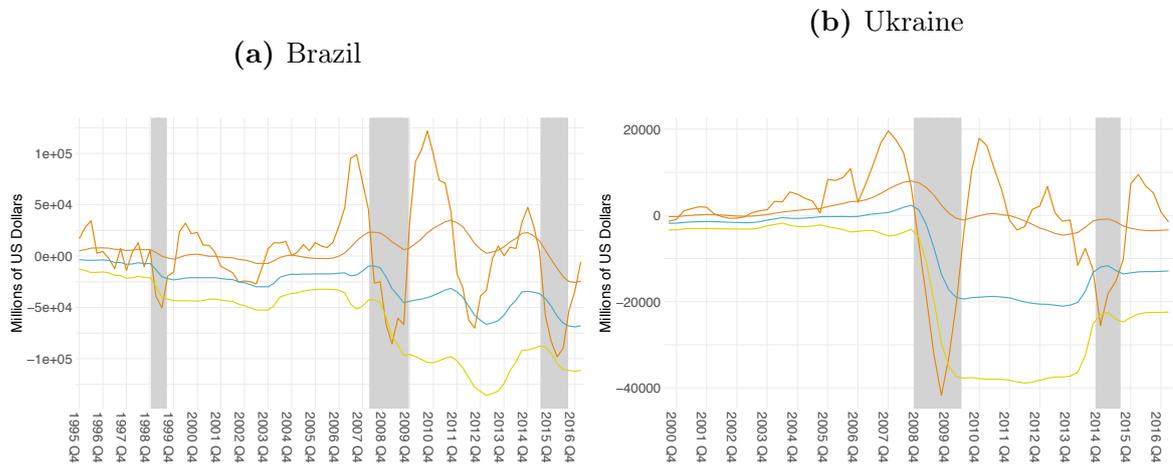
¹For a full list, see [Table XXV](#) in the appendix.

foreign capital inflows that persists over a prolonged period of time. For all the countries in our sample, we obtain total gross inflows in a quarter by summing up the total liabilities flows in the Financial Account i.e. Foreign Direct Investments (FDIs), portfolio investments and other investments liabilities. Data on inflows are retrieved from the International Financial Statistics (IFS) database of the IMF.

In practice, let us define $C_{i,t}$ as the total inflows for a country i in a quarter t . We calculate the cumulative sum of yearly inflows for each country as $C_{i,t}^{sum} = \sum_{s=0}^3 C_{i,t-s}$ and calculate the yearly growth rate $\Delta C_{i,t}^{sum} = C_{i,t}^{sum} - C_{i,t-4}^{sum}$ to remove seasonality issues. We then compute the rolling mean and standard deviation for $\Delta C_{i,t}^{sum}$ over a period of 5 years. We identify a sudden stop when $\Delta C_{i,t}^{sum}$ drops by more than 2 standard deviations below its rolling mean. The episode, however, begins when the drop is greater than one standard deviation from the mean.²³ Figure 5 shows a graphical example of the algorithm.

²We exclude episodes that last only one quarter and collapse adjacent sudden stops into the same episode if the gap among the end of the former and the start of the latter is equal or lower than two quarters.

³We also create an alternative longer definition of ending: in this case inflows have to come back to their rolling mean in order to mark the end of an episode. We use this longer definition of duration in the robustness part.

Figure 5: Examples of Sudden Stops Identification

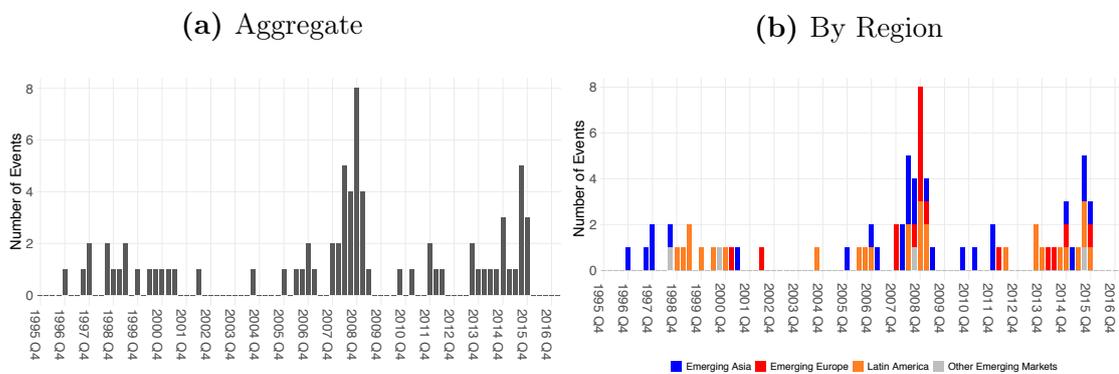
Note: The figure shows the algorithm proposed by Forbes and Warnock (2012) for the identification of sudden stops applied to our sample. A sudden stop begins when the y-o-y gross capital inflows (dark orange line) go below their rolling mean minus one standard deviation (light blue line) conditional on crossing the rolling mean minus two standard deviations (yellow line). The episode ends when y-o-y gross inflows come back above their rolling mean minus one standard deviation. The duration is highlighted by the grey shaded area.

We identify a total of 75 sudden stops ($SS_{i,t}$).⁴ Figure 6a shows the number of countries experiencing a sudden stop throughout the sample period. Clusters of sudden stops generally correspond to very well known macroeconomic and financial events e.g. the East Asian crisis, the

⁴For 10 episodes, the two thresholds criterion is missed by a whisker. Nevertheless, after consulting the IMF Article IV, they are included.

Global Financial Crisis (GFC) and the turbulence following the normalization of US monetary policy in the post-GFC period. Panel 6b, instead, highlights the characteristic of regional contagion in sudden stops: episodes tend to occur temporally closely in the same EM region.

Figure 6: Number of Sudden Stops over time



3.3.2 Explanatory Variables

We test an extensive set of indicators that includes both domestic and global factors, drawing from the literature on financial crises.⁵ Starting from domestic factors, we evaluate the significance of real economic developments through growth and inflation. A low growth pre-crisis can spark some doubts on the willingness of the monetary authority to raise the policy

⁵See for example [Frankel and Saravelos \(2012\)](#).

rate in response to capital outflows and fuel a self-fulfilling speculative attack (Obstfeld, 1986). Similarly, low growth may undermine fiscal solvency and spread fear of repayment across external creditors. On the other hand, high growth can create problems in the financial sector of the economy through higher risk appetite, credit growth and the formation of asset bubbles (Gourinchas and Obstfeld (2012)). Likewise, periods of high inflation often signal excesses on the monetary and fiscal side, but low inflation may be dangerous as well, especially for small open economies, warning of a surge in inflows and a rapidly appreciating currency. The pre-2008 EWSs literature has also highlighted the importance of external sector variables, in particular real exchange rate, international reserves and current account (Bussière and Fratzscher, 2006). Other variables belonging to this category are the trade balance, as an alternative to the current account, and short-term liabilities, that expose countries to roll-over risk and have been cited as a key factor in the Asian meltdown of 1997-98 (Rodrik et al., 1999). We also test the significance of bilateral trade contagion: crisis countries "infect" their main trading partners through import compression and higher competitiveness of their products, given the real devaluation that often follows a sudden stop.⁶ The interplay between domestic banking problems and capital flows (Kaminsky and Reinhart, 1999) is captured through measures of credit developments. Lastly,

⁶For a detailed review of contagion variables and their transmission mechanism, see Caramazza et al. (2000). In theory, trade contagion can also occur through competition in third markets, but capturing this channel properly is extremely difficult as it would require bilateral trade data dis-aggregated at the product level: two countries may have a common trading partner, but sell two entirely different and unrelated products.

we include both trade and capital openness measures for which we do not have a clear prior on the direction of the impact.

Moving to global factors, after the influential paper by [Rey \(2015\)](#), the VIX has become the standard proxy for the Global Financial Cycle.⁷ Another measure that we use to capture global risk is the TED spread.⁸ Inter alia, [Fratzscher \(2012\)](#) finds an important negative relationship between liquidity risk and flows to EMEs in the pre-2008 period. We also try different rates: the 10-years global and US bonds yield and the 3-months T-bills rate. Historically, there has been a strong negative correlation between gross capital inflows to EME and interest rates in the financial centres ([Reinhart, 2018](#)). High money growth in centre countries can flag rising vulnerabilities in the banking and financial sector and be positively correlated with sudden stops. That said, it can also portray the monetary and debt management stance in AE and thus, be negatively correlated. Other interesting variables are measures of global economic activity. [Broner et al. \(2013\)](#) find that capital flows both in EME and AE are procyclical: they flow out in good times and flow in during bad times. We test both global growth and inflation as proxies.

⁷While many papers confirm the central role this variable plays in outflows from EME ([Forbes and Warnock \(2012\)](#), [Comelli \(2015\)](#), [Eichengreen and Gupta \(2016\)](#)), its importance has been recently challenged ([Cerutti et al. \(2017\)](#) and [Avdjiev et al. \(2017b\)](#)).

⁸The spread rises when either the inter-bank market is fragmented and banks prefer to sit idle on their excess liquidity or when the demand for safe assets increases, driving down their return

3.3.3 Data Transformation

Different variables previously listed are non stationary. To avoid cases of spurious relationship, we need to remove their deterministic and/or stochastic trend. This process is carried out in two distinct manners: by way of a one-sided HP filter or through the calculation of growth rates. When dealing with EWSs, one must also be careful to not include future information in the out-of-sample forecasting exercise: the one-sided filter ensures the fulfillment of this criterion. We compute absolute deviations from the HP trend (gaps) using two different values for λ , the smoothing parameter. In one case we set $\lambda = 1600$ while in the other $\lambda = 400000$, allowing for a more slowly updating trend and more ample fluctuations.⁹ Equivalently, growth rates are also calculated on two frequencies: year-on-year and the four-years horizon. Without imposing any *a priori* constraint on the model, we keep the best performers in the final specification of the EWS. We report all variables with the respective transformations tested in [Table I](#).¹⁰

⁹Some financial variables like banking credit to the private sector might exhibit a lower frequency cycle ([Drehmann et al., 2012](#)). In the same way, real exchange rates deviations might be quite persistent, especially for countries adopting a fixed exchange rate regime ([Beutel et al., 2018](#)).

¹⁰To further verify that there is a true relationship between the underlying variables and the probability of a sudden stop, in other words that our results do not hinge on the de-trending approach chosen, we substitute the variable of interest with its alternative transformation in the robustness part of the paper.

TABLE I: List of variables tested

Variable	Level	Year-on-year growth rate	Four-years growth rate	HP Filter ($\lambda = 1600$)	HP Filter ($\lambda = 400000$)
<i>Global Factors:</i>					
VIX	x				
TED Spread	x				
Global 10-years nominal rate	x				
US 10-years nominal rate	x				
US Federal Funds Rate	x				
Global Real GDP		x	x		
Global CPI		x	x		
Global Liquidity (M2)		x	x		
<i>Domestic Factors:</i>					
Real GDP		x	x	x	x
Current Account over GDP (%)	x				
Private Credit over GDP (%)		x	x	x	x
Real Exchange Rate		x	x	x	x
International Reserves over GDP (%)	x				
International Reserves		x	x		
Short-term Liabilities to BIS Banks over GDP (%)	x				
Reserves (% of ST Liabilities)	x				
CPI		x	x		
Trade Balance over GDP (%)	x				
Trade Openness (% GDP)	x				
Trade Contagion	x				
Capital Controls Measures	x				
Macroprudential Measures (Loan-to-Value) - IMaPP	x				

Note: The table reports the main indicators tested in the construction of the Early Warning. The first column shows the description of the variable; the "Level" column indicates whether the level of the variable has been tested; the "Year-on-year growth rate" and "Four-years growth rate" columns indicate whether, respectively, yearly and four years growth rates of the variable have been tested; the "HP Filter ($\lambda = 1600$)" and "HP Filter ($\lambda = 400000$)" indicate whether percentage deviations from Hodrick–Prescott trends of the variable have been tested: "short" ("long") Hodrick–Prescott trend is computed with the smoothing parameter λ set to 1600 (400000).

Finally, to normalise the scale of the regressors and address the problem of large outliers, we convert all variables in country specific percentiles (Berg et al., 2005): the fundamental assumption behind this normalisation is that it is not the value of the indicators *per se* that matters, but rather their position with respect to their historical distribution.

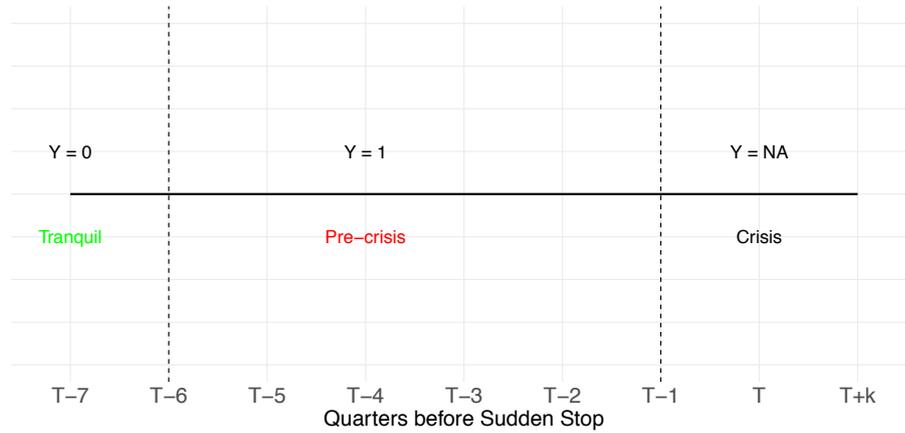
3.4 Methodology

3.4.1 Dependent Variable

We focus on the identification of the six quarters before the actual sudden stop i.e. the pre-crisis period. This kind of flexibility is usually allowed in EWSs given the intrinsic difficulty in predicting with precision the timing of a crisis: while the build-up of domestic vulnerabilities and the worsening of global macroeconomic conditions is observable (the “causa remota”), the “causa proxima” that drives foreign investors away is random and not foreseeable.¹¹ We drop the start of the sudden stop and the quarters of its duration before the estimation. If not removed, these observations would generate a post-crisis bias for variables that during the crisis magnify the movement pre-crisis.¹² Figure 7 shows in a stylized way the definition of our dependent variable.

¹¹We leave out a gap of one quarter between the window and the start of the crisis to mitigate endogeneity concerns and measurement errors on the starting date of the sudden stop: the resulting dependent variable is equal to 1 between 2 and 6 quarters before the actual sudden stop.

¹²See Bussière and Fratzscher (2006).

Figure 7: Dependent Variable Definition

Note: T and K indicate, respectively, the start of the sudden stop and the duration of each individual episode.

3.4.2 Estimation Strategy

We employ a standard Logit model to estimate the model and compute in and out-of-sample probabilities. The non-linear properties of this kind of models is ideal for our classification problem. Most probably, indeed, the effect of the relevant indicators is not linear, but follows an S-shape.

While scholars have recently applied new techniques that exploit Machine Learning (ML) algorithms to solve these classification problems, clear-cut evidence on the best performing method

does not exist.¹³ Interpretability is also a concern. Logit models provide a clear ranking of the predictors, enhancing the understanding of crises by policymakers. This is particularly important in the context of our study since we include both domestic and global variables and has deep implications. If local factors are found to matter, the usefulness of predicting a crisis rises: policymakers receiving the signal, indeed, have the possibility to target the fundamentals responsible for the rise in probability and lower the overall risk, possibly avoiding the materialization of the event. On the other hand, if capital outflows are only determined by policy decisions in the centre countries, even if a crisis is predicted well in advance, the scope for reaction is heavily limited. Compared to parametric methods, causal inference in ML is not clear.¹⁴

For the estimation, we pool observations across time and countries.¹⁵ The final specification is:

$$P(Y_{i,t} = 1) = F(X_{i,t}\beta) = \frac{e^{X_{i,t}\beta}}{1 + e^{X_{i,t}\beta}} \quad (3.1)$$

¹³Beutel et al. (2018) and Comelli (2014), for banking and currency crises respectively, find that standard Logit models outperform different ML algorithms in the out-of-sample forecasting, while Holopainen and Sarlin (2016) find exactly the opposite for banking crises.

¹⁴Although some steps have been recently taken in this direction, see for example Suss and Treitel (2019).

¹⁵In the robustness part, we also control for the sensitivity of our results to the introduction of country fixed effects.

3.4.3 Evaluation

To classify predicted probabilities into binary signals, we need to impose a threshold: if the predicted probability crosses this value, a signal is sent, otherwise not. The signals are then compared to the actual value and the performance of the model is evaluated. This choice implies a trade-off between Type 1 error i.e. missing a crisis and Type 2 error i.e. issuing a fake alarm. The lower the threshold, the more fake alarms are issued and the other way around. The four possible outcomes of a classification problem are shown in [Table II](#).

TABLE II: Example of Confusion Matrix

		Predicted	
		$Signal_{i,t} = 0$	$Signal_{i,t} = 1$
Actual	$Y_{i,t} = 0$	True Negative (TN)	Fake Positive (FP)
	$Y_{i,t} = 1$	Fake Negative (FN)	True Positive (TP)

It follows that for any fixed threshold τ the loss function of a policy-maker can be written as:

$$L(\theta, \tau) = \theta \frac{FN(\tau)}{FN(\tau) + TP} + (1 - \theta) \frac{FP(\tau)}{FP(\tau) + TN} \quad (3.2)$$

where θ indicates the preference for Type 2 errors as compared to Type 1 errors. A θ higher than 0.5 indicates that missed crises weigh more than fake alarms on the policy-maker loss function. It has become standard to set the optimal threshold τ^* to maximize the relative usefulness of a model:

$$U_r(\theta, \tau) = 1 - \frac{L(\theta, \tau)}{\min(\theta, 1 - \theta)} \quad (3.3)$$

This function compares the usefulness of an EWS with a naive rule. The rationale is that policy-makers can always realize a loss of $\min(\theta, 1 - \theta)$ disregarding any model by always or never signalling an alarm. If θ is smaller than 0.5, policy-makers give more weight to Type 2 errors: the benchmark is obtained by ignoring the EWS, which amounts to never having any signals issued so that $TP = FP = 0$. The resulting loss according to equation [Equation 3.2](#) is θ . If θ exceeds 0.5, they give more weight to Type 1 errors. The benchmark is to assume there is always a sudden stop: in this case a signal is always issued so that $FN = TN = 0$. The resulting loss is $1 - \theta$. When $\theta = 0.5$, independently from the naive rule chosen, the loss is the same and equal to 0.5. From [Equation 3.3](#), an EWS is the more useful, the lower the loss it generates with respect to a completely uninformed decision. In this context, not only this

function is appropriate to find an optimal threshold, but also furnishes a natural and simple way to evaluate and compare the overall performance of different models.¹⁶

3.4.4 Forecasting Procedure

We must spend some words on the out-of-sample forecasting procedure. Our analysis is conducted in a quasi real-time manner and the evaluation period goes from 2006Q1 to 2017Q1, a time span that corresponds to half of our original sample. Hereafter we list all the steps of the exercise:

- (i) At each quarter t of the evaluation period, we divide between a training sample that goes from 1995Q4 (the beginning of the original sample) to quarter $t - 1$ and a test sample composed exclusively by quarter t . The indicators are transformed into country-specific percentiles for the training sample.
- (ii) We estimate the model on the training sample and save the optimal threshold i.e. the one that maximizes the in-sample relative usefulness function.

¹⁶The parameter θ is unobservable and must be set exogenously. For our benchmark forecasts, we choose a standard value of $\theta = 0.5$ that indicates a policymaker weighing equally Type 1 and Type 2 errors. In reality, Type 2 errors may be less worrisome from a policy-maker's perspective for two reasons. First, Type 2 errors tend to be less costly from a welfare perspective than Type 1 errors. The cost of Type 2 errors may be the cost of taking pre-emptive policy measures. By contrast, missing a crisis has often a higher welfare cost, for example large output contractions and rising unemployment. Second, Type 2 errors may not always be due to the predictive failure of the model, but simply reflect the fact that although fundamentals were indeed vulnerable, appropriate policy initiatives were taken to improve the resilience of the economy and prevent a crisis.

- (iii) Re-calculating the percentiles, we compute the pre-crisis probability for quarter t . We store it together with the respective optimal threshold and recursively repeat these three steps for every quarter t until 2017Q1.¹⁷
- (iv) *Ex post*, we compare the collected probabilities with the respective threshold, count the number of missed signals and fake alarms and evaluate the model.

The whole procedure is designed so as to mimic as closely as possible the information available to policy-makers in each quarter and at the same time, we are careful to not introduce future information in the forecasts produced and bias the results in favor of our model.¹⁸¹⁹

¹⁷This passage is needed to have the position of the new observation with respect to the historical distribution.

¹⁸A pitfall of this exercise is the forward-looking nature of the dependent variable $Y_{i,t}$. This tricky point can be better explained with an example. Imagine we are in 2005Q4 and want to estimate the pre-crisis probability for the first quarter of the out-of-sample exercise, 2006Q1. The training model will be estimated with data from 1995Q4 to 2005Q4. If between 2006Q1 and 2007Q2 a sudden stop occurs, the dependent variable in the training sample will identify some pre-crisis observations with value 1, hence incorporating future knowledge in the model. To correct, at each recursive update of the training sample, we set the last 6 quarters observations equal to 0 before estimating the model. This correction is consistent with the noise in the information set of the policy-maker: they do not know whether the build-up observed in other countries will materialize in a sudden stop.

¹⁹Step (ii) and (iii) entail that the out-of-sample optimal threshold is time-varying and step (iii) that the estimation sample is an expanding-window.

3.5 Results

3.5.1 Determinants of Sudden Stops

[Table III](#) shows the result for our preferred specification considering the whole sample period (1995-2017): we include in our benchmark model only indicators that are significant at the 5% level and have the expected sign. For variables that are highly collinear e.g. alternative transformations or overlapping definitions, we include the one that maximizes the goodness of fit as measured by the relative usefulness function. Robustness checks and alternative variables are reported in [Table XXVIII](#).

TABLE III: Full Sample Logit

	<i>Dependent variable:</i>
	Pre-crisis
TED Spread	1.504*** (0.236)
Global Liquidity Growth	-1.258*** (0.264)
Credit-Gap	1.059*** (0.245)
RER-Gap	-1.912*** (0.251)
ST Liabilities to BIS Banks/GDP	0.725*** (0.237)
CA/GDP	-1.108*** (0.240)
Trade Contagion	0.700*** (0.170)
Controls on Capital Inflow	-0.432** (0.198)
Constant	-1.281*** (0.291)
Observations	1,753
Log Likelihood	-716.762
Akaike Inf. Crit.	1,451.524

Note: The sample consists of 31 Emerging Markets over the period 1995Q1-2017Q1. Robust standard errors in parentheses. Global liquidity growth is calculated on the four-years horizon. RER-Gap and Credit-Gap are deviations from Hodrick-Prescott trends with $\lambda = 1600$. Controls on capital inflows are from [Fernández et al. \(2015\)](#). * Statistical significance at 10% level. ** Statistical significance at 5% level. *** Statistical significance at 1% level.

Table IV reports the relative goodness-of-fit statistics. The model calls correctly 62.6 % of the pre-crisis quarters with only 20% of fake alarms. The conditional probability of a crisis given a signal from the model is 40%, double the unconditional probability of experiencing a crisis (23%). Similarly, compared to a completely uninformed decision (see section 3.4.3), the model generates a relative usefulness for a policymaker of roughly 40%.

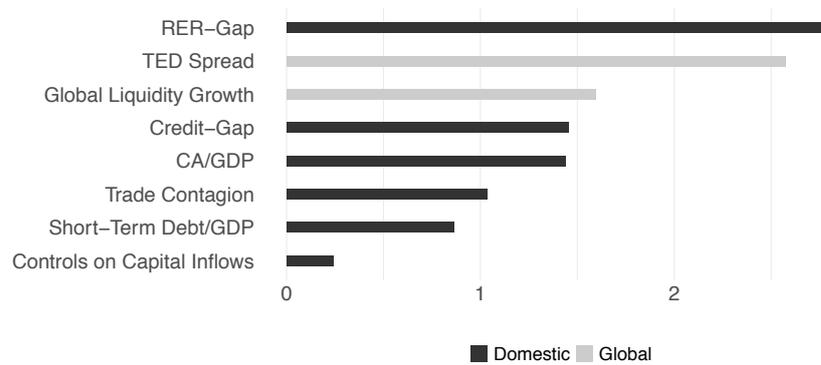
TABLE IV: In-sample Performance

			True Positives rate: 62.6%
			True Negatives rate: 78.1%
	$S_{i,t} = 0$	$S_{i,t} = 1$	Prob. pre-crisis given a signal: 40.2%
$Y_{i,t} = 0$	1108	311	Prob. pre-crisis given no signal: 10.1%
$Y_{i,t} = 1$	125	209	Relative Usefulness: 40.7%

Note: The table reports the results for the in-sample performance of the benchmark logit. The forecast horizon is 1-6 quarters ahead and the preference parameter θ is equal to 0.5. The evaluation is carried out through the above measures: True Positives rate = $TP/(TP+FN)$, True Negatives rate = $TN/(TN+FP)$, Prob. sudden stop given a signal = $TP/(TP+FP)$, Prob. sudden stop given no signal = $FN/(FN+TN)$ and Relative Usefulness U_r (see Equation 3.3). Threshold optimized in-sample to maximize the relative usefulness and equal to 23.4%.

Are crises the byproduct of domestic shortcomings or are emerging markets solely at the mercy of policy decisions and economic conditions in the global financial centres? Since our model is non-linear, we cannot directly interpret its coefficients. To calculate marginal effects and understand the relative importance of each indicator, we have, instead, to set precise values for all variables. [Figure 8](#) shows marginal effects under a specific scenario i.e. moving the variable of interest from its tranquil time to its pre-crisis average, while other indicators are kept equal to their tranquil time average.

Figure 8: Marginal Effects Covariates



Note: Marginal effects calculated increasing the value of each variable individually from its tranquil time average to its pre-crisis average and keeping other covariates at their tranquil time average. Coefficients retrieved from the benchmark specification. X-axis in percentage points.

Exchange rate overvaluation and rise in the TED spread have the strongest impact, increasing the probability of a sudden stop by more than 2.5 percentage points (pp). Global liquidity tightening, credit booms and current account deficits compose a second group and have a smaller impact (1.5 pp). At the end of the spectrum follow trade contagion and short-term debt (1 pp) and last, controls on capital inflows (less than 0.5 pp). Even though these effects may appear relatively small, one must bear in mind different points: first, the pre-crisis period is long, starting a year and a half prior to the crisis, thus influencing the value of the pre-crisis average. Second, the reference point matters: while to isolate the effects we have kept all other indicators to their tranquil average, in practice a rise in the TED spread might have a much larger effect, for example, when also the current account deficit is large.

To control for the presence of these pairing effects, we introduce an additional specification including the significant interactions between domestic and global factors: a prominent view is that, indeed, during periods of global turmoil, investors first pull out funds from countries with weak fundamentals. Our results endorse this view as the predictive power of the model increases after the introduction of significant interaction terms by roughly 5 pp compared to the benchmark specification. (Table V). In terms of causality, although the interpretation of the estimated coefficients is not straightforward due to non-linearities, the combination of tight global lending conditions, high level of short-term maturities and domestic credit booms increases substantially the probability of a sudden stop.

TABLE V: Full Sample Logit - Benchmark Model With Interactions

	<i>Dependent variable:</i>
	Pre-crisis
TED Spread & Short Term Debt/GDP	3.461*** (0.813)
TED Spread & Private Credit-Gap	3.975*** (0.853)
Observations	1,753
True Positive rate	70.1%
True Negative rate	74.6%
Relative Usefulness	45%

Note: The table reports the same benchmark specification of [Table III](#) with the addition of the interaction between global and domestic factors significant at the 5% level (only variables displayed). * Statistical significance at 10% level. ** Statistical significance at 5% level. *** Statistical significance at 1% level. For the evaluation of the in-sample performance, the forecast horizon is 1-6 quarters ahead and the preference parameter θ is equal to 0.5. The evaluation is carried out through the above measures: True Positive rate = True Positives/(True Positives + Fake Negatives), True Negative rate = True Negatives/(True Negatives + Fake Positives), and Relative Usefulness U_r (see [Equation 3.3](#)). Threshold optimized in-sample to maximize the relative usefulness and equal to 20.4%.

All in all, we do not find strong evidence of predominance by neither of the two group of factors: it is rather the combination of the two that is extremely pernicious for EMs. This is

suggestive that policy- makers have at least some leeway to act targeting weak fundamentals when confronted with a newly issued signal.

3.5.2 Out-of-Sample Performance and Forecast Horizon

The paper by [Berg and Patillo \(1999\)](#) pointed to a large divergence between the in and out-of-sample performance for EWSs. Since then it has become standard to evaluate the predictive power of these models on the base of their out-of-sample performance. The framework developed in section [3.4.4](#) allows us to do so in a quasi-real time manner i.e. having the same information set of the policymaker at the time of the prediction and without introducing future knowledge in the model.²⁰ The results of the out-of-sample estimation are reported in [Table VI](#).

²⁰Choosing an appropriate de-trending approach and correcting for the forward-looking nature of the dependent variable.

TABLE VI: Out-of-sample Performance

	$S_{i,t} = 0$	$S_{i,t} = 1$	
			True Positives rate: 47.4%
			True Negatives rate: 84.7%
			Prob. sudden stop given a signal: 50%
$Y_{i,t} = 0$	648	117	Prob. sudden stop given no signal: 16%
$Y_{i,t} = 1$	129	116	Relative Usefulness: 32%

Note: The table reports the results for the quasi real-time out-of-sample performance of the benchmark logit. The forecast horizon is 1-6 quarters ahead and the preference parameter θ is equal to 0.5. The evaluation is carried out through the above measures: True Positives rate = $TP/(TP+FN)$, True Negatives rate = $TN/(TN+FP)$, Prob. sudden stop given a signal = $TP/(TP+FP)$, Prob. sudden stop given no signal = $FN/(FN+TN)$ and Relative Usefulness U_r (see [Equation 3.3](#)).

The EWS predicts almost 50% of the pre-crisis episodes while sending relatively few alarms, about 15% of the total tranquil quarters. This means every time a signal is sent, there is a 50% probability of a correct call, a percentage higher than its in-sample counterpart. Thus, the uncertainty involved by our model in the out-of-sample predictions is limited. Overall, the EWS would result in a 32% gain compared to a completely uninformed decision for a policy-maker with balanced preferences between Type 1 and Type 2 errors.

The result is robust to the choice of a different time horizon as shown in [Table VII](#). The relative usefulness statistics always remains around 30% and reaches its maximum for our 6

quarters benchmark, validating *ex post* our choice. Moving to the 8 quarters specification, the number of correctly identified pre-crisis periods increases (TP Rate), but also does the amount of fake alarms (1-TN Rate). This is suggestive that for policymakers with a greater value for θ the choice of a longer horizon span would be preferable.

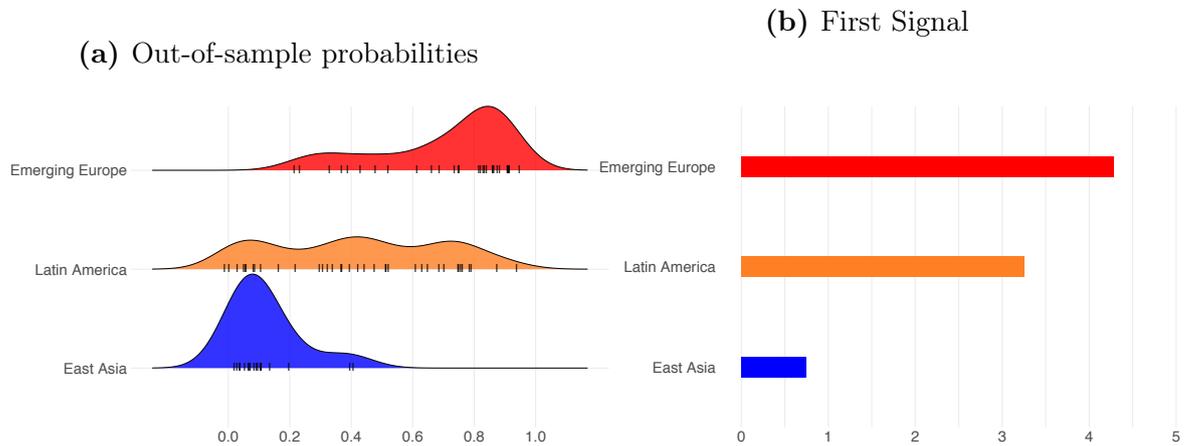
TABLE VII: Different forecast horizons out-of-sample performance

Forecast Horizon	TP	TN	FP	FN	TP Rate	TN Rate	P(Pre-crisis Signal)	P(Pre-crisis No Signal)	$mU_r(\theta)$
4 Quarters	73	719	137	81	47.4%	84%	34.8%	10.1%	31.4%
6 Quarters	116	648	117	129	47.4%	84.7%	50%	16%	32%
8 Quarters	163	538	142	163	49.4%	79.1%	53%	23.7%	28%

Note: The table reports the results for the quasi real-time out-of-sample performance of the benchmark logit changing the forecast horizon. The preference parameter θ is equal to 0.5. The evaluation is carried out through the above measures: True Positives rate = $TP/(TP+FN)$, True Negatives rate = $TN/(TN+FP)$, Prob. sudden stop given a signal = $TP/(TP+FP)$, Prob. sudden stop given no signal = $FN/(FN+TN)$ and Relative Usefulness U_r (see [Equation 3.3](#)).

3.5.3 Timing

How did our model work with particular reference to the 2008 crisis? With how much certainty were the signal sent? And were they timely? In panel [9a](#) we show the distribution of the out-of-sample predicted probabilities for the GFC episodes during the individual pre-crisis periods, conditioning on the EM regional group.

Figure 9: GFC Sudden Stops

Note: Panel (a) shows the distribution of the predicted probabilities for the out-of-sample recursive exercise in the pre-crisis period of each country-specific Global Financial Crisis sudden stop. We consider Global Financial Crisis sudden stops those episodes occurring in the time window 2006Q4 - 2008Q4. Panel (b) shows for the aforementioned episodes the advance, on average, of the model in issuing a signal. X-axis indicates quarters before the sudden stop.

While for Emerging Europe the bulk of the distribution i.e. the probability of being in a pre-crisis quarter, is around 80%, for East Asia most of the individual probabilities reach only 20%. For Latin American the distribution is, instead, more uniform. For the first group of countries, the first signal was sent, on average, more than one year before the sudden stop (Figure 9b). For the second, instead, episodes were not signalled at all or just with a small advance. Latin America, as before, lies in between.

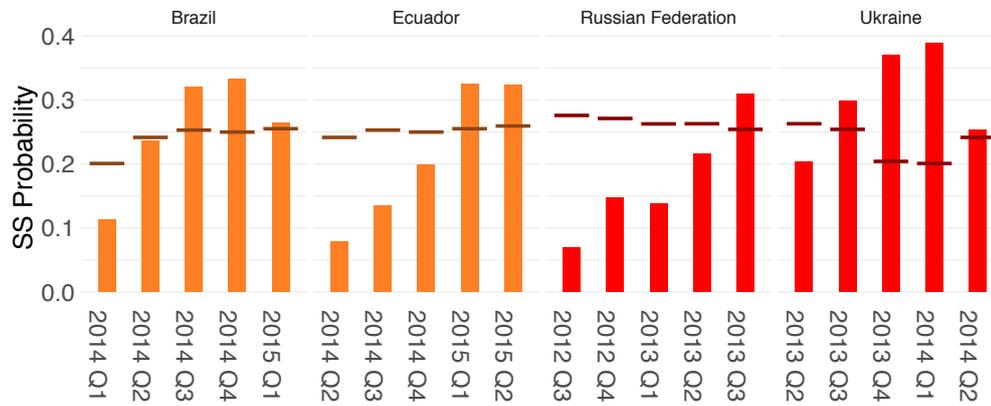
Domestic factors largely explain this heterogeneity of results between regional groups. East Asian countries enforced counter cyclical fiscal and monetary policy in the years that followed the regional crisis of 1997-98, approaching 2008 with large current account surpluses, a competitive and flexible exchange rate and a solid financial sector.²¹²² EECA countries, instead, neared the GFC with extremely flawed fundamentals. Pre-2008 capital inflows financed exceptionally large current account deficits with a considerable part of this foreign capital that was channeled into short-term maturities: real exchange rates appreciated steeply and there was a lending boom operated by the banking sector (Gourinchas and Obstfeld, 2012): this surge was allowed by a simultaneous liberalization of capital markets.

We then move to the post-GFC period and ask the same question. While sudden stops in the GFC are more or less synchronized, afterwards they are more distributed across the test period: instead of a regional aggregation, we proceed on a case-by-case basis. In particular, we study the behavior of fitted probabilities before four interesting crises: Russia 2014 Q1, Ukraine 2014 Q3, Brazil 2015 Q3 and Ecuador in 2015 Q4 (Figure 10).²³

²¹See for example Park et al. (2013).

²²South Korea and Indonesia are the countries in the group with the highest predicted probabilities. The first had sound macroeconomic fundamentals, but a large level of short-term liabilities in the banking sector. The second was the last country in the region to experience a reversal of gross inflows and as such probabilities are greatly influenced by shattered global factors and regional contagion.

²³We focus on these occurrences because they are associated with at least a quarter of recession throughout the sudden stop duration, while other episodes are not.

Figure 10: Post GFC Sudden Stops

Note: Out-of-sample fitted probabilities in the pre-crisis period for four different sudden stops episodes: Brazil 2015Q3, Ecuador 2015 Q4, Russian Federation 2014 Q1 and Ukraine 2014 Q4. The thick line corresponds to the time-varying optimal threshold.

Our EWS sends a signal half a year before for Russia, the minimum allowed. For the other crises, the advance widens: Ukraine is called more than one year in advance, Brazil one year and Ecuador three quarters. Even though probabilities are much lower compared to the GFC, owing mostly to improved global conditions, is extremely encouraging that all the four episodes would have been signalled with advance.²⁴ Moreover, this result is achieved without considering important factors that analyst have linked to these crises: worsening political landscape (Ukraine

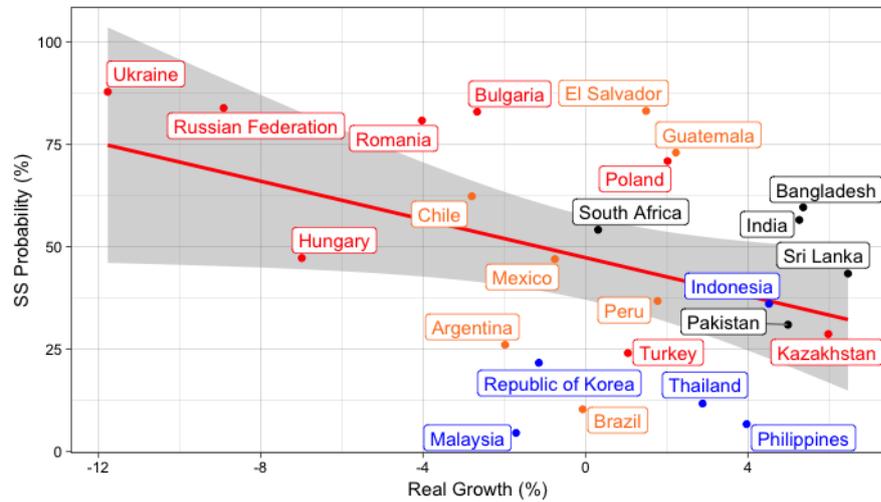
²⁴Another interesting feature is that fitted probabilities rise next to monotonically in the pre-crisis period.

and Brazil) and the occurrence of natural disasters (Ecuador). This means that even if the last two may have contributed to the drop in inflows, multiple factors exerted pressure on these countries.

3.5.4 Sudden Stop Impact and Fitted Probabilities

Hitherto we based the evaluation of our EWS solely on missed crises (Type 1 errors) and fake alarms (Type 2 errors), in line with the literature. The underlying assumption of such an evaluation framework is that external crises are alike episodes and produce the same effects on the economies hit. Nevertheless, it is a well-known fact that some crises are more painful than others. This owes to different factors: the magnitude of the external shock, the conditions of the domestic economy at the time and of course, the policy response that follows as well as the behaviour of domestic investors. These elements are not mutually exclusive, but rather complementary.

Therefore, when evaluating an EWS, policymakers should also be concerned about which specific events are predicted and which not. In doing so, the novelty of this paper is that we link two parts of the literature that have so far been kept separated: one is the classic EWS literature that tries to predict in advance the occurrence of a crisis, while the second, instead, tries to predict its incidence. [Figure 11](#) shows the relationship between the estimated pre-crisis median probability and the median growth experienced during the episode for every country-specific sudden stop in the GFC.

Figure 11: Fitted Probabilities and Ex-Post Growth

Note: The figure shows the relationship between the median out-of-sample probability in the pre-crisis period for GFC related sudden stops and the median growth during the associated sudden stop. Red line is the regression line with 95% confidence intervals.

At a first glance, EECA countries monopolize the north-western quadrant. This means these are the countries that suffered more during the GFC sudden stops and contemporaneously those that exhibit the highest median probabilities. Further, from Figure 9b, they are also the sudden stops for which a signal was sent with large advance by our model. On the other hand, East Asian countries occupy mostly the south eastern quadrant i.e. countries that suffered less and with the lowest median probabilities. Latin American countries are, once again, highly heterogeneous. We consider two countries during the GFC, one from the EECA and the other

from the East-Asian group e.g. Romania and Thailand. The first suffered from a full-fledged recession with a median quarterly contraction of 4% of GDP, while the second continued to grow at a moderate pace, around 2.5%. For the first a signal would have been sent early and with extreme certainty, while the second belongs to the group of missed crises.²⁵

We further proceed investigating formally the issue: we pool all the out-of-sample observations together and estimate the cross-sectional relationship. [Table VIII](#) shows the regression results. We find that an increase in median pre-crisis probability by a percentage point decreases significantly growth during the sudden stop by 0.07%. Sudden stops predicted with more certainty by our EWS are also the most destructive ones in terms of output losses, while those that are not identified (Type 1 error) or identified with relatively low probabilities are those with mild consequences for the real economy. Against this new evidence, reporting standard evaluation metrics without investigating for which sudden stops a signal would have been sent and for which not, would highly underestimate the true value of an Early Warning for policymakers.

²⁵While median growth rate during the sudden stop is a simple proxy of output impact, this measure, however, ignores that countries might have different trend growth rates before the event: therefore, we test robustness of the relationship using a different metric. We construct this new measure as the median growth rate during the sudden stop minus the median growth rate in the preceding tranquil period: the resulting scatter plot displays a similar pattern (See [Figure 38](#) in the appendix).

TABLE VIII: Predicted Probability and Sudden Stop Incidence

<i>Dependent variable: Median Growth</i>				
	Coeff.	Std. error	t-statistic	P> t
Pre-crisis probability	-0.070	0.027	-2.54	0.02
Observations	41			

Note: The sample comprises the 41 sudden stops episodes occurred after 2006 Q1, the beginning of the out-of-sample period. Dependent variable is the median growth calculated over the duration of each episode. Independent variable is the median out-of-sample probability estimated over the whole pre-crisis period. Intercept omitted and robust standard error reported.

3.6 Conclusion

This paper contributes to the financial crisis literature investigating the predictability of sudden stops in emerging markets. We extend the existing literature in different ways. First, we test a large variety of domestic and global indicators and evaluate their relative importance in the materialization of sudden stops: if emerging markets are only at the mercy of policy decisions in the global financial centres, the scope for intervention by domestic policymakers results highly limited even when a signal is issued. Second, we propose a framework to evaluate the performance of the model in quasi-real time, taking heed not to include future information in the recursive exercise. Third, we study the relationship between the probabilities estimated by our model and the output loss associated with the ensuing sudden stop.

We find a near equivalence between the marginal impact of domestic and global factors on the probability of a crisis: this result highlights the role EWSs can play not only as a surveillance, but also as a stability tool available to policymakers. We then proceeded evaluating the out-of-sample performance of the model. The recursive exercise yields encouraging results with a parsimonious specification: the pre-crisis periods correctly called are close to 50% of the total with fake alarms corresponding to less than 15%. This means that the uncertainty involved with the predictions is low: compared to the unconditional probability of being in a pre-crisis quarter (roughly 20%), the conditional probability given a signal rises to 50%. Finally, we brought forward a new argument in “defense” of EWSs. We show that there is a negative, statistically and economically significant relationship between the median probability predicted for the whole pre-crisis period and median economic growth experienced during the associated sudden stop:

in other words, the model works well in predicting catastrophic events and less so for rather innocuous ones.

All in all, even if the prediction of rare events like sudden stops remains a humbling task, our model would have sent reliable, timely and relevant signals. This is especially promising in view of the lengthy out-of-sample horizon chosen and the different non-economic factors that exerted pressure on emerging markets in the recent decade.

CHAPTER 4

WHEN (WHERE AND WHY) FORECASTERS GET IT WRONG?

30 YEARS OF GROWTH FORECASTS

Abstract: We construct a database that contains short-term real growth projections from all major institutions and private sector and examine the quality of these forecasts for a large sample of countries over the last three decades. Our findings are three-fold. First, the inability to forecast recessions and in particular financial crises determine the optimism usually ascribed by economic agents to growth forecasts. Second, different forecasts closely resemble each other, exhibiting a high degree of collinearity. Third, turning from surveillance to program forecasts, IMF-supported program forecasts tend to be more optimistic for large programs. Nevertheless, a comparison with the private sector thwarts the hypothesis of a deliberate bias. Our findings suggest that some caution is in order when consulting growth forecasts and all the products that build on them (fiscal and debt forecasts in primis) and have first-order implications for policymakers during times of global and regional turmoil.

Contributions: Prakash Loungani shared his decennial knowledge on the field of forecasts evaluation: he directed toward the relevant literature, added constant institutional insight and helped with result formulation and framing. Zidong An worked closely on data gathering and wrangling.

Code: Full replication code is on-line and can be accessed from a public Github repository:

https://github.com/umbertocollodel/When_where_and_why.

4.1 Introduction

We have two classes of forecasters: Those who don't know — and those who don't know they don't know.

John Kenneth Galbraith

Growth forecasts are the main product of leading institutional publications such as the IMF *World Economic Outlook* (WEO) and the World Bank *Global Economic Prospects* (GEP). Different economic actors rely heavily on these forecasts to take better informed decisions: in particular, forecasts of economic activity play a crucial role in shaping governments' policies and guiding investors' allocation of capital. Moreover, growth forecasts constitute the building stone for other important products that guide policy advice to countries, such as fiscal and debt forecasts. Their quality has, however, been repeatedly called into question by the academic world, specialized news agencies and individual governments alike (e.g. Sandefur and Subramanian (2020)). Often official forecasts are perceived to be overly optimistic and not matching their relative narratives: this optimism can, in turn, lead to slack economic policies or conversely, inadequate support, depending on the current state of the economy. Moreover, if the institution is directly financially involved with a country, this optimism raises questions relative to its vested interest in producing biased forecasts.

In this chapter, we take stock on all the main questions related to the quality of short-term growth forecasts: i) Are forecasts really too optimistic? How much and what does this imply in

practice? Is this optimism shared across different forecasters? What are its underlying causes?
 ii) How do the forecasts produced by different forecasters relate to each other? Is there evidence of regional specialization? iii) Are forecasts published in the context of financial assistance any different? Specifically, is there a positive relationship between inflated forecasts and greater assistance?

These questions are not new to the forecast evaluation literature (IEO (2014); Loungani (2001); Timmermann (2007)), but authors have mostly focused on an early sample (pre-2010) and solely on IMF WEO forecasts. Building on this literature, we assemble a large dataset of growth forecasts that incorporates all main institutional forecasters as well as the private sector. This allows us to draw general conclusions about the performance of the economic forecasting profession in the last decades. In addition, extending the time dimension to incorporate the post-GFC period is particularly relevant for the third question we ask since the amount extended through Fund programs soared in the last decennium.

We find that i) optimism differs in extension and magnitude according to the geographical group and forecaster considered. Nevertheless, we confirm it as a hallmark of growth forecasts. We ascribe this optimism to a general inability of the forecasting world to predict recessions: in this way, we complement and extend the evidence presented in An et al. (2018). Furthermore, we show that the volatility in forecast errors during recessions is critically linked to the occurrence of financial crises: therefore, we demonstrate that the problem of forecasting economic activity and predicting financial crises are closely related ones ii) Although we find marginal evidence of regional specialization in forecasts, these differences are economically small. Instead, we

document an almost perfect collinearity among different forecasters iii) IMF program forecasts are more optimistic for large programs. However, we show that Consensus forecasts exhibit, for the same set of programs, a statistically equivalent relationship. Therefore, we deem the claim that forecasts are inflated deliberately ill-founded.

The rest of the chapter is organized as follows. Section 4.2 illustrates the different data sources employed in the analysis and how we combine them. Section 4.3 focuses on surveillance forecasts, examining optimism and its determinants and comparing the performance of different forecasters. Section 4.4 investigates the relationship between high level of assistance and program forecasts. Lastly, Section 4.5 concludes.

4.2 Data

In this paper, we examine annual real GDP growth forecasts at the short-term horizon i.e. current year and year-ahead. To this end, we assemble a large dataset incorporating the forecasts of the main public and private sector economic institutions. Since we draw from and compare a variety of different sources, avoiding any confusion relative to the underlying raw data is essential. In this section, we illustrate the different databases we combine in our analysis, clarify the terminology we employ in the rest of the paper and discuss the kind of issues that can arise when comparing forecasts from different sources as well as the solutions we propose.

The forecasts published by the IMF in its most known publication, the *World Economic Outlook* (WEO), are the starting point and constitute the foundation for our study. The main WEO publication is released twice a year, in April and October; hence, we refer to them, as, respectively, Spring and Fall publication. Forecasts produced range from current-year and one-year ahead (short-term forecasts) to three, four and five-years ahead (medium-term forecasts). We retrieve short-term WEO forecasts covering the time span 1990-2019 for a nearly balanced sample of 192 countries representing the whole Fund membership.

An evaluation of forecasts requires the identification of an actual value of the variable being forecast. Nevertheless, data revisions can lead to differences between the value ascribed to a variable when it is first published and the value ascribed, for the same time period, some years later, after data revisions have been made.¹ Following [Timmermann \(2007\)](#), we use the earliest

¹[Genberg and Martinez \(2014\)](#) shows how this difference can be substantial in practice.

available actual i.e. the actual real GDP growth in the Fall WEO issue released after the year being forecasted.² This choice ensures a uniformity of treatment between actuals: the latest available set of data is not homogeneous in vintage, early data are revised many times, while latest data are, perhaps, still preliminary or partial estimates. Moreover, in this way we can conduct a consistent comparison with the results found previously on the earlier samples.

In the second part of our work, we compare Fund forecasts with other institutional forecasters, namely the World Bank, for emerging and developing economies, and the European Commission, for the European Union and emerging Europe, and private forecaster reported in the Consensus survey. In this respect, differences in release dates among forecasters can influence the determination of relative forecast performance, especially when a later forecast can incorporate an earlier forecast's information. Thus, to compare the accuracy of different forecasters, it is crucial to minimize this timing difference. We retrieve European Commission forecasts from the AMECO database: forecasts are available starting 2010.³ The Fall and Spring forecasts are released only one month later compared to the WEO: respectively, in the first week of May and November. The World Bank, instead, releases its forecasts in the *Global Economic Prospects* (GEP) publication in January and July starting from 2010.⁴ In this case, there is a big timing

²To clarify, the actual GDP growth in 2001 will be the real GDP growth reported for 2001 in the 2002 Fall issue.

³https://ec.europa.eu/info/business-economy-euro/indicators-statistics/economic-databases/macro-economic-database-ameco/ameco-archive_en.

⁴These forecasts are not publicly available in a structured format.

difference with respect to the WEO Fall and Spring publication. To obviate the issue, we resort to the WEO updates that are released in January and June. There is less of a publication timing issue when comparing WEO forecasts with private forecasts such as those issued by Consensus Economics. This is largely because private forecasters produce their forecasts monthly and thus the publication date can be selected so as to minimize the timing differences (IEO, 2014). For each country, the number of forecasters reported in the Consensus survey varies between 20 and 30. These may use different models and assumptions to form point estimates forecasts: the bulk of our analysis focuses on the arithmetic mean of these forecasts. We obtain Consensus forecasts covering the whole period 1990-2019.⁵⁶ In addition, we provide also more granular results based on individual forecasters for a subgroup of 45 countries starting 2007.⁷ Table IX reports for each forecaster the release month of forecasts and our availability in terms of years.

⁵Figure 39 shows the evolution of country availability for WEO and Consensus data. We do not report the same for World Bank and European Commission since they are almost constant throughout the sample period.

⁶We choose the April and September vintage of Consensus as in Timmermann (2007).

⁷The subgroup of countries is dictated by data availability: Consensus survey does not report data on individual forecasters for all countries.

TABLE IX: Forecasts Availability

Forecaster	Release month	Availability
International Monetary Fund	April & October	1990-2019
	Update: January & June	Update: 2010-2019
World Bank	January & July	2010-2019
European Commission	May & November	2010-2019
Consensus - Mean	Every month	1990-2019
Individual	Every Month	2007-2019

Finally, in the context of program forecasts, we use data from the Monitoring of Fund Arrangements (MONA) database: this source contains macroeconomic forecasts produced at the inception of each program and their revision at each program review. The data cover all Fund programs in the period 2002-2018.⁸ In the rest of the paper we refer to surveillance forecasts for all forecasts with the exception of MONA forecasts, that we refer to, instead, as programs forecasts.

⁸The full database is divided into two periods: 1993–2003 and 2002 to present. The reason behind this distinction is the reclassification and restructuring of several economic variables that occurred in the early 2000s (Luna, 2014). For this reason, we concentrate only on the second part of the database.

This rich variety of data ensures a well-rounded view on the forecast performance of all major institutional and private economic actors allowing us to draw general conclusions about the economic forecasting profession in the last 30 years.

4.3 Surveillance Forecasts

4.3.1 Surveillance Forecasts and Optimism

Bias is the presence of an error “that does not balance itself out on average” (Genberg and Martinez, 2014). The most common test of forecasts biasedness is a test of significance on the average forecast error for individual countries. We define the forecast error as the difference between the actual value of variable y at time $t + h$ and the forecast formulated at time t for period $t + h$:

$$FE_{t,t+h} = y_{t+h} - \hat{y}_{t,t+h} \quad (4.1)$$

We then regress the forecast error on an intercept α that corresponds to the average value of the forecast error

$$FE_{t,t+h} = \alpha + \epsilon_{t+h} \quad (4.2)$$

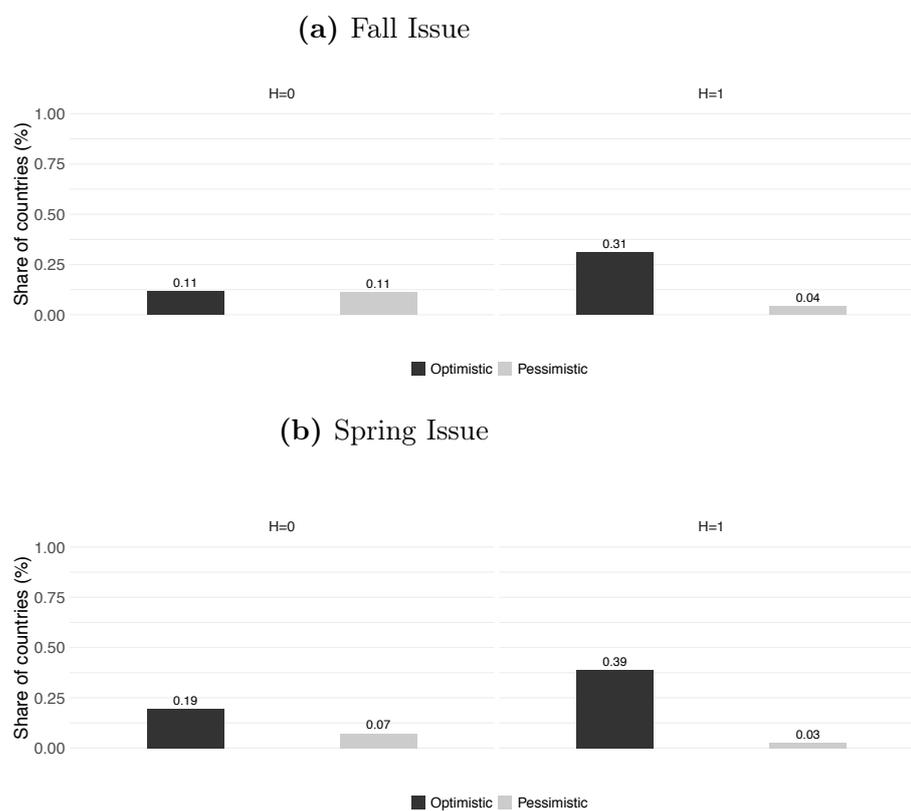
A negative significant value indicates an optimistic bias (over prediction) while, conversely, a positive one, indicates a pessimistic bias (under prediction).

Comprehensive assessments of growth biases in WEO forecasts frequently find evidence of optimistic biases (Artis, 1996; Timmermann, 2007) with, however, many nuances indicating that

biases often differ in sign between type of countries. Other individual studies present similar results: for example, [Beach \(1999\)](#) concludes that IMF growth forecasts are pessimistic for developed countries, but optimistic for developing countries. By the same token, [Takagi and Kucur \(2006\)](#) find evidence of widespread optimism for African and Latin America countries, pessimism for industrial countries and the Middle East, and lack of a systematic bias in emerging Asia. It is worth noticing that, however, these studies are relatively dated and focus on a early sample. In addition to a required update, the use of a longer time span mitigates the concerns of statistical problems associated with short samples.⁹

[Figure 12](#) shows the share of countries for which we detect the presence of a bias. For current-year Fall forecasts we find a that optimistic and pessimistic biases are equally distributed and they each correspond to 11 % of the countries. Nevertheless, as the forecast horizon increases, the former increases roughly 4-fold, moving from 11% to 39%, and the latter shrinks, going from 11% to 3%. The total proportion of biases doubles from around 20% for current-year Fall forecasts to more than 40% for year-ahead Spring forecasts.

⁹I.e. invalid inference based on asymptotic distributions.

Figure 12: Share of Countries with Short-Term Biases

Note: The figure shows the share of countries for each forecast horizon and issue of the World Economic Outlook with a 5% statistically significant negative or positive bias. H=0 and H=1 indicate, respectively, current year and year-ahead. Test of statistical significance is run individually with country-by-country regressions.

Although measuring the proportion of countries with biases is an important first step, it is also essential to quantify their magnitude: indeed, these could range from relatively small mistakes to real blunders. [Table X](#) show summary statistics for the significant intercepts.

TABLE X: Magnitude of Short-Term Biases

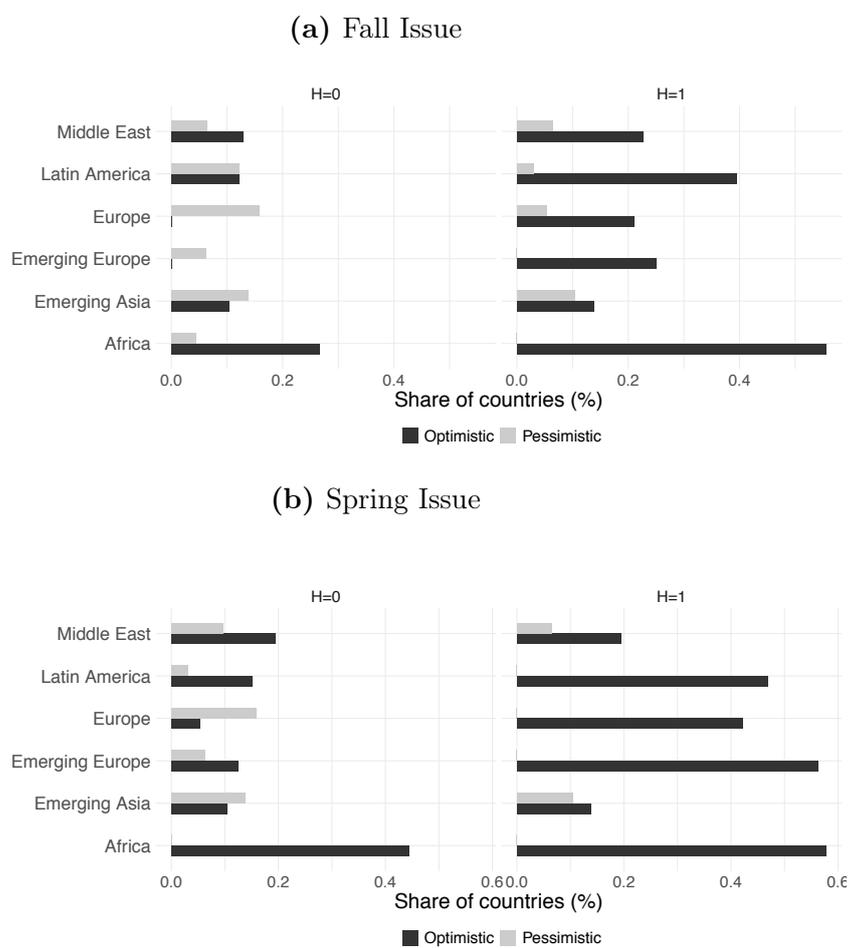
Horizon	Type of bias	Mean	Median	Min.	Max.
H=0	Optimistic	-1.19	-0.88	-0.32	-3.66
H=1	Optimistic	-1.83	-1.31	-0.49	-18.09
H=0	Pessimistic	0.84	0.62	2.03	0.23
H=1	Pessimistic	1.12	1.1	2.49	0.22

Note: Summary statistics of country-by-country intercepts significant at 5% level for the WEO sample. H=0 and H=1 indicate, respectively, current year and year-ahead. Fall and Spring issues of WEO pooled together by horizon.

The average optimistic bias is -1.2 % for current-year forecasts and -2% for year-ahead forecasts. Since the distribution of biases is, however, characterized by negative skewness and fat-tails (see Max. column), we rely on the median value as a better estimate: the median optimistic bias is -0.89% for current-year and -1.34% for year-ahead. Interestingly, the difference between mean and median of the distribution widens with the forecast horizon indicating the presence of larger outliers at the year-ahead horizon. The average pessimistic bias is, instead, 0.97% and 1.36% respectively and the corresponding median is equal to 0.65% and 1.1%. Even though the distributions of pessimistic biases also exhibit skewness and fat-tails, outliers are, however,

relatively small compared to optimistic biases. To summarise, not only optimistic biases are far more frequent, but they are also about 0.2% larger in absolute value and characterized by more extreme outliers.

Hitherto, we focused only on the aggregate picture. However, another question worth investigating is what group of countries exhibits the largest number of biases and in which direction. [Figure 13](#) shows the percentage of countries with a bias conditional on their geographical group. Africa has the highest number of countries with optimistic biases both for Fall and Spring current year forecasts and Fall year-ahead forecasts, with a proportion ranging from 20% to more than 50%. Although it is not anymore the worst-ranking region, the proportion increases and hovers around 60% for Spring year-ahead forecasts. Moreover, within the group we do not find any country with a pessimistic bias at any horizon. Emerging Europe has the largest proportion of countries with optimistic biases in the year-ahead Spring forecasts (more than 60%) and the second largest proportion for Fall year-ahead forecasts (around 40%). Surprisingly, however, optimistic biases are much lower in current-year Spring (10%) and especially, Fall forecasts (2%), indicating that optimism for these countries is only a feature of longer forecast horizons. Emerging Asia shows at the same time the highest proportion of countries with pessimistic biases in current Fall forecasts and year-ahead Spring and Fall and the lowest proportion of optimistic biases in year-ahead Spring and Fall. For Latin America, Europe and Middle East we obtain mixed results, with the presence of both pessimistic and optimistic biases at the current-year horizon that gradually tilts toward optimism at the year-ahead.

Figure 13: Share of Countries with Short-Term Biases - Region

Note: The figure shows the share of countries for each forecast horizon and issue of the World Economic Outlook with a 5% statistically significant negative or positive bias. H=0 and H=1 indicate, respectively, current year and year-ahead. Test of statistical significance is run individually with country-by-country regressions.

In addition to the highest proportion of optimistic biases, Africa exhibits also the greatest average and median bias for current year forecasts (respectively, -1.5% and -1.35%) and greatest average at the year-ahead horizon(-2.77%), both well above the overall mean and median (Table XI).¹⁰ Similar results hold for Emerging Europe. Instead, European countries show the smallest biases at both horizons. For a more detailed view, Table XXIX in the Appendix reports the result of individual estimations for single countries.

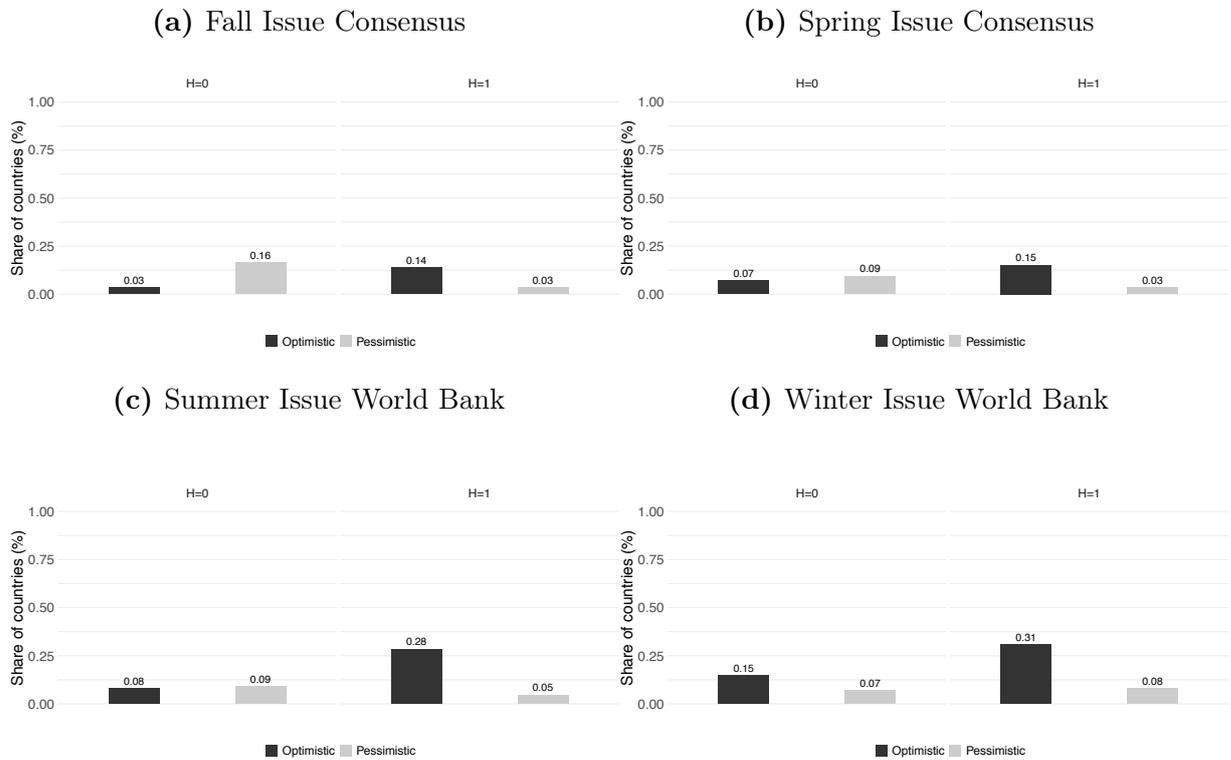
TABLE XI: Magnitude of Short-Term Biases - Region

Horizon	Geo. group	Mean	Median	Max.
H=0	Africa	-1.45	-1.16	-3.66
H=0	Emerging Asia	-0.78	-0.82	-0.97
H=0	Emerging Europe	-1.29	-1.29	-1.33
H=0	Europe	-0.33	-0.33	-0.33
H=0	Latin America	-1	-0.84	-2.24
H=0	Middle East	-0.87	-0.72	-2.45
H=1	Africa	-2.59	-1.64	-18.09
H=1	Emerging Asia	-1.32	-1.21	-1.72
H=1	Emerging Europe	-1.88	-1.99	-2.78
H=1	Europe	-0.79	-0.74	-1.18
H=1	Latin America	-1.33	-1.27	-3.5
H=1	Middle East	-1.68	-1.25	-4.45

Note: Summary statistics of country-by-country intercepts significant at 5% level for the WEO sample. H=0 and H=1 indicate, respectively, current year and year-ahead. Fall and Spring issues of pooled together by horizon. Only negative (optimistic) intercepts reported.

¹⁰For space reasons, in Table XI we report only summary statistics for optimistic biases.

We find that optimism, however, is not an exclusive feature of Fund forecasts, but is pervasive across institutional and private forecasters. [Figure 14](#) shows the share of countries for which we detect the presence of a bias in different issues of World Bank and Consensus forecasts. Since in the case of World Bank the number of observations for each country is limited, we design a simple bootstrap experiment to prevent at best distortions in the individual test statistics: specifically, for each country, horizon and issue, we repeat the estimation of an intercept on 1,000 bootstrapped samples and then calculate the standard deviation of the estimate. Lastly, we compute individual test statistics and draw inferences based on the empirical standard deviation.

Figure 14: Share of Countries with Short-Term Biases - Consensus and World Bank

Note: The figure shows the share of countries for each forecast horizon and issue of Consensus survey and World Bank Global Economic Prospects with a 5% statistically significant negative or positive bias. H=0 and H=1 indicate, respectively, current year and year-ahead. Test of statistical significance is run individually with country-by-country regressions; for World Bank, the standard errors are obtained from a 1,000 replications bootstrap for each country, issue and horizon.

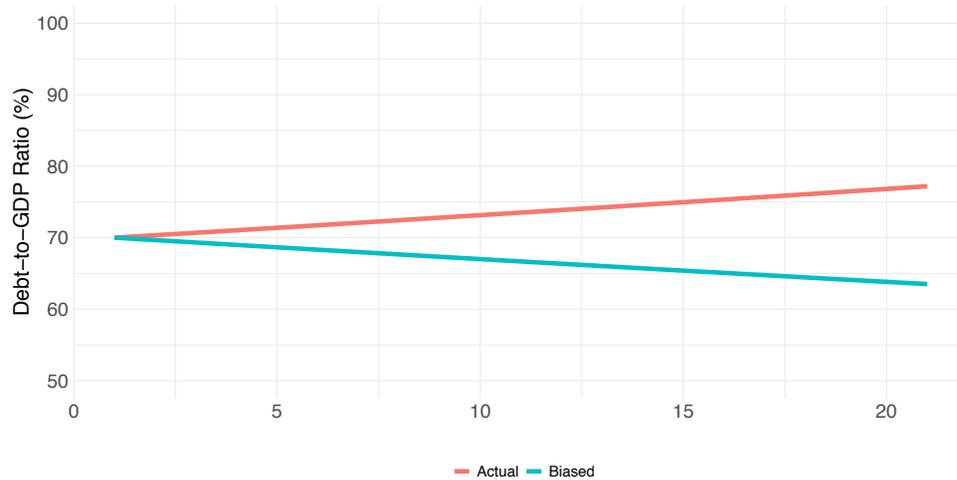
For both World Bank and Consensus, the share of countries with optimistic biases grows monotonically with the forecast horizon, with the proportion increasing 5-fold, from 3% to 15%,

for Consensus forecasts, and 4-fold, from 8% to 31%, for World Bank forecasts.¹¹ At the same time, pessimistic biases shrink at longer forecast horizons.

Optimistic biases are, in turn, pivotal, because of their tight connection with policy advice: indeed, they can hinder effective policy reaction in periods of high turmoil and incentive shortsighted fiscal policy during expansion times, thus jeopardizing future debt sustainability. [Frankel \(2011\)](#) examines growth forecasts made by 33 official government agencies and finds that overoptimism helps to explain excessive budget deficits and in particular, the failure to run surpluses during periods of high output. [Beaudry et al. \(2021\)](#) exploit an instrumental variable approach to mitigate endogeneity concerns and show that Balance of Payments-difficulties are more likely to arise in economies for which past WEO growth forecasts have been overly optimistic. To complement this evidence and provide a simple intuition of the possible pernicious ramifications optimistic biases can have, we rely on a straightforward simulation exercise.

¹¹The smaller share of optimistic biases for Consensus stems from the near absence of African countries: only Nigeria and South Africa are included in the survey.

Figure 15: Simulated Debt-to-GDP: with and without bias



Note: The simulation exercise follows closely De Resende (2014). Actual debt path is simulated with growth rate equal to 2%, interest rate 2.5% and initial Debt-to-GDP 70%. Biased forecast for growth rate is 3%. Government assumed to run balanced budget.

In [Figure 15](#) we simulate the evolution of the debt-to-GDP ratio based on the actual growth rate and a biased forecast over 20 periods.¹² While the actual debt path would be increasing, Fund staff would estimate downward sloping dynamics and hence, with high likelihood, incentive the use of expansionary macroeconomic policies: this misguided policy advice, in turn, could result in explosive debt dynamics.

¹²The biased growth forecasts is calibrated based on the average optimistic WEO bias (see [Table XVI](#)).

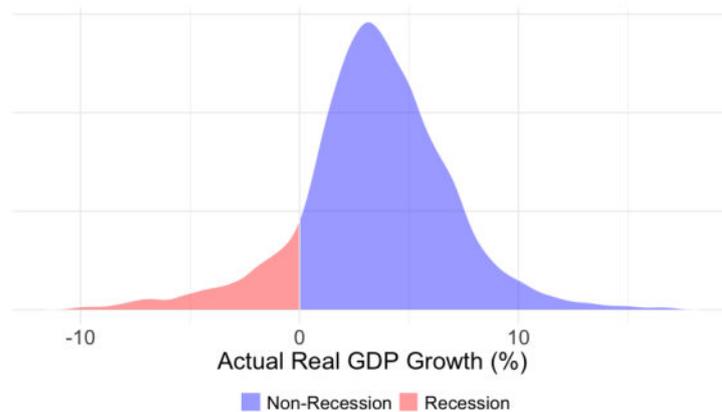
4.3.2 Optimism, Recessions and Financial Crises

Widespread and sizeable optimistic biases characterise the short-term forecasts of real GDP growth of different institutions. Reporting solely the unconditional mean of forecast errors may, however, shroud important heterogeneity of the last over the state of the business cycle: specifically, over predictions might occur symmetrically during recessions and expansions or not.

[Loungani \(2001\)](#) shows that forecast errors for Consensus Forecasts of 63 industrialized and developing countries in the period 1989-1999 are largely negative during recessions. Private forecasters fail to predict nearly all recessions at the year-ahead horizon, with a quarter of them remaining unpredicted in current-year October forecasts. [Batchelor \(2007\)](#) analyzes Consensus Forecasts for G7 countries in the years 1990-2005. He finds evidence of an optimistic bias for the forecasts of real GDP growth: he ascribes this optimism to a rational adaptation of forecasts to the fall in the trend growth rate in the aforementioned group of countries. [Lewis and Pain \(2014\)](#) assess the performance of OECD projections for GDP growth over the period 2007-12: they report a general optimism with the largest errors occurring for the vulnerable euro area economies and at the height of the financial crisis in 2009. [Dovern and Janssen \(2017\)](#) document how Consensus Forecasts errors for 19 advanced economies over the period 1990–2013 are state-dependent: forecasts for recession years exhibit large and systematic negative forecast errors, while errors are considerably smaller during recoveries and expansions. Finally, [An et al. \(2018\)](#) describe the evolution of growth forecasts in the run-up to recessions. For a panel of 63 countries in the years 1992 to 2014, they find that WEO forecasts and Consensus Forecasts fail to predict recessions in a similar way.

Following [Jordà et al. \(2013\)](#), we set recessions equal to periods of actual negative growth.¹³ The total number of recessions corresponds to around 12% for the WEO sample ([Figure 16](#)).¹⁴ Although this percentage is much lower compared to the number of expansion periods, recessions are far from being “black swan” episodes.

Figure 16: Distribution of Actual Real GDP Growth



Note: Distribution of actual real GDP growth for the WEO sample over the period 1990-2019. Actual growth is the one reported in the October WEO issue of the following year. Recessions are periods of negative growth.

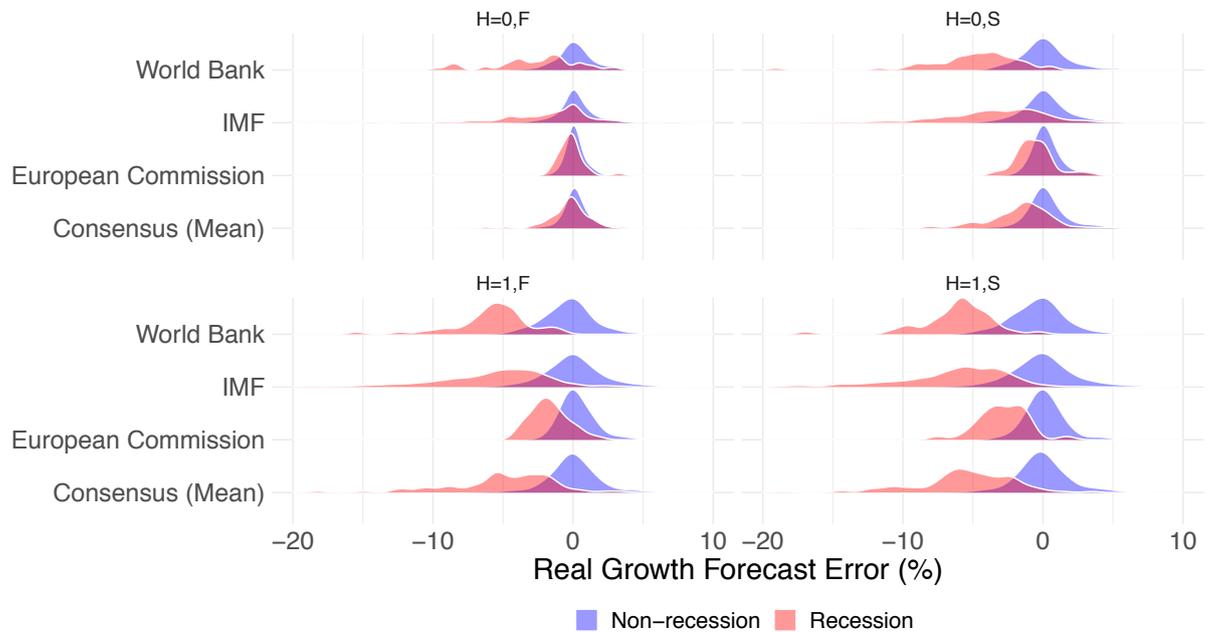
¹³Although there exist more sophisticated methods to date business cycles (e.g. [Calderón and Fuentes \(2014\)](#); [Harding and Pagan \(2002\)](#)), they are mainly employed for data at lower frequency.

¹⁴For all different samples the percentage is similar.

Figure 17 shows the distribution of forecast errors for all the main institutional and private forecasters, discriminating between recession and non-recession years. Three striking patterns emerge. First of all, forecast errors in non-recession years are perfectly symmetric around 0 for all forecasters: this means that during periods of positive growth, forecasters have no tendency toward either optimism or pessimism. In contrast, the whole distribution of forecast errors during recessions is shifted to the left of 0 for year-ahead forecasts and partially, for current year forecasts: not only forecasters considerably underplay recessions on average, but they do so virtually in every instance. Second, the longer the forecast horizon, the more pronounced the difference between the two distributions: intuitively, this reflects the smaller information set available to forecasters at longer horizons. Third, while the variance of the distribution in non-recession years is relatively contained, the distribution widens considerably for recession years, indicating a significant heterogeneity in the errors.¹⁵¹⁶

¹⁵Although one may be tempted to compare distributions across forecasters and conclude on their relative performance in predicting recessions, this would be a flawed comparison: the sample used for each forecaster corresponds to data availability, thus the underlying volatility of the business cycle is different.

¹⁶Given that we have also data for individual private forecasters from Consensus survey, we also test the “rocking the boat” hypothesis: although institutions and the average private forecaster might not be able to predict recessions, there still might be some individual bank that outperforms all others. Figure 41 shows forecast errors conditional on the state of business cycle for the “best” private forecaster: we find that results are qualitatively similar.

Figure 17: Major Forecasters and Recessions

Note: Distribution of real GDP growth forecast errors for main institutional and private forecasters at different horizons. Actual real GDP growth rate from October WEO issue of following year. H=0 and H=1 indicate, respectively, current year and year-ahead forecast errors. F and S indicate, respectively, Fall and Spring issue forecast errors with the exception of World Bank (respectively, Summer and Winter). The sample for each forecaster reflects data availability. Recessions are periods of negative growth

Furthermore, since different institutions report growth forecasts for different income groups, e.g. World Bank for emerging markets and low-income economies, European Commission for

advanced economies, [Figure 17](#) confirms that the inability to forecast correctly during downturns is not a feature of specific income groups, but rather a general result.

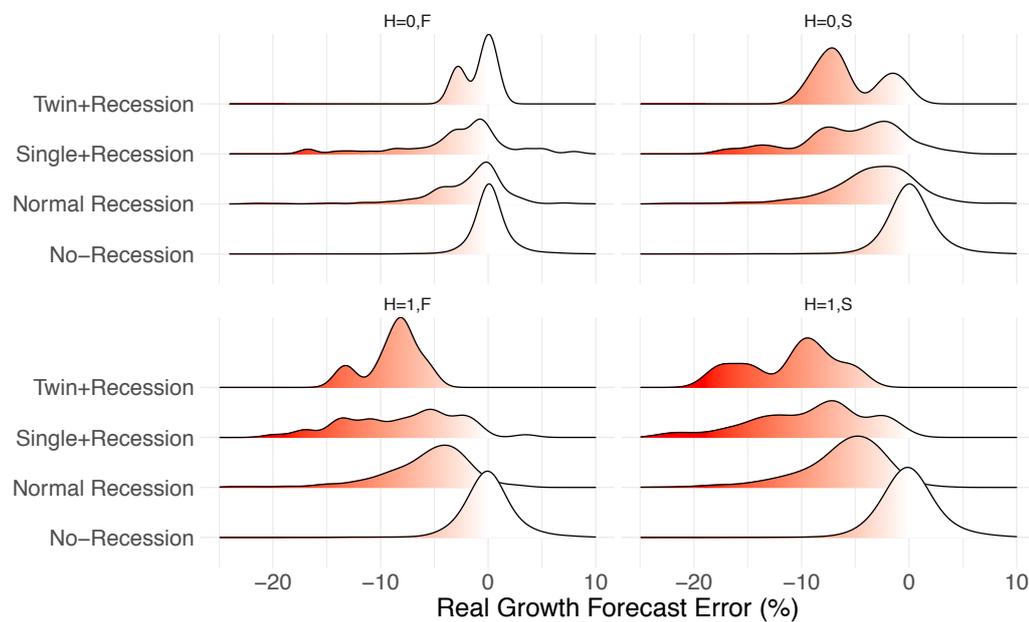
The figure does not, however, provide an explanation for why recessions are not forecasted ahead of time or their size is, in any case, underplayed. A number of different explanations have been put forward by the literature: first, poor data and models might hamper the prediction of recessions. Second, there might be a lack of incentives for forecasters, with an asymmetric loss functions characterizing their behaviour: the weight they place on Type 1 errors, or fake alarms, might be larger than the one placed on Type 2 errors. Third, forecasters may hold on to their priors and only revise them slowly and insufficiently in response to incoming information ([Nordhaus, 1987](#)). Finally, recessions might occur as the result of random shocks that are difficult if not impossible, by definition, to forecast. We add to this literature, reconducting the large heterogeneity in the size of forecast errors during recessions years to the occurrence of financial crises.

[Figure 18](#) shows the distribution of forecast errors in periods of no-recession, simple recession and recession accompanied by a single or multiple financial crises.¹⁷ At all horizons, forecast errors are larger if the recession is accompanied by financial crises. This behaviour is particularly marked for year-ahead forecasts: while the median forecast errors for simple recessions (in absolute value) hovers, respectively, around 5% and 6% for year-ahead Fall and Spring forecasts,

¹⁷We show only results for WEO forecasts because of the significantly larger time span available compared to the other forecaster and do not display results for financial crises without recessions and triple crises because of the few observations available.

it increases by two percentage points for episodes accompanied by a single financial crisis, roughly 7% and 8%, and more than three percentage points for episodes accompanied by twin crises.

Figure 18: IMF Forecasts, Recessions and Financial Crises



Note: Distribution of real GDP growth WEO forecast errors for episodes of no-recession, only recession and recession accompanied by a single or twin financial crises. Actual real GDP growth rate from October WEO issue of following year. Recessions are periods of negative growth. Financial crises correspond to currency, banking and sovereign debt crises. The corresponding dummy is from [Laeven and Valencia \(2018\)](#). $H=0$ and $H=1$ indicate, respectively, current year and year-ahead forecast errors. F and S indicate, respectively, Fall and Spring issue forecast errors. Brighter red indicates more optimistic errors.

In a nutshell, the extensive optimistic biases found for institutional and private forecasters in section 4.3.1 are the result of large negative errors during recessions while, instead, errors during periods of positive growth are relatively small and balance themselves out on average. This inability to predict the onset and extent of recessions is not unique to the Fund, but is ubiquitous to all main institutional and private forecasters. We explain the high variability in the size of forecast errors for recession periods by the joint occurrence of recessions accompanied by financial crises: this result, in turn, can help directing future research on the determinants of this dismal performance.

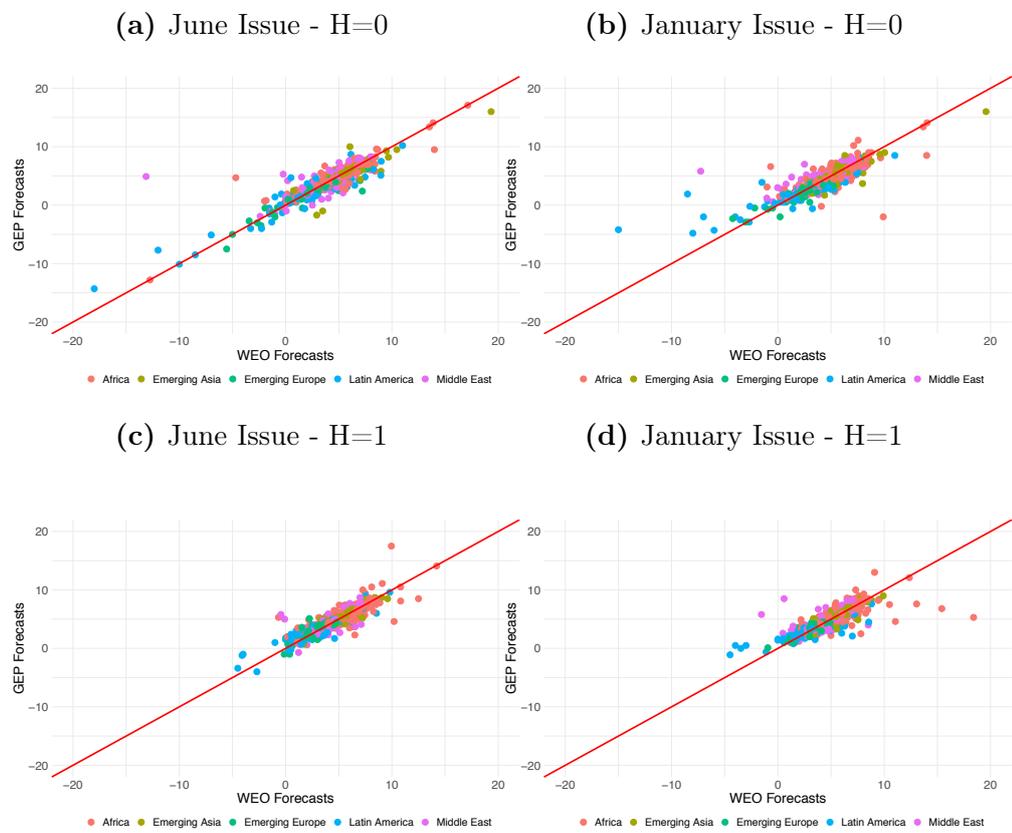
4.3.3 Comparison between Forecasters

The second step in our quality assessment of surveillance forecasts is the comparison between the accuracy of different institutional and private forecasters. First of all, different institutions present a different focus: specifically, the World Bank prioritises the surveillance of low income countries, while the European Commission monitors closely European countries. This could result in better forecasts compared to the Fund for the aforementioned. In turn, better performance may reflect divergence in the underlying information set and the econometric model deployed, making it worthwhile to investigate these differences. Finally, a comparable performance between different forecaster is also an important result, as it is indicative of an unexplicable component in the realization of GDP growth.

Figure 19 shows the scatter plots of current year and year-ahead forecasts for the Fund and the World Bank. As evident, the correlation between the two in the period 2010-2018 is close to or higher than 0.9, with forecasts clustering around the 45 degrees line. Nevertheless, the

forecasts for some country groups, i.e. Africa and Latin America, are less collected compared to the others.

Figure 19: Comparison between World Bank and IMF Forecasts



Note: Red line is 45 degrees line. Each dot represents a country-year forecast.

In [Table XII](#) we report for each country group the share of countries for which the Fund produces a lower Root Mean Squared Error (RMSE) compared to the Bank and the result for a two-sided Diebold-Mariano test where the null is of equal predictive accuracy. As expected, we find few statistical differences between the two institutions forecasts: the only striking differences we notice is for Latin America, where the Fund performs significantly better than the Bank and has lower RMSE for basically all countries at the current year horizon.¹⁸ A more granular view with a detailed country by country comparison ([Table XXXI](#)) reveals, that, however, even for this group of countries, differences, although statistically significant, are not economically large: the ratio of the GEP RMSE to the WEO RMSE for the countries in question is never below 50% and rarely even above 25%.

We repeat a similar exercise, this time including in the comparison Fund and European Commission forecasts ([Figure 20](#)). The correlation remains high, hovering around 0.9, suggesting also in this case almost perfect collinearity. Some marked differences between the two are, however, are visible in the bottom left part of each panel. This suggests that in the case of some countries, the European Commission has produced more pessimistic forecasts compared to the Fund.¹⁹ [Table XXXIII](#) reports the percentages and test of equal accuracy for the two largest country groups. The European Commission is significantly better than the Fund for both

¹⁸The result is robust to the use of a one-sided test where the null hypothesis is that Fund forecasts are no more accurate than Bank forecasts.

¹⁹These countries are mainly three, Cyprus, Greece and Portugal, and the periods correspond to well-known episodes of financial crises.

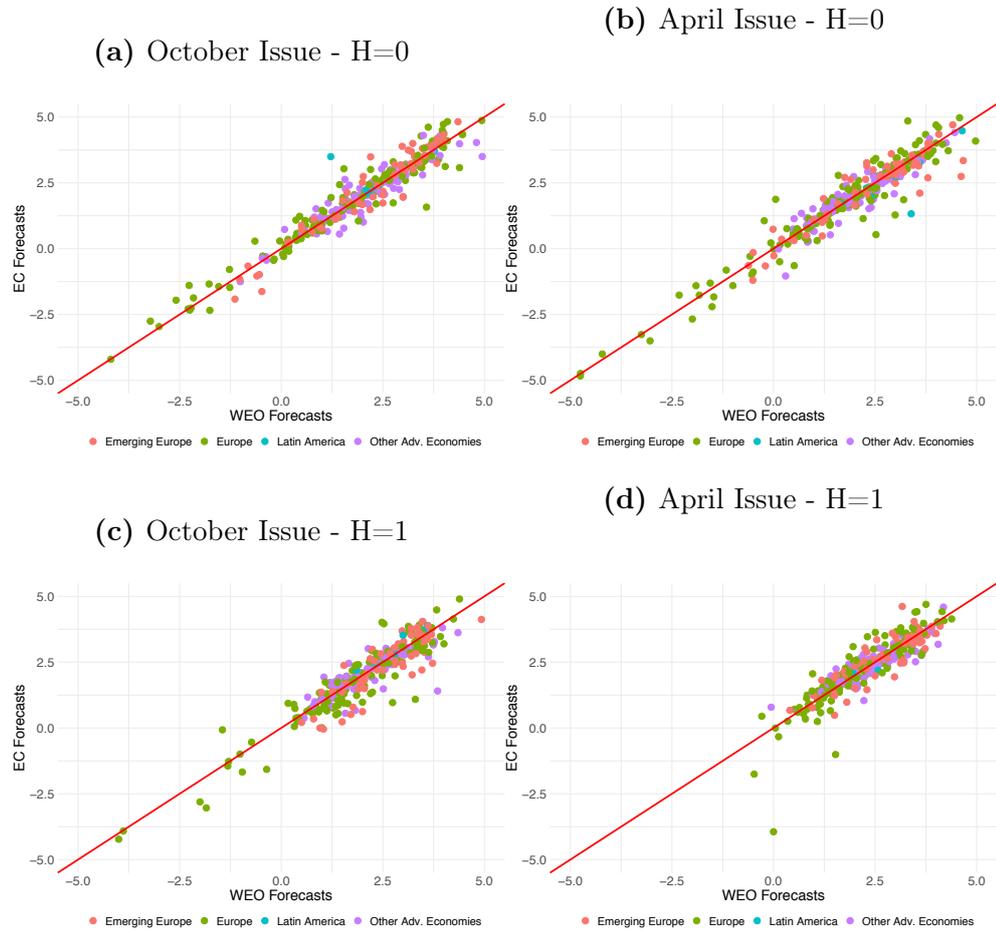
TABLE XII: Regional Specialization - World Bank and IMF

Geo.group	Variable	H=0,Jun.	H=0,Jan.	H=1,Jun.	H=1,Jan.
Africa	Percentage	0.57	0.43	0.43	0.43
	DM Test	-1.11	0.48	1.16	1.7*
Emerging Asia	Percentage	0.77	0.69	0.15	0.38
	DM Test	0.24	0.64	0.77	0.76
Emerging Europe	Percentage	0.62	0.75	0.38	0.38
	DM Test	-1.49	-2.89**	-1.46	1.02
Latin America	Percentage	0.81	0.81	0.52	0.57
	DM Test	-2.72**	-3.3**	-2.52**	-1.5
Middle East	Percentage	0.69	0.44	0.62	0.38
	DM Test	-1.33	-1.07	-0.28	-0.21

Note: Percentage refers to the share of countries with a lower root mean squared error for WEO forecasts compared to GEP forecasts. DM Test is the test statistic associated with a two-sided Diebold-Mariano test where the null is of equal accuracy between forecasts. ***: significant at 1*: significant at 10

country groups at all horizons, producing a lower RMSE for the majority of countries. Also in this case, however, improvements are marginal and in the order of magnitude of one fourth ([Table XXXIII](#)). The most drastic improvement we observe is for Portugal at the year-ahead horizon: in this case, the better performance from the European Commission stems from the higher accuracy for the period of the sovereign debt crisis.

Figure 20: Comparison between European Commission and IMF Forecasts



Note: Red line is 45 degrees line. Each dot represents a country-year forecast.

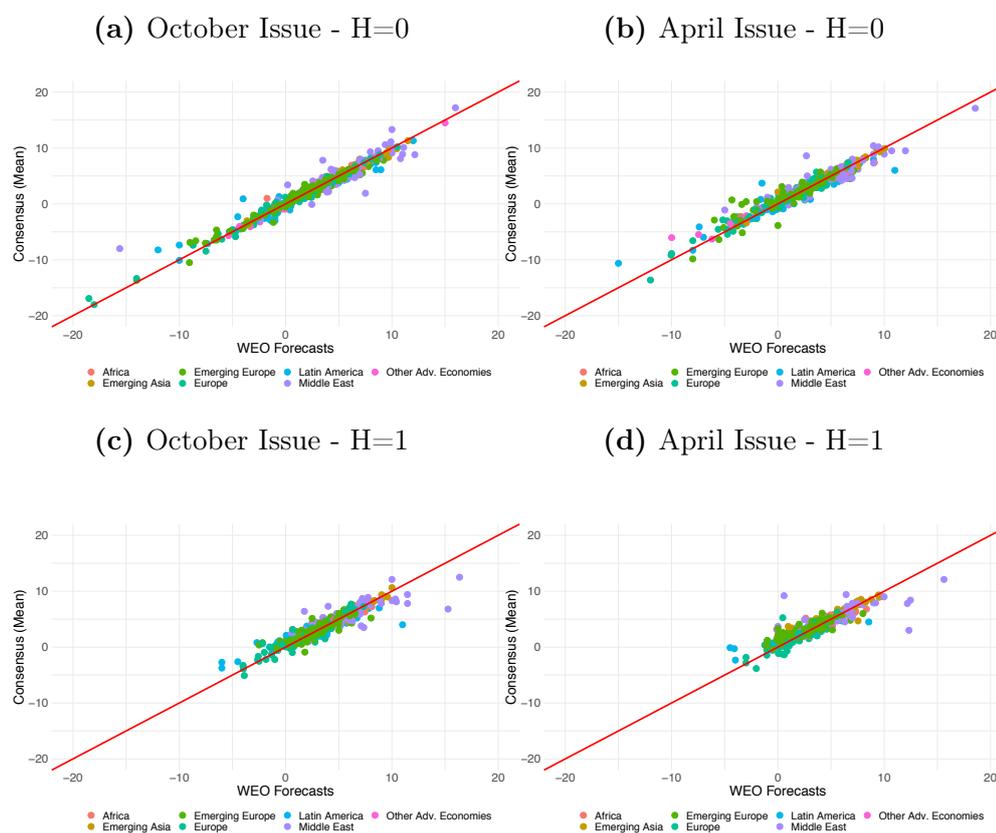
TABLE XIII: Regional Specialization - European Commission and IMF

Geo.group	Variable	H=0,Oct.	H=0,Apr.	H=1,Oct.	H=1,Apr.
Emerging Europe	Percentage	0.33	0.11	0.11	0.33
	DM Test	1.92*	2.95**	2.21**	0.4
Europe	Percentage	0.37	0.21	0.16	0.47
	DM Test	1.8*	2.41**	3.8**	1.7*

Note: Percentage refers to the share of countries with a lower root mean squared error for WEO forecasts compared to AMECO forecasts. DM Test is the test statistic associated with a two-sided Diebold-Mariano test where the null is of equal accuracy between forecasts. ***: significant at 1*: significant at 10

Finally, we compare Fund forecasts with the private sector: we find that the two exhibit a similar degree of collinearity as in the previous cases ([Figure 21](#)). In the same way as before, statistical differences among country groups are few and sparse ([Table XIV](#)).²⁰

²⁰In the case of Africa, WEO produces a lower RSME for 100% of the countries in the sample. However, the sample is composed only by two countries, Nigeria and South Africa.

Figure 21: Comparison between Consensus and IMF Forecasts

Note: Red line is 45 degrees line. Each dot represents a country-year forecast.

In short, IMF WEO forecasts display a correlation of 75% or higher with World Bank, European Commission and private sector at different horizons: although some regional differences at the statistical level hold, they remain economically small.

TABLE XIV: Regional Specialization - Consensus and IMF

Geo.group	Variable	H=0,Oct.	H=0,Apr.	H=1,Oct.	H=1,Apr.
Africa	Percentage	1	0.5	1	1
	DM Test	-0.32	0.39	-1.33	0.03
Emerging Asia	Percentage	0.5	0.6	0.7	0.4
	DM Test	-1.11	0.85	0.29	2.33**
Emerging Europe	Percentage	0.43	0.64	1	0.86
	DM Test	-0.43	0.05	-1.82*	-0.2
Europe	Percentage	0.71	0.76	1	1
	DM Test	-1.68	-1.89*	-3.23**	-1.61
Latin America	Percentage	0.83	0.44	0.5	0.39
	DM Test	-2.67**	-1.4	-2.04**	-0.81
Middle East	Percentage	0.78	0.78	0.44	0.33
	DM Test	-2.36**	-0.77	0.72	0.48
Other Adv. Economies	Percentage	0.81	0.38	1	0.5
	DM Test	-2.49**	1.23	-2.48**	1.32

Note: Percentage refers to the share of countries with a lower root mean squared error for WEO forecasts compared to Consensus forecasts. DM Test is the test statistic associated with a two-sided Diebold-Mariano test where the null is of equal accuracy between forecasts. ***: significant at 1*: significant at 10

4.4 Program Forecasts

In this section, we focus on growth program forecasts. We study different issues, such as the presence of biases and the difference between concessional and non-concessional programs, but, in particular, our focus goes on the relationship between the amount of credit granted by the Fund and the size of the ensuing forecast error.

A common claim is that forecasts are more optimistic for countries where the amount of financial assistance granted through an IMF-funded program is higher. This claim stems from three underlying reasons: first, a more optimistic forecast may serve as justification to provide a higher level of resources to a country suffering severe external pressure.²¹ Second, since the Fund becomes a “de facto” creditor when intervening in a country with a program, a higher forecast increases the likelihood the country will attract anew foreign capital, thus increasing its chances of repayment. Third, in context of programs and in contrast with the surveillance case, forecasts are the result of discussions and a subsequent agreement between Fund officials and country authorities: it follows that “if the authorities are leaning towards very optimistic projections, the final result of the process will be biased in that direction [too].” (Luna, 2014).

The literature suggests mixed evidence on the matter, with results highly dependent on the time period and sample chosen. Musso and Phillips (2002) study a sample of 69 programs granted between 1993 and 1997, not finding evidence of a general optimistic bias. Nevertheless,

²¹The same argument, however, can also apply conversely, with a more optimistic forecasts used to not intervene.

restricting the analysis to solely big programs, they find that projections were significantly optimistic for this subsample during the period. [Baqir et al. \(2005\)](#), based on data for 94 programs between 1989 and 2002, find an optimistic bias in forecasts of GDP growth. Nevertheless, they show that actually “high-profile” programs are less biased than “normal” programs.²² The most recent study, [Luna \(2014\)](#), analyses a sample of 103 programs approved in the period 2002-2011: growth forecasts exhibit only a non-significant optimistic bias. At the same time, they find that for the countries with exceptional access arrangements, the forecast bias at program inception was optimistic and significant.

We use data for a sample of 214 programs approved in the period 2002-2018: a full list of the programs with date of Board approval, type of program and amount is available in [Table XXXVI](#).²³ [Table XV](#) reports, instead, summary statistics for the variable of interest, the amount of the program: there is a large heterogeneity in the size of programs approved, ranging from a minimum of 5% (Rwanda 2002) to a maximum of 3211.8% (Greece 2010) with a standard deviation of roughly 360%.

²²Their paper uses “high profile” to identify “large access” programs with lending exceeding 2 billion SDRs.

²³We include also precautionary programs in the sample, albeit the relative forecast errors might come from a different distribution compared to normal programs. All results are robust to the exclusion of precautionary programs.

TABLE XV: Descriptive Statistics - Amount Programs

Statistic	N	Mean	St. Dev.	Min	Max
Amount (% of quota)	225	228.9	363.8	5.0	3,211.8

Note: Table shows descriptive statistics for IMF programs size (in % of member country quota at Board approval) over the period 2002-2018. Data retrieved from MONA database.

Table XVI shows 5 different regression specifications for current and year-ahead forecasts. First, as in section 4.3.1, we check for the presence of biases in programs forecasts: we find in both cases the presence of optimistic biases, ranging from -0.3% for current year to -0.7%, on average, for year-ahead forecasts. In the second specification, we examine whether the Global Financial Crisis (GFC) constitutes a structural break in the sample and if there is any difference in bias between programs granted prior to 2008 and in the period post GFC. We find somewhat limited evidence of this: if anything, optimistic biases are smaller for current-year forecasts, but not for year-ahead. In the third specification, we introduce the variable of interest, the amount of the program: we find that a 1% higher amount corresponds to a forecast error lower by -0.001%. If, however, large programs are granted early in the year and therefore present more optimistic forecast errors, the coefficient would be negatively biased. We control for the possible effect of the time difference introducing a variable equal to the number of months remaining between the approval of the program and the end of the year. Reassuringly, the control has no influence on the result and is not significant. Finally, we check differences in forecast errors

between concessional and non-concessional programs.²⁴ If larger programs are concessional and at the same time present larger forecast errors, the coefficient would be negatively biased. The introduction of the dummy has no effect for current year forecasts while is negative, significant and equal to -0.5% for year-ahead, indicating that non-concessional programs exhibit more optimistic forecasts. Nevertheless, the coefficient estimated for the variable of interest does not lose magnitude and remains significant: this means that the relationship holds independently from the difference concessional/non-concessional.

The evidence reported in [Table XVI](#) supports the hypothesis of more optimistic forecasts for bigger programs. Nevertheless, this result does not necessarily entail that biases are discretionary: for instance, it might be intrinsically more difficult to forecast the growth realization in economies that require heavy monetary intervention by the Fund. A good test for “discretion” is to compare MONA forecasts with private sector forecasts: if private forecasters have been equally or more optimistic than the Fund in the case of large programs, then the hypothesis of a voluntary bias becomes increasing unlikely.²⁵ [Figure 22](#) compares the forecasts produced by MONA and Consensus for three well-known post-GFC programs: if anything, Consensus seems more optimistic in some instances (e.g. Argentina 2018).

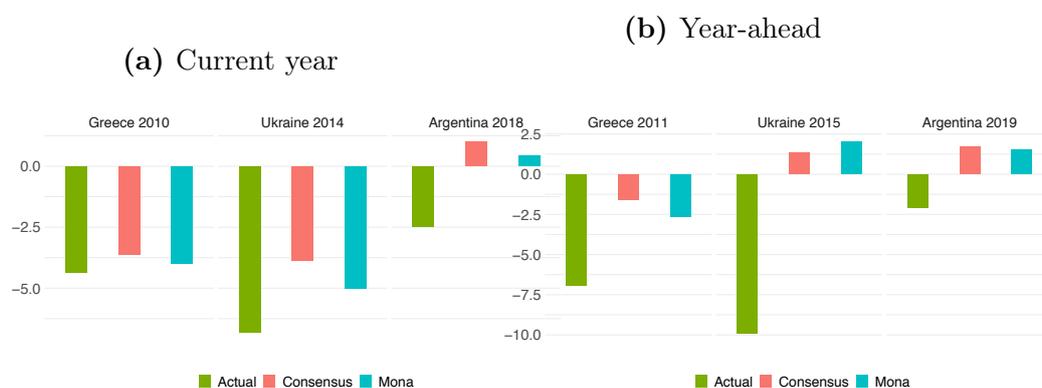
²⁴Non-concessional programs come from the General Resources Account (GRA) and are SBAs, EFFs, PCLs and PLLs.

²⁵Indeed, the same incentives to produce better forecasts are not shared by the private sector.

TABLE XVI: Amount Program and Bias

	<i>Dependent variable:</i>									
	GDP forecast error (current year)					GDP forecast error (year ahead)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Post-GFC		0.340 (0.225)					-0.275 (0.314)			
Total amount (% quota)			-0.001** (0.0003)	-0.001** (0.0003)	-0.001* (0.0003)			-0.001** (0.0004)	-0.001** (0.0004)	-0.001*** (0.0005)
Remaining months				0.024 (0.033)					0.034 (0.046)	
Concessional					0.211 (0.250)					-0.563* (0.340)
Constant	-0.348*** (0.113)	-0.525*** (0.162)	-0.174 (0.132)	-0.303 (0.219)	-0.323 (0.221)	-0.665*** (0.154)	-0.504** (0.240)	-0.425** (0.179)	-0.610** (0.303)	-0.026 (0.300)
Observations	225	225	225	225	225	216	216	216	216	216
F Statistic		2.282	6.081**	3.305**	3.394**		0.767	6.467**	3.514**	4.635**

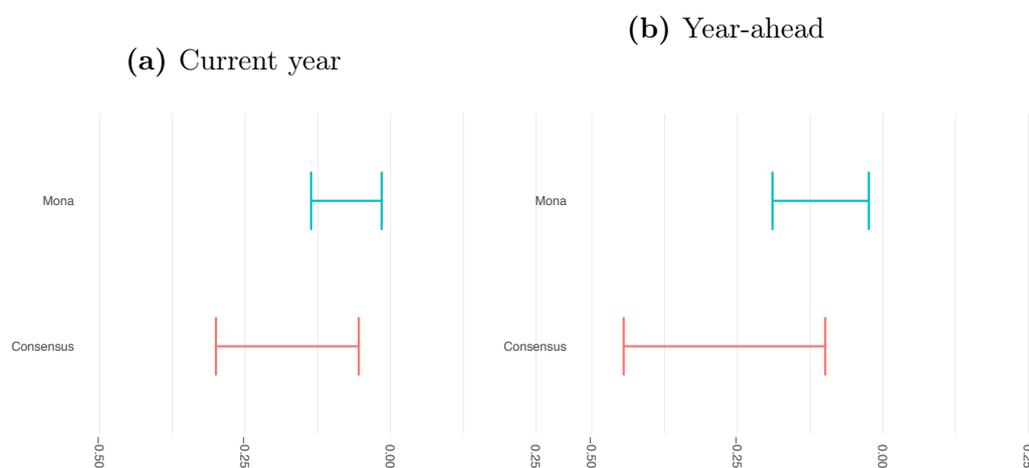
Note: Dependent variable winsorized at the 10% level. For columns 1-5 the dependent variable is equal to the current year forecast error, while for column 5-10 is equal to the year-ahead forecast error. Post-GFC is a dummy equal to 1 for programs approved after 2009. Remaining months is a variable that represents the number of months remaining before the end of the year from the date of program approval. Non-concessional is a dummy equal to 1 for non-concessional GRA programs (see note 24). Heteroskedasticity robust standard errors in parentheses. ***: significant at 1% level, **: significant at 5% level, *: significant at 10% level.

Figure 22: Comparison with Consensus: Anecdotal Evidence

Note: The figure shows the real GDP growth forecasts formulated at program approval for Greece 2010, Ukraine 2014 and Argentina 2018 SBAs from MONA database and Consensus survey (mean value). Actual growth rate is the real GDP growth from the October WEO issue of the following year.

We test formally this equivalence substituting MONA forecasts with Consensus forecasts produced in the same month, calculate the forecast error and estimate the same regressions of [Table XVI](#) (columns 4-9). [Figure 23](#) shows the estimated relationship between forecast errors and the total amount of the program for MONA and Consensus.²⁶

²⁶Since Consensus forecasts are not available for all programs, we estimate both regressions on a subset of 70 programs.

Figure 23: Comparison with Consensus: Statistical Relationship

Note: The figure shows 95% confidence intervals obtained regressing the forecast errors for programs on the amount of the program (in % of country quota) and multiplying by 100. Data for forecasts are, respectively, from MONA and Consensus. The sample of programs included in the regressions corresponds to data availability for Consensus (70 programs).

We find that the negative relationship holds in a statistically equivalent way for private forecasters and Fund alike, thus thwarting the hypothesis of a deliberate bias.

4.5 Conclusion

This paper investigates the predictability of real economic activity in the last 30 years, analyzing the short-term performance of the main institutional and private sector forecasters. To this end, we assemble a large sample of growth projections coming from different multilateral institutions, specifically IMF, World Bank and European Commission, and private forecasts from Consensus survey. We study three main aspects. First, we examine the extent and magnitude of possible forecasts biases and how these relate to specific regions of the world and forecaster. Second, we compute and compare the accuracy of forecasts produced by different institutions. Third, shifting our attention toward program forecasts, we analyze the claim that larger IMF-funded programs correspond to greater forecast errors and explore the possibility of a vested interest of the Fund in producing biased forecasts.

We find evidence of widespread optimism in growth forecasts: optimism is much more prevalent than pessimism and is ubiquitous in all forecasters analyzed. Inspecting more closely forecasts errors, we find that they are highly asymmetric over the state of the business cycle: errors during expansion are, on average, balanced and small. Errors during recessions are, instead, optimistic and much larger in magnitude. This result points to a sheer inability of the economic forecasting world to predict recessions, hence allowing us to generalize the results found by previous literature (An et al., 2018; Loungani, 2001). In terms of accuracy, although some regional differences still remain across forecasters, official sector forecasts are so similar between themselves and to the private sector that a statistical horse race to assess which one is better ends

with a photo finish for all horizons. These findings suggest a large degree of cross-pollination across different sources and a minimal incentive to deviate from the “consensus”.

Lastly, we confirm that larger IMF programs exhibit larger forecast errors for forecasts formulated at program inception. The estimated effect is similar for current year and year-ahead forecasts and corresponds to a 0.1% decrease in the forecast error for a 100% increase in the amount of the program. Substituting forecasts errors from IMF with Consensus ones, we show that, however, a statistically equivalent relationship holds for the forecasts formulated by the private sector. If anything, private sector forecasts in the case of highly covered programs were more optimistic than Fund staff ones. This result hinders the hypothesis of a vested interest or of political nudges influencing IMF program forecasts. In light of the conditionality of program forecasts, future research should aim at evaluating this evidence against some measures of program success, with some steps already been taken into this direction (e.g. [IMF \(2019\)](#)).

Taken altogether, our findings have first-order implications for both domestic policymakers and lender of last resort institutions during these times of global turmoil: if history is of any indication, the aforementioned should bear in mind that the economic outlook presented by growth forecasts is likely to be more sanguine than its actual realization. These considerations may avoid excessive complacency when considering the extension or withdrawal of domestic stimulus and constitute the basis for effective communication to private capital markets about debt developments.

CHAPTER 5

A (MODERN) NARRATIVE ON 70 YEARS OF MACROECONOMIC CRISES

A DATABASE FOR COMPLEXITY

Abstract: While the recent empirical literature on macroeconomic crises focused on a limited subset of events (e.g. banking, currency and sovereign), macroeconomic crises are usually characterized by large scale domino effects that involve a much wider and heterogeneous array of sectors and transform them into highly complex events. This data limitation, in turn, hampers the understanding of these chaotic and painful episodes for researchers and policymakers alike. In this paper, after building a raw corpus of roughly 23,000 International Monetary Fund country reports, we harness the power of text mining to produce a new database on crises discussion: this database covers 20 different types of economic, financial and non economic events for a sample of 181 countries over the period 1950-2019. We document a substantial rise in complexity of macroeconomic crises throughout the *XX* and *XXIth* century and a higher centrality of the non-fundamental channel in the system.

Contributions: Both authors contributed equally in all aspects of the work.

Code: Code for calculation of the IMF corpus term-frequencies and replication of chapter material is available from a public Github repository: https://github.com/manuelbetin/Text_mining_IMF. An *R* package to extend the methodology to other corpora is also available in a public Github repository:

<https://github.com/manuelbetin/TextMiningCrisis>. Scraping and OCR implementation codes available upon request.

5.1 Introduction

"As research methods advance, and as more social media data accumulate, textual analysis will be a stronger field in economics in coming years. It may allow us to move beyond 1930s-style models of feedback, the "multiple rounds of expenditure", and get closer to all the kinds of feedback that really drive economic events."

Shiller (2019)

Macroeconomic crises originate in different areas of the economic system, propagate through various channels and ultimately, disrupt different sectors of economic activity. A peripheral and localized shock e.g. the collapse of the housing market at the onset of the GFC, can start a domino effect that extends far beyond the initial disruption and breeds into highly complex outcomes. While in the last decades the empirical crises literature has been growing substantially, it has only devoted particular attention to a limited subset of crises: mainly, banking crises, sovereign debt default and currency crashes (e.g. Laeven and Valencia (2013), Reinhart and Rogoff (2013)). Although these crises are central and critical components of macroeconomic dynamics, it would be deceptive to isolate them from non economic and less conventional events. These events, such as epidemics, political uncertainty, violent conflicts or migration crises might act as causes, consequences and amplification mechanisms, whose occurrence determines the speed, intensity and duration of economic and financial downturns. Reconstructing the complex narratives of periods of high macroeconomic volatility in a quantitative and coherent framework

is an herculean challenge that can both enrich the historical understanding of crises and provide empirical support to highlight specific mechanisms in theoretical frameworks.

In this chapter, exploiting the recent technological advances in terms of computational power, image recognition and text mining techniques, and the overseer role the International Monetary Fund (IMF) played over the last 70 years for its membership, we provide new and rich material for the analysis of macroeconomic crises. First, we provide accessibility to a raw text database of roughly 23,000 documents - country reports and program related - covering the whole IMF membership throughout the period 1950-2019. Second, we manually compile an IMF crisis-specific dictionary and propose a simple term-frequency approach to capture and quantify Fund discussions about a large variety of economic and non economic crisis events for each country and year. The large time span (70 years) and country coverage (181 countries) as well as the scope of crises covered within a comparable framework complement and extend standard datasets of macroeconomic crises and provide useful material for a deeper understanding of the complexity at play during these highly chaotic events.

We provide evidence that crises complexity, measured as the number and intensity of correlations between the different term-frequencies, has increased massively starting from the first wave of financial globalization, shaping a sparse network with mostly real and domestic crises in the Bretton Woods era into a highly dense one, financially dominated, in the recent decade. Furthermore, we highlight another connected trend: the rise in centrality of non-fundamental drivers, expectations, in the unfolding of complex events. All in all, this evidence calls for new strategies of crises prevention and mitigation by policymakers.

The chapter is structured as follows. Section 5.2 summarizes the relevant literature on text mining and details the construction of the corpus and of the vocabulary and the method used to compute the 20 crisis indicators. Section 5.3 presents a general overview of the dataset, validating it against standard benchmarks and reporting examples of non-standard indicators. Section 5.4 shows evidence of rising crises complexity throughout the sample period. Lastly, section 5.5 concludes.

5.2 From Qualitative Judgment to Quantitative Measures

The analysis of macroeconomic crises necessarily suffers from important data limitations that often limit and bias the general understanding of these highly chaotic and painful episodes. Sometimes, however, “[d]ifferent terrains [simply] call for different vehicles” (Akerlof, 2020), with the emergence of new techniques that contribute to the rise of novel perspectives and findings. Natural Language Processing (NLP) and text analysis have gained great popularity over the recent period in academia: this development has been fueled by the evolution of information technologies and the booming of the big data area. These techniques have permitted to transform large quantities of text into numerical data, extracting valuable insights and avoiding labor-intensive reading and manual coding. In the social sciences, this type of empirical approach has been used to analyse a large spectrum of subjects, ranging from the political slant of media to drivers of consumer decision-making.¹

In macroeconomics, it has proved a useful tool to capture the perceptions of economic agents as well as a good complement to the traditional economic and financial data. A first strand of literature harnesses sentiment-analysis i.e. the interpretation and classification of emotions (positive, negative and neutral) to enhance the forecasting of economic fluctuations and upcoming crises as well as providing an additional understanding of the swings in assets prices. Fratzscher and Reynaud (2011) assess the degree of favorableness in the Public Information Notices (PINs) issued after Executive Board discussions of IMF *Article IV Consultations* with

¹See Gentzkow et al. (2019) for a comprehensive review.

member countries. The sentiment classification depends on the authors' interpretation of IMF information and results in a quantitative score that goes from -2 to +2. For a set of 36 emerging market economies over the period 2001-2007, they find that the degree of favorableness significantly influences sovereign spreads. [García \(2013\)](#) constructs a sentiment index from the financial columns of *The New York Times*. The author uses a dictionary approach to classify positive and negative words in each article. He finds that news content helps predict stock returns at the daily frequency, especially during recessions. Exploiting a similar dictionary approach, [Fraiberger \(2016\)](#) constructs a sentiment index over the period 1987 -2013 across 12 countries: to do so, he analyzes a corpus of economic news articles produced by *Reuters*. He finds that information from news articles is not incorporated into Consensus forecasts. [Huang et al. \(2019\)](#) build monthly sentiment indices for 20 countries from 1980 to 2019 using *Financial Times* news articles. Instead of a predefined dictionary, they use a *word2vec* algorithm ([Mikolov et al., 2013](#)), an unsupervised technique focusing on the distribution of words, to map them into a high-dimensional space and then count the occurrence of precise semantic clusters across the articles. Following this initial clustering, they classify each group according to its general sentiment (fear, risk, hedging and crisis). They find that their sentiment indices spike ahead of financial crises and conclude that these new indices could complement traditional forecasting methods with early real time data. [Fayad et al. \(2020\)](#), working on a corpus of IMF *Article IV Consultations* including all member countries in the period 2000-2018, develop a sentiment index measuring the reception of policy advice at the time of the consultations: they find

that, although authorities of member countries largely agree with Fund advice, there is sizeable heterogeneity connected with different country-specific economic and political characteristics.

Uncertainty and risk measures are the second avenue of research. These studies, rather than relying on the emotional intent of words, exploit solely their presence or frequency within a document. In their seminal work, [Baker, Scott R. et al. \(2016\)](#) create an economic policy uncertainty (EPU) index for the United States from 1985 onwards counting the number of articles in the 10 leading US newspapers with words related to the economy, uncertainty and policymaking. They then extend the same methodology to include all G10 economies. [Ahir et al. \(2018\)](#) use the same approach with quarterly *Economist Intelligence Unit* (EIU) country reports: they produce an uncertainty index for 143 individual countries on a quarterly basis from 1996 onwards. Recently, a modified versions of this index, the World Pandemic Uncertainty index, has been published to improve the understanding of the economic consequences of epidemics.² [Ghirelli et al. \(2019\)](#) refine this methodology for the Spanish case extending both the newspaper coverage and enriching the set of keywords to search for. [Engle et al. \(2020\)](#) construct a climate change news index relying on a corpus from *The Wall Street Journal* (WSJ) covering the time span 1980-2017. The authors convert WSJ term counts into “term frequency–inverse document frequency” and compare the resulting scores to a corpus of authoritative texts on the subject

²<https://blogs.imf.org/2020/04/04/global-uncertainty-related-to-coronavirus-at-record-high/>

of climate change: they use this new measure and a portfolio approach to build climate change hedge portfolios.³

Finally, different authors rely on a simple reading methodology for the identification of financial crisis episodes. [Romer and Romer \(2017\)](#) create a new semiannual measure of financial distress in a sample of 24 advanced economies from 1967 to 2007. This measure is based on the manual coding of the *OECD Economic Outlook*: after reading all the documents, the authors classify the degree of financial distress for a certain country/half-year on a scale from 0 to 15. They then use this new series to explore the behavior of economic activity following financial crises. [Vannier \(2020\)](#) develops a conceptual framework to guide the choice of key elements entailed in such a methodology and applies it to date the start of currency crises: he proposes a narrative taxonomy of currency crises for 54 countries based on 315 IMF publications – mainly article IV consultations - covering the time span 1970-2020.⁴

Compared to the previous literature, our contribution is twofold. First, we significantly improve accessibility to IMF documents for economists and social scientists alike: we provide a raw text database of roughly 23,000 documents - country reports and program related - covering the whole IMF membership throughout the period 1950-2019. Previously, [Mihalyi and Mate \(2019\)](#) introduced a text dataset of country reports published by the IMF between 2004 and

³Another notable mention goes to [Choi and Varian \(2012\)](#) and [Scott and Varian \(2014\)](#) that use *Google Trends* data and a term-frequency approach to nowcast economic activity.

⁴Manual coding has also been used for other purposes e.g. [Hernandez \(2020\)](#) characterizes the policy discourse in IMF–Argentine *Article IV Consultations*.

2018 for 201 countries. We build on their work and extend the sample of documents backwards exploiting Optical Character Recognition (OCR) and auto correction techniques to overcome the accuracy hurdle that hindered previous work.⁵ Second, we manually compile an IMF crisis-specific dictionary and propose a simple term-frequency approach to capture and quantify Fund discussions about a large variety of economic and non economic crisis events for each country and year. This algorithmic approach allows us to process in a computationally feasible way the large volume of data available and ultimately, to capture the entire complexity of events associated with episodes of macroeconomic volatility.

In the remainder of the section we present the source, country and time coverage of the corpus, explaining in detail the data acquisition and processing part, describe the construction of the lexicon and illustrate the empirical method used to extract crisis discussion indices from IMF texts.

5.2.1 The Corpus of IMF Documents

The IMF is an international organization created in July 1944 at the Bretton Woods conference. Its primary mandate is to preserve the stability of the international monetary system i.e. the system of exchange rates and international payments. Although the institution is mostly known for its role of financial assistance for countries experiencing balance of payments difficulties, its mandate is larger and ranges from the provision of technical assistance and the strengthening of local capacity to the production of regular forecasts for its member countries.

⁵“We choose 2004 as our starting year because...from earlier periods...the majority are scanned PDF which make text recognition difficult and imprecise”(Mihalyi and Mate, 2019).

In this paper, we exploit the surveillance activity of the Fund and concentrate on drawing a quantitative synthesis of country-specific outlooks.

The Fund regularly monitors and evaluates the situation of economic and financial systems in order to identify contingent sources of risk. Surveillance is carried out at different levels: (i) from a global and wide perspective in publications such as the World Economic Outlook, (ii) on more specific topics/regions in recurrent publications such as the Global Financial Stability Report, Fiscal Monitor, External Sector Report, Regional Economic Outlook or (iii) at a country level in the Article IV and other country reports e.g. the Recent Economic Developments series. Moreover, surveillance is often also a key condition associated with the Fund financing programs. Hence, program-related documents such as requests for assistance or program reviews also contain important surveillance elements. To produce credible and comprehensive information about the economic outlook of its member countries, the Fund relies on an evolving conceptual framework for assessing country risks (Ahuja et al., 2017) and large teams of economic experts that work in close relationship with national authorities and main economic actors, collecting and analysing a large variety of quantitative and qualitative information. This research culminates in the production of regular and formal economic reports which provide background information for political, economic and financial decisions.

The privileged relationship of the Fund with national authorities of almost all countries in the world, the close interactions of its staff with leading scholars and policy makers as well as the rigorous editorial process ensure an evident comparative advantage compared to other textual sources such as newspapers and tweets. Fund documents exhibit a number of desirable

characteristics; first of all, they are long and detailed: they depict meticulously the real time outlook of the economic situation. While newspapers might neglect some elements in favor of others or not correctly grasp the situation at hand, this risk is minimized by the Fund. Second, they are extremely cautious: each word is discussed, weighed and negotiated and follows a strict protocol of revision and publishing. This revision system, commonly supervised by the Strategy, Policy and Review (SPR) department leads to a situation of homogeneous linguistic i.e. a high likelihood that linguistic findings based on one document apply to another (Kilgarriff, 2001). Homogeneous wording is the kingpin of our lexicon approach: it allows us to compile consistently a dictionary of expressions the Fund uses to refer to different occurrences of the same type of events. The heterogeneous wording that characterizes other textual sources would render this methodology non-viable.

The scraping of the IMF archives for all documents for each of the IMF member countries provides around 250,000 references produced between 1947 and 2016. In addition to the references of documents available in the archives, we also scrape the current IMF website to obtain references for the documents published after 2016.⁶⁷ The reference provides rich metadata: mainly, the title of the document, the day of publication and the link to the attached document in PDF format. We perform a semantic analysis on the title to retain only the relevant types

⁶Respectively, <https://archivescatalog.imf.org> and <https://www.elibrary.imf.org>.

⁷We do not include documents produced after 2019 as for many countries they are still not disclosed. Moreover, these recent documents may present important changes in the structure.

of documents i.e. documents related to individual country surveillance.⁸ This leaves 39,000 references for which we download the associated PDF. Sometimes, however, metadata exhibit inconsistencies: thus, from this original sample, we further remove a number of misclassified documents.⁹¹⁰ This further cleaning leaves us with a final corpus of 23,465 documents covering the time span 1950-2019.

To read and convert these documents into text suitable for statistical analysis, we have to overcome a substantial accuracy hurdle: most of the early reports feature wandering baselines (horizontal lines on which the letters “sit”) and ink splodges, resulting in a highly imprecise text recognition with standard libraries available in modern statistical softwares. At the same time, more advanced image recognition technologies i.e. OCR require a much longer computational time to process the enormous amount of information. In order to obtain the maximum accuracy

⁸We consider these as country reports and program-related documents. See section C.1.1.1 in the Appendix for more details on the procedure.

⁹Titles may contain reference to Article IV or programs and be of different nature e.g. a Working Paper on the evaluation of program outcome. In turn, these documents are problematic since they discuss specific topics and confuse information about different countries.

¹⁰Technically, to automate the process and not revise one-by-one the documents, we compile a list of keywords that commonly appear on the first page of these misclassified documents and exclude the document from the term-frequency calculation (see section 5.2.3) if we detect one of them on the first page. A full list of these keywords is available in Table Table XXXVII.

while making the task computationally feasible, we decide to harness the power of cloud services: in particular, we choose to use Google Cloud's Vision API (see [Figure 24](#)).¹¹

¹¹For the details, see section [C.1.1.3](#).

Figure 24: Converting Documents to Text Data

(a) Original Document (b) Normal Extraction

- 10 -

The impact of the new exchange system will also affect the domestic monetary situation since the increase in price, in domestic currency, of the imported materials is reflected in the price of the final product and in the cost of living. This in turn may give rise to an inflationary factor, i.e., possible increases in the level of money wages and the initiation of a wage-price spiral. To some extent such developments will depend on the behavior of the banking system in extending additional credit, and in this connection the importance of general credit and monetary policies should be re-emphasized.

Another factor is the importance the Brazilian Government may attach, for psychological reasons, to (1) keeping the various exchange rates at a level regarded as tolerable and (2) to the avoidance of violent fluctuations in such rates. It is therefore important to evaluate the intensity of private demand and of the possible reactions of the business community to the operation of the exchange system. A factor of importance in this connection is that in recent months, due to the severity of import restrictions, inventories have been running down at a significant pace. Import demand for various products may therefore be influenced by the need for reconstituting inventories as well as for current use.

Availabilities of exchange for the auction markets may also be limited by the value of import licenses issued in the recent period, the amount of which is unknown. If this is large, the supply of exchange allocated to new imports would be correspondingly limited.

A third element which could affect the intensity of import demand is the public expectations created by measures the Government might take to regulate the auction market, such as shifting of commodities, changes in allocations and possibly a lowering of the maximum exchange allowance for individual importers. Some of these measures might introduce a speculative element in import demand and in this connection the importance of restrictive credit and monetary measures should also be re-emphasized.

In view of these considerations, import demand will probably be sufficiently great with relation to availabilities (after deduction of Government needs and allocation for past licenses) that the tendency will be for effective rates to go to relatively high levels. Should this lead to gradual increases in the percentages of exchange allocated to the most essential categories, or to increasing downward shifts of commodities (including shifts to the category entitled to the official rate without auction premium), the effectiveness of the system as a means of monetary disinflation and price adjustment will be minimized.

An additional danger may be noted in connection with the problem of readjusting the domestic cost and price structure to price conditions in the outside world. The exchange system creates a separate exchange rate relationship with each country. It is difficult to see how orderly cross rates can be maintained since the levels of the rates are affected by two elements influencing the demand and supply of exchange which are unavoidably arbitrary, i.e., the schedules of import and export commodities in the case of trade agreement countries, and the percentage allocations of the various currencies within each category.

The impact of the new exchange system will also affect the domestic monetary situation since the increase in price, in domestic currency, of the imported materials is reflected in the price of the final product and in the cost of living. This in turn may give rise to an inflationary factor, i.e., possible increases in the level of money wages and the initiation of a wage-price spiral. To some extent such developments will depend on the behavior of the banking system in extending additional credit, and in this connection the importance of general credit and monetary policies should be re-emphasized. Another factor is the importance the Brazilian Government may attach, for psychological reasons, to (1) keeping the various exchange rates at a level regarded as tolerable and (2) to the avoidance of violent fluctuations in such rates. It is therefore important to evaluate the intensity of private demand and of the possible reactions of the business community to the operation of the exchange system. A factor of importance in this connection is that in recent months, due to the severity of import restrictions, inventories have been running down at a significant pace. Import demand for various products may therefore be influenced by the need for reconstituting inventories as well as for current use. Availabilities of exchange for the auction markets may also be limited by the value of import licenses issued in the recent period, the amount of which is unknown. If this is large, the supply of exchange allocated to new imports would be correspondingly limited. A third element which could affect the intensity of import demand is the public expectations created by measures the Government might take to regulate the auction market, such as shifting of commodities, changes in allocations and possibly a lowering of the maximum exchange allowance for individual importers. Some of these measures might introduce a speculative element in import demand and in this connection the importance of restrictive credit and monetary measures should also be re-emphasized. In view of these considerations, import demand will probably be sufficiently great with relation to availabilities (after deduction of Government needs and allocation for past licenses) that the tendency will be for effective rates to go to relatively high levels. Should this lead to gradual increases in the percentages of exchange allocated to the most essential categories, or to increasing downward shifts of commodities (including shifts to the category entitled to the official rate without auction premium), the effectiveness of the system as a means of monetary disinflation and price adjustment will be minimized. An additional danger may be noted in connection with the problem of readjusting the domestic cost and price structure to price conditions in the outside world. The exchange system creates a separate exchange rate relationship with each country. It is difficult to see how orderly cross rates can be maintained since the levels of the rates are affected by two elements influencing the demand and supply of exchange which are unavoidably arbitrary, i.e., the schedules of import and export commodities in the case of trade agreement countries, and the percentage allocations of the various currencies within each category.

(c) OCR Extraction

- 10 -

The impact of the new exchange system will also affect the domestic monetary situation since the increase in price, in domestic currency of the imported materials is reflected in the price of the final product and in the cost of living. This in turn may give rise to an inflationary factor, i.e., possible increases in the level of money wages and the initiation of a wage-price spiral. To some extent such developments will depend on the behavior of the banking system in extending additional credit, and in this connection the importance of general credit and monetary policies should be re-emphasized.

Another factor is the importance the Brazilian Government may attach, for psychological reasons, to (1) keeping the various exchange rates at a level regarded as tolerable; and (2) to the avoidance of violent fluctuations in each rate. It is therefore important to evaluate the intensity of private demand and of the possible reactions of the business community to the operation of the exchange system. A factor of importance in this connection is that in recent months, due to the severity of import restrictions, inventories have been running down at a significant pace. Import demand for various products may therefore be influenced by the need for reconstituting inventories as well as for current use.

Availabilities of exchange for the auction markets may also be limited by the value of import licenses issued in the recent period, the amount of which is unknown. If this is large, the supply of exchange allocated to new imports would be correspondingly limited.

A third element which could affect the intensity of import demand is the public expectations created by measures the Government might take to regulate the auction market, such as shifting of commodities, changes in allocations and possibly a lowering of the maximum exchange allowance for individual importers. Some of these measures might introduce a speculative element in import demand and in this connection the importance of restrictive credit and monetary measures should also be re-emphasized.

In view of these considerations, import demand will probably be sufficiently great with relation to availabilities (after deduction of Government needs and allocation for past licenses) that the tendency will be for effective rates to go to relatively high levels. Should this lead to gradual increases in the percentages of exchange allocated to the most essential categories, or to increasing downward shifts of commodities (including shifts to the category entitled to the official rate without auction premium), the effectiveness of the system as a means of monetary disinflation and price adjustment will be minimized.

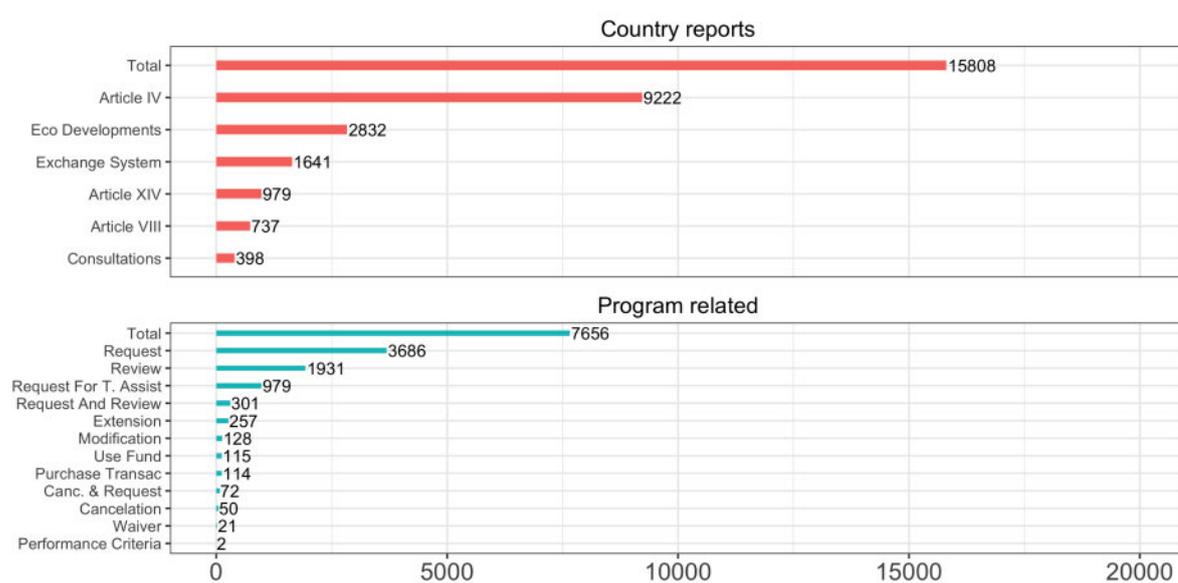
An additional danger may be noted in connection with the problem of readjusting the domestic cost and price structure to price conditions in the outside world. The exchange system creates a separate exchange rate relationship with each country. It is difficult to see how orderly cross rates can be maintained since the levels of the rates are affected by two elements influencing the demand and supply of exchange which are unavoidably arbitrary, i.e., the schedules of import and export commodities in the case of trade agreement countries, and the percentage allocations of the various currencies within each category.

Note: Panel (a) shows an extract from the document *1953 Consultations - Brazil*, Panel (b) the associated text extraction with a standard library (*pdf_text* function from the *pdftools* R package) and Panel (c) the associated OCR extraction performed using Google Cloud Vision.

Figure 25 shows the composition of the final corpus, dividing between country reports and program-related documents. Country reports comprise two thirds of the total: the most common documents are Article IVs, followed by Recent Economic Developments. While the former are the last step of the *Article IV consultation* and are published with the prior agreement of country authorities, the latter are internal documents that provide background for the whole process and prepare in advance their analysis (Vannier, 2020). While *Article IVs consultations* should be conducted annually, in practice the final publication often has biennial frequency, especially in earlier years. In addition, countries that are in disagreement with Fund analysis may refuse the release of Article IVs. Hence, including Recent Economic Developments in the corpus is fundamental to compensate the otherwise inevitable loss of information. Article XIV and Article VIII documents appraise the motivation behind the introduction and maintenance of exchange rate restrictions, and thus are particularly useful to capture episodes of currency volatility. Lastly, simple Consultations are an “archaic” version of Article IVs, mostly present in the 50s and 60s and replaced afterwards. Program-related documents are, instead, mainly composed by Requests and Reviews, with other minor documents, e.g. Modifications, Waivers etc., completing the picture. In Appendix C.1.1.4, we present an overview of the different types of documents: their purpose, whether or not they are currently issued by the Fund and whether their production is or was regular (Table XXXVIII). Furthermore, we also show the

evolution in the number of individual country reports and program-related documents over time (Figure 42).¹²

Figure 25: Size of the Corpus



¹²We find a strong correspondence between the description of the documents and their evolution over time. For instance, Article VIII and Article XIV documents that are published for countries maintaining exchange restrictions, disappear at the turn of the 21st century (Figure 42a).

The documents in the corpus exhibit an extensive coverage, covering almost every country in the world (Figure 26).¹³ While Latin American countries have been widely covered by the Fund, developments in some countries of Africa and Central Asia are less documented. Rather than a shortcoming, however, this geographical distribution of the corpus reflects deep historical reasons: most nations in Africa were colonized by European states in the early modern era and gained independence relatively late compared to Latin American countries.¹⁴ Similarly, different Central Asian nations gained independence only after the fall of the Soviet Union in 1992.¹⁵ North-American and European countries reveal, instead, a uniform pattern.

¹³With the exception of non-IMF members, 11 in total: Cuba, East Timor, North Korea, Liechtenstein, Monaco, Taiwan, and Vatican City.

¹⁴Many countries in the early 1960s, but some after 1970 e.g. Guinea-Bissau from Portugal in 1973 and even afterwards e.g. Zimbabwe from Britain in 1980.

¹⁵Obviously, this is only part of the explanation: the other part is that large Latin American countries requested a high number of programs compared with other geographical groups.

Figure 26: Country coverage

Reports on high income countries appear, on average, more than a decade earlier in the sample compared to lower income groups ([Table XVII](#)): this result is coherent with advanced economies being the earliest clients of the Fund.¹⁶ In [Appendix C.1.1.4](#), we report the detailed number of documents by country and the year the first document was published ([Table XXXIX](#)).¹⁷

¹⁶“Advanced economies had been [the Fund] earliest and largest clients before the emerging market economies started to dominate its activity in the 1980s.” ([Reinhart and Trebesch, 2016](#)).

¹⁷From the table other patterns emerge: for example, small islands such as Antigua & Barbuda and Vanuatu are the least covered in the sample.

TABLE XVII: Start Date by Income Group

Income Group	Avg. Start Year
High income	1959
Low income	1971
Upper middle income	1974

Note: Avg. Start Year corresponds to the average publication year of the first document in the final corpus of IMF reports.

To the best of our knowledge, in this work we provide accessibility and consider the to-date broadest corpus of relevant IMF documents both in time, country coverage and type of report. This rich material allows us to adopt a holistic stance toward the analysis of crisis events and to ensure a general perspective for all income groups, thus re-balancing the crisis literature toward low-income countries. Furthermore, since it encompasses a long time horizon, it allows us to draw comparisons between the early Bretton Woods era and the most recent period: in this way, it corrects the usual focus bias on financially dominated crises for the post 1980 decades.

5.2.2 The Lexicon of Crises

A proper lexicon is a dictionary of words, based on prior knowledge that provides a list of synonyms or near-synonyms describing the occurrence of an event. The rising interest for unsupervised and more complex text mining techniques is largely motivated by the necessity to by-pass the construction of a lexicon, a long and cumbersome process. Moreover, a large number of corpora displays little information on its actual content that is highly heterogenous, making prior knowledge largely irrelevant. In this respect, the format of IMF documents and

the homogeneous language of the corpus largely eased the identification of the patterns and recurrent expressions commonly used by Fund staff.

In practice, the lexicon has been constructed by a first identification of each category according to prior knowledge on the dates and locations of the major disruptions to the macroeconomic outlook.¹⁸ A careful reading of these well-known events formed the building stone of the vocabulary identification for each category and the first iteration in the process. Following this, preliminary term frequencies (see section 5.2.3 below) provided indices for each country and document that, in turn, served as guideline toward finding the most relevant documents. In this second step, we both control the veracity of the identification and add or correct the words and expression to include in the lexicon. This two-step process was repeated as long as evident marginal improvements were noticeable. In order to have a comprehensive and accurate lexicon, capturing extensively the occurrence of specific events, but also limiting potential Type 1 and Type 2 errors, we establish and follow a number of coding rules. The main guidelines are the following. First, we refrain from adopting a predetermined length for n-grams: no fixed number of words was defined. The lexicon includes both single words e.g. *epidemia*, *flood* or *rainfall* and longer expressions such as *large real depreciation* or *slowdown of economic activity*. We limit words with multiple meanings including the associated adjective in the expression, for example *trade war* indicates instances of trade conflicts, while *civil war* indicates armed battles. The couple noun+adjective was constructed using expert knowledge combined with manual reading

¹⁸Wikipedia provided a complementary source of initial information, in particular for non-economic events e.g. https://en.wikipedia.org/wiki/List_of_natural_disasters_by_death_toll.

of the documents. Regular expressions and anchors have been extensively applied to capture plural forms and avoid other matching errors.

The final dictionary counts almost 700 expressions organized in 20 categories. [Table XVIII](#) provides an extract of the vocabulary included in the lexicon.¹⁹ The heterogeneity in the number of expressions for each category is large, ranging from 64 words for contagion to 7 for migration crises. Nevertheless, this difference reflects some structural characteristics of the corpus itself: namely, the diversity of expressions increases with the degree of economic relevance of the category. Economists easily find a number of synonyms for well-established economic phenomena, but less so for more unconventional events that are, hence, characterized by redundancies. Moreover, some categories contain by construction shorter terms that can refer to a multiplicity of complicated expressions. For example, *Paris club* is directly associated with debt rescheduling and thus, useful to capture a variety of phrases. No generic expression of the sort exists for regional crises, forcing us to include the different variations (*regional crisis, crisis in the region, crisis in neighboring countries* etc.), therefore increasing the number of expressions in the category. The political crises group is also particularly large. We ascribe this to the diplomatic tone necessarily adopted by Fund staff: numerous euphemisms are often used to substitute harsher terms, requiring a careful tracking of the different variants of an expression (e.g. *political atmosphere, political instability, political turmoil, political uncertainty*, etc.).

¹⁹[Table XL](#) reports the full vocabulary.

TABLE XVIII: Lexicon summary

Category	Total number	Examples
Contagion	64	regional crisis, crisis in the region, spillovers from the global crisis, systemic crisis ...
Political	60	political turmoil, internal security situation, political atmosphere, political crisis ...
Expectations	60	crisis risks, market reversal, economic sentiment remains poor, market sentiment has collapsed ...
BoP	54	Shortage of foreign exchange, bop crisis, balance of payment crisis, capital account crisis ...
Epidemics	42	epidemic, epidemia, pandemia, pandemic ...
Sovereign	41	rescheduled debt, external payments crisis, difficulties in servicing its external debt, difficult time in rolling over its debt ...
Commodity	40	oil crisis, rice crisis, crop crisis, crop failure ...
Banking	38	bank resolution, bank crisis, Banking sector restructuring, restructuring of nonperforming loans ...
Housing	35	home prices have been declining, drops in real estate prices, house price trends, home-price overvaluation ...
Sev. Recession	34	severe economic crisis, very difficult economic circumstances, Severe recession, severe crisis ...
Wars	28	war damage, insurgency crisis, security crisis, civil conflict ...
World	28	world-wide recession, global economic crisis, global crisis, world recession ...
Soft recession	28	slowdown in the economic activity, slowdown in economic growth, slowdown of the economy, slowdown of output ...
Inflation	26	inflation pressure, inflationary pressure, high inflation, high rate of inflation ...
Trade	26	trade war, trade policy tension, trade tension, trade conflict ...
Financial	20	financial stability crisis, international monetary crisis, crisis in financial market, financial risks ...
Currency	15	exchange rate crisis, large real depreciation, foreign exchange crisis, severe disruption of exchange markets ...
Nat. disaster	14	flood, drought, rainfall, torrential rains ...
Social	13	social risk, social strain, social turmoil, social disruption ...
Migration	7	refugee, migrant, inward migration, population inflow ...

Note: Authors' own elaboration.

In the selection of the categories, we gave particular attention to cover both economic and non-economic crises, domestic and non-domestic, in the real and the financial sector. Although the macroeconomic literature has devoted much less attention to the specific role of political crises, epidemics, violent conflicts, social tensions or migration outcomes, we deem these events of great importance in the general macroeconomic dynamics.²⁰ The categories include real domestic perturbations such as economic slowdowns and recessions, but also financial market related disruptions e.g. financial crises, currency crises and banking crises. Nonetheless, identi-

²⁰The Covid-19 crisis showed how little prepared are traditional macroeconomic models to deal with such events and fostered a new literature seeking to integrate epidemiologic and macroeconomic models (Bodenstein et al., 2020; Kaplan et al., 2020; Martin S. Eichenbaum et al., 2020).

fying the nature of the crisis is not always clear-cut. Since the keyword approach made further refinements difficult, we had to accept this ambiguity for a number of crises: for example, the category contagion refers to instances of both trade and/or financial contagion.²¹ Similarly, a number of authors in the sovereign crisis literature have isolated precisely the domestic instances of default from external ones ((Bordo et al., 2000; Reinhart and Rogoff, 2009)) but this distinction is difficult to implement with the keyword approach employed in this paper.

The present lexicon includes different nuances of intensity within the same category. For instance, the Contagion category includes different shades of external crisis: either contagion from a crisis in other countries ("*crisis spillovers*", "*crisis in the region*", etc.), slowdown of activity in partners ("*regional situation turned adverse*", "*adverse exogenous events*", etc.) or regionally located crisis ("*Mexican crisis*", "*Crisis in Argentina*", "*2002 crisis*", "*Asian crisis*", etc.).²² The only categories for which a clear separation between expressions referring to moderate intensity events and large collapses was necessary are those related to output dynamics. The first category, Soft Recession, refers to the "peak to trough" moment in the business cycle and denotes the regime shifting from expansion to recession ("*slowdown in economic activity*", "*economic decline*", "*slowdown of output*", etc.), a sluggish recovery ("*low rate of economic growth*", "*activity remains weak*", "*sluggish recovery*", etc.) or a moderate contraction ("*contraction of output*",

²¹See Fratzscher (2003) for a review.

²²World-wide crises have their dedicated category to capture contagion and shocks not regionally located, but concerning major economies ("*World-wide recession*", "*international crisis*", "*turbulence in international markets*", "*worsening international environment*", "*ongoing global downturn*", etc)

"decline in economic activity", "output is estimated to have contracted", etc.). In contrast, the category Severe recession includes instances of effective output collapses i.e. deep troughs of the business cycle (*"severe economic crisis", "sharp decline in output", "deep recession", "very difficult economic circumstances"*). Non-fundamental drivers of economic activity and crises have been largely documented in the macroeconomic literature: business confidence, panics, euphoria or adverse expectations are well-known drivers of economic activity (e.g. [Akerloff and Shiller \(2009\)](#)). To capture these non-fundamental drivers, we include a specific category, Expectations, that includes the perception of general risk (*"crisis risk", "potential risk", "upward risk", etc*) the shift in expectations from economic agents (*"market confidence", "general uncertainty", "uncertainty among market participants", etc*), clear non-fundamental crises (*"self-fulfilling crisis", "speculative attacks"*) and call for forward guidance (*"signals to market", "bolster confidence", "restore market confidence", etc*).

The lexicon is highly dependent on the corpus on which it is applied. In the present case, omissions of some categories may not denote the absence of an event, but rather an irrelevance for macroeconomic stability with respect to the main stream economic models: not all non-conventional crises produce potential or realized consequences for economic activity. They may thus not enter as a risk in the outlook produced by Fund staff.²³ The lexicon for these categories produces an interesting sub-sample of the more economically relevant occurrences with a cross-country comparability that is often missing in such areas.

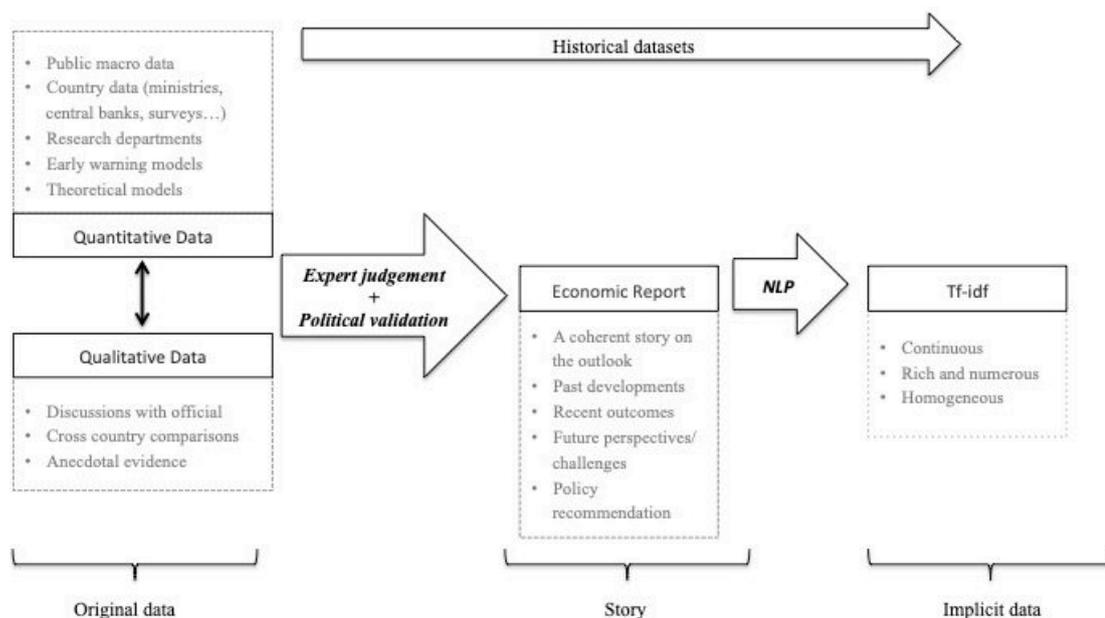
²³Non-conventional in the sense of macroeconomic theory.

5.2.3 The Extraction Method

The calculations of the indices used throughout the paper follows a term-frequency approach. After compiling the lexicon for each category, we count the number of times each document contains keywords belonging to a single category and divide by the total number of characters in the text: in this way, we gauge the importance of the given category in the specific document. Formally, we define the term frequency of document i for category j : $tf_{ij} = \frac{N_{ij}}{N_i}$. In practice, this process of term matching depends on the pre-processing method chosen: in particular, the unit of tokenization can be different.²⁴ We decide to tokenize the documents by sentence to not impose any predetermined length on the keywords we will search for.

²⁴A token is an instance of a sequence of characters in some particular document that are grouped together as a useful semantic unit for processing.

Figure 27: Sketch of the Data Generating Process



Note: The diagram shows the whole process of text analysis in a stylized way. IMF reports are a combination of quantitative and qualitative data concerning different events that are properly ranked in terms of importance and then transformed into words. The purpose of the NLP method is to extract from the story these initial data.

The choice of this naive approach rather than more advanced text analysis techniques stems from the nature of our corpus and the research question raised in this paper. While in a large number of corpora there is little information on the actual content of the documents, in this case we have already a predefined list of the main topics we seek. Thus, looking at the whole

distribution of words and, in particular, frequent words is less informative than focusing on the lower frequencies of specific expressions.

The implicit process of data construction is summarized in [Figure 27](#) that lays out the different components and the sequence leading to the organisation of the original data into economic reports and ultimately, the transformation into quantitative indices. The data presented in this paper is more numerous than those of standard datasets not only in the scope of crises covered, but also in the quality and number of sources underlying the expert judgement: the Fund has real-time access to a vast amount of information and resources that comprises both quantitative and qualitative data, public and private. These sources are compiled, analyzed and summarized by “experts” explicitly in charge of rendering coherent facts on the economic outlooks and undergoing events.

The narrative nature is the main characteristic of these new data. While it grants us a clear advantage with respect to the country and time coverage as well as the richness of information available, it may raise doubts about the correspondence between Fund discussions and reality because of potential editorial biases and methodological shortcomings. The first concern stems from the stability of the methodology and the constancy of the concept of crisis over time. For instance, a banking crisis identified in 1970 might not be the same as one detected in 2012. This critic is, however, not specific to narrative data, but applies also to the national accounts, financial records and accounting rules; any work that covers a long time span will be subject to these statistical discrepancies across time and it is not clear whether in this respect a non-narrative methodology would be superior. If anything, the interesting feature of a text-based

indicator is that it is much more resilient to structural breaks.²⁵ A second issue is raised by the role of member countries in the drafting of Fund documents: as long as they have some input in the preparation, a worsening economic outlook may not be revealed to avoid self-fulfilling spells. This, in turn, would introduce a discrepancy between the real and the text outlook. Nevertheless, we do not believe this to be an alarming concern: major economic and non-economic events are discussed thoroughly in these documents.²⁶ Although suspicions of systematic biases in Fund activity may not be necessarily unfounded (e.g. [Barro and Lee \(2005\)](#); [Dreher et al. \(2009\)](#); [Hernandez \(2020\)](#)), we believe them to be limited to the choice and design of intervention in member countries through programs rather than in the surveillance of countries.

Last but not least, text analysis and large scale data transformation may lead to important noise and potential Type 1 and Type 2 errors, hence undermining the validity of the data.²⁷ In this specific case, the errors derive from the complexity of language and semantics that may not be perfectly captured through a predefined lexicon. Yet, given the length of the documents, the degree of detail and the emphasis on risky outcomes, the erroneous assessment from misleading sentences is unlikely to change dramatically the information conveyed by our indices. While a false positive will definitely produce a non-zero term-frequency, it is unlikely that in the same

²⁵For instance, the IMF staff knows how to interpret differently a current account deficit depending on the exchange-rate regime and capital mobility regime.

²⁶See [Romer and Romer \(2017\)](#) for the same point on the *OECD Economic Outlook*.

²⁷Detecting discussions relative to a category where there is none and the other way around.

document these errors are repeated often enough to blur reality: by the same token, in the case of adverse outcomes, several sentences will be devoted to its description and assessment.²⁸

5.3 A Database on Crises Discussion

The empirical literature on macroeconomic crises is vast and characterized by a large variety of identification methods (mainly, non-parametric, parametric and qualitative), data sources, country and time coverage, frequency and features of the resulting data (binary or continuous measures) as well as focuses (e.g. from real activity drops to exchange rate crashes). We first shortly review the literature on macroeconomic crises, discuss the definitions and main features of the benchmark data-sets for some key indices (sovereign default and economic recession) and then, in the first part of the section, compare them to our term-frequencies, highlighting the correspondence between the two and detailing the major differences.²⁹³⁰ In the second part of the section, instead, we zoom in on the behaviour of the narrative indicators for non-economic crises and assert their relevance to understand the economic outlook of individual countries.

Identifying economic crises, defined as a drop in domestic output benefit from the compiling of long GDP series for most countries and several proxies for periods prior to the compilation of standardized national accounts (e.g. [Feenstra et al. \(2015\)](#)). In addition, specific work on the

²⁸Furthermore, to ensure soundness, in the next section we validate our term-frequencies against standard benchmarks found in the literature.

²⁹For a comprehensive literature review on macroeconomic crises databases, see [Vannier \(2020\)](#).

³⁰Even though, for feasibility reasons, we constrain the comparison to some key indices, most arguments we put forward apply to all categories.

dating procedure in the spirit of [Harding and Pagan \(2002\)](#) and [Kose et al. \(2020\)](#) have provided convincing dating of individual and global economic downturns. For financial crises the literature combined long, qualitative, narrative studies ([Kindleberger \(1975\)](#), [Diaz-Alejandro \(1985\)](#)) with more rigorous, quantitative investigations. Among the latter, [Reinhart and Rogoff \(2009\)](#) stand out as the most comprehensive source of information on the timing of financial crises: the authors date for 68 countries, mostly advanced economies and major emerging markets, going back to the nineteenth century, sovereign defaults, inflation spikes, exchange rate crashes, stock market bursts and banking crises. Similarly, [Laeven and Valencia \(2013\)](#) expand the country coverage to 160 countries focusing on a shorter time period (1970-2017) and identifying, specifically, episodes of systemic banking crises. Finally, [Beers and Mavalwalla \(2017\)](#) report a continuous measure of sovereign crises: for all countries that underwent a default in the period 1960-2016 the authors provide an estimation of the amount of outstanding debt in default.

The current vintage of the crisis discussion database contains 20 variables, one for each category of the lexicon, where each variable corresponds to the term-frequency computed as detailed in section [5.2.3](#). [Table XIX](#) provides a classification of the term-frequencies across different dimensions.

TABLE XIX: Typology of Crises

	Variable	Type	Nature
1	Banking crisis	Economic	Financial
2	Financial crisis	Economic	Financial
3	Inflation crisis	Economic	Real
4	Trade crisis	Economic	Real
5	World crisis	Economic	Real and financial
6	Contagion crisis	Economic	Real and financial
7	Expectation crisis	Economic	Financial
8	Housing crisis	Economic	Real
9	B.o.P. crisis	Economic	Financial
10	Currency crisis	Economic	Financial
11	Eco. recession	Economic	Real
12	Eco. slowdown	Economic	Real
13	Sovereign crisis	Economic	Financial
14	Violence crisis	Non economic	Real
15	Nat. disaster	Non economic	Real
16	Commodity crisis	Non economic	Real
17	Political crisis	Non economic	Real
18	Social crisis	Non economic	Real
19	Epidemic crisis	Non economic	Real
20	Migration crisis	Non economic	Real

Note: Authors' own elaboration.

Since the frequency of reports' publication is uneven, with several documents per year in most cases, to harmonize across countries and period we aggregate the term-frequencies at the yearly frequency.³¹ This aggregation brings the final dataset to 7,788 observations distributed across 181 countries over the period 1950-2019. The current vintage of the database can be downloaded here ([link](#)).

5.3.1 Comparison with Benchmarks: Stylized Facts

The comparison of our database with standard benchmarks of macroeconomic crises requires us to take heed of some important *a priori* differences between the aforementioned. These differences, in turn, help to explain discrepancies between the measures, that may not necessarily reflect contradictions, but rather convey different information.

First, a fundamental difference between our term-frequencies and the benchmark originates in their distinct nature. While we aim at capturing crises discussions by the Fund, the literature has usually tried to pinpoint their exact timing. The paramount example is that of policy reactions: if policymakers intervene effectively in a country experiencing debt distress, ultimately avoiding any missing payment or rescheduling, the country/date would not appear in standard sovereign crises data sets. Nevertheless, the debt problem, most probably, has received considerable attention by Fund staff and is thoroughly discussed in their reports. For example, the euro-area sovereign crisis, while generating acute pressure on the debt of several

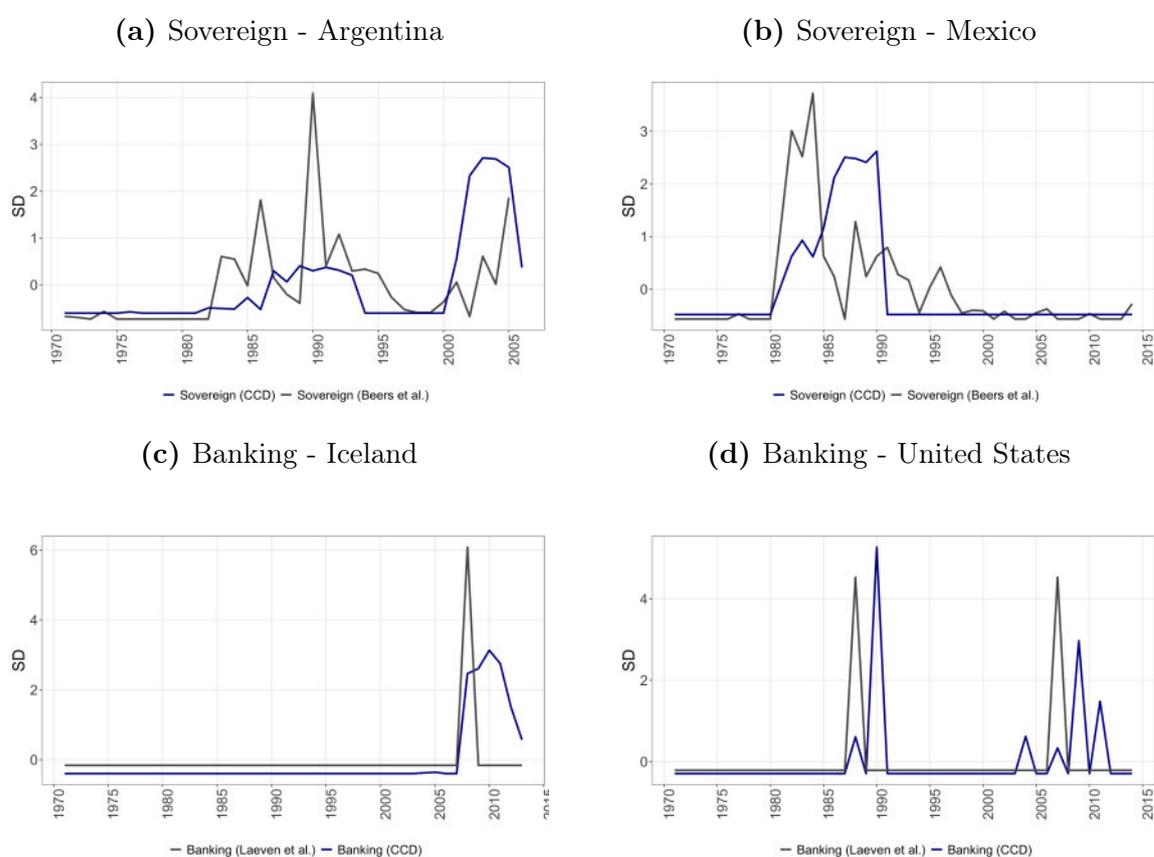
³¹This aggregation ensures a homogeneous frequency and is more suitable for quantitative analysis: however, it comes at a cost of a big information loss. The use of complementary information will be explored in further versions of the data-set.

countries (Greece, Spain, Italy, Portugal), did not materialize in widespread defaults (excepted for Greece). Other similar cases are the Tequila crisis in Mexico (1994-1995), when the IMF's and FED's external assistance compensated the rollover pressure experienced by the country, and Italy during the collapse of the European Exchange Rate Mechanism, when the speculative attack that weighted on the value of the Lira was fueled by high levels of public debt and rising interest rates. The same argument applies also to crises that have been often foretold, but have never materialized e.g. the United States balance of payments crisis. Furthermore, even if the comparison database does not try to pinpoint the exact timing of a crisis, but conveys another type of information, there still might be differences in the nature of the measure: for example, while both our sovereign term-frequency and the [Beers and Mavalwalla \(2017\)](#) index are continuous measures of default intensity, they lend themselves to different interpretations: the former will tend to peak earlier than the latter, with the most acute moment of discussion likely to be anterior to the default itself (legal procedure and agreement).

Second, one must be cautious about the underlying features of the data. Most of the crises databases employ a binary measure rather than a continuous one to indicate the first year of occurrence of a crisis: in addition, the authors usually decide a time window after the first signal is issued to avoid capturing multiple instances of the same episode. For instance, [Reinhart and Rogoff \(2013\)](#) define currency crises as years of exchange rate devaluations higher than 15%. After the first year, signals are filtered over a 3-year windowss, with all positive signals muted. [Laeven and Valencia \(2013\)](#) use a similar methodology (30% depreciation threshold) with a 5-year window. It follows that a mechanical comparison of our term-frequencies with databases

identifying the start of a crisis would result in unsatisfactory low correlation values. [Figure 28](#) illustrates the points mentioned above.

Figure 28: Difference with Benchmark - Sovereign and Banking



Note: In the first two panels, the grey line represents the normalized amount of debt in default or restructuring from [Beers and Mavalwalla \(2017\)](#), the blue line is the sovereign term-frequency. In the second panel, the grey line represents a dummy variable for the first year of banking crisis from [Laeven and Valencia \(2013\)](#), the blue line shows the banking term-frequency.

The first panel of the figure compares our sovereign term-frequency to [Beers and Mavalwalla \(2017\)](#) for Argentina and Mexico. For Argentina, in the early 1980s, concerns about default rose, sharply starting from 1982, with local peaks in 1983, 1986 and 1990. Nevertheless, the default becomes effective only from 1987 on-wards: the Argentinean case exemplifies a slow-moving default, anticipated years before its occurrence. For Mexico, in 1982, the outlook is very different, with the term-frequency and the amount in default peaking up simultaneously with little anticipation: while the term-frequency spikes in the first years of the default, capturing widespread Fund discussion, the actual amount, however, reaches its maximum only 4 to 5 years later.³² The second panel, instead, compares the banking term-frequency to the banking crises start dates from [Laeven and Valencia \(2013\)](#) for Iceland and the United States. In both countries, banking crisis are sudden and unexpected, but their resolution is more gradual. Simple correlation would yield low correspondence irrespective of the validity of the two approaches.

While perfect matching of our index and the benchmark is neither possible nor relevant, a certain degree of correspondence is nonetheless necessary to validate our approach. [Table XX](#) shows the confusion matrix between the [Beers and Mavalwalla \(2017\)](#) and the sovereign term-frequency for different income groups.³³

³²In addition, as mentioned earlier, while for the Tequila crisis in 1995 our term-frequency reaches almost one standard deviation, the [Beers and Mavalwalla \(2017\)](#) remains flat.

³³We convert both indices to binary measures before the comparison to obviate their different meaning.

TABLE XX: Confusion Matrix: Sovereign discussion and Default

Crisis	Default	All	High income	Middle income	Low income	Obs
0	0	34.88%	65.73%	24.88%	11.62%	2324
0	1	11.26%	1.28%	15.12%	16.98%	750
1	0	17.63%	27.16%	16.53%	4.86%	1175
1	1	36.23%	5.83%	43.47%	66.53%	2414

Note: Default is defined as 1 if the amount of debt in default from [Beers and Mavalwalla \(2017\)](#) is strictly positive. Crisis is equal to 1 if the sovereign term-frequency is strictly positive.

Pooling across all countries, 71% of the country/years observations display a correspondence between the two measures: the result is homogeneous across income groups, although slightly higher for low income countries (78%). Among high income countries, two third of the sample correspond to normal times, defined as periods with neither discussion nor default on debt. The lion's share of of the mismatch come from the occurrence of default discussion without an effective default (18%): as explained earlier, this result mostly stems from the nature of our term-frequency that also captures latent episodes of default.

To check the time stability of the lexicon, in [Table XXI](#) we also compare the sovereign debt term-frequency with the Beers & Maravella over different time chunks . The proportion of correctly classified periods hovers stable around 80%.³⁴

TABLE XXI: Confusion Matrix: Sovereign Discussion and Default Over Time

Crisis	Default	1950-1976	1976-1992	1992-2003	2012-2019
0	0	40.8%	38.2%	37.2%	20.5%
0	0	5.6%	11.3%	9.5%	15.5%
1	1	12.4%	12.7%	14%	24%
1	1	41.2%	37.8%	39.3%	40%

Note: Default is defined as 1 if the amount of debt in default from [Beers and Mavalwalla \(2017\)](#) is strictly positive. Crisis is equal to 1 if the sovereign term-frequency is strictly positive.

5.3.2 Comparison with Benchmarks: Econometric Estimation

To better understand what our term-frequencies capture and how it relates to several measure of recessions and default, we run different regressions on the sovereign default and economic recessions benchmarks. [Table XXII](#) presents the result for our measure of economic recession.

³⁴The only doubt is for the recent period where the proportion of fake alarms (positive sovereign term frequency and zero amount of debt rescheduled) rises considerably: it could indicate a change in the structure of the reports over the last ten years.

TABLE XXII: Comparison with Benchmark: Severe Recession

	<i>Dependent variable:</i>						
	g	g	g<0	g<0	Phase B	Large B	Phase B2
Y	-0.19 (0.20)	-0.31** (0.14)					
Y>0			0.07*** (0.02)	0.13*** (0.02)	-0.01 (0.02)	0.02* (0.01)	0.03 (0.02)
Constant	3.97*** (0.45)			0.06*** (0.01)			
Country FE	No	Yes	Yes	No	Yes	Yes	Yes
Time FE	No	Yes	Yes	No	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes	Yes
Robust se	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F	0.97	3.13***	9.17***	50.15***	3.01***	2.47***	3.61***
Observations	2,061	2,060	2,061	2,061	2,098	2,021	2,098

Note: Y is the severe recession term-frequency. G corresponds to real GDP growth rate. Y>0 and g<0 are dummies equal to 1 when the condition is satisfied. Phase B, Large B and Phase B2 are dummies indicating the cyclical component of real GDP obtained following [Harding and Pagan \(2002\)](#). Phase B is equal to 1 for all the years in between the peak and the trough of the cycle. Large B indicates the downturn phases with the largest amplitude. Phase B2 refers to the second half of the downturn. GDP data are from the Analytical Database of the OECD. ***: significant at 1% level, **: significant at 5% level, *: significant at 10% level.

First, we regress the normalized severe recession term-frequency (Y) on the growth rate of real GDP (g) (column 1 and 2). The results show that the term-frequency is significantly negatively correlated with the growth rate of real GDP after controlling for the other term-frequencies, country and time fixed effects (column 2). For robustness, we regress the term-frequency on a dummy variable equal to 1 when the real GDP growth rate falls below -1%: across both specifications (column 3 and 4), the correlation is positive and significant. Lastly, to show that our measure captures effectively the occurrence of particularly severe economic recessions rather than slowdowns, we observe the correlation between our term-frequency and different dummies indicating the state of the business cycle: Phase B, equal to 1 for all the years in between the peak and the trough of the cycle, Large B, indicating the downturn phases with the largest amplitude and Phase B2, referring solely to the second half of the downturn.³⁵ The term-frequency does not significantly correlate with the peak to trough, but only with the most severe slumps in economic activity (column 5, 6 and 7).³⁶ Over the sample of countries and time periods where both output measures and the severe recession term-frequency are available we observe that the narrative indicator significantly matches the economic outcome. Moreover,

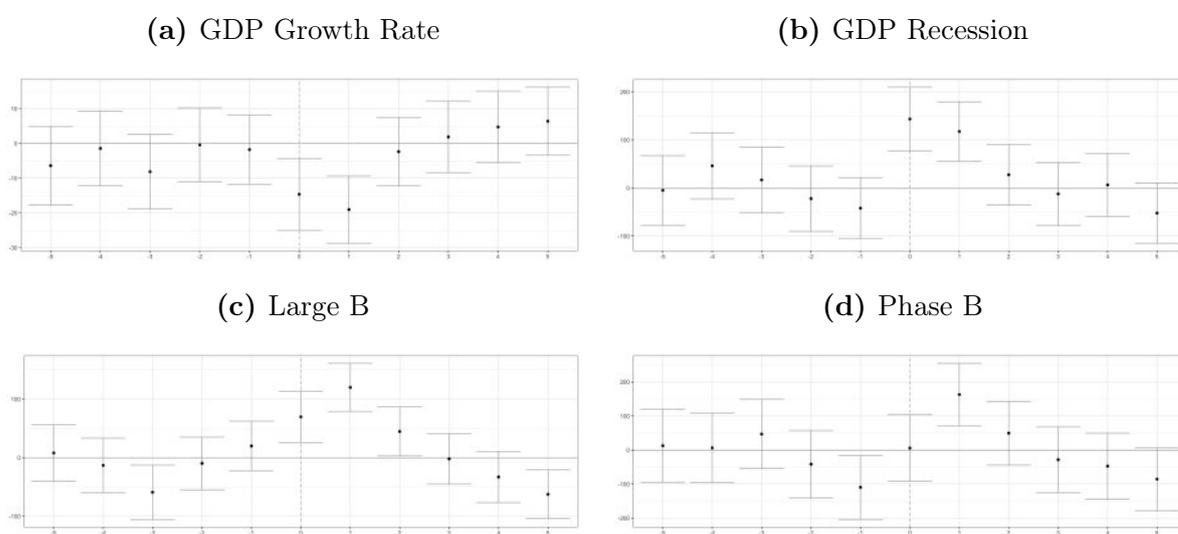
³⁵For a visual representation of the [Harding and Pagan \(2002\)](#) algorithm, see [Figure 43](#).

³⁶[Table XLI](#) shows similar specifications for the soft recession term-frequency that should, instead, captures the slowdown of economic activity. While the index is correlated with the both the GDP growth rate and the recession dummy, it is no longer associated with large downturns (Phase B) but only with the second part of the downturn (Phase B2).

we confirm that the latter captures the difference between slowdowns and particularly dire recessions.

To enhance the understanding of the timing of our narrative indicator, we also regress the different output measures on the lags and leads of the severe recession term-frequency (Figure 29). We find that output measures are significantly correlated and with the expected sign contemporaneously and for the following one to two years, without any evidence of anticipated warning. This results highlights the real time and backward looking nature of our term-frequencies.

Figure 29: Contemporaneous, Backward or Forward looking?



Note: The figure displays estimates obtained regressing the different output measures on five lags and leads of the severe recession term-frequency. Estimates are computed controlling for all other term-frequencies, country and time fixed effects.

We perform a similar validation for the sovereign term-frequency with respect to the benchmark measure of sovereign default. [Table XXIII](#) shows the results for the different specifications. We first regress the term-frequency (S) on the amount of debt in default ($D.Default$) [unit] (column 1 and 2): we find that a one standard deviation increase in the term-frequency is unconditionally associated to an increase of 225.39 [unit] of debt in default (column 1), with the result robust to different controls (column 2). Similarly, we show the correlation between a dummy for a strictly positive term-frequency ($S > 0$) and the probability of default ($D.Default > 0$): a one standard deviation increase in the term-frequency of crises increases the probability of default by 7%.

TABLE XXIII: Comparison with Benchmark: Sovereign

	<i>Dependent variable:</i>				
	D.Default	D.Default	D.Default>0	D.Default>0	D.Default>0
S	222.72*** (46.29)	167.77*** (39.23)			0.07*** (0.01)
S>0			0.44*** (0.02)	0.11*** (0.02)	
Constant	362.11*** (73.79)		0.22*** (0.02)		
Country FE	No	Yes	No	Yes	Yes
Time FE	No	Yes	No	Yes	Yes
Robust se	Yes	Yes	Yes	Yes	Yes
F	23.15***	1.85*	388.9***	16.63***	16.79***
Observations	6,680	6,678	6,677	6,677	6,678

Note: S is the sovereign term-frequency. D.Default is the amount of debt in default or rescheduling from [Beers and Mavalwalla \(2017\)](#). S>0 and D.Default>0 are dummies equal to 1 when the respective variable is strictly positive. ***: significant at 1% level, **: significant at 5% level, *: significant at 10% level

All in all, although comparison with other crises database is complicated by differences in the nature and features of our term-frequencies and that perfect matching between the two

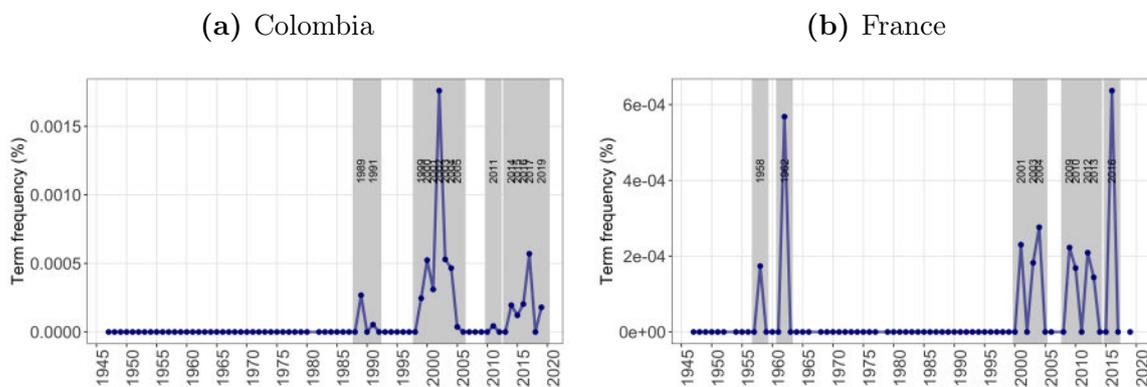
is neither possible nor relevant, we find that the economic recession and sovereign narrative indicators are highly correlated with their respective benchmark and that they constitute a real-time and backward looking economic assessments of the country's outlook rather than a forward looking measure.

5.3.3 Zeroing in on Non-Economic Indicators

The main contribution of the crises discussion database is to provide an extended, comprehensive and comparable set of narrative indicators also for non-economic crises. While detailed data covering specific non-economic events have been already made available (e.g. [Global Terrorism Database](#) (GTD), [EM-DAT](#) (The Emergency Events Database)), they often incorporate different countries and time periods, hence lacking comparability. Moreover, rather than a proxy for the intensity of the event *per se*, differently from the aforementioned, our indicators signal the relevance of the event for the economic outlook of the country under scrutiny.³⁷

[Figure 30](#) provides an illustrative example of the term-frequency for the Violence category.

³⁷In the present case, omissions of some categories may not denote the absence of an event, but rather an irrelevance for macroeconomic stability: not all non-conventional crises produce potential or realized consequences for economic activity. Thus, they may not enter in Fund staff discussions.

Figure 30: Examples of Violence Indicator

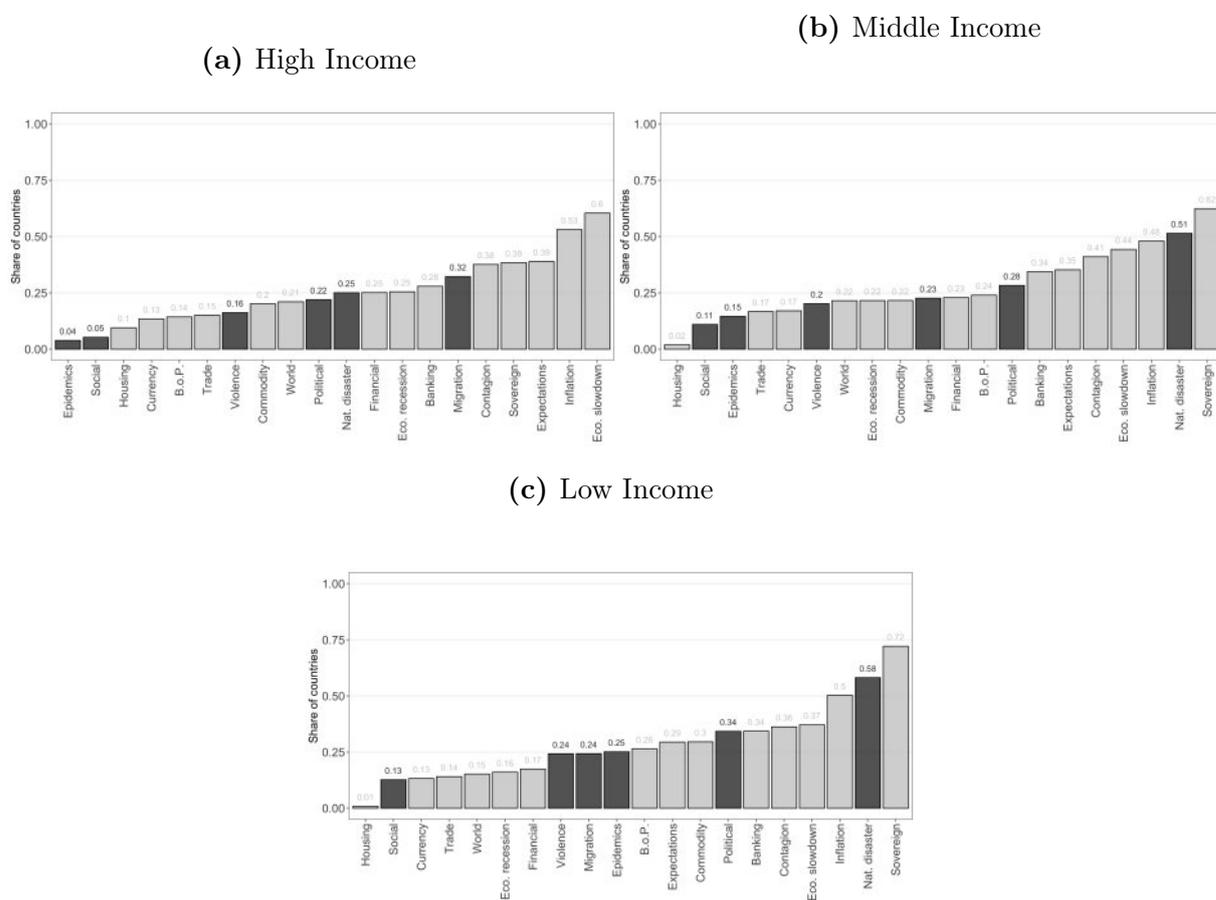
Note: The blue line corresponds to the violence term-frequency for, respectively, Colombia and France. Shaded gray areas are years of strictly positive term-frequency.

Although the conflict between the Colombian government and the Revolutionary Armed Forces of Colombia (FARC) as well as other guerilla forces started in 1960, it intensified in the mid-1990s as a consequence of the higher wealth accumulated by terrorist groups through drug-related activities. The indicator peaks again in 2017 when the peace referendum between the government and FARC rebels failed as the “No” gained the majority (Figure 30a). For France, the indicator peaks at the end 1950s-early 1960s, in correspondence of the Algerian war, and then shows a turbulent behaviour throughout the XXI^{th} century when the country experienced

a series of violent terrorist attacks: it peaks again in 2016 after the November 2015 *Bataclan* attack (Figure 30b).³⁸

While a strict focus on non-economic crises is rarely part of the macroeconomic academic debate, that usually emphasizes purely financial outcomes, Fund staff discusses frequently these topics in their reports and especially so for precise income groups (Figure 31). For high income countries, migration issues appear in more than one fourth of total annual observations (32%), followed by natural disasters (25%) and political crises (22%). For middle and low income countries, natural disasters are discussed habitually: more than half of total annual observations (51% and 58% respectively), becoming the second most widely considered issue. Similarly, political instability occupies a greater role in lower income groups (28% and 34%). Further, epidemics and violence issues also fill up a substantial part of the analysis in low income countries (roughly 25%).

³⁸For an event study of the natural disaster indicator, see Figure 44.

Figure 31: Non-economic crises: Breadth of Discussion

Note: The bars denote the unconditional frequencies of the occurrence of crises discussions. Formally, it is the proportion of periods with strictly positive term-frequency. Dark-gray bars represent non-economic categories.

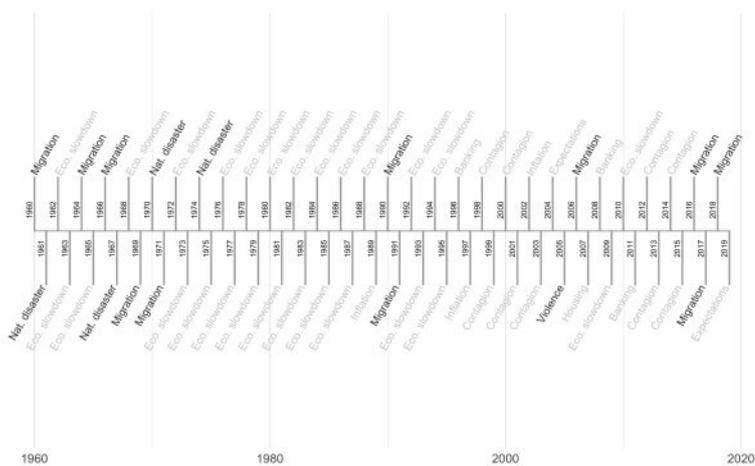
One might argue that, however, breadth does not coincide with depth of topic discussion: Fund staff might still discuss some categories in a shallow way, but do so in every report. If

this is the case, non-economic events may be less critical for the economic outlook of a country than we are trying to argue. In [Figure 32](#) we compute for each year of the sample the mean of the 20 term-frequencies, pooling across different countries, and then report the category corresponding to the highest value: the pattern that emerges contrasts for different income groups. While for high income countries non-economic events are rarely the most discussed category, outmatched by financial outcomes and output slowdowns ([Figure 32a](#)), for low income countries, non-economic events and specifically natural disasters are covered painstakingly in Fund publications ([Figure 32b](#)).³⁹

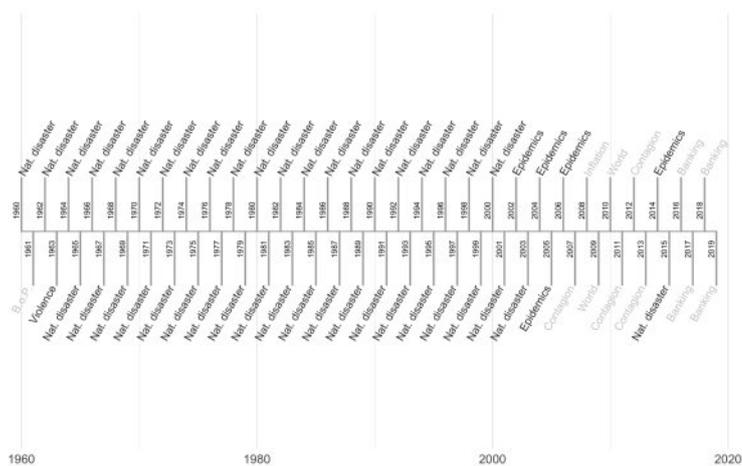
³⁹Another interesting pattern is that for high income countries, when non-economic events are the year priority, the relevant category is migration crises: for instance, migration concerns were at the forefront of the economic discourse following the fall of the Berlin wall and the uncertainty regarding the stability of the the URSS entering the 1990s.

Figure 32: Non-economic crises: Depth of Discussion

(a) High Income



(b) Low Income



Note: The figures displays, for each year, the category with the highest unconditional mean between the 20 term-frequencies. Dark-gray labels represent non-economic categories.

In short, although non-economic events usually fall into the cracks of the macroeconomic discourse, they can be pivotal to comprehend the economic outlook of a country and in particular, of middle and low-income groups. The richness of this database and the potential for rapid and flexible extension constitute an element of novelty in the empirical crises literature.⁴⁰

5.4 Looking Back: A Rising Complexity

The new material presented so far allows us to adopt a holistic stance toward macroeconomic crises and explore a dimension, complexity, left, because of the data limitations we already discussed, relatively untouched by previous empirical literature. In this section, borrowing from network theory (Jackson, 2010) and looking through the Fund lens, we highlight some compelling patterns in the evolution of the “crisis system” over the last 70 years. In particular, we underline the rising co-occurrence of crises and the exponential rise in importance of the non-fundamental *expectations* channel.

Figure 33 presents a visual representation of the evolution of the “crisis system”, constructed considering each category in the database as a node in the network and the contemporaneous correlations across term-frequencies as their edges.⁴¹ The resulting pattern provides prima facie evidence on the rising complexity of the system: links between term-frequencies have both

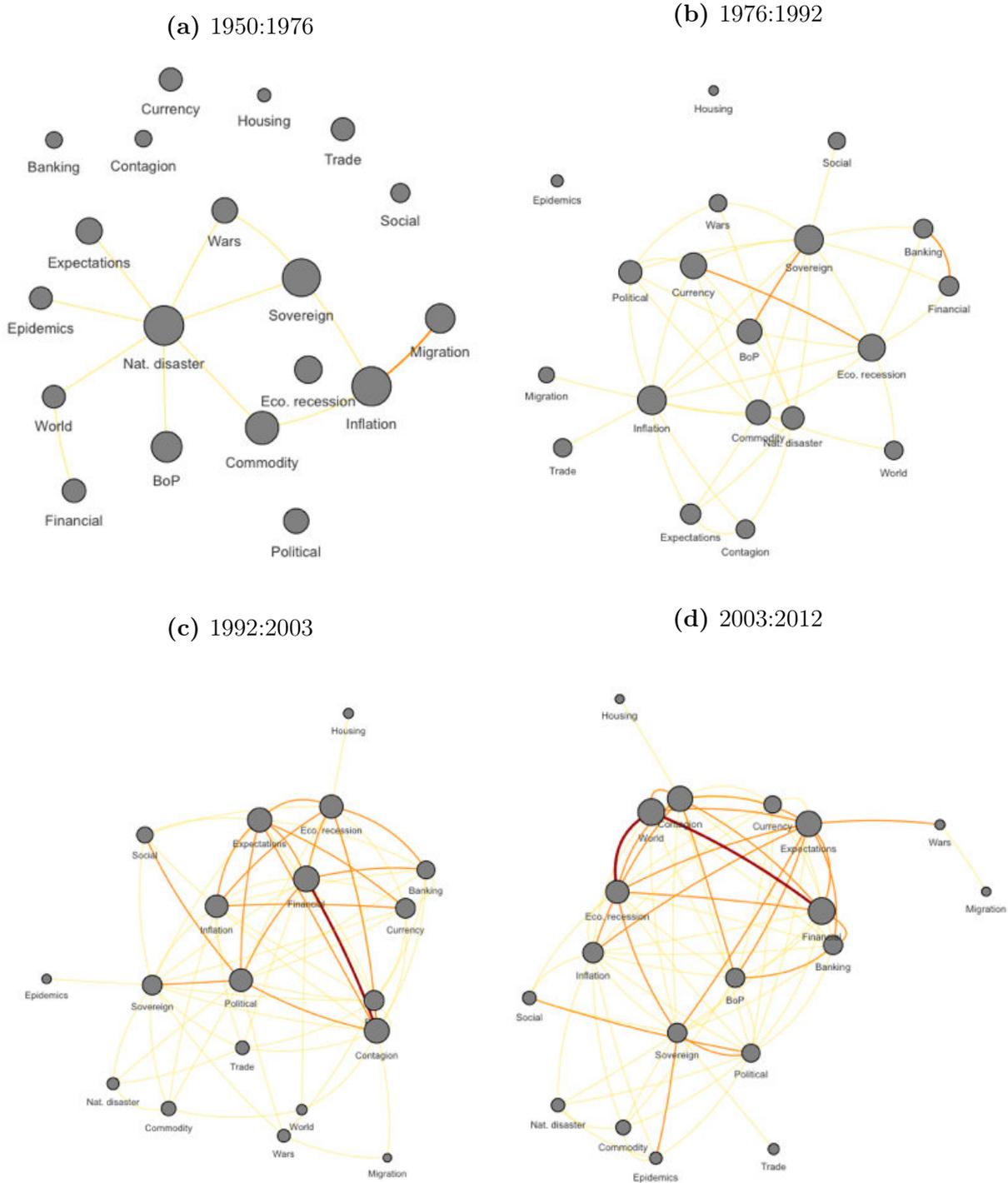
⁴⁰Other types of non-economic events might be of utter importance in the future e.g. cyber attacks. In light of this, it is essential to have a framework that is easily expandable.

⁴¹The division in time bins is based on previous knowledge and corresponds to well-known events: the Bretton Woods system, the first wave of financial globalization, the second wave of financial globalization, the run-up to the GFC and the recent post-GFC period.

become more frequent and thicker, increasing approximately 6-fold. During the earliest period (1950-1976), the network is relatively sparse and mainly organized around real crises, natural disasters and inflation crises in primis (Figure 33a). Figure 33b portrays the initial stretch after the collapse of the Bretton-Woods system: numerous novel connections appear and several sub-networks emerge. While the sub-network around natural disaster persists as well as the one around inflation, a new cluster around sovereign crises appears. Among the others, we note the strong interconnection of sovereign and balance of payment crises, currency crashes and deep economic recessions and banking and financial turmoil. The early 1980s clearly stand out as a period of structural change from a shallow system to a deeper network where financially related elements gradually take over.⁴² This trend consolidates at the end of the century: this decade is characterized by the persistence of a cluster around sovereign crises, the move to the periphery of the natural disaster node and the clear emergence of a financial *clique* composed by contagion, financial turmoil, banking crises and expectations. Finally, starting from the 21st century, all nodes become connected to the network, with the *clique* between the financial components of the systems consolidating while real shocks move further away to the periphery.

⁴²See [Diaz-Alejandro \(1985\)](#) for a detailed chronicle of the first wave of financial globalization and deregulation.

Figure 33: A Network of Economic Complexity



(e) 2012:2019



Note: Adjacency matrix built from pairwise correlations between term-frequencies: minimum correlation to display edge equal to 0.1. Size of nodes proportional to their eigencentrality. Legend indicates correlations between categories. Visualization of the network through the ForceAtlas2 algorithm (Jacomy et al., 2014).

Table XXIV summarizes the previous visual observation calculating the average shortest path by time period.⁴³ We find a roughly 3-fold reduction in the shortest path that holds for all income groups: we confirm that the overall evolution of the system concerns all income groups and underline the profound and lasting structural shift toward a more dense and financially

⁴³The average shortest path of a network indicates how far, on average, are all pairs of nodes based on the geodesic distance (i.e. shortest path): thus, a lower shortest path indicates that perturbations transmit more rapidly across the network

dominated system. Similarly to the international financial network (Haldane, 2009), the “crises system” displays increasingly the features of a “small world” where particular disturbances spread quickly across the whole system.

TABLE XXIV: Average Shortest Path

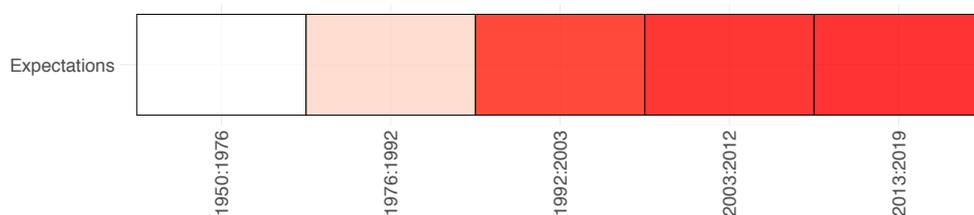
Income Group	Min. Corr	1950:1976	1976:1992	1992:2003	2003:2013	2013:2019
High income	0.2	17.78	16.15	13.49	13.35	5.48
Low income	0.2	-	17.37	13.28	7.27	5.54
Upper middle income	0.2	18.39	13.43	9.8	11.18	4.02

Note: Average shortest path is the mean shortest distance (number of links) between any single pair of nodes. Adjacency matrix built from pairwise correlations between term-frequencies. Minimum correlation indicates that pairwise correlations lower than the respective value are set equal to 0 when building the adjacency matrix. If two nodes are not connected, their shortest distance is set equal to the number of nodes in the network. Algorithm does not converge for Low Income countries in the period 1950-1976 and is replaced by missing value.

Lastly, we focus on a specific category, expectations, and track its behaviour over time within the system. This category is of particular interest given its the only one characterized by a non-fundamental attribute. Moreover, while the role of manias and panics has been deemed central for the unraveling of macroeconomic crises by different strands of the narrative literature (e.g. Kindleberger (1975), Akerloff and Shiller (2009)), its actual contribution has been difficult to quantify given the intrinsic challenge in measuring this channel and comparing it to the fundamental one. Figure 34 shows the evolution of the eigenvector centrality for the expectations category over the different time periods: while expectations are mostly peripheral until the early

1990s, they gain an increasingly prominent role in the last 30 years, thus confirming their present key role as complexifying element.⁴⁴

Figure 34: Centrality of Expectations Channel



Note: The eigenvector centrality of a node is the associated i_{th} element of the eigenvector with the largest eigenvalue for the given adjacency matrix. It indicates the importance of a node based on the number of connections it has with other “important” (well connected) nodes. Scales of red indicate the eigenvector centrality during a precise time period, where a brighter red indicates higher eigenvector centrality. The adjacency matrix is built from the correlation matrix of all categories within the period under consideration.

⁴⁴The basic idea of eigenvector centrality is that a node importance is not only determined by the fact that the node is directly connected to many other nodes, but also by whether or not it is connected to well-connected nodes. For a more thorough explanation, see [Jackson \(2010\)](#).

5.5 Conclusion

In this paper, we provide accessibility to researchers to a raw text database of roughly 23,000 documents covering the whole IMF membership throughout the period 1950-2019, building on and improving significantly over the closest paper in the field (Mihalyi and Mate, 2019). Moreover, to capture and quantify Fund discussions about a multiplicity of adverse economic and non economic events, we manually compile an IMF crisis-specific dictionary and propose a simple term-frequency approach. The large time span (70 years) and country coverage (181 countries) of the resulting database as well as the scope of crises covered within a comparable framework complement and extend standard datasets of macroeconomic crises and provide useful material for a deeper understanding of macroeconomic volatility episodes.

Comparing some key economic indicators (severe recession and sovereign) to standard benchmarks found in the literature, we confirm that the term-frequencies constitute an accurate real-time and backward looking economic assessment of the countries' outlook. In addition, we show that, while non economic events tend to fall into the cracks of the academic macroeconomic discourse, they occupy a substantial amount of discussion in Fund reports for all income groups and especially so for lower income groups. Finally, exploiting the vast amount of data at our disposal, we introduce the notion of crises complexity, defining it as the co-occurrence of crises. We study its evolution within the last 70 years and find that it has risen considerably: in particular, the system underwent a clear structural break starting from the early 1980s, shaping from a simple network dominated by real crises in the Bretton Woods era to a highly complex,

financially dominated one, in the recent post-GFC period. Within the system, we highlight the rise in centrality of the non-fundamental expectations channel.

Taken altogether, these last findings have far-reaching implications for domestic policymakers and Lenders of Last Resort (LOLR) institutions. First, they highlight the intrinsic difficulty in forecasting. While the economic system can remain stable for long periods, a small perturbation can spread quickly across sectors and breed into complex outcomes. Instead of point estimates, it would be better to provide alternative scenarios based on an assessment of emerging threats to systemic stability. Second, the rising complexity of crises calls for an adequately diversified program toolbox available from LOLR institutions: while the IMF has recently expanded its emergency facilities in face of the Covid-19 pandemic, the question of adequacy remains open for future research and discussion. Similarly, the rising centrality of the non-fundamental channel calls for a heavy focus on the managing of expectations by policymakers.

CHAPTER 6

CONCLUSION

This dissertation focused on the relationship between economic forecasting, financial crises, and rising macroeconomic complexity. In particular, the three main chapters delved into (I) the prediction of a specific type of financial crises, sudden stops of foreign inflows, in EM economies, (II) the evaluation of short-term forecasts of economic activity produced by the main economic institutions and private sector and (III) the analysis of macroeconomic crises as complex episodes brought to surface by the entangling of different economic and non-economic individual events.

I find that (I) although forecasting with certainty the precise timing of a sudden stop is virtually impossible, detecting the underlying economic weaknesses that usually mark its outbreak is a viable undertaking. In line with [Gourinchas and Obstfeld \(2012\)](#), I show that few domestic indicators, such as the overvaluation of the exchange rate and buoyant private credit, have a high predictive power for sudden stops. Moreover, combining two different strands of the international finance literature, I demonstrate that the probability of a sudden stop is highly correlated with the ensuing output loss: this result strongly validates the adoption of this class of models for policy-making purposes.

We show that (II) the accuracy of forecasts of economic activity strongly hinges on the state of the country's business cycle: while forecasts for expansions periods are unbiased and the

underlying uncertainty contained, during recessions projections are nearly always optimistic, exhibiting high variance. This result is robust to the forecaster chosen. We also find that the largest errors correspond to periods of complex crises when real and multiple financial events entwine. Lastly, we analyze the political economy dimensions of forecast errors and IMF programs and do not find evidence of a deliberate bias.

In the end, we contribute (III) to the understanding of macroeconomic crises providing a new text-based dataset to researchers covering bilateral surveillance archival IMF documents. We design and implement a simple term-frequency approach to capture discussions about a large variety of crises events and exploit its comprehensiveness to study the correlation between them. We show that the “crisis system” has significantly complexified over the last 70 years, becoming both more dense and formed by stronger connections.

The themes discussed in this dissertation and the relative results will direct my future research agenda. In Chapter 3 we employed a Logit model for forecasting purposes. While this class of models ensures better causal interpretation of the covariates and allows us to compare the relative contribution of local and global indicators, in the recent years new forecasting methods from the field of Machine Learning have been applied extensively for purely forecasting purposes, especially in the prediction of banking crises (Beutel et al., 2018; Fouliard et al., 2020), with encouraging results. Moreover, while we focused solely on the de-leveraging by foreign investors, a potentially interesting endeavor is the analysis of its antecedent. How do the length

and size of the foreign-fueled boom correlate to the severity of the resulting sudden stop?

The recent Covid-19 crisis calls for an update and comparison of the results reported in Chapter 4. How did forecasters behave in face of the unprecedented health shock? Do the patterns of forecast errors observed during previous recessions held still? Are economic and non-economic shocks similar in terms of forecasting performance? Does large emergency financing correspond to larger optimistic forecast errors? We will explore all these questions in an upcoming papers with the same coauthors.

At last, in Chapter 5 we furnished new rich material for empirical research and introduced the theme of crises complexity. There is a variety of dimensions in which the dataset could be expanded and harnessed e.g. capturing the sectoral heterogeneity of shocks and revisiting the old question of the Mundell-Fleming model, the relationship between exchange rate regime and crises. Furthermore, in the same spirit of [Reinhart and Rogoff \(2009\)](#), we could work on a summary indicator of crises complexity, with the final purpose to study the relationship between complexity and crises' severity.

Bibliography

- Ahir, Hites, Nicholas Bloom, and Davide Furceri, “The World Uncertainty Index,” SSRN Electronic Journal, 2018.
- Ahuja, Ashvin, Murtaza Syed, and Kevin Wiseman, “Assessing Country Risk— Selected Approaches—Reference Note,” IMF Technical Notes and Manuals, 2017, pp. 1–28.
- Akerlof, George A., “Sins of omission and the practice of economics,” *Journal of Economic Literature*, 2020, 58 (2), 405–418.
- Akerlof, George and Robert Shiller, “Animal spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism,” 2009.
- Aldasoro, Iñaki, Claudio Borio, and Mathias Drehmann, “Early Warning Indicators of Banking Crises: Expanding the Family,” *BIS Quarterly Review*, 2018.
- Alessi, Lucia and Carsten Detken, “‘Real Time’ Early Warning Indicators for Costly Asset Price Boom/Bust Cycles - A role for Global Liquidity,” ECB Working Paper, 2009.
- and —, “Identifying excessive credit growth and leverage,” *Journal of Financial Stability*, 2018.
- An, Zidong, João Tovar Jalles, and Prakash Loungani, “How well do economists forecast recessions?,” *International Finance*, 2018, 21 (2), 100–121.
- Artis, Michael J., “How Accurate Are the Imf’s Short-Term Forecasts? Another Examination of the World Economic Outlook,” *Staff Studies for the WEO*, 1996.

- Avdjiev, Stefan, Bryan Hardy, Sebnem Kalemli-Ozcan, and Luis Servén, “Gross Capital Inflows to Banks, Corporates and Sovereigns,” NBER Working Paper, 2017.
- , Leonardo Gambacorta, Linda S Goldberg, and Stefano Schiaffi, “The Shifting Drivers of Global Liquidity,” BIS Working Papers, 2017.
- Babecký, Jan, Tomáš Havránek, Jakub Matějů, Marek Rusnák, Katerina Smidková, and Vašicek Borek, “Banking, Debt, and Currency Crises Early Warning Indicators for Developed Countries,” ECB Working Paper, 2012.
- Baker, Scott R., Bloom, Nicholas, and Steven J. Davis, “Measuring Economic Policy Uncertainty,” *The Quarterly Journal of Economics*, 2016, 131 (4), 0–52.
- Baqir, Reza, Rodney Ramcharan, and Ratna Sahay, “IMF programs and growth: Is optimism defensible?,” *IMF Staff Papers*, 2005, 52 (2), 260–286.
- Barro, Robert J. and Jong Wha Lee, “IMF programs: Who is chosen and what are the effects?,” *Journal of Monetary Economics*, 2005, 52 (7), 1245–1269.
- Basu, Suman S, Roberto A Perrelli, and Weining Xin, “External Crisis Prediction Using Machine Learning : Evidence from Three Decades of Crises Around the World,” IMF Working Paper, 2019.
- Batchelor, Roy, “Bias in macroeconomic forecasts,” *International Journal of Forecasting*, 2007, 23 (2), 189–203.
- Beach, W., “How Reliable Are IMF Economic Forecasts?,” Heritage Foundation Report, 1999.

- Beaudry, Paul, Tim Willems, Matthieu Bellon, Paolo Cavallino, Mark Flanagan, Klaus Hellwig, and Jean-Paul L 'huillier, "On the Macroeconomic Consequences of Over-Optimism," *American Economic Journal: Macroeconomics*, 2021.
- Beers, David T. and Jamshid Mavalwalla, "Database of Sovereign Defaults, 2017," Bank of Canada Technical report, 2017, (101).
- Berg, Andrew and Catherine Patillo, "Are Currency Crises Predictable? A Test," *IMF Staff Papers*, 1999, 42 (2), 107–138.
- , Eduardo Borensztein, and Catherine A. Pattillo, "Assessing Early Warning Systems: How Have they Worked in Practice?," *Imf Staff Papers*, 2005.
- Beutel, Johannes, Sophia List, and Gregor Von Schweinitz, "An evaluation of early warning models for systemic banking crises: Does machine learning improve predictions?," *Bundesbank Discussion Paper*, 2018.
- Blanchard, Olivier J, Hamid Faruquee, and Mitali Das, "The Initial Impact of the Crisis on Emerging Market Countries," *Brookings Papers on Economic Activity*, 2010.
- Bodenstein, Martin, Giancarlo Corsetti, and Luca Guerrieri, "Social Distancing and Supply Disruptions in a Pandemic," *Finance and Economics Discussion Series*, 2020, 2020 (031).
- Bordo, Michael, Barry Eichengreen, Daniela Klingebiel, and Maria Soledad Martinez-peria, "Is the Crisis Problem Growing More Severe?," *Economic Policy*, 2000.
- Broner, Fernando, Tatiana Didier, Aitor Erce, and Sergio L Schmukler, "Gross Capital Flows: Dynamics and Crises," *Journal of Monetary Economics*, 2013, 60 (1), 113–133.

Bussière, M and M Fratzscher, “Towards a new early warning system of financial crises,” *Journal of International Money and Finance*, 2006.

Bussière, Matthieu, “Balance of Payment Crises in Emerging Markets : How Early Were the "Early" Warning Signals?,” *ECB Working Paper*, 2007.

—, “In Defense Of Early Warning Signals,” *Banque de France Working Paper*, 2013.

—, “For a Few Dollars More: Reserves and Growth in Times of Crises,” 2014.

Calderón, César and J. Rodrigo Fuentes, “Have business cycles changed over the last two decades? An empirical investigation,” *Journal of Development Economics*, 2014, 109, 98–123.

Calvo, GA, A Izquierdo, and LF Mejia, “On the empirics of sudden stops: the relevance of balance-sheet effects,” *NBER Working Paper*, 2004.

Caramazza, Francesco, Luca Ricci, and Ranil Salgado, “Trade and Financial Contagion in Currency Crises,” *IMF Working Paper*, 2000.

Catão, Luis A V and Gian Maria Milesi-Ferretti, “External Liabilities and Crises,” *Journal of International Economics*, 2014.

Cerutti, Eugenio, S. Claessens, and Andrew K. Rose, “How Important is the Global Financial Cycle? Evidence from capital flows,” *BIS Working Papers*, 2017.

Chinn, Menzie D. and Hiro Ito, “A New Measure of Financial Openness,” *Journal of Comparative Policy Analysis: Research and Practice*, 2008.

- Choi, Hyunyoung and Hai Varian, "Predicting the Present with Google Trends," *The Economic Record*, 2012, pp. 193–204.
- Clark, Todd and Michael McCracken, "Advances in forecast evaluation," *Handbook of Economic Forecasting*, 2013, 2, 1107–1201.
- Comelli, Fabio, "Comparing parametric and non-parametric early warning systems for currency crises in emerging market economies," *Review of International Economics*, 2014.
- , "Estimation and out-of-sample Prediction of Sudden Stops: Do Regions of Emerging Markets Behave Differently from Each Other?," IMF Working Paper, 2015.
- Diaz-Alejandro, Carlos, "Good-bye financial repression, hello financial crash," *Journal of Development Economics*, 1985, 19 (1-2), 1–24.
- Diaz-Alejandro, Carlos F., "Latin American Debt: I Don't Think We are in Kansas Anymore," *Brookings Papers on Economic Activity*, 1984, 1984 (2), 335.
- Diebold, Francis X. and Jose A. Lopez, "Forecast evaluation and Combination," *Handbook of Statistics*, 1996, 14, 241–268.
- Dovern, Jonas and Nils Jannsen, "Systematic errors in growth expectations over the business cycle," *International Journal of Forecasting*, 2017, 33 (4), 760–769.
- Dreher, Axel, Jan Egbert Sturm, and James Raymond Vreeland, "Development aid and international politics: Does membership on the UN Security Council influence World Bank decisions?," *Journal of Development Economics*, 2009, 88 (1), 1–18.

- Drehmann, Mathias, Claudio Borio, and Kostas Tsatsaronis, “Characterising the Financial Cycle: Don’t Lose Sight of the Medium Term!,” BIS Working Papers, 2012.
- Duca, Marco Lo and Tuomas A. Peltonen, “Assessing systemic risks and predicting systemic events,” *Journal of Banking and Finance*, 2013.
- Eichengreen, B and P Gupta, “Managing sudden stops,” World Bank Policy Research Working Paper, 2016.
- Eichengreen, Barry, Poonam Gupta, and Oliver Masetti, “Are Capital Flows Fickle? Increasingly? and Does the Answer Still Depend on Type?,” *Asian Economic Papers*, 2018.
- EM-DAT (The Emergency Events Database), “Université catholique de Louvain (UCL) - CRED.”
- Engle, Robert F., Stefano Giglio, Bryan Kelly, Heebum Lee, and Johannes Stroebel, “Hedging climate change news,” *Review of Financial Studies*, 2020, 33 (3), 1184–1216.
- Fayad, Ghada, Chengyu Huang, Yoko Shibuya, and Peng Zhao, “How Do Member Countries Receive IMF Policy Advice: Results from a State-of-the-art Sentiment Index,” IMF Working Paper, 2020.
- Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer, “The Next Generation of the Penn World Table,” *American Economic Review*, 2015, 105 (10), 3150–3182.
- Fernández, Andrés, Michael W Klein, Alessandro Rebucci, Martin Schindler, and Martín Uribe, “Capital Control Measures: A New Dataset,” IMF Working Paper, 2015.

- Forbes, Kristin J and Francis E Warnock, “Capital Flow Waves: Surges, Stops, Flight, and Retrenchment,” *Journal of International Economics*, 2012.
- Fouliard, Jeremy, Michael Howell, and H el ene Rey, “Answering the Queen: Machine Learning and Financial Crises,” 2020.
- Fraiberger, Samuel P., “News Sentiment and Cross-Country Fluctuations,” *SSRN Electronic Journal*, 2016, pp. 125–131.
- Frankel, Jeffrey, “Over-optimism in forecasts by official budget agencies and its implications,” *Oxford Review of Economic Policy*, 2011, 27 (4), 536–562.
- Frankel, Jeffrey A. and Andrew K. Rose, “Currency Crashes in Emerging Markets: an Empirical Treatment,” *International Finance Discussion Papers*, 1996.
- Frankel, Jeffrey and George Saravelos, “Can leading indicators assess country vulnerability ? Evidence from the 2008 – 09 global financial crisis ,” *Journal of International Economics*, 2012, 87 (2), 216–231.
- Fratzscher, Marcel, “On currency crises and contagion,” *International Journal of Finance and Economics*, 2003.
- , “Push versus pull factors and the global financial crisis,” *Journal of International Economics*, 2012.
- and Julien Reynaud, “IMF surveillance and financial markets-A political economy analysis,” *European Journal of Political Economy*, 2011, 27 (3), 405–422.

- Fritz-Krockow, Bernhard and Ramlogan Parmeshwar, *International Monetary Fund Handbook: Its Functions, Policies, and Operations* 2007.
- García, Diego, “Sentiment during Recessions,” *Journal of Finance*, 2013, 68 (3), 1267–1300.
- Genberg, Hans and Andrew Martinez, “On the Accuracy and Efficiency of IMF Forecasts: A Survey and Some Extensions,” IEO Background Paper, 2014.
- Gentzkow, Matthew, Bryan Kelly, and Matt Taddy, “Text as data,” *Journal of Economic Literature*, 2019, 57 (3), 535–574.
- Ghirelli, Corinna, Javier J. Pérez, and Alberto Urtaun, “A new economic policy uncertainty index for Spain,” *Economics Letters*, 2019, 182, 64–67.
- Global Terrorism Database (GTD), “National Consortium for the Study of Terrorism and Responses to Terrorism (START),” 2017.
- Gourinchas, Pierre-Olivier and Maurice Obstfeld, “Stories of the Twentieth Century for the Twenty-First,” *American Economic Journal: Macroeconomics*, 2012.
- Haldane, Andrew, “Rethinking the Financial Network,” Speech at the Financial Student Association, Amsterdam, 2009.
- Harding, Don and Adrian Pagan, “Dissecting the cycle: A methodological investigation,” *Journal of Monetary Economics*, 2002, 49 (2), 365–381.
- Hernandez, Christian, “IMF flexibility or neoliberal adaptation: A discursive content analysis of Article IV policy biases in Argentina,” *Governance*, 2020, 33 (1), 135–154.

- Holopainen, Markus and Peter Sarlin, “Toward robust early-warning models: a horse race, ensembles and model uncertainty,” ECB Working Paper, 2016.
- Huang, Chengyu, Sean Simpson, Daria Ulybina, and Agustin Roitman, “News-based Sentiment Indicators,” IMF Working Papers, 2019, 19 (273).
- IEO, “IMF Forecasts - Process, Quality, and Country Perspectives,” IEO Evaluation Report, 2014.
- IMF, “2018 Review of Program Design and Conditionality,” IMF Policy Papers, 2019.
- Jackson, Matthew O., *Social and Economic Networks* 2010.
- Jacomy, Mathieu, Tommaso Venturini, Sebastien Heymann, and Mathieu Bastian, “ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software,” PLoS ONE, 2014, 9 (6), 1–12.
- Johnson, Robert C. and Guillermo Noguera, “A portrait of trade in value-added over four decades,” *Review of Economics and Statistics*, 2017, 99 (5), 896–911.
- Kalemli-Ozcan, Sebnem, “Emerging Market Capital Flows under COVID : What to Expect Given What We Know,” IMF Research, 2020.
- Kaminsky, Graciela, “Leading Indicators of Currency Crises,” 1998.
- Kaminsky, Graciela L., “Varieties of currency crises,” *Annals of Economics and Finance*, 2003.
- and Carmen M. Reinhart, “The Twin Crises: The Causes of Banking and Balance-Of-Payments Problems,” *American Economic Review*, 1999.

- Kaplan, Greg, Benjamin Moll, and Giovanni Violante, “The Great Lockdown and the Big Stimulus: Tracing the Pandemic Possibility Frontier for the U.S.,” SSRN Electronic Journal, 2020.
- Kilgarriff, Adam, “Comparing Corpora,” *International Journal of Corpus Linguistics*, 2001, 6 (1), 97–133.
- Kindleberger, Charles P, *Manias, Panics, and Crashes* 1975.
- Kose, M. Ayhan, Naotaka Sugawara, and Marco E Terrones, “Global Recessions,” SSRN Electronic Journal, 2020.
- Laeven, Luc and Faián Valencia, “Systemic banking crises database,” *IMF Economic Review*, 2013.
- and —, “Systemic Banking Crises Revisited,” *IMF Economic Review*, 2018.
- Lane, Philip R. and Gian Maria Milesi-Ferretti, “The external wealth of nations mark II: Revised and extended estimates of foreign assets and liabilities, 1970-2004,” *Journal of International Economics*, 2007, 73 (2), 223–250.
- Levchenko, Andrei A. and Paolo Mauro, “Do some forms of financial flows help protect against "sudden stops"?” *World Bank Economic Review*, 2007.
- Lewis, Christine and Nigel Pain, “Lessons from OECD forecasts during and after the financial crisis,” *OECD Journal: Economic Studies*, 2014, 1, 9–39.
- Loungani, Prakash, “How accurate are private sector forecasts? Cross-country evidence from consensus forecasts of output growth,” *International Journal of Forecasting*, 2001, 17 (3), 419–432.

- Luna, Francesco, "IMF Forecasts in the Context of Program Countries," IEO Background Paper, 2014.
- Manasse, Paolo and Nouriel Roubini, "Rules of thumb" for sovereign debt crises," *Journal of International Economics*, 2009.
- , —, and Axel Schimmelpfennig, "Predicting Sovereign Debt Crises," IMF Working Paper, 2003.
- Martin S. Eichenbaum, Sergio Rebelo, and Mathias Trabandt, "Epidemics in the Neoclassical and New Keynesian Models," NBER Working Paper Series, 2020, p. 36.
- Mihalyi, David and Akos Mate, "Text-mining IMF country reports - an original dataset," MPRA Paper, 2019.
- Mikolov, Thomas, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean, "Distributed Representations of Words and Phrases and their Compositionality," *Proceedings of the 26th International Conference on Neural Information Processing Systems*, 2013.
- Musso, Alberto and Steven Phillips, "Comparing Projections and Outcomes of IMF-Supported Programs," IMF Working Paper, 2002, 49 (1), 22–48.
- Nordhaus, William D., "Forecasting Efficiency: Concepts and Applications," *The Review of Economics and Statistics*, 1987, 69 (4), 667.
- Obstfeld, Maurice, "Rational and Self-Fulfilling Balance-of-Payments Crises," NBER Working Paper, 1986.

- Park, Donghyun, Arief Ramayandi, and Kwanho Shin, “Why Asia Fare Better during the Global Financial Crisis than during the Asian Financial Crisis ?,” in “Responding to Financial Crisis: Lessons from Asia Then, the United States and Europe Now” 2013.
- Reinhart, Carmen, Graciela Kaminsky, and Saul Lizondo, “Leading Indicators of Currency Crises,” Staff Papers (International Monetary Fund), 1998.
- Reinhart, Carmen M., “Financial crises: past and future,” AEI Economics Working Paper, 2018.
- and Christoph Trebesch, “The International Monetary Fund: 70 Years of reinvention,” *Journal of Economic Perspectives*, 2016, 30 (1), 3–28.
- Reinhart, Carmen M and Kenneth S Rogoff, *This Time Is Diferent - Eight Centuries of Financial Folly* 2009.
- Reinhart, Carmen M. and Kenneth S. Rogoff, “Banking Crises: An Equal Opportunity Menace,” *Journal of Banking and Finance*, 2013.
- Rey, Hélène, “Dilemma not Trilemma: The Global Financial Cycle and Monetary Policy Independence,” NBER Working Paper, 2015.
- Rodrik, Dani, Andrés Velasco, and John F Kennedy, “Short Term Capital Flows,” NBER Working Paper, 1999.
- Romer, Christina D. and David H. Romer, “New Evidence On The Impact Of Financial Crises In Advanced Countries,” *American Economic Review*, 2017.
- Rose, Andrew K. and Mark M. Spiegel, “Cross-Country Causes And Consequences Of The 2008 Crisis: Early Warning,” *Global Journal of Economics*, 2010.

- and —, “Cross-country causes and consequences of the crisis: An update,” *European Economic Review*, 2011.
- Sandefur, Justin and Arvind Subramanian, “The IMF’s Growth Forecasts for Poor Countries Don’t Match Its COVID Narrative,” 2020, (CDG Working Paper).
- Scott, Steven L. and Hal R. Varian, “Predicting the present with Bayesian structural time series,” *International Journal of Mathematical Modelling and Numerical Optimisation*, 2014, 5 (1-2), 4–23.
- Shiller, Robert, “Narrative Economics: How Stories Go Viral and Drive Major Events,” 2019.
- Suss, Joel and Henry Treitel, “Predicting bank distress in the UK with machine learning,” *Bank of England, Working Paper*, 2019.
- Takagi, Shinji and Halim Kucur, “Testing the Accuracy of IMF Macroeconomic Forecasts, 1994–2003,” *IEO Background Paper*, 2006.
- Timmermann, Allan, “An evaluation of the World Economic Outlook forecasts,” *IMF Working Paper*, 2007, 54 (1), 1–33.
- Vannier, Brendan Harnoys, “The Four Horsemen of the Apocalypse : New narrative evidence on economic , currency , banking and sovereign crises ,” 2020.
- Òscar Jordà, Moritz Schularick, and Alan M. Taylor, “When credit bites back,” *Journal of Money, Credit and Banking*, 2013, 45 (SUPPL2), 3–28.

APPENDICES

Appendix A

Finding A Needle In A Haystack: Do Early Warning System For Sudden Stops Work?

A.1 Data

TABLE XXV: Countries List

Country	Region
Argentina	Latin America
Bangladesh	Other Emerging Markets
Belarus	EECA
Bolivia (Plurinational State of)	Latin America
Brazil	Latin America
Bulgaria	EECA
Chile	Latin America
Colombia	Latin America
Ecuador	Latin America
El Salvador	Latin America
Guatemala	Latin America

Appendix A (Continued)

Country	Region
Hungary	EECA
India	Other Emerging Markets
Indonesia	East Asia
Kazakhstan	EECA
Malaysia	East Asia
Mexico	Latin America
Pakistan	Other Emerging Markets
Peru	Latin America
Philippines	East Asia
Poland	EECA
Republic of Korea	East Asia
Romania	EECA
Russian Federation	EECA
South Africa	Other Emerging Markets
Sri Lanka	Other Emerging Markets
Thailand	East Asia
Turkey	EECA
Ukraine	EECA
Uruguay	Latin America

Appendix A (Continued)

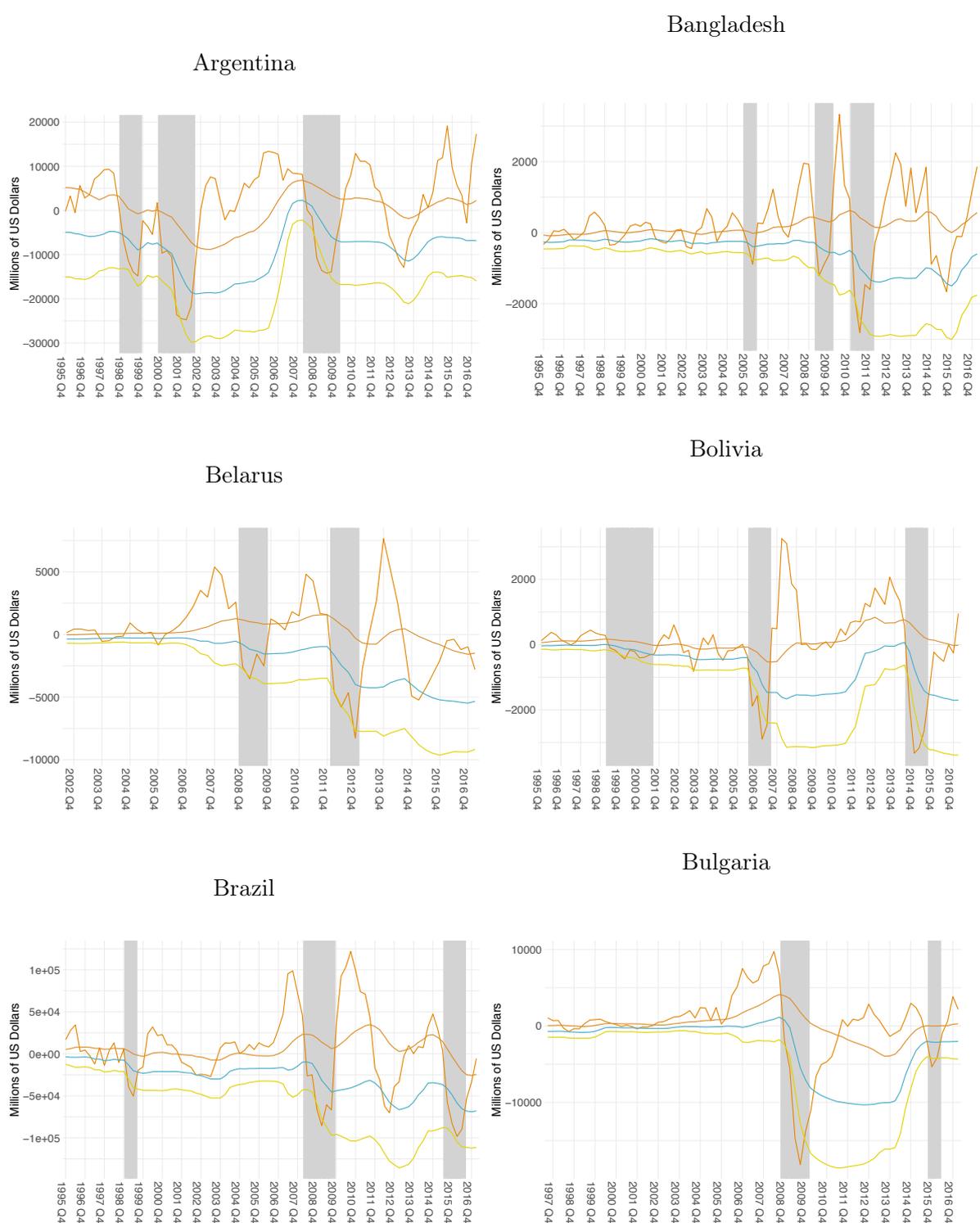
Country	Region
Venezuela, Bolivarian Republic of	Latin America

Appendix A (Continued)

Appendix A (Continued)

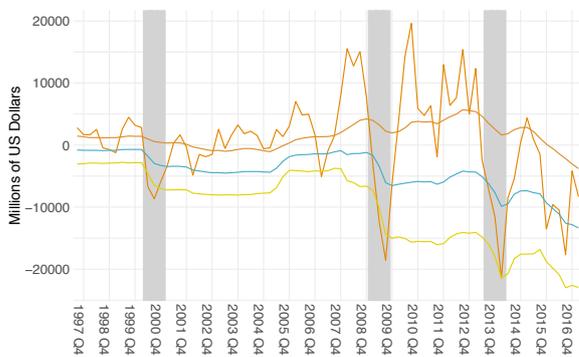
A.1.1 Sudden Stops

Figure 35: Sudden Stops Identification

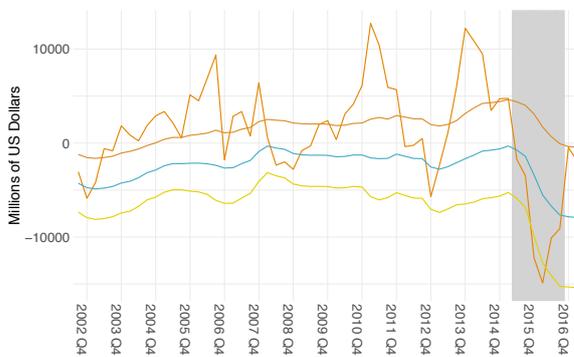


Appendix A (Continued)

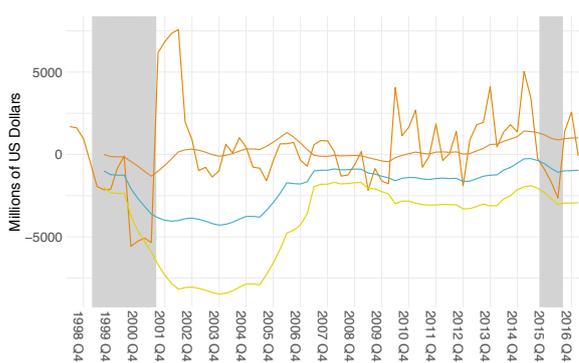
Chile



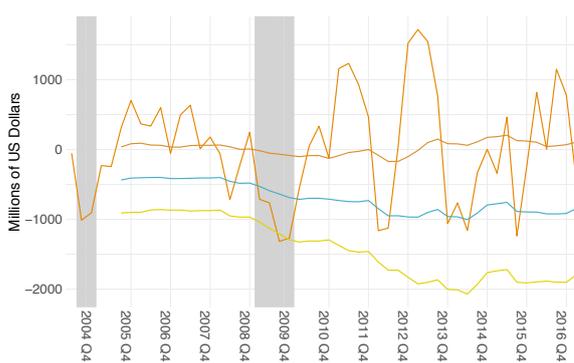
Colombia



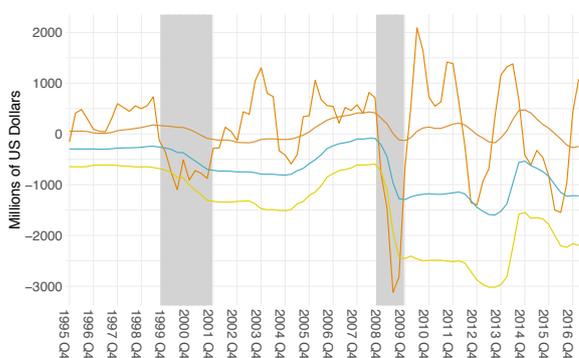
Ecuador



El Salvador



Guatemala



Hungary

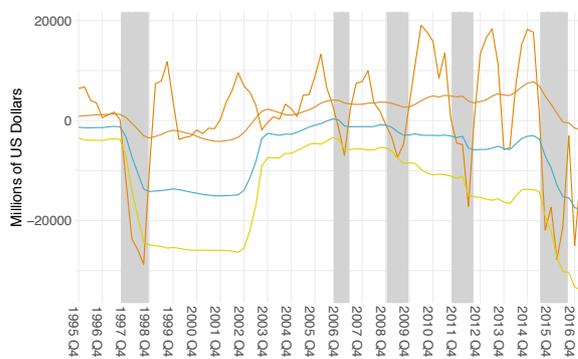


Appendix A (Continued)

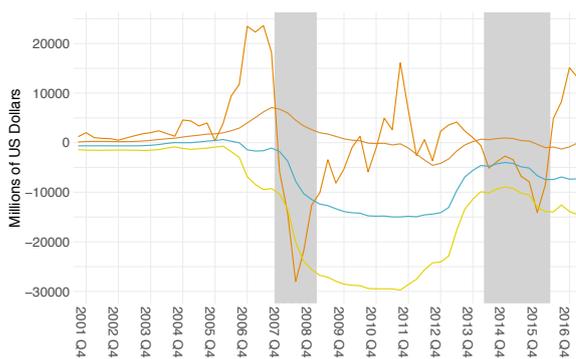
India



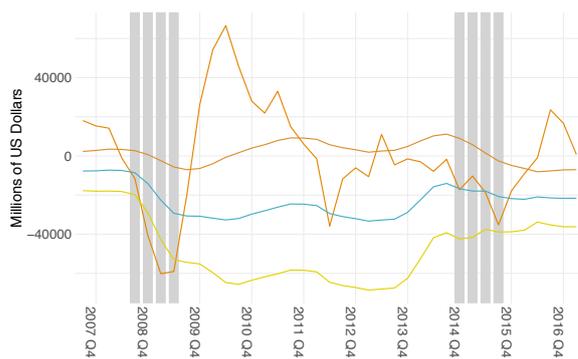
Indonesia



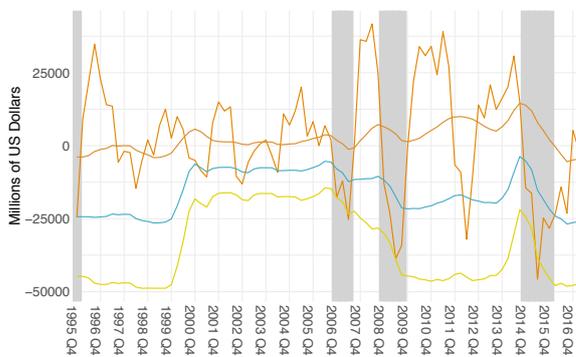
Kazakhstan



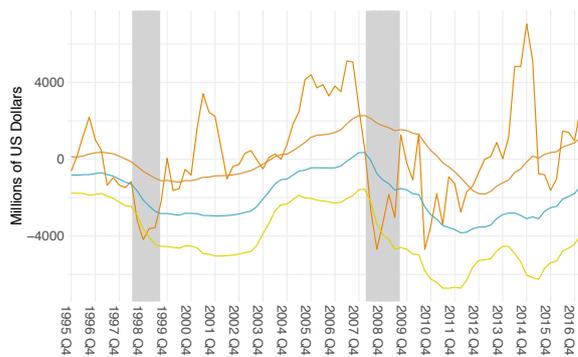
Malaysia



Mexico

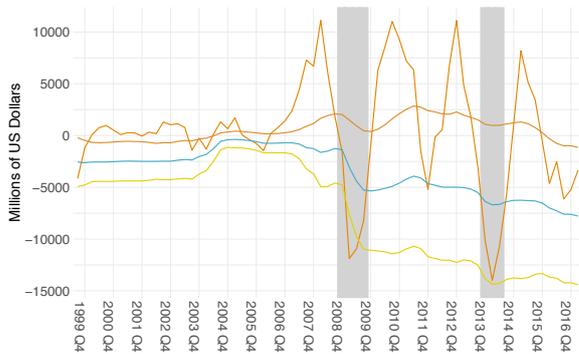


Pakistan

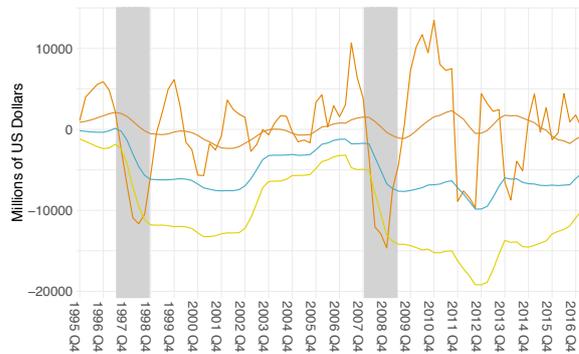


Appendix A (Continued)

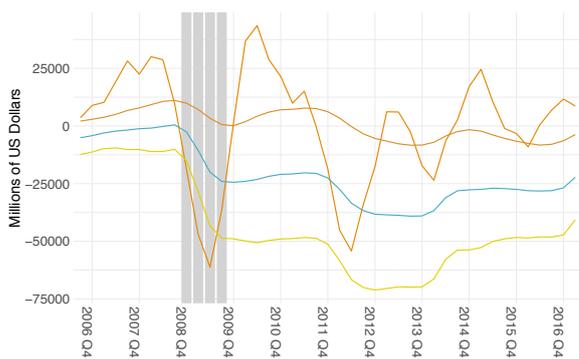
Peru



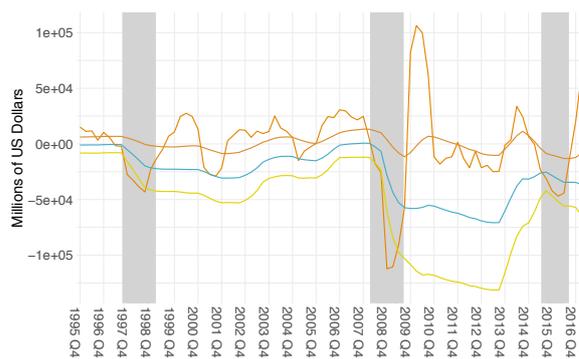
Philippines



Poland



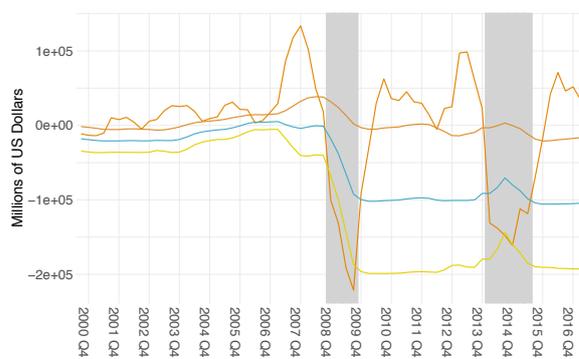
Republic of Korea



Romania

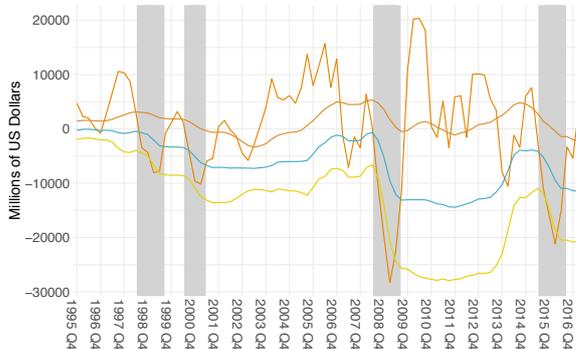


Russian Federation

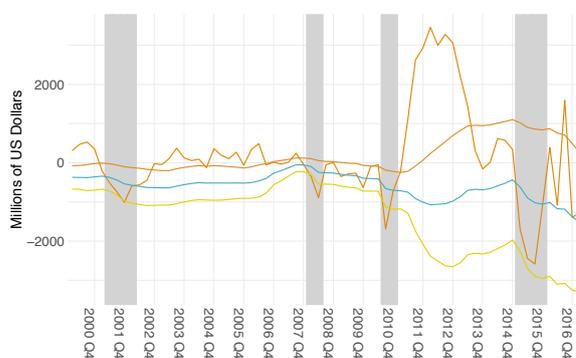


Appendix A (Continued)

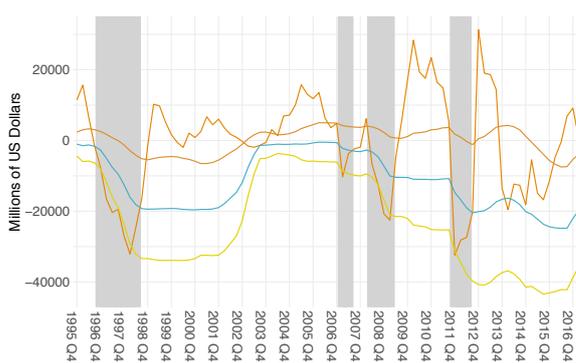
South Africa



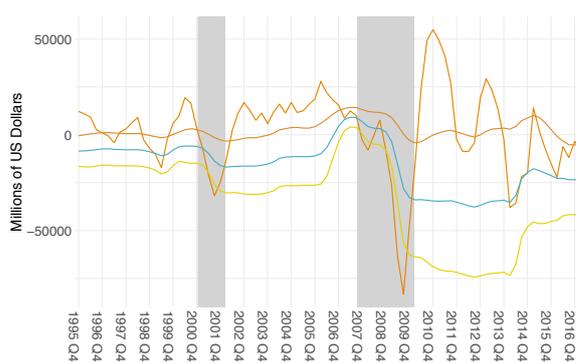
Sri Lanka



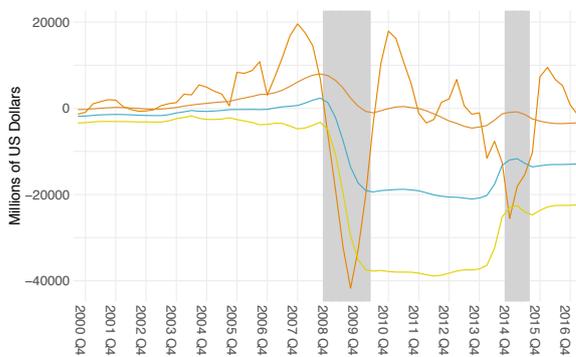
Thailand



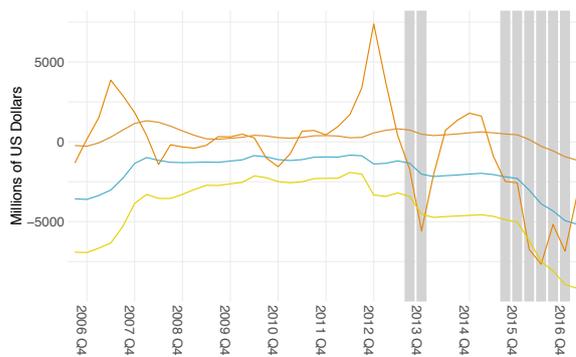
Turkey



Ukraine

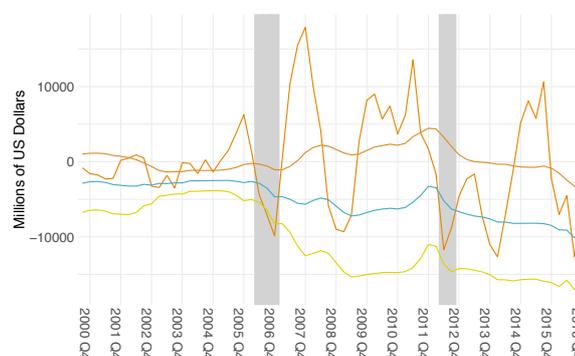


Uruguay



Appendix A (Continued)

Venezuela



Note: The figure shows the algorithm proposed by Forbes and Warnock (2012) for the identification of sudden stops applied to our sample. A sudden stop begins when the y-o-y gross capital inflows (dark orange line) go below their rolling mean minus one standard deviation (light blue line) conditional on crossing the rolling mean minus two standard deviations (yellow line). The episode ends when y-o-y gross inflows come back above their rolling mean minus one standard deviation. The duration is highlighted by the grey shaded area.

TABLE XXVI: List of Sudden Stops

Country	Quarter	Duration (in quarters)
Argentina	1998 Q4	4
Argentina	2000 Q4	7
Argentina	2008 Q2	7
Bangladesh	2005 Q4	2

Appendix A (Continued)

Country	Quarter	Duration (in quarters)
Bangladesh	2009 Q2	3
Bangladesh	2011 Q1	4
Belarus	2008 Q4	4
Belarus	2012 Q1	4
Bolivia (Plurinational State of)	1999 Q2	9
Bolivia (Plurinational State of)	2006 Q3	4
Bolivia (Plurinational State of)	2014 Q3	4
Brazil	1999 Q1	2
Brazil	2008 Q2	6
Brazil	2015 Q3	4
Bulgaria	2008 Q4	5
Bulgaria	2015 Q4	2
Chile	2000 Q2	3
Chile	2009 Q1	3
Chile	2013 Q3	3
Colombia	2015 Q2	6
Ecuador	1999 Q2	9
Ecuador	2015 Q4	3
El Salvador	2004 Q3	2

Appendix A (Continued)

Country	Quarter	Duration (in quarters)
El Salvador	2009 Q1	4
Guatemala	1999 Q4	8
Guatemala	2008 Q4	4
Hungary	2002 Q2	2
Hungary	2009 Q1	5
India	2008 Q3	5
India	2015 Q4	4
Indonesia	1997 Q4	4
Indonesia	2006 Q4	2
Indonesia	2009 Q1	3
Indonesia	2011 Q4	3
Indonesia	2015 Q3	4
Kazakhstan	2007 Q4	5
Kazakhstan	2014 Q2	8
Malaysia	2008 Q3	4
Malaysia	2014 Q4	4
Mexico	2006 Q4	3
Mexico	2008 Q4	4
Mexico	2014 Q4	5

Appendix A (Continued)

Country	Quarter	Duration (in quarters)
Pakistan	1998 Q3	4
Pakistan	2008 Q2	5
Peru	2008 Q4	4
Peru	2013 Q4	3
Philippines	1997 Q3	5
Philippines	2008 Q1	5
Poland	2008 Q4	4
Republic of Korea	1997 Q4	5
Republic of Korea	2008 Q2	5
Republic of Korea	2015 Q3	4
Romania	2008 Q3	7
Russian Federation	2008 Q4	4
Russian Federation	2014 Q1	6
South Africa	1998 Q3	4
South Africa	2000 Q3	3
South Africa	2008 Q3	4
South Africa	2015 Q3	4
Sri Lanka	2001 Q2	4
Sri Lanka	2008 Q1	2

Appendix A (Continued)

Country	Quarter	Duration (in quarters)
Sri Lanka	2010 Q3	2
Sri Lanka	2015 Q1	4
Thailand	1996 Q4	7
Thailand	2007 Q1	2
Thailand	2008 Q2	4
Thailand	2011 Q4	3
Turkey	2001 Q1	4
Turkey	2007 Q4	9
Ukraine	2008 Q4	6
Ukraine	2014 Q4	3
Uruguay	2013 Q3	2
Uruguay	2015 Q3	6
Venezuela, Bolivarian Republic of	2006 Q2	3
Venezuela, Bolivarian Republic of	2012 Q2	2

Appendix A (Continued)

A.1.2 Explanatory Variables and Data Transformation

TABLE XXVII: Raw Data - Description and Sources

Serie	Description	Source
TED Spread	Difference between 3-months USD LIBOR and 3-months T-Bills rate	FRED
VIX	CBOE Volatility Index	FRED
Global Liquidity Growth	Year-on-year and four-years growth rate of global money supply - sum of M2 in United States, euro area and Japan	IFS Statistics
Global Real GDP Growth	Median year-on-year growth rate in the United States, euro area, Japan and UK	IFS Statistics
Global 10-years interest rate	Median yield 10-years government bonds United States, euro area, Japan and UK	IFS Statistics
10-years US interest rate		IFS Statistics
T-bills rate	3-months US Treasury bills rate	FRED
Global Inflation	Median year-on-year CPI inflation in the United States, euro area, Japan and UK	IFS Statistics
Nominal GDP		IFS Statistics and national sources. When not available, interpolated annual from WEO
Real GDP growth	Year-on-year growth rate real GDP	IFS Statistics and national sources. When not available, interpolated annual from WEO
CPI Inflation	Year-on-year growth rate CPI	IFS Statistics
International Reserves	Gold excluded	IFS Statistics
Real Exchange Rate	The bilateral US dollar real exchange rate calculated as the nominal exchange rate against the US dollar times the US CPI and divided by the domestic CPI	
Private Credit	Deposit money banks and other financial institutions claims on private sector	IFS Statistics. To have full sample availability we extend we the non-standardized presentation (line 22d) using the growth rate of the standardized one
ST Liabilities to BIS reporting banks		Joint External Debt Hub (JEDH)
Current Account		IFS Statistics
Capital Controls	Overall restrictions on capital mobility; inflows and outflows specific measures	Different versions: Chinn and Ito (2008), Fernández et al. (2015)
Macprudential Indicators	Loan-to-Value threshold	iMaPP Database
Trade Contagion	$\sum_{i=1}^j \frac{E_{i,j,t} + I_{i,j,t}}{E_{i,t} + I_{i,t}} + SS_{j,t}$	Aggregate and bilateral trade data from Direction of Trade (DOT) Statistics

A.2 Results

A.2.1 Determinants of Sudden Stops

- **De-trending (1):** We replace the HP-filtered RER and private credit over GDP ($\lambda = 1600$) by their year-on-year growth rate. We want to make sure the significance is not only a by-product of

Appendix A (Continued)

the de-trending approach chosen, but that there exists a true relationship between these indicators and dependent variable. The coefficients remain significant and similar in magnitude. The two fit measures, however, are lower than in the benchmark suggesting that the use of an HP filter is better than simple growth rates for forecasting purposes.

- **Short-term indicators (2):** We remove observations up to 1 year before the sudden stop. If some external crises start before our dating or there are some expectation mechanisms at play, the indicators could be affected by endogeneity issues. Results are similar to the benchmark, but with one main difference: trade contagion loses significance. This means contagion variables act as really short-term indicators and may be less useful from a policy-maker perspective.
- **Post-crisis bias (3):** Employing the alternative definition of sudden stop duration we can check for the presence of a post-crisis bias. Since few variables would be affected by the latter in our benchmark specification, we include two other domestic factors, real growth and inflation. The latter have not been found significant with the normal definition, but often appear in EWSs: these two variables, in turn, can be heavily affected by the aforementioned bias. We find, however, no relationship between the two and the dependent variable.
- **Fixed effects (4):** We introduce country fixed effects. This sensitivity check is particularly important for the result on capital controls: if some countries have established a better relationship with markets and this allows them to maintain a high level of capital controls on inflows, and the other way around, the associated coefficient would be downward biased. Nevertheless, the introduction of country dummies does not change the capital controls coefficient and there is no other important coefficient variation. Notice that the fixed effect model shows a better fit than its pooled counterpart. This means country dummies are actually capturing some characteristics

Appendix A (Continued)

correlated with our dependent variable, but the latter are orthogonal to our indicators. This result is the opposite out-of-sample: the fixed effect model exhibits a typical problem of overfitting.

Appendix A (Continued)

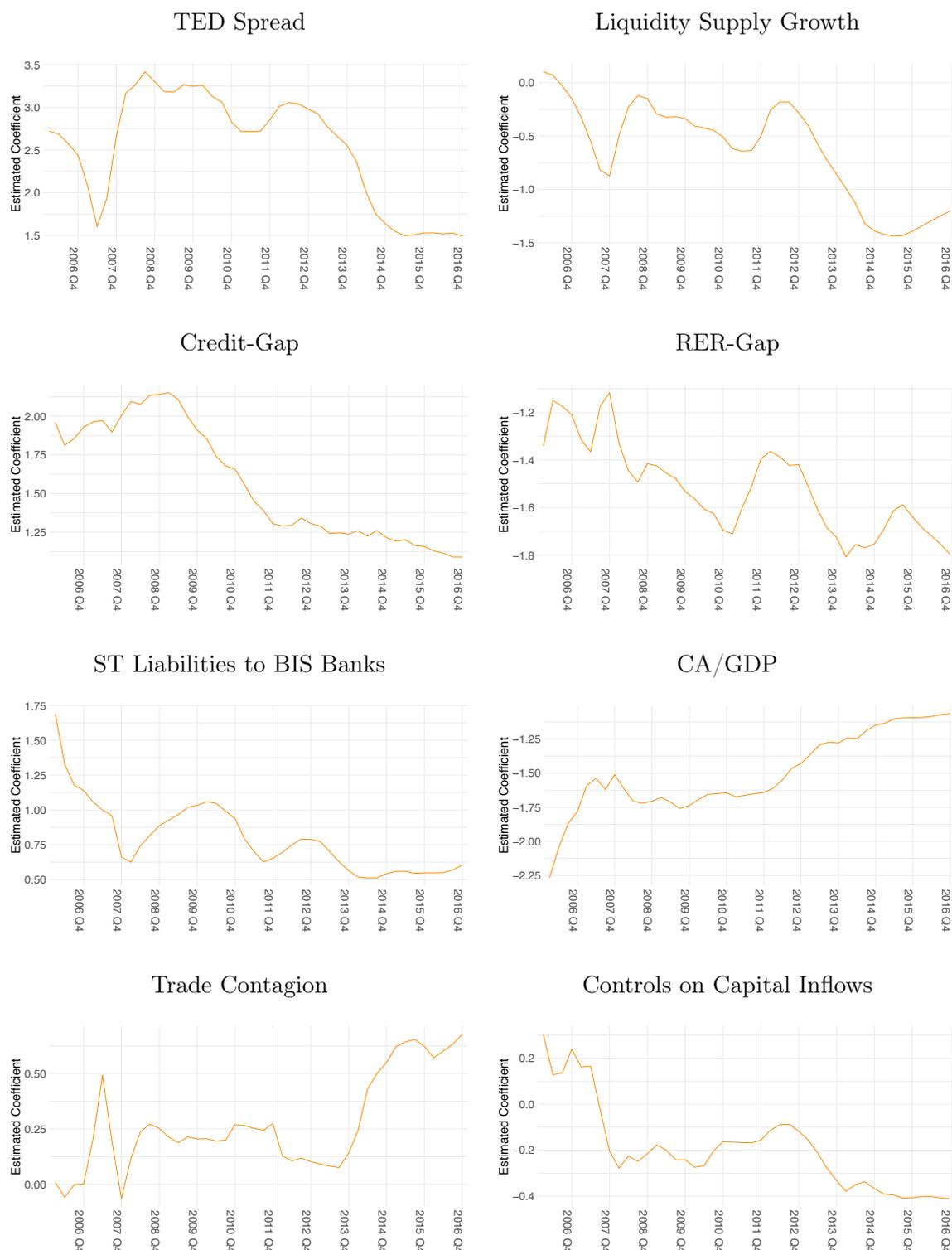
TABLE XXVIII: Robustness Analysis

	De-trending	Pre-crisis Period 4-6	Post-Crisis Bias	Country FE
	(1)	(2)	(3)	(4)
TED Spread	1.495*** (0.238)	1.540*** (0.294)	1.348*** (0.253)	1.579*** (0.245)
Global Liquidity Growth	-1.780*** (0.276)	-1.477*** (0.320)	-1.400*** (0.274)	-1.549*** (0.280)
Real GDP Growth (y-o-y)			0.324 (0.259)	
Inflation (y-o-y)			-0.075 (0.267)	
Private Credit-Gap	1.018*** (0.241)	0.865*** (0.297)	1.069*** (0.252)	1.111*** (0.259)
RER-Gap	-1.249*** (0.256)	-1.517*** (0.312)	-1.717*** (0.259)	-2.095*** (0.268)
ST Liab. to BIS Banks/GDP	0.817*** (0.237)	0.770*** (0.294)	0.748*** (0.249)	0.977*** (0.260)
CA/GDP	-1.145*** (0.237)	-1.177*** (0.294)	-1.027*** (0.248)	-1.152*** (0.260)
Trade Contagion	0.804*** (0.167)	0.341 (0.209)	0.640*** (0.174)	0.807*** (0.177)
Capital Controls on Inflows	-0.511*** (0.196)	-0.507** (0.241)	-0.362* (0.206)	-0.662*** (0.228)
Observations	1,748	1,560	1,618	1,753
Relative Usefulness	38%	38%	38.5%	45.8%

Note: Robust standard errors in parentheses. The "De-trending" column substitutes from the benchmark specification the HP-Filtered RER and Credit-Gap with the year-on-year growth rate. The "Pre-crisis Period 4-6" column is equivalent to the benchmark specification with a pre-crisis period that goes from 4 quarters to 6 quarters before the sudden stop. The "Post-crisis bias" column employs an alternative definition of sudden stop duration (see section 3.3.1 for details) and includes in the benchmark specification the yearly growth rate of GDP and yearly inflation. The "Country FE" column is equivalent to the benchmark specification with the addition of country fixed effects. * Statistical significance at 10% level. ** Statistical significance at 5% level. *** Statistical significance at 1% level.

Appendix A (Continued)

Figure 36: Rolling Coefficients



Note: The figure shows coefficients from the benchmark specification estimated through an expanding window over the period 2006Q1-2017Q1.

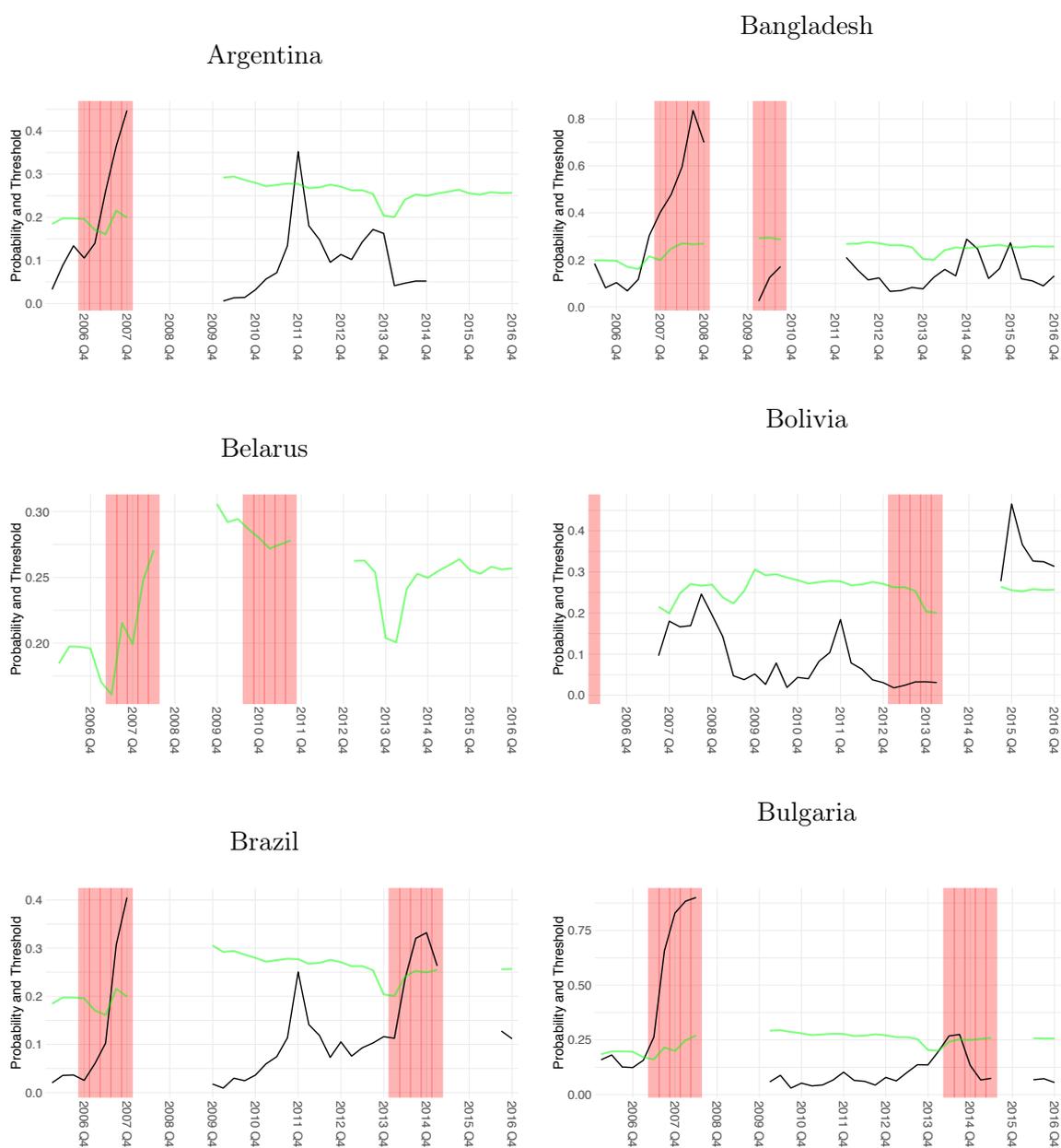
Appendix A (Continued)

Most of the indicators' estimates remain significant and stable in magnitude throughout the whole sample period. Interestingly, the coefficients associated to the TED spread decrease significantly (halves) after the GFC. The opposite trend holds for the most prominent local indicator, the RER-Gap. This points to a decreasing importance of the spread as a driver of capital flows: this result is in line with recent research on the decreasing importance of global factors after the GFC [Avdjiev et al. \(2017a\)](#).

Appendix A (Continued)

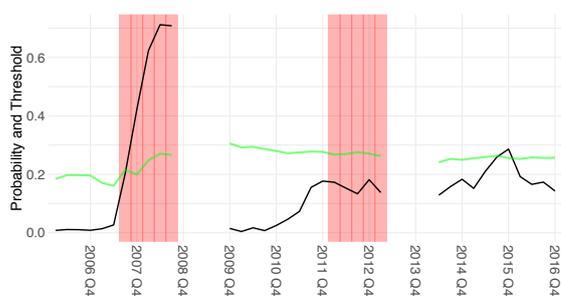
A.2.2 Out-of-Sample Performance and Forecast Horizon

Figure 37: Out-of-sample Estimated Probabilities

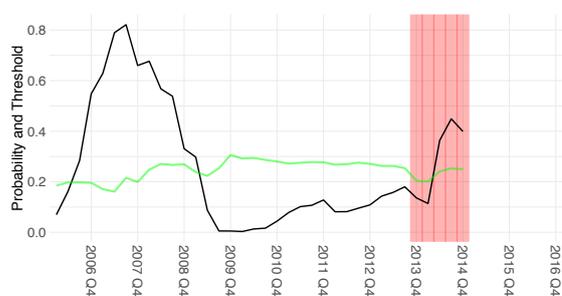


Appendix A (Continued)

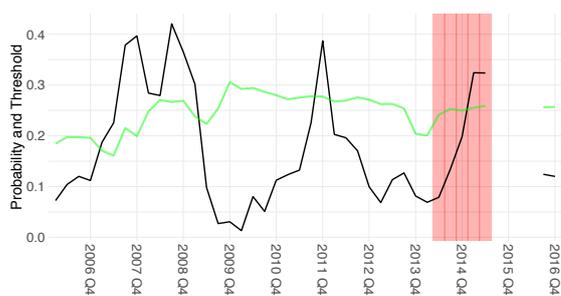
Chile



Colombia



Ecuador



El Salvador



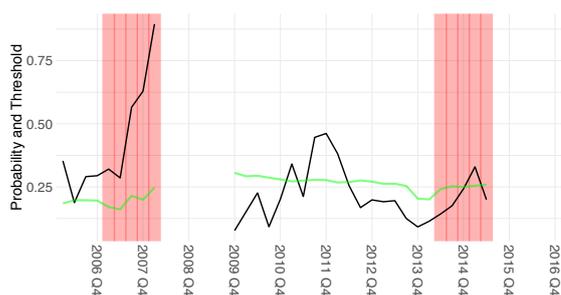
Guatemala



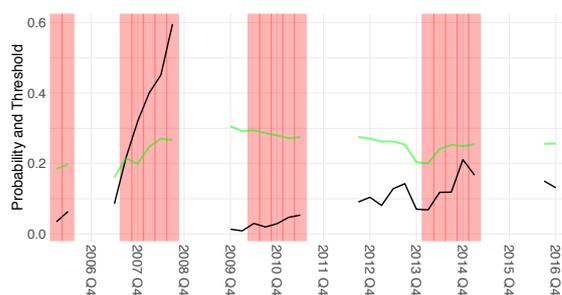
Hungary



India

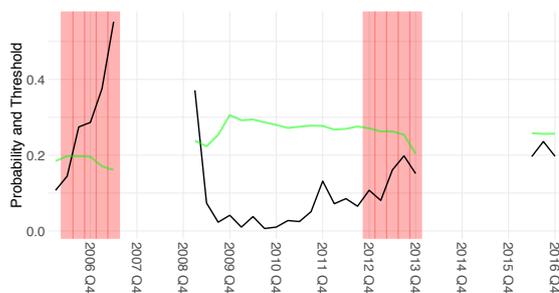


Indonesia

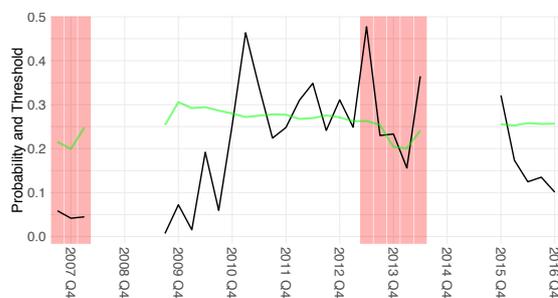


Appendix A (Continued)

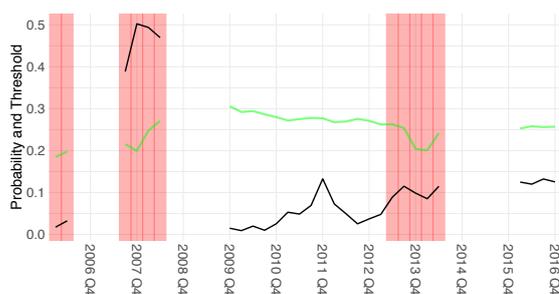
Kazakhstan



Malaysia



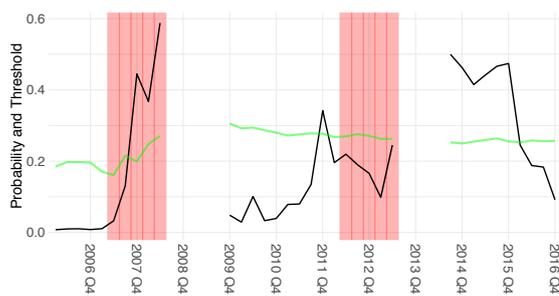
Mexico



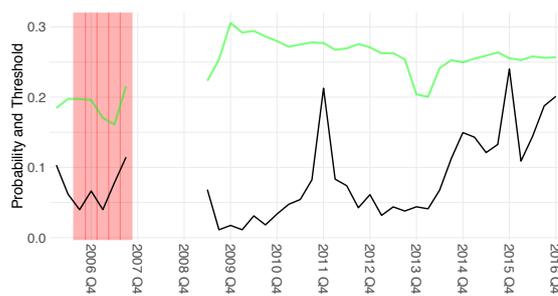
Pakistan



Peru



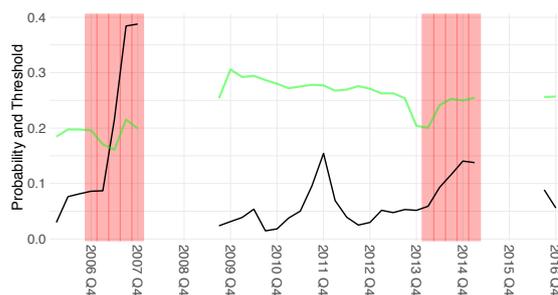
Philippines



Poland

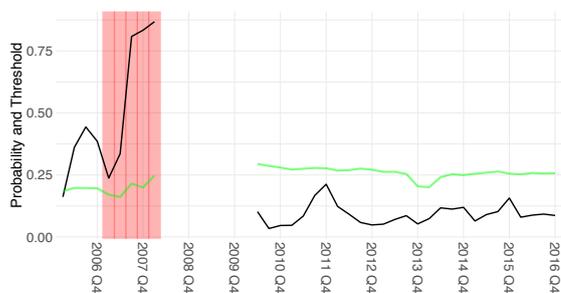


Republic of Korea

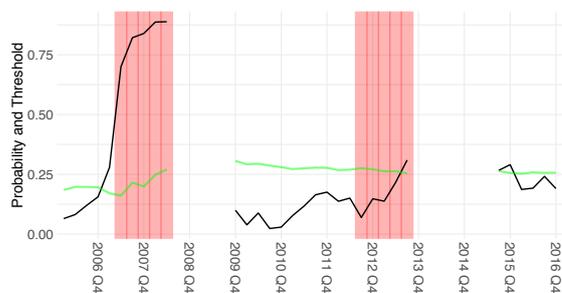


Appendix A (Continued)

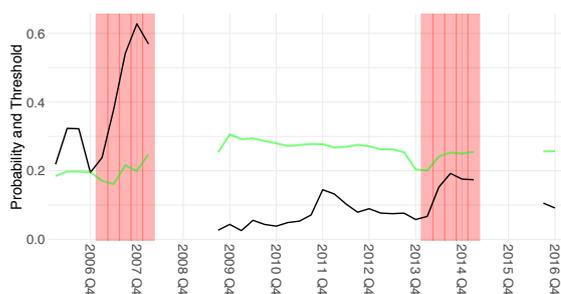
Romania



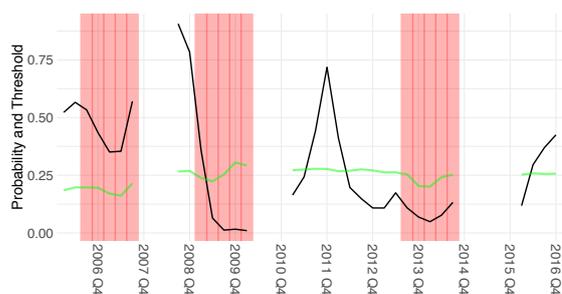
Russian Federation



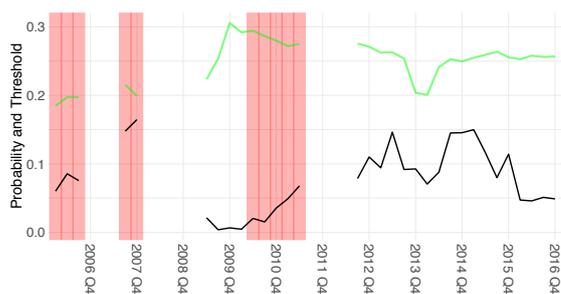
South Africa



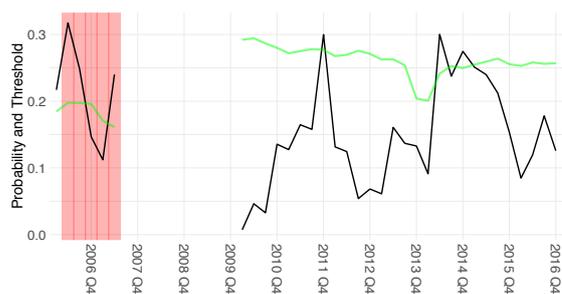
Sri Lanka



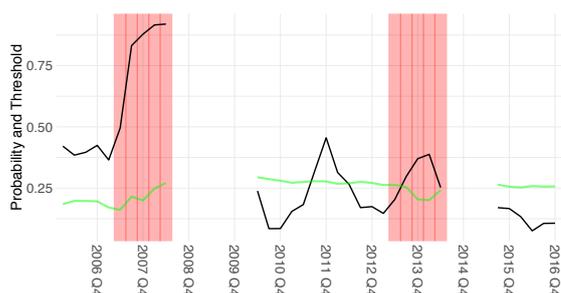
Thailand



Turkey



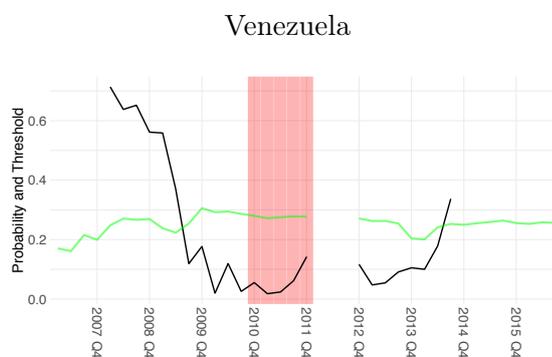
Ukraine



Uruguay



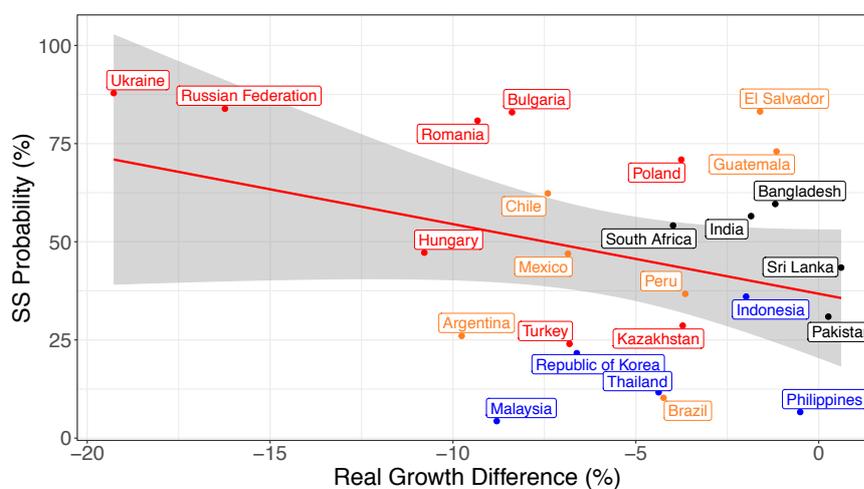
Appendix A (Continued)



Note: The figure shows the estimated probabilities from the out-of-sample recursive exercise (black line), the time-varying optimal threshold above which a signal is sent by the model (green line) and the pre-crisis period (red area). Blank areas correspond to the quarter before and the duration of the sudden stop.

Appendix A (Continued)

A.2.3 Sudden Stops Impact and Fitted Probabilities

Figure 38: Fitted Probabilities and Ex-Post Growth - Robustness

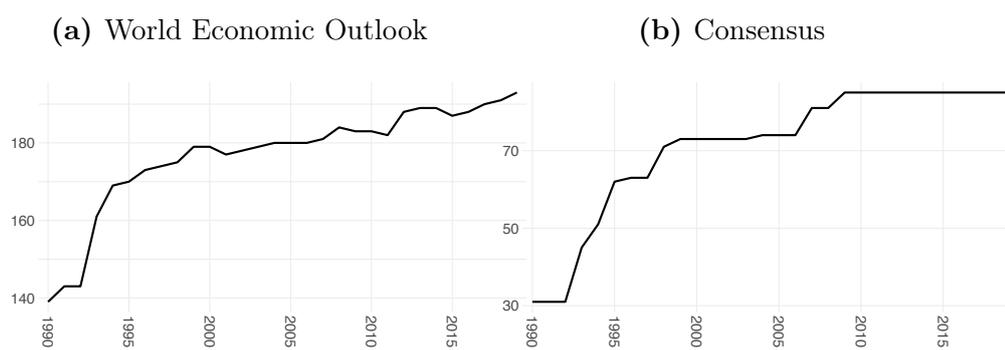
Note: The figure shows the relationship between the median out-of-sample probability in the pre-crisis period for GFC related sudden stops and a measure of output impact constructed as the difference between median growth during the sudden stop and median growth in the preceding tranquil period. Red line is regression line with 95% confidence intervals.

Appendix B

When (Where and Why) Forecasters Get It Wrong?

B.1 Data

Figure 39: Country availability



Appendix B (Continued)

B.2 Surveillance Forecasts

B.2.1 Surveillance Forecasts and Optimism

TABLE XXIX: H=0 & H=1

Country	Geo. Group	H=0,Fall	H=0,Spring	H=1,Fall	H=1,Spring
Angola	Africa	-0.03	-1.92**	-3.17**	-2.33**
Benin	Africa	-0.04	-0.1	-0.32	-0.49**
Botswana	Africa	0.21	0.43	-0.08	-0.39
Burkina Faso	Africa	-0.09	-0.01	-0.4	-0.52
Burundi	Africa	-1.25*	-3.29**	-3.58**	-3.36**
Côte d'Ivoire	Africa	-0.39	-0.37	-1.72**	-1.62**
Cameroon	Africa	0.12	-0.07	-0.59*	-0.86**
Cape Verde	Africa	-0.21	-0.29	-0.43	-0.49
Central African Republic	Africa	-2.26**	-3.46**	-3.61**	-3.84**
Chad	Africa	-0.52	-1.65**	-1.99**	-2.27**
Comoros	Africa	-0.38**	-1.02**	-1.43**	-1.72**
Congo - Brazzaville	Africa	-1.41**	-2.79**	-3.24**	-3.07**
Congo - Kinshasa	Africa	-1.46**	-3.12**	-4.29**	-5.05**
Equatorial Guinea	Africa	-1.23	-0.13	3.99	5.45
Eritrea	Africa	-0.06	-0.33	-0.71	-0.54
Ethiopia	Africa	0.45	0.45	-0.09	0.14
Gabon	Africa	0.07	-0.41	-0.52	-0.57
Gambia	Africa	-0.52	-0.56	-1.1*	-1.24*
Ghana	Africa	0.01	-0.22	-0.4	-0.96**
Guinea	Africa	-0.42*	-0.7**	-1.31**	-1.43**

Appendix B (Continued)

Guinea-Bissau	Africa	-1.42	-2.31*	-2.19*	-2.28*
Kenya	Africa	-0.33*	-0.51*	-1.31**	-1.4**
Lesotho	Africa	0.32	-0.3	-0.14	-0.04
Liberia	Africa	-0.24	-0.31	-1.62*	-1.39
Madagascar	Africa	-0.48	-1.61**	-2.35**	-2.34**
Malawi	Africa	-0.91	-1.11	-1.17	-1.01
Mali	Africa	-0.16	-0.4	-0.92**	-0.97**
Mauritius	Africa	-0.17	-0.32**	-0.31	-0.3
Mozambique	Africa	0.09	-0.24	-0.39	-0.27
Namibia	Africa	-0.46*	-0.86**	-1.3**	-1.5**
Niger	Africa	-0.16	-0.32	-0.45	-0.61
Nigeria	Africa	0.59**	0.06	-0.47	-0.56
Rwanda	Africa	-1.2	-2.18	-1.41	-1.35
São Tomé & Príncipe	Africa	-0.64**	-1.06**	-1.19**	-1.64**
Senegal	Africa	-0.5**	-0.83**	-0.8**	-0.73**
Seychelles	Africa	0.23	0.19	-0.46	-0.67
Sierra Leone	Africa	-1.5	-2.9**	-4.17**	-2.54*
South Africa	Africa	-0.14	-0.51**	-0.8**	-1.06**
South Sudan	Africa	3.19	-3.66**	-15.37**	-18.09**
Swaziland	Africa	0.07	0.22	-0.03	-0.13
Tanzania	Africa	0.13	0.05	-0.22	-0.27
Togo	Africa	-1.34**	-1.63**	-1.64**	-1.73**
Uganda	Africa	-0.01	0.14	-0.21	-0.25
Zambia	Africa	-0.7*	-1.03**	-1.12**	-1.06**
Zimbabwe	Africa	0.62**	-0.96	-3.19**	-4.35**
Bangladesh	Emerging Asia	0.23**	0.39**	0.27**	0.22*

Appendix B (Continued)

Bhutan	Emerging Asia	-0.55*	-0.3	-0.74	-0.66
Brunei	Emerging Asia	-0.92*	-0.88*	-1.68**	-1.72**
Cambodia	Emerging Asia	0.41	0.47	0.2	0.3
China	Emerging Asia	0.44**	0.95**	1.15**	1.23**
Fiji	Emerging Asia	-0.27	-0.41	-0.57	-0.46
India	Emerging Asia	0.16	-0.04	-0.01	-0.17
Indonesia	Emerging Asia	0.18	0	-0.64	-0.77
Kiribati	Emerging Asia	-0.15	-0.01	0.15	0.3
Laos	Emerging Asia	0.26	0.33	0.36	0.35
Malaysia	Emerging Asia	0.46**	0.23	-0.15	-0.29
Maldives	Emerging Asia	0.65	0.9	0.35	0.48
Marshall Islands	Emerging Asia	0.02	-0.19	0.01	0
Micronesia (Federated States of)	Emerging Asia	-0.08	-0.53	-0.52	-1.37*
Mongolia	Emerging Asia	0.36	-0.09	-0.01	0.17
Myanmar (Burma)	Emerging Asia	0.54	0.54	0.49	0.75
Nauru	Emerging Asia	0.37	1.37**	2.49**	2.04*
Nepal	Emerging Asia	0.03	-0.54	-0.86	-0.83
Palau	Emerging Asia	0.02	-0.33	-1.35	-2.25
Papua New Guinea	Emerging Asia	-0.48	-0.94	-0.59	-0.53
Philippines	Emerging Asia	0.16	0.02	-0.25	-0.39
Samoa	Emerging Asia	0.02	-0.25	-0.33	-0.43
Solomon Islands	Emerging Asia	0.14	-0.08	-0.14	-0.3
Sri Lanka	Emerging Asia	-0.01	-0.21	-0.4	-0.46
Thailand	Emerging Asia	-0.21	-0.61	-1.15*	-1.33
Timor-Leste	Emerging Asia	-0.75	-1.81	-0.03	-0.32
Tonga	Emerging Asia	-0.53	-0.97**	-1.11**	-1.26**

Appendix B (Continued)

Vanuatu	Emerging Asia	-0.58*	-0.75*	-1.09**	-1.15**
Vietnam	Emerging Asia	0.49**	0.49**	0.43	0.31
Albania	Emerging Europe	0.47	0.01	-0.59	-0.5
Belarus	Emerging Europe	1.59**	1.77**	1.35	1.11
Bosnia & Herzegovina	Emerging Europe	-0.01	-0.57	-1.17	-1.49*
Bulgaria	Emerging Europe	-0.2	-1.33*	-2.48**	-2.58**
Croatia	Emerging Europe	-0.09	-0.27	-0.89*	-1.05**
Hungary	Emerging Europe	-0.27	-0.69	-1.04	-1.37**
Kosovo	Emerging Europe	-0.08	-0.3	-0.46	-0.84**
Macedonia	Emerging Europe	-0.25	-1.24**	-1.66**	-1.99**
Moldova	Emerging Europe	-0.91	-1.76	-2.23	-2.54*
Montenegro	Emerging Europe	0.2	0.06	-1.03	-0.84
Poland	Emerging Europe	0.23	0.07	-0.22	-0.34
Romania	Emerging Europe	-0.05	-1	-1.73	-2*
Russia	Emerging Europe	0.29	0.39	-1.01	-1.14
Serbia	Emerging Europe	-0.09	-0.32	-1.42	-1.71
Turkey	Emerging Europe	0.4	0.58	0.12	-0.23
Ukraine	Emerging Europe	0.15	-0.62	-2.71**	-2.78**
Austria	Europe	0.05	0.03	-0.34	-0.47
Belgium	Europe	0.04	0.02	-0.37	-0.57**
Cyprus	Europe	0.07	-0.2	-0.45	-0.74*
Estonia	Europe	0.66*	0.36	-0.24	-0.41
Finland	Europe	-0.14	-0.37	-0.85	-0.79
France	Europe	-0.03	-0.12	-0.57**	-0.79**
Germany	Europe	-0.03	0.09	-0.5	-0.66*
Greece	Europe	0.07	-0.3	-0.61*	-1.07**

Appendix B (Continued)

Ireland	Europe	1.83**	2.03**	1.79*	1.68
Italy	Europe	-0.11	-0.33**	-0.95**	-1.18**
Latvia	Europe	0.3	-0.33	-0.56	-0.72
Lithuania	Europe	0.51	0.17	-0.42	-0.57
Luxembourg	Europe	0.43	0.52*	0.18	-0.04
Malta	Europe	0.2	0.17	0	0.13
Netherlands	Europe	0.11	0.05	-0.39	-0.45
Portugal	Europe	-0.1	-0.16	-0.75**	-1.03**
Slovakia	Europe	0.62**	0.56**	0.01	-0.02
Slovenia	Europe	0.12	0.15	-0.31	-0.44
Spain	Europe	-0.02	0.04	-0.26	-0.54*
Antigua & Barbuda	Latin America	0.23	0.1	-0.11	-0.24
Argentina	Latin America	0.7*	0.62	0.08	-0.33
Aruba	Latin America	-0.1	-0.3	-0.6	
Bahamas	Latin America	-0.56*	-0.84*	-1.34**	-1.68**
Barbados	Latin America	-0.35	-0.55	-1.13**	-1.36**
Belize	Latin America	0.07	0.07	-0.23	-0.32
Bolivia	Latin America	-0.23	-0.23	-0.49*	-0.59**
Brazil	Latin America	0.04	-0.03	-0.72*	-1.09**
Chile	Latin America	0.07	-0.08	-0.41	-0.65
Colombia	Latin America	-0.05	-0.23	-0.42	-0.69
Costa Rica	Latin America	0.39	0.45	0.39	0.37
Dominica	Latin America	-0.41	-0.46	-1.13**	-1.29**
Dominican Republic	Latin America	1.05**	1.1**	1.03*	0.87
Ecuador	Latin America	0.63**	0.46	-0.17	-0.32
El Salvador	Latin America	-0.01	-0.31	-0.4	-0.78**

Appendix B (Continued)

Grenada	Latin America	-0.27	-0.81	-1.25*	-1.48**
Guatemala	Latin America	0.04	-0.15	-0.34	-0.55**
Guyana	Latin America	0.24	0.05	-0.15	-0.07
Haiti	Latin America	-1.56**	-2.24**	-3.35**	-3.5**
Honduras	Latin America	-0.24	-0.29	-0.42	-0.53
Jamaica	Latin America	-0.56**	-0.89**	-1.31**	-1.44**
Mexico	Latin America	-0.01	-0.43	-1.33**	-1.5**
Nicaragua	Latin America	-0.13	-0.49	-0.91*	-1.15**
Panama	Latin America	0.48	0.39	0.19	0.38
Paraguay	Latin America	-0.17	-0.11	-0.72	-0.57
Peru	Latin America	0.23	0.21	-0.28	-0.29
St. Kitts & Nevis	Latin America	-0.32	-0.59	-0.54	-0.62
St. Lucia	Latin America	-0.36	-0.54	-1.08**	-0.99*
St. Vincent & Grenadines	Latin America	-0.61**	-0.84**	-1.25**	-1.28**
Suriname	Latin America	-0.16	-0.71	-0.51	-0.87
Trinidad & Tobago	Latin America	-0.56	-0.88*	-1.45**	-1.72**
Uruguay	Latin America	0.59	0.38	0.22	-0.12
Venezuela	Latin America	0.99*	0.43	-1.25	-2.44
Afghanistan	Middle East	0.18	0.49	0.57	0.45
Algeria	Middle East	-0.48*	-0.51*	-0.99**	-1.07**
Armenia	Middle East	0.84	-0.19	-1.79	1.53
Azerbaijan	Middle East	0.15	-0.66	-2	-2.41
Bahrain	Middle East	0.21	0.47**	0.55**	0.44*
Djibouti	Middle East	-0.72**	-1.08**	-1.16**	-1.37**
Egypt	Middle East	0.25	0.33**	-0.13	-0.28
Georgia	Middle East	-1.57	-1.72	-1.58	-1.4

Appendix B (Continued)

Iran	Middle East	0.29	0.05	-0.48	-0.6
Iraq	Middle East	-0.5	-0.99	-1.24	-1.55
Jordan	Middle East	0.25	0.05	-0.03	-0.41
Kazakhstan	Middle East	0.16	0.07	-0.27	-0.36
Kuwait	Middle East	-0.16	-4.2	-3.32	-0.67
Kyrgyzstan	Middle East	-0.94	-1.73	-1.81	-1.96
Lebanon	Middle East	-0.12	-0.52	-0.68	-1.19*
Libya	Middle East	0.1	-0.4	-5.25	-5.68
Mauritania	Middle East	-0.38	-0.72*	-1.55**	-2.17
Morocco	Middle East	-0.41	-0.96**	-1.25**	-1.09
Oman	Middle East	1.01**	0.71	0.54	0.45
Pakistan	Middle East	-0.54**	-0.39	-0.47	-0.8**
Qatar	Middle East	-0.43	-0.17	-0.22	0.52
Saudi Arabia	Middle East	0.22	0.54	0.14	0.47
Somalia	Middle East	-0.15	-0.18	-0.67**	0.05
Sudan	Middle East	0.48	0.3	-0.03	-0.09
Syria	Middle East	0.12	-0.07	-0.07	0.04
Tajikistan	Middle East	0.48	0.45	0.59	0.44
Tunisia	Middle East	-0.45**	-0.83**	-1.41**	-1.55**
Turkmenistan	Middle East	-1.21	-0.53	-0.42	0.02
United Arab Emirates	Middle East	0.25	0.32	0.43	0.63
Uzbekistan	Middle East	1.5**	1.42**	1.2**	1.06**
Yemen	Middle East	-0.5	-2.45**	-4.36**	-4.45**

Appendix B (Continued)

Figure 40: Largest Optimistic Biases by Region

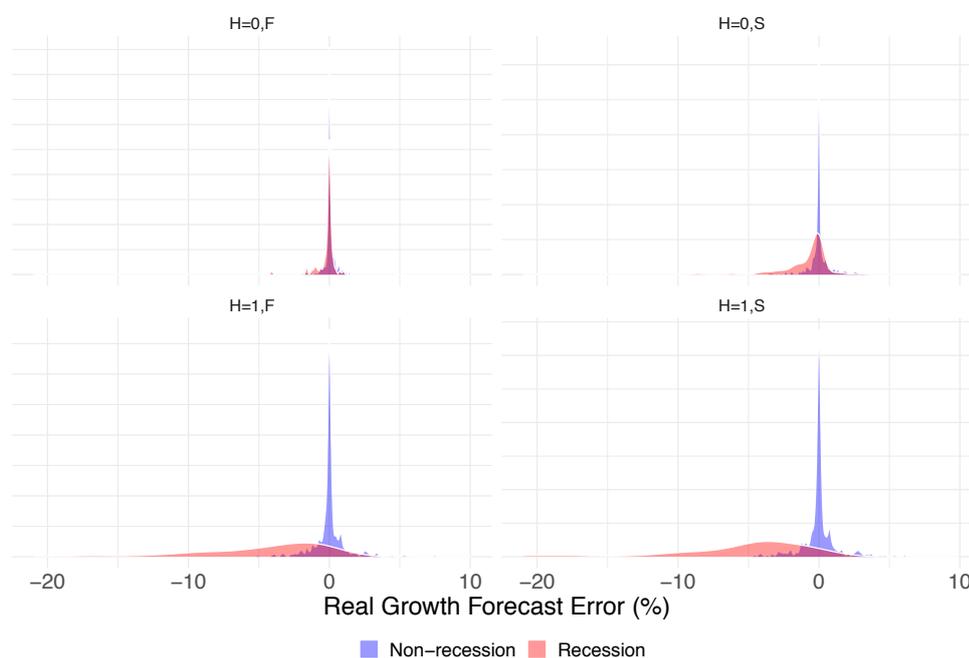


Note: The figure shows the name of the countries with the three largest optimistic biases for each geographical group and forecast horizon. Y-axis corresponds to the magnitude of the bias. If horizon not displayed, no optimistic bias for the respective region.

Appendix B (Continued)

B.2.2 Optimism, Recessions and Financial Crises

Figure 41: “Best” Private Forecasters and recessions



Note: Distribution of real GDP growth forecast errors for a subsample of individual private forecasters. The subsample is composed by forecasters that for each country, year and horizon produce the forecast closest to the actual value. Recessions are periods of negative growth.

Appendix B (Continued)

B.2.3 A Comparison Between Forecasters

TABLE XXX: List of countries - IMF/World Bank

Country	Geo. group
Angola	Africa
Benin	Africa
Botswana	Africa
Burkina Faso	Africa
Burundi	Africa
Cameroon	Africa
Comoros	Africa
Ethiopia	Africa
Gabon	Africa
Ghana	Africa
Guinea	Africa
Guinea-Bissau	Africa
Kenya	Africa
Lesotho	Africa
Madagascar	Africa
Malawi	Africa
Mali	Africa
Mauritius	Africa
Mozambique	Africa
Namibia	Africa
Niger	Africa

Appendix B (Continued)

Country	Geo. group
Nigeria	Africa
Rwanda	Africa
Senegal	Africa
Sierra Leone	Africa
South Africa	Africa
Tanzania	Africa
Togo	Africa
Uganda	Africa
Zambia	Africa
Bangladesh	Emerging Asia
Cambodia	Emerging Asia
China	Emerging Asia
Fiji	Emerging Asia
India	Emerging Asia
Indonesia	Emerging Asia
Malaysia	Emerging Asia
Nepal	Emerging Asia
Papua New Guinea	Emerging Asia
Philippines	Emerging Asia
Sri Lanka	Emerging Asia
Thailand	Emerging Asia
Vietnam	Emerging Asia
Albania	Emerging Europe
Belarus	Emerging Europe

Appendix B (Continued)

Country	Geo. group
Bulgaria	Emerging Europe
Moldova	Emerging Europe
Romania	Emerging Europe
Russia	Emerging Europe
Turkey	Emerging Europe
Ukraine	Emerging Europe
Argentina	Latin America
Belize	Latin America
Bolivia	Latin America
Brazil	Latin America
Colombia	Latin America
Costa Rica	Latin America
Dominican Republic	Latin America
Ecuador	Latin America
El Salvador	Latin America
Guatemala	Latin America
Guyana	Latin America
Haiti	Latin America
Honduras	Latin America
Jamaica	Latin America
Mexico	Latin America
Nicaragua	Latin America
Panama	Latin America
Paraguay	Latin America

Appendix B (Continued)

Country	Geo. group
Peru	Latin America
St. Lucia	Latin America
Venezuela	Latin America
Algeria	Middle East
Armenia	Middle East
Azerbaijan	Middle East
Georgia	Middle East
Iran	Middle East
Jordan	Middle East
Kazakhstan	Middle East
Lebanon	Middle East
Mauritania	Middle East
Morocco	Middle East
Pakistan	Middle East
Qatar	Middle East
Sudan	Middle East
Tajikistan	Middle East
Tunisia	Middle East
Uzbekistan	Middle East

TABLE XXXI: Full RMSE comparison - IMF/World Bank

Country	H=0,Jul.	H=0,Jan.	H=1,Jul.	H=1,Jan.
Turkey	-0.13	-0.15	0.37	0.29

Appendix B (Continued)

Country	H=0,Jul.	H=0,Jan.	H=1,Jul.	H=1,Jan.
South Africa	-0.26	-0.15	-0.16	0
Argentina	-0.35	-0.12	0.17	0.07
Bolivia	-0.16	-0.3	-0.28	-0.4
Brazil	-0.21	-0.21	-0.04	-0.04
Colombia	-0.31	-0.06	-0.1	-0.04
Costa Rica	-0.13	-0.27	0.96	-0.05
Dominican Republic	-0.21	-0.07	0.43	-0.06
Ecuador	-0.46	-0.09	-0.21	0.09
El Salvador	-0.28	-0.12	-0.04	-0.05
Guatemala	-0.05	-0.23	0.18	0.05
Haiti	0.02	-0.19	0.31	0.41
Honduras	-0.21	-0.1	-0.24	-0.06
Mexico	-0.24	-0.1	-0.03	0.2
Nicaragua	-0.47	0.01	0.02	-0.05
Panama	-0.28	-0.34	-0.01	-0.29
Paraguay	-0.33	-0.28	0.32	0.28
Peru	-0.01	-0.27	0.31	0.01
Venezuela	0.02	-0.28	-0.26	-0.24
Guyana	0.13	0.02	0.57	0.26
Belize	0.22	0.09	-0.06	-0.12
Jamaica	-0.14	-0.16	-0.18	-0.03
St. Lucia	-0.12	0.02	0	0.03
Iran	-0.01	0.01	-0.02	-0.12
Jordan	-0.1	-0.09	0.1	0.16

Appendix B (Continued)

Country	H=0,Jul.	H=0,Jan.	H=1,Jul.	H=1,Jan.
Lebanon	-0.48	-0.46	-0.09	-0.11
Qatar	0.63	0.63	0.68	2.63
Bangladesh	-0.27	-0.28	-0.25	-0.25
Cambodia	0.11	-0.17	1	-0.08
Sri Lanka	-0.18	-0.07	0.24	0.18
India	-0.18	-0.03	-0.03	-0.02
Indonesia	-0.46	-0.17	0.09	0.32
Malaysia	-0.25	-0.12	1.35	0.91
Nepal	-0.35	0.34	0.11	0.06
Pakistan	-0.34	0.1	-0.24	0.06
Philippines	-0.47	0.05	0.98	0.56
Thailand	-0.21	-0.17	0.22	0.3
Vietnam	0.01	-0.17	0.14	-0.08
Algeria	-0.26	-0.37	-0.08	-0.07
Angola	-0.13	0.08	0.13	0.17
Botswana	-0.05	-0.14	0.14	-0.01
Burundi	-0.01	0.17	0.02	0.08
Cameroon	-0.36	-0.38	-0.32	-0.41
Comoros	0.11	0.23	0.06	0.14
Benin	0.02	-0.02	-0.06	0.39
Ethiopia	0.02	0.1	0.15	0.24
Gabon	-0.14	0.01	0.32	0.11
Ghana	-0.11	0.02	-0.01	0.29
Guinea-Bissau	-0.32	0.11	0.42	0.3

Appendix B (Continued)

Country	H=0,Jul.	H=0,Jan.	H=1,Jul.	H=1,Jan.
Guinea	0	0.15	-0.21	-0.19
Kenya	0.26	0.15	0.56	0.74
Lesotho	0.03	0	0.08	0.09
Madagascar	0.22	-0.36	-0.08	0.38
Malawi	-0.15	-0.04	0	-0.05
Mali	0.21	0.04	-0.06	-0.05
Mauritania	-0.15	0.26	0.27	0.61
Mauritius	0.32	0.02	0.6	0.03
Morocco	0.36	0.17	-0.08	0.04
Mozambique	0.08	0.42	-0.05	-0.03
Niger	0.02	-0.05	0.1	-0.12
Nigeria	-0.38	-0.17	0.05	-0.06
Rwanda	0.05	-0.11	-0.06	0
Senegal	-0.13	-0.09	-0.11	-0.18
Sierra Leone	-0.25	0.11	0.36	0.06
Namibia	0.05	0.02	0.13	0
Sudan	-0.12	-0.2	-0.2	0.01
Tanzania	-0.15	-0.41	0.24	-0.17
Togo	-0.17	-0.17	-0.1	-0.13
Tunisia	-0.26	-0.3	-0.06	0.05
Uganda	-0.13	-0.2	-0.17	-0.21
Burkina Faso	-0.01	0.04	0.15	0.03
Zambia	0.17	0.16	0.25	0.18
Fiji	-0.04	0.01	0	0.04

Appendix B (Continued)

Country	H=0,Jul.	H=0,Jan.	H=1,Jul.	H=1,Jan.
Papua New Guinea	0.34	0.23	0.03	-0.03
Armenia	0.09	0.06	-0.2	0.19
Azerbaijan	-0.44	0.11	0.31	-0.26
Belarus	-0.43	-0.04	0.14	0.13
Albania	0.11	0.35	-0.33	-0.19
Georgia	-0.16	-0.08	0.23	-0.14
Kazakhstan	-0.48	-0.08	-0.01	0.17
Bulgaria	-0.09	-0.29	0.13	-0.06
Moldova	-0.12	-0.2	0.16	0.09
Russia	-0.08	0.09	-0.07	0.62
Tajikistan	0.15	0.25	0.3	0.12
China	-0.53	-0.44	0.32	1.39
Ukraine	0.25	-0.17	-0.1	-0.05
Uzbekistan	0.42	0.09	-0.1	-0.09
Romania	0.08	-0.02	0.2	0.17

Note: This table reports the ratio of the estimated RMSE for the WEO real GDP growth forecasts versus the RMSE for the Global Economic Prospect (GEP) forecasts. We have subtracted one, so that values greater than zero suggest that the WEO forecasts are less accurate than the GEP forecasts, while values below zero suggest that the WEO forecasts are more accurate.

TABLE XXXII: List of countries - IMF/EC

Country
Albania
Australia

Appendix B (Continued)

Country
Austria
Belgium
Bulgaria
Canada
Croatia
Cyprus
Czechia
Denmark
Estonia
Finland
France
Germany
Greece
Hungary
Iceland
Ireland
Italy
Japan
Latvia
Lithuania
Luxembourg
Macedonia
Malta
Mexico

Appendix B (Continued)

Country
Montenegro
Netherlands
New Zealand
Norway
Poland
Portugal
Serbia
Slovakia
Slovenia
South Korea
Spain
Sweden
Switzerland
Turkey
United Kingdom
United States

TABLE XXXIII: Full RMSE comparison - IMF/EC

Country	H=0,F	H=0,S	H=1,F	H=1,S
Albania	-0.18	-0.21	0.35	0.68
Australia	-0.17	0.61	0.58	0.84
Austria	0.62	0.17	0.33	0.1
Belgium	0.63	-0.36	0.18	-0.18

Appendix B (Continued)

Country	H=0,F	H=0,S	H=1,F	H=1,S
Bulgaria	-0.18	0.06	-0.23	-0.12
Canada	0.04	0.13	-0.13	-0.06
Croatia	0.18	0.35	0.35	0.11
Cyprus	0.05	0.26	0.12	0
Czechia	0.43	0.07	0.14	0.04
Denmark	0.18	-0.11	0.11	0.05
Estonia	0.1	0.17	0.41	0.25
Finland	-0.22	0.1	0.08	-0.05
France	-0.05	0.04	0.29	-0.18
Germany	0.11	0.13	0.26	0.06
Greece	0.47	0.3	0.3	0.12
Hungary	0.29	0.33	0.21	0.09
Iceland	-0.17	-0.06	-0.21	-0.19
Ireland	0.08	0	-0.03	-0.04
Italy	-0.26	-0.02	-0.01	-0.14
Japan	-0.02	0.18	0.15	0.38
Latvia	0.22	0.05	-0.01	0.22
Lithuania	0.11	0.21	0.61	0.12
Luxembourg	0.21	0.2	0.11	-0.01
Macedonia	0.1	0.13	0.02	0
Malta	0.29	0.26	0.06	0.1
Mexico	-0.8	1.25	-0.16	-0.02
Montenegro	0.5	0.39	0.37	-0.03
Netherlands	-0.05	-0.02	0.36	-0.12

Appendix B (Continued)

Country	H=0,F	H=0,S	H=1,F	H=1,S
New Zealand	-0.05	0.72	0.57	0.87
Norway	-0.35	0.27	-0.1	-0.09
Poland	-0.04	0.2	0.1	0.06
Portugal	-0.25	0.1	1.14	0.45
Serbia	0.59	0.01	0.37	0.78
Slovakia	-0.09	0.36	0.31	-0.02
Slovenia	0.49	0.18	0.29	-0.08
South Korea	-0.07	0.1	0.34	-0.02
Spain	-0.43	0.88	0.16	0.14
Sweden	0.23	-0.05	0.6	0.55
Switzerland	0	0.43	-0.07	0.1
Turkey	0.12	0.19	0.44	0.17
United Kingdom	-0.39	0.45	0.32	0.34
United States	-0.14	0.15	-0.12	0.01

Note: This table reports the ratio of the estimated RMSE for the WEO real GDP growth forecasts versus the RMSE for the AMECO forecasts. We have subtracted one, so that values greater than zero suggest that the WEO forecasts are less accurate than the AMECO forecasts, while values below zero suggest that the WEO forecasts are more accurate.

TABLE XXXIV: List of countries - IMF/Consensus

Country	Geo. group	Individual Level
Nigeria	Africa	No
South Africa	Africa	No
Bangladesh	Emerging Asia	No

Appendix B (Continued)

Country	Geo. group	Individual Level
China	Emerging Asia	Yes
India	Emerging Asia	Yes
Indonesia	Emerging Asia	Yes
Malaysia	Emerging Asia	Yes
Myanmar (Burma)	Emerging Asia	No
Philippines	Emerging Asia	Yes
Sri Lanka	Emerging Asia	No
Thailand	Emerging Asia	Yes
Vietnam	Emerging Asia	No
Albania	Emerging Europe	No
Belarus	Emerging Europe	No
Bosnia & Herzegovina	Emerging Europe	No
Bulgaria	Emerging Europe	Yes
Croatia	Emerging Europe	Yes
Hungary	Emerging Europe	Yes
Macedonia	Emerging Europe	No
Moldova	Emerging Europe	No
Poland	Emerging Europe	Yes
Romania	Emerging Europe	Yes
Russia	Emerging Europe	Yes
Serbia	Emerging Europe	No
Turkey	Emerging Europe	Yes
Ukraine	Emerging Europe	Yes
Austria	Europe	No

Appendix B (Continued)

Country	Geo. group	Individual Level
Belgium	Europe	No
Cyprus	Europe	No
Estonia	Europe	Yes
Finland	Europe	No
France	Europe	Yes
Germany	Europe	Yes
Greece	Europe	No
Ireland	Europe	No
Italy	Europe	Yes
Latvia	Europe	Yes
Lithuania	Europe	Yes
Netherlands	Europe	Yes
Portugal	Europe	No
Slovakia	Europe	Yes
Slovenia	Europe	Yes
Spain	Europe	Yes
Argentina	Latin America	Yes
Bolivia	Latin America	No
Brazil	Latin America	Yes
Chile	Latin America	Yes
Colombia	Latin America	Yes
Costa Rica	Latin America	No
Dominican Republic	Latin America	No
Ecuador	Latin America	No

Appendix B (Continued)

Country	Geo. group	Individual Level
El Salvador	Latin America	No
Guatemala	Latin America	No
Honduras	Latin America	No
Mexico	Latin America	Yes
Nicaragua	Latin America	No
Panama	Latin America	No
Paraguay	Latin America	No
Peru	Latin America	Yes
Uruguay	Latin America	No
Venezuela	Latin America	Yes
Armenia	Middle East	No
Azerbaijan	Middle East	No
Egypt	Middle East	No
Georgia	Middle East	No
Kazakhstan	Middle East	No
Pakistan	Middle East	No
Saudi Arabia	Middle East	No
Turkmenistan	Middle East	No
Uzbekistan	Middle East	No
Australia	Other Adv. Economies	Yes
Canada	Other Adv. Economies	Yes
Czechia	Other Adv. Economies	Yes
Denmark	Other Adv. Economies	No
Hong Kong SAR China	Other Adv. Economies	Yes

Appendix B (Continued)

Country	Geo. group	Individual Level
Israel	Other Adv. Economies	No
Japan	Other Adv. Economies	Yes
New Zealand	Other Adv. Economies	Yes
Norway	Other Adv. Economies	Yes
Singapore	Other Adv. Economies	Yes
South Korea	Other Adv. Economies	Yes
Sweden	Other Adv. Economies	Yes
Switzerland	Other Adv. Economies	Yes
Taiwan	Other Adv. Economies	Yes
United Kingdom	Other Adv. Economies	Yes
United States	Other Adv. Economies	Yes

TABLE XXXV: Full RMSE comparison - IMF/Consensus

Country	H=0,F	H=0,S	H=1,F	H=1,S
United States	-0.07	-0.01	-0.2	-0.06
United Kingdom	-0.11	-0.2	-0.14	-0.03
Austria	-0.13	-0.13	-0.16	-0.13
Belgium	-0.16	0.19	-0.24	-0.22
Denmark	0.16	-0.23	-0.1	-0.14
France	-0.15	-0.03	-0.17	-0.06
Germany	0	0.16	-0.08	-0.03
Italy	-0.18	-0.23	-0.13	-0.09
Netherlands	-0.02	-0.04	-0.06	-0.11

Appendix B (Continued)

Country	H=0,F	H=0,S	H=1,F	H=1,S
Norway	-0.5	-0.24	-0.24	-0.19
Sweden	-0.11	0.03	0	-0.04
Switzerland	0.39	0.4	-0.04	-0.07
Canada	-0.03	0.1	-0.15	0.01
Japan	-0.1	0.13	-0.11	-0.03
Finland	-0.31	-0.28	-0.07	-0.12
Greece	0.09	0.24	-0.13	-0.08
Ireland	-0.04	-0.12	-0.09	-0.11
Portugal	-0.44	-0.09	-0.2	-0.15
Spain	-0.06	-0.17	-0.08	-0.01
Turkey	-0.12	-0.04	-0.11	-0.03
Australia	-0.08	0.5	0	0.19
New Zealand	0.2	0.27	-0.02	0.1
South Africa	0	-0.25	-0.17	-0.22
Argentina	-0.06	0.01	0	-0.02
Bolivia	-0.32	-0.29	-0.26	-0.18
Brazil	0	0.01	-0.03	-0.05
Chile	0.13	0.14	-0.05	-0.08
Colombia	-0.12	0.09	0.01	0.16
Costa Rica	-0.25	0.17	0.02	0.09
Dominican Republic	-0.23	-0.03	0.11	0.05
Ecuador	-0.27	0.08	0.01	0.22
El Salvador	-0.1	0.25	3.55	2.98
Guatemala	-0.26	-0.08	0.63	1.16

Appendix B (Continued)

Country	H=0,F	H=0,S	H=1,F	H=1,S
Honduras	-0.36	0.37	2.07	2.31
Mexico	-0.2	-0.07	-0.14	-0.13
Nicaragua	-0.53	-0.02	0.1	0.09
Panama	-0.19	-0.03	0.03	0.1
Paraguay	-0.3	-0.29	-0.19	-0.15
Peru	-0.08	0.03	-0.02	0.09
Uruguay	-0.07	-0.27	-0.05	0.12
Venezuela	0.1	0	-0.14	-0.13
Cyprus	0.03	0.17	-0.25	-0.24
Israel	-0.35	0.19	-0.23	0.12
Saudi Arabia	0.08	-0.14	0.05	0.27
Egypt	-0.44	-0.03	-0.08	-0.04
Bangladesh	0.15	-0.13	-0.17	-0.26
Myanmar (Burma)	0.24	-0.33	1.98	-0.09
Sri Lanka	0.01	0.01	-0.06	-0.03
Taiwan	-0.09	0	-0.09	0.13
Hong Kong SAR China	-0.04	-0.04	-0.11	0.02
India	-0.22	-0.12	-0.01	-0.09
Indonesia	0.86	0.14	-0.03	0.19
South Korea	-0.08	0.02	-0.1	0.05
Malaysia	0.23	0.2	0.01	0.01
Pakistan	0.31	-0.05	0	0.76
Philippines	-0.05	-0.03	-0.07	0.26
Singapore	-0.21	0.2	-0.03	0.11

Appendix B (Continued)

Country	H=0,F	H=0,S	H=1,F	H=1,S
Thailand	-0.09	-0.03	-0.01	0.1
Vietnam	-0.16	-0.07	-0.07	0.01
Nigeria	-0.06	0.26	-0.1	-0.02
Armenia	-0.27	-0.23	0.04	-0.11
Azerbaijan	-0.3	-0.12	0.31	0.11
Belarus	-0.29	0.01	-0.05	0.03
Albania	0.35	1.75	-0.07	-0.31
Georgia	-0.58	0.06	-0.09	0.01
Kazakhstan	-0.06	-0.03	-0.12	-0.03
Bulgaria	0.54	-0.24	-0.16	-0.21
Moldova	0.03	-0.12	-0.03	-0.04
Russia	-0.2	-0.32	-0.1	-0.1
China	-0.24	0.27	0.29	0.6
Turkmenistan	0	0.28	0.11	0.02
Ukraine	0.07	0.24	-0.11	-0.13
Uzbekistan	-0.2	-0.21	0.19	0.26
Czechia	-0.11	-0.1	-0.1	-0.1
Slovakia	-0.01	-0.22	-0.06	-0.12
Estonia	-0.14	-0.08	-0.09	-0.04
Latvia	-0.21	-0.13	-0.17	-0.18
Serbia	-0.24	-0.07	-0.03	0.02
Hungary	0.14	0.31	-0.06	-0.11
Lithuania	0.21	-0.17	-0.15	-0.12
Croatia	-0.12	-0.18	-0.08	-0.17

Appendix B (Continued)

Country	H=0,F	H=0,S	H=1,F	H=1,S
Slovenia	0.07	-0.14	-0.03	-0.17
Macedonia	-0.07	-0.11	-0.02	-0.03
Bosnia & Herzegovina	0.16	-0.44	-0.05	-0.15
Poland	0.01	0.1	-0.06	-0.02
Romania	0.26	-0.22	-0.07	-0.12

Note: This table reports the ratio of the estimated RMSE for the WEO real GDP growth forecasts versus the RMSE for the Consensus Economics (CE) forecasts. We have subtracted one, so that values greater than zero suggest that the WEO forecasts are less accurate than the CE forecasts, while values below zero suggest that the WEO forecasts are more accurate.

Appendix B (Continued)

B.3 Program Forecasts

TABLE XXXVI: List of IMF programs - 2002-2018

Program ID	Country	Date Approval	Type of program	Total amount (% quota)
570	Afghanistan	2006-06-26	PRGF	50.03
675	Afghanistan	2011-11-14	ECF	52.5
745	Afghanistan	2016-07-20	ECF	10
507	Albania	2002-06-21	PRGF	57.49
565	Albania	2006-01-27	PRGF-EFF	17.5
623	Angola	2009-11-23	SBA	300
774	Angola	2018-12-07	EFF	361.17
644	Antigua & Barbuda	2010-06-07	SBA	500
510	Argentina	2003-01-24	SBA	102.71
508	Argentina	2003-09-20	SBA	424.21
770	Argentina	2018-06-20	SBA	1277.38
557	Armenia	2005-05-25	PRGF	25
602	Armenia	2008-11-17	PRGF	10
611	Armenia	2009-03-06	SBA	580
649	Armenia	2010-06-28	ECF-EFF	145
710	Armenia	2014-03-07	EFF	89.36
511	Bangladesh	2003-06-20	PRGF	75.07
682	Bangladesh	2012-04-11	ECF	120.01
772	Barbados	2018-10-01	EFF	289.95
608	Belarus	2009-01-12	SBA	587.35
560	Benin	2005-08-05	PRGF	40.02

Appendix B (Continued)

Program ID	Country	Date Approval	Type of program	Total amount (% quota)
647	Benin	2010-06-14	ECF	120
754	Benin	2017-04-07	ECF	151.4
512	Bolivia	2003-04-02	SBA	85
506	Bosnia & Herzegovina	2002-08-02	SBA	39.98
618	Bosnia & Herzegovina	2009-07-08	SBA	600
692	Bosnia & Herzegovina	2012-09-26	SBA	329.88
747	Bosnia & Herzegovina	2016-09-07	EFF	167.06
513	Brazil	2002-09-06	SBA	901.65
549	Bulgaria	2004-08-06	SBA	15.62
514	Burkina Faso	2003-06-11	PRGF	49.97
578	Burkina Faso	2007-04-23	PRGF	80
645	Burkina Faso	2010-06-14	ECF	136.66
708	Burkina Faso	2013-12-27	ECF	92.43
767	Burkina Faso	2018-03-14	ECF	90
538	Burundi	2004-01-23	PRGF	90
596	Burundi	2008-07-07	PRGF	66.49
678	Burundi	2012-01-27	ECF	51.95
612	Côte d'Ivoire	2009-03-27	PRGF	115
674	Côte d'Ivoire	2011-11-04	ECF	159.93
750	Côte d'Ivoire	2016-12-12	ECF-EFF	43.25
563	Cameroon	2005-10-24	PRGF	10
760	Cameroon	2017-06-26	ECF	175
515	Cape Verde	2002-04-10	PRGF	90
575	Central African Republic	2006-12-22	PRGF	124.99

Appendix B (Continued)

Program ID	Country	Date Approval	Type of program	Total amount (% quota)
686	Central African Republic	2012-06-25	ECF	75.4
746	Central African Republic	2016-07-20	ECF	120
553	Chad	2005-02-16	PRGF	45
717	Chad	2014-08-01	ECF	210.51
761	Chad	2017-06-30	ECF	160
503	Colombia	2003-01-15	SBA	200
555	Colombia	2005-05-02	SBA	52.33
769	Colombia	2018-05-25	FCL	383.86
621	Comoros	2009-09-21	PRGF	152.51
607	Congo - Brazzaville	2008-12-08	PRGF	10
625	Congo - Kinshasa	2009-12-11	PRGF	65
614	Costa Rica	2009-04-11	SBA	300
517	Croatia	2003-02-03	SBA	29
548	Croatia	2004-08-04	SBA	27.12
698	Cyprus	2013-05-15	EFF	563.21
598	Djibouti	2008-09-17	PRGF	140
518	Dominica	2002-08-28	SBA	40
519	Dominica	2003-12-29	PRGF	93.76
539	Dominican Republic	2003-08-29	SBA	200
552	Dominican Republic	2005-01-31	SBA	200
622	Dominican Republic	2009-11-09	SBA	500
509	Ecuador	2003-03-21	SBA	49.95
752	Egypt	2016-11-11	EFF	422
609	El Salvador	2009-01-16	SBA	300

Appendix B (Continued)

Program ID	Country	Date Approval	Type of program	Total amount (% quota)
635	El Salvador	2010-03-17	SBA	300
629	Ethiopia	2009-08-26	ESF	115
540	Gabon	2004-05-28	SBA	45
759	Gabon	2017-06-19	EFF	215
577	Gambia	2007-02-21	PRGF	80
685	Gambia	2012-05-25	ECF	61.09
541	Georgia	2004-06-04	PRGF	65.2
597	Georgia	2008-09-15	SBA	497.07
683	Georgia	2012-04-11	SBA-SCF	83.17
716	Georgia	2014-07-30	SBA	66.53
755	Georgia	2017-04-12	EFF	230.23
521	Ghana	2003-05-09	PRGF	50
725	Ghana	2015-04-03	ECF	180
638	Greece	2010-05-09	SBA	3211.77
680	Greece	2012-03-15	EFF	2158.77
566	Grenada	2006-04-17	PRGF	140
713	Grenada	2014-06-26	ECF	120
522	Guatemala	2002-04-01	SBA	39.96
523	Guatemala	2003-06-18	SBA	39.96
616	Guatemala	2009-04-22	SBA	300
589	Guinea	2007-12-21	PRGF	64.82
679	Guinea	2012-02-24	ECF	162.59
765	Guinea	2017-12-11	ECF	56.25
637	Guinea-Bissau	2010-05-07	ECF	157.5

Appendix B (Continued)

Program ID	Country	Date Approval	Type of program	Total amount (% quota)
730	Guinea-Bissau	2015-07-10	ECF	160
572	Haiti	2006-11-20	PRGF	220
651	Haiti	2010-07-21	ECF	50.06
727	Haiti	2015-05-18	ECF	60
591	Honduras	2008-04-07	SBA	30
654	Honduras	2010-10-01	SBA-SCF	50.19
720	Honduras	2014-12-03	SBA-SCF	60
600	Hungary	2008-11-06	SBA	1014.78
603	Iceland	2008-11-19	SBA	1190.48
564	Iraq	2005-12-23	SBA	40
588	Iraq	2007-12-19	SBA	39.97
633	Iraq	2010-02-24	SBA	200
741	Iraq	2016-07-07	SBA	230.26
697	Jamaica	2013-05-01	EFF	224.86
748	Jamaica	2016-11-11	SBA	312.17
505	Jordan	2002-07-03	SBA	50.02
690	Jordan	2012-08-03	SBA	800
744	Jordan	2016-08-24	EFF	150
524	Kenya	2003-11-21	PRGF	55.27
661	Kenya	2011-01-31	ECF	180
722	Kenya	2015-02-02	SBA-SCF	50
650	Kosovo	2010-07-21	SBA	157.63
684	Kosovo	2012-04-27	SBA	154.24
732	Kosovo	2015-07-29	SBA	250

Appendix B (Continued)

Program ID	Country	Date Approval	Type of program	Total amount (% quota)
671	Kyrgyzstan	2011-06-20	ECF	75
726	Kyrgyzstan	2015-04-08	ECF	75
605	Latvia	2008-12-23	SBA	1200.02
641	Lesotho	2010-06-02	ECF	145.01
590	Liberia	2008-03-14	PRGF-EFF	191.87
693	Liberia	2012-11-19	ECF	86.43
562	Macedonia	2005-08-31	SBA	75
658	Macedonia	2011-01-19	PCL	599.42
571	Madagascar	2006-07-21	PRGF	60.01
742	Madagascar	2016-07-27	ECF	205.03
559	Malawi	2005-08-05	PRGF	70
627	Malawi	2008-12-03	ESF	75
632	Malawi	2010-02-19	ECF	75
689	Malawi	2012-07-23	ECF	199.86
768	Malawi	2018-04-30	ECF	56.25
543	Mali	2004-06-23	PRGF	10
677	Mali	2011-12-27	ECF	32.15
707	Mali	2013-12-18	ECF	200
525	Mauritania	2003-07-18	PRGF	10
574	Mauritania	2006-12-18	PRGF	25
634	Mauritania	2010-03-15	ECF	120
764	Mauritania	2017-12-06	ECF	90
763	Mexico	2017-11-29	FCL	600
567	Moldova	2006-05-05	PRGF	90

Appendix B (Continued)

Program ID	Country	Date Approval	Type of program	Total amount (% quota)
630	Moldova	2010-01-29	ECF-EFF	150
749	Moldova	2016-11-07	ECF-EFF	24.99
613	Mongolia	2009-04-01	SBA	300
757	Mongolia	2017-05-24	EFF	435
691	Morocco	2012-08-03	PLL	700
743	Morocco	2016-07-22	PLL	279.96
775	Morocco	2018-12-17	PLL	240.47
544	Mozambique	2004-07-06	PRGF	10
733	Mozambique	2015-12-18	SCF	180
526	Nepal	2003-11-19	PRGF	70
586	Nicaragua	2007-10-05	PRGF	60
551	Niger	2005-01-31	PRGF	40
681	Niger	2012-03-16	ECF	182.5
751	Niger	2017-01-23	ECF	90
604	Pakistan	2008-11-24	SBA	700
703	Pakistan	2013-09-04	EFF	424.98
528	Paraguay	2003-12-15	SBA	50.05
568	Paraguay	2006-05-31	SBA	30.03
545	Peru	2004-06-09	SBA	45
576	Peru	2007-01-26	SBA	27
753	Poland	2017-01-13	FCL	158.71
670	Portugal	2011-05-20	EFF	2305.72
531	Romania	2004-07-07	SBA	24.27
617	Romania	2009-05-04	SBA	1110.76

Appendix B (Continued)

Program ID	Country	Date Approval	Type of program	Total amount (% quota)
662	Romania	2011-03-31	SBA	300.04
704	Romania	2013-09-27	SBA	169.97
529	Rwanda	2002-08-12	PRGF	4.99
739	Rwanda	2016-06-08	SCF	90
561	São Tomé & Príncipe	2005-08-01	PRGF	40
610	São Tomé & Príncipe	2009-03-02	PRGF	35
688	São Tomé & Príncipe	2012-07-20	ECF	35
731	São Tomé & Príncipe	2015-07-13	ECF	60
530	Senegal	2003-04-28	PRGF	15
606	Serbia	2009-01-16	SBA	559.99
673	Serbia	2011-09-29	SBA	199.91
723	Serbia	2015-02-23	SBA	200
601	Seychelles	2008-11-14	SBA	200
626	Seychelles	2009-12-23	EFF	300
712	Seychelles	2014-06-04	EFF	105
573	Sierra Leone	2006-05-10	PRGF	50.03
705	Sierra Leone	2013-10-21	ECF	180
758	Sierra Leone	2017-06-05	ECF	78
773	Sierra Leone	2018-11-30	ECF	60
640	Solomon Islands	2010-06-02	SCF	120
676	Solomon Islands	2011-12-06	SCF	50
695	Solomon Islands	2012-12-07	ECF	10
504	Sri Lanka	2003-04-18	PRGF-EFF	65.07
620	Sri Lanka	2009-07-24	SBA	400

Appendix B (Continued)

Program ID	Country	Date Approval	Type of program	Total amount (% quota)
738	Sri Lanka	2016-06-03	EFF	185
672	St. Kitts & Nevis	2011-07-27	SBA	590
737	Suriname	2016-05-27	SBA	265.32
502	Tajikistan	2002-12-11	PRGF	74.71
615	Tajikistan	2009-04-21	PRGF	120
534	Tanzania	2003-08-16	PRGF	9.85
687	Tanzania	2012-07-06	SCF	74.91
592	Togo	2008-04-21	PRGF	129.99
756	Togo	2017-05-05	ECF	168.7
699	Tunisia	2013-06-07	SBA	400
735	Tunisia	2016-05-20	EFF	358.08
556	Turkey	2005-05-11	SBA	691.08
501	Uganda	2002-09-13	PRGF	7.48
546	Ukraine	2004-03-29	SBA	30
599	Ukraine	2008-11-05	SBA	801.75
652	Ukraine	2010-07-28	SBA	728.86
711	Ukraine	2014-04-30	SBA	800
724	Ukraine	2015-03-11	EFF	900
776	Ukraine	2018-12-18	SBA	139.18
535	Uruguay	2002-04-01	SBA	648.78
558	Uruguay	2005-06-08	SBA	250
653	Yemen	2010-07-30	ECF	100
718	Yemen	2014-09-02	ECF	150
547	Zambia	2004-06-16	PRGF	45

Appendix B (Continued)

Program ID	Country	Date Approval	Type of program	Total amount (% quota)
594	Zambia	2008-06-04	PRGF	45

Appendix C

A (Modern) Narrative On 70 Years of Macroeconomic Crises:

A Database For Complexity

C.1 From Qualitative Judgements to Quantitative Measures

C.1.1 The Corpus of IMF Documents

C.1.1.1 Scraping and Semantic Analysis

1. Scrape the URLs of all documents country by country using the form from the IMF archives website and consolidate into a single database of 250,000 documents that contains the title of the file, the date of publication and the country of interest
2. Find documents related to programs using a text analysis of the title. Specifically, find the occurrence of the following list of expressions: "arrangement under the flexible credit line", "letter on economic policy", "stand-by arrangement", "extended arrangement", "extended fund facility", "enhanced structural adjustment", "poverty reduction and growth", "structural adjustment facility".
3. Separate "requests" from "reviews" and isolate the number of the review excluding misleading expressions such as "request for increase in quotas", "request for enhanced article iv" or "request for postponement"
4. Find the name of the countries in the title and check for mismatch with the metadata information, correct when necessary by considering the country name specified in the title.

Appendix C (Continued)

5. Find consultations and surveillance documents: "article iv consultation", "article xiv consultation", "recent economic developments", "selected issues", "article viii", "background papers", "consultations", "exchange system", , "economic report".
6. Find technical assistance documents

This initial cleaning of the metadata to maintain only relevant documents resulted in approximately 39,000 files remaining. After downloading all the PDFs, an additional cleaning was performed to remove files with less than 5 pages and files containing specific terms in the first page.

C.1.1.2 First Page Analysis

TABLE XXXVII: List of Keywords for Problematic Documents

Keyword
minutes of executive board meeting
minutes of executive board minutes
executive board attendance
final minutes of executive board meeting
this is a working paper
working paper
a working paper of the international monetary fund
background paper
provides background to the paper
attached paper provides background information
background documentation for

Appendix C (Continued)

Keyword

draft issues paper

selected issues(?! (financial stability system assessment)_{0,1}the imfs transparency ...)

poverty reduction strategy paper

enhanced heavily indebted poor countries initiative

individual economy assessments

global financial stability report

debt sustainability analysis

triennial surveillance review

interim surveillance review

report on the observance of standard and codes

C.1.1.3 PDF Extraction

1. Convert PDFs to images: this step is necessary because Google OCR works only on images (*pdf2img* library).
2. Image Preprocessing: Remove noise from images and make text prominent to improve accuracy when applying OCR (*OpenCV* library).
3. OCR Text extraction
4. Text Preprocessing: Apply auto correction techniques to increase the accuracy and erase spelling mistakes (*spellchecker* library)
5. Saving Output Text

Appendix C (Continued)

C.1.1.4 Final Corpus

TABLE XXXVIII: Description of IMF Documents

Type of document	Purpose	Currently Issued	Regular production	Details
Article IV	Main tool of bilateral surveillance: inform about developments, vulnerabilities, outlook and furnish recommendations	Y	Y	Annually (not always)
Article VII & Article XIV	Bilateral surveillance of country maintaining exchange rate restrictions: motivate the measures and notify of any change	Y	Y	Required every 12 months: normally included in Article IV, but can be issued independently
Consultations	Bilateral surveillance	N	No info	Old version of Article IV
Exchange system	Communicate changes in par value of exchange rate	N	N	/
Recent Economic Developments	Economic and financial developments and trends in member countries	Y	No info	Internal paper for background analysis of <i>Article IV consultations</i>
Request	Set out the agreed policy goals and strategies in the economic program as well as conditionality and how observance will be monitored	Y	N	/
Review	Ascertain whether the relevant conditions for a purchase have been observed by member. If it is the case, the purchase becomes available	Y	Y (Conditional on program) ²	Prior to purchase by member

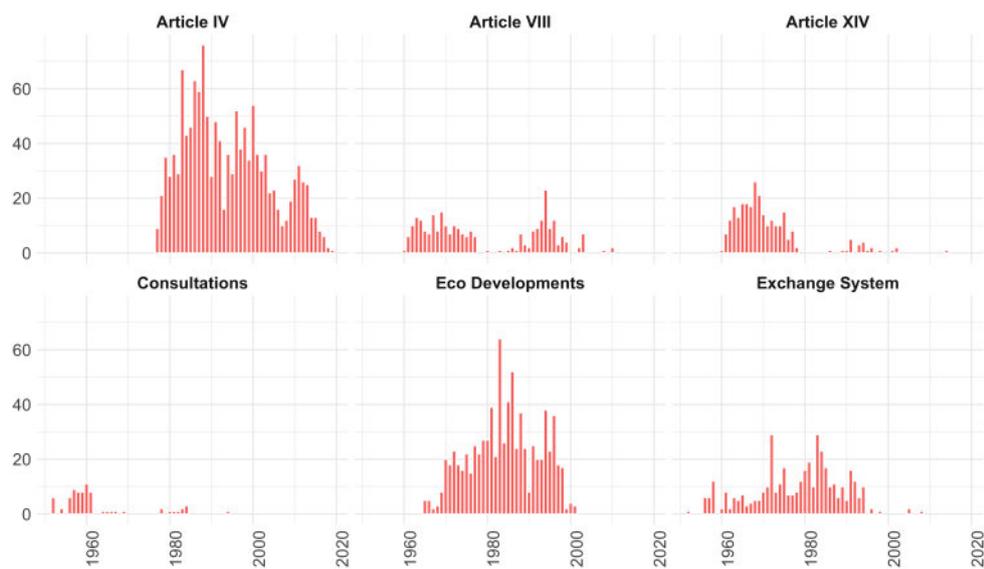
Source: Fritz-Krockow and Parmeshwar (2007)

Note: Some purchases do not require a review, e.g., in case of a Stand-By Arrangement that has quarterly purchases but semi-annual reviews

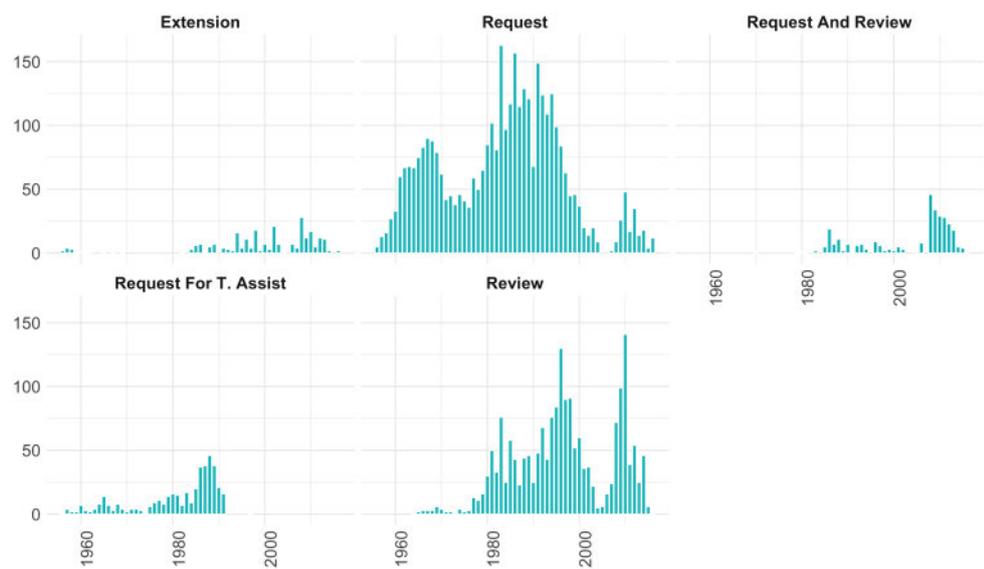
Appendix C (Continued)

Figure 42: Number of IMF Documents per Year

(a) Country Reports



(b) Program Related



Appendix C (Continued)

TABLE XXXIX: Country Coverage

	ISO3	Country name	First document	N. of documents
1	AFG	Afghanistan	1956	204
2	AGO	Angola	1989	135
3	ALB	Albania	1991	178
4	ARE	United Arab Emirates	1974	125
5	ARG	Argentina	1957	406
6	ARM	Armenia	1992	204
7	ATG	Antigua & Barbuda	1982	2
8	AUS	Australia	1949	187
9	AUT	Austria	1949	176
10	AZE	Azerbaijan	1992	133
11	BDI	Burundi	1965	223
12	BEL	Belgium	1949	220
13	BEN	Benin	1976	209
14	BFA	Burkina Faso	1984	217
15	BGD	Bangladesh	1972	268
16	BGR	Bulgaria	1990	185
17	BHR	Bahrain	1983	59
18	BIH	Bosnia & Herzegovina	1995	110

Appendix C (Continued)

	ISO3	Country name	First document	N. of documents
19	BLR	Belarus	1992	103
20	BLZ	Belize	1970	132
21	BOL	Bolivia	1949	349
22	BRA	Brazil	1946	402
23	BRB	Barbados	1971	167
24	BRN	Brunei	1973	65
25	BTN	Bhutan	1981	92
26	BWA	Botswana	1971	146
27	CAF	Central African Republic	1968	212
28	CAN	Canada	1955	191
29	CHE	Switzerland	1975	109
30	CHL	Chile	1946	335
31	CHN	China	1981	172
32	CIV	Côte d'Ivoire	1964	206
33	CMR	Cameroon	1967	221
34	COD	Congo - Kinshasa	1972	198
35	COG	Congo - Brazzaville	1969	95
36	COL	Colombia	1947	376
37	COM	Comoros	1978	144

Appendix C (Continued)

	ISO3	Country name	First document	N. of documents
38	CPV	Cape Verde	1978	175
39	CRI	Costa Rica	1947	294
40	CYP	Cyprus	1961	141
41	CZE	Czechia	1990	134
42	DEU	Germany	1953	212
43	DJI	Djibouti	1978	132
44	DMA	Dominica	1978	184
45	DNK	Denmark	1946	133
46	DOM	Dominican Republic	1959	190
47	DZA	Algeria	1966	198
48	ECU	Ecuador	1948	320
49	EGY	Egypt	1946	252
50	ERI	Eritrea	1994	31
51	ESP	Spain	1949	190
52	EST	Estonia	1992	143
53	ETH	Ethiopia	1947	236
54	FIN	Finland	1951	164
55	FJI	Fiji	1971	128
56	FRA	France	1947	243

Appendix C (Continued)

	ISO3	Country name	First document	N. of documents
57	GAB	Gabon	1967	210
58	GBR	United Kingdom	1947	307
59	GEO	Georgia	1992	187
60	GHA	Ghana	1958	339
61	GIN	Guinea	1965	242
62	GMB	Gambia	1977	196
63	GNQ	Equatorial Guinea	1971	120
64	GRC	Greece	1947	191
65	GRD	Grenada	1975	182
66	GTM	Guatemala	1948	216
67	GUY	Guyana	1967	321
68	HND	Honduras	1947	307
69	HRV	Croatia	1993	106
70	HTI	Haiti	1947	295
71	HUN	Hungary	1982	242
72	IDN	Indonesia	1949	414
73	IND	India	1946	263
74	IRL	Ireland	1958	233
75	IRN	Iran	1948	205

Appendix C (Continued)

	ISO3	Country name	First document	N. of documents
76	IRQ	Iraq	1951	127
77	ISL	Iceland	1948	240
78	ISR	Israel	1953	224
79	ITA	Italy	1949	226
80	JAM	Jamaica	1962	380
81	JOR	Jordan	1953	263
82	JPN	Japan	1953	248
83	KAZ	Kazakhstan	1992	147
84	KEN	Kenya	1965	310
85	KGZ	Kyrgyzstan	1992	182
86	KHM	Cambodia	1970	122
87	KIR	Kiribati	1986	69
88	KNA	St. Kitts & Nevis	1987	101
89	KOR	South Korea	1956	306
90	KWT	Kuwait	1963	136
91	LAO	Laos	1957	201
92	LBN	Lebanon	1952	126
93	LBR	Liberia	1963	392
94	LBY	Libya	1983	63

Appendix C (Continued)

	ISO3	Country name	First document	N. of documents
95	LCA	St. Lucia	1979	92
96	LKA	Sri Lanka	1972	279
97	LSO	Lesotho	1970	192
98	LTU	Lithuania	1992	127
99	LUX	Luxembourg	1971	80
100	LVA	Latvia	1992	130
101	MAR	Morocco	1959	315
102	MDA	Moldova	1992	130
103	MDG	Madagascar	1976	236
104	MDV	Maldives	1978	138
105	MEX	Mexico	1947	296
106	MHL	Marshall Islands	1992	41
107	MKD	Macedonia	1993	125
108	MLI	Mali	1964	312
109	MLT	Malta	1970	132
110	MMR	Myanmar (Burma)	1953	107
111	MNE	Montenegro	2006	41
112	MNG	Mongolia	1990	131
113	MOZ	Mozambique	1985	237

Appendix C (Continued)

	ISO3	Country name	First document	N. of documents
114	MRT	Mauritania	1977	193
115	MUS	Mauritius	1968	190
116	MWI	Malawi	1967	302
117	MYS	Malaysia	1964	215
118	NAM	Namibia	1989	100
119	NER	Niger	1967	246
120	NGA	Nigeria	1963	177
121	NIC	Nicaragua	1947	303
122	NLD	Netherlands	1983	127
123	NOR	Norway	1948	139
124	NPL	Nepal	1963	238
125	NZL	New Zealand	1955	165
126	OMN	Oman	1972	106
127	PAK	Pakistan	1950	397
128	PAN	Panama	0001	289
129	PER	Peru	1946	423
130	PHL	Philippines	1949	390
131	PLW	Palau	1999	23
132	PNG	Papua New Guinea	1974	148

Appendix C (Continued)

	ISO3	Country name	First document	N. of documents
133	POL	Poland	1986	203
134	PRT	Portugal	1962	208
135	PRY	Paraguay	1946	275
136	QAT	Qatar	1973	94
137	ROU	Romania	1973	298
138	RUS	Russia	1992	224
139	RWA	Rwanda	1962	258
140	SAU	Saudi Arabia	1958	54
141	SDN	Sudan	1958	505
142	SEN	Senegal	1963	274
143	SGP	Singapore	1966	146
144	SLB	Solomon Islands	1978	127
145	SLE	Sierra Leone	1964	379
146	SLV	El Salvador	1957	235
147	SMR	San Marino	1992	45
148	SOM	Somalia	1964	62
149	STP	São Tomé & Príncipe	1978	41
150	SUR	Suriname	1978	104
151	SVK	Slovakia	1993	86

Appendix C (Continued)

	ISO3	Country name	First document	N. of documents
152	SVN	Slovenia	1993	64
153	SWE	Sweden	1952	167
154	SWZ	Swaziland	1983	84
155	SYC	Seychelles	1972	160
156	SYR	Syria	1948	196
157	TCD	Chad	1967	190
158	TGO	Togo	1965	211
159	THA	Thailand	1949	268
160	TJK	Tajikistan	1992	107
161	TKM	Turkmenistan	1992	46
162	TON	Tonga	1985	105
163	TTO	Trinidad & Tobago	1962	172
164	TUN	Tunisia	1957	234
165	TUR	Turkey	1947	406
166	TUV	Tuvalu	2011	14
167	TZA	Tanzania	1965	269
168	UGA	Uganda	1965	301
169	UKR	Ukraine	1992	210
170	URY	Uruguay	1949	408

Appendix C (Continued)

	ISO3	Country name	First document	N. of documents
171	USA	United States	1947	221
172	UZB	Uzbekistan	1992	74
173	VCT	St. Vincent & Grenadines	1980	108
174	VEN	Venezuela	1946	182
175	VNM	Vietnam	1958	274
176	VUT	Vanuatu	1981	83
177	WSM	Samoa	1973	165
178	YEM	Yemen	1970	237
179	ZAF	South Africa	1948	211
180	ZMB	Zambia	1967	424
181	ZWE	Zimbabwe	1980	237

C.1.2 The Lexicon of Crises

TABLE XL: Full Lexicon

Category	Keyword
B.o.P.	shortage of foreign exchange
B.o.P.	bop crisis
B.o.P.	balance of payment crisis

Appendix C (Continued)

Category	Keyword
B.o.P.	capital account crisis
B.o.P.	balance of payment crisis
B.o.P.	balance of payment problem
B.o.P.	balance of payment difficulties
B.o.P.	cessation of official foreign capital inflows
B.o.P.	decline in net international reserves
B.o.P.	pressures in the official foreign exchange market
B.o.P.	external account came under pressure
B.o.P.	external account came under severe pressure
B.o.P.	external account came under serious pressure
B.o.P.	balance of payments problems
B.o.P.	shortage of international reserves
B.o.P.	sharp reduction in international reserves
B.o.P.	strong decline in international reserves
B.o.P.	international reserves exhausted
B.o.P.	decline in reserves
B.o.P.	drop in reserves
B.o.P.	loss of official reserves

Appendix C (Continued)

Category	Keyword
B.o.P.	decline in net capital inflows
B.o.P.	decline in international reserves
B.o.P.	decline in official reserves
B.o.P.	official international reserves exhausted
B.o.P.	major loss in net international reserves
B.o.P.	foreign exchange scarcity
B.o.P.	decline in receipts of official foreign loans
B.o.P.	exhaustion of the disposable official international reserves
B.o.P.	capital flight
B.o.P.	flight of capital
B.o.P.	pull-back of capital
B.o.P.	capital flow reverse
B.o.P.	capital flow reversal
B.o.P.	pressure on capital flows
B.o.P.	large capital outflows
B.o.P.	strong balance of payment pressures
B.o.P.	balance of payment assistance
B.o.P.	depleted international reserves

Appendix C (Continued)

Category	Keyword
B.o.P.	large external financing needs
B.o.P.	substantial capital outflows
B.o.P.	unforeseen balance of payments contingencies
B.o.P.	tail risks to the balance of payments
B.o.P.	large balance of payments imbalances
B.o.P.	exhausted official international reserves
B.o.P.	pressure on the capital account
B.o.P.	exceptional balance of payments need
B.o.P.	balance of payment sustainability
B.o.P.	reversal in the flow of private capital
B.o.P.	sharpe reduction in access to international capital markets
B.o.P.	sharp fall in private inflows
B.o.P.	decline in net inflow
B.o.P.	severe external imbalances
B.o.P.	severe internal and external imbalances
Banking	bank resolution
Banking	bank crisis
Banking	banking sector restructuring

Appendix C (Continued)

Category	Keyword
Banking	restructuring of nonperforming loans
Banking	undercapitalized banking system
Banking	weak bank capitalization
Banking	reorganization of the banking sector
Banking	restructuring of the banking
Banking	fragility of the banking sector
Banking	fragile banking sector
Banking	banking crisis
Banking	banking system restructuring
Banking	insolvent banks
Banking	insolvent banking sector
Banking	bailout
Banking	crisis in the banking sector
Banking	take over of private banks
Banking	private banks taken over
Banking	recapitalize private banks
Banking	collapse of the banking sector
Banking	increase in nonperforming loans

Appendix C (Continued)

Category	Keyword
Banking	recapitalization of the banks
Banking	recapitalizing the banking system
Banking	recapitalizing the banking sector
Banking	banking system collapsed
Banking	additional nonperforming loans
Banking	collapsed in the banking system
Banking	banking system stability
Banking	pressure on the banking
Banking	bankrun
Banking	bank recapitalization
Banking	deteriorating credit quality
Banking	recapitalization
Banking	bank restructuring
Banking	recapitalize private financial institutions
Banking	confidence in the domestic banking system
Banking	strengthen bank supervision
Banking	financial support package
Commodity	oil crisis

Appendix C (Continued)

Category	Keyword
Commodity	rice crisis
Commodity	crop crisis
Commodity	crop failure
Commodity	commodity crisis
Commodity	energy crisis
Commodity	cotton crisis
Commodity	crisis in the cotton
Commodity	severe shortages ^{1} of rice
Commodity	fall in prices of raw materials
Commodity	price of copper continue to drop
Commodity	swing in copper price
Commodity	weakness in the copper price
Commodity	adverse movement in the price of copper
Commodity	decline in coffee prices
Commodity	decline in international coffee prices
Commodity	drop in world coffee price
Commodity	fell of agricultural prices
Commodity	tourism.*suffer

Appendix C (Continued)

Category	Keyword
Commodity	terms-of-trade shock
Commodity	deterioration in the terms of trade
Commodity	deteriorating terms of trade
Commodity	adverse terms of trade
Commodity	terms of trade loss
Commodity	unfavorable terms of trade
Commodity	severe drop in terms of trade
Commodity	severe terms of trade drop
Commodity	severe terms of trade shock
Commodity	significant terms of trade loss
Commodity	sharp fall in its terms of trade
Commodity	large terms of trade loss
Commodity	adverse movement in the terms of trade
Commodity	terms of trade were adversely affected
Commodity	dependence on oil-related revenue
Commodity	budgetary dependency on oil revenue
Commodity	increase in world oil prices
Commodity	drop in world coffee price

Appendix C (Continued)

Category	Keyword
Commodity	oil price increase
Commodity	fluctuations in oil prices
Commodity	increase in petroleum price
Contagion	regional crisis
Contagion	crisis in the region
Contagion	spillovers from the global crisis
Contagion	systemic crisis
Contagion	crisis in emerging economies
Contagion	regional financial crisis
Contagion	spillovers from the global crisis
Contagion	vulnerable to external shocks
Contagion	crisis spillover
Contagion	regional economic situation turned adverse
Contagion	contagion from the crisis in neighboring
Contagion	external shocks
Contagion	external shock
Contagion	adverse exogenous events
Contagion	external vulnerability

Appendix C (Continued)

Category	Keyword
Contagion	exogenous events
Contagion	contagion
Contagion	fears of contagion
Contagion	spillovers
Contagion	vulnerability to international
Contagion	russian debt crisis
Contagion	asian currency crisis
Contagion	crisis in southeast asia
Contagion	southeast asia crisis
Contagion	crisis in russia
Contagion	crisis in libya
Contagion	libya crisis
Contagion	regional currency crisis
Contagion	kosovo crisis
Contagion	cyprus crisis
Contagion	crisis in ukraine
Contagion	regional dimension of the crisis
Contagion	mexican exchange crisis

Appendix C (Continued)

Category	Keyword
Contagion	gulf crisis
Contagion	middle east crisis
Contagion	mexican crisis
Contagion	crisis in argentina
Contagion	crisis in russia
Contagion	argentine crisis
Contagion	crisis in mexico
Contagion	the crisis of 1994
Contagion	the 1997 crisis
Contagion	the crisis in 2002
Contagion	2002 crisis
Contagion	euro area crisis
Contagion	eurozone contagion
Contagion	eurozone crisis
Contagion	crisis in europe
Contagion	world financial crisis
Contagion	greek crisis
Contagion	brazil crisis

Appendix C (Continued)

Category	Keyword
Contagion	asian and russian crisis
Contagion	asia crisis
Contagion	crisis in turkey
Contagion	argentinan crisis
Contagion	crisis in argentina
Contagion	crisis in greece
Contagion	asian crisis
Contagion	global economic crisis
Contagion	global financial shock
Contagion	international systemic spillover
Contagion	crisis in brazil
Contagion	linkage with the us
Contagion	contagion effects of the thai crisis
Currency	exchange rate crisis
Currency	large real depreciation
Currency	foreign exchange crisis
Currency	severe disruption of exchange markets
Currency	major devaluation

Appendix C (Continued)

Category	Keyword
Currency	currency crisis
Currency	currency crash
Currency	large devaluation
Currency	large depreciation
Currency	sharp depreciation
Currency	sharp depreclatlon
Currency	currency attack
Currency	exchange rate crisis
Currency	unsuccessful attempt to.*maintain the exchange rate unchanged
Currency	foreign currency turmoil
Eco. recession	severe economic crisis
Eco. recession	very difficult economic circumstances
Eco. recession	severe recession
Eco. recession	severe crisis
Eco. recession	economic crisis
Eco. recession	steep recession
Eco. recession	strong recessionary headwinds
Eco. recession	sharp slowdown

Appendix C (Continued)

Category	Keyword
Eco. recession	sharp declines in output
Eco. recession	significant loss of output
Eco. recession	economic collapse
Eco. recession	deeper recession
Eco. recession	deepening recession
Eco. recession	painful recession
Eco. recession	prolonged recession
Eco. recession	lengthening recession
Eco. recession	severity of the recession
Eco. recession	economic recession
Eco. recession	sharp contraction of economic activity
Eco. recession	strong contraction of economic activity
Eco. recession	large contraction of economic activity
Eco. recession	deep recession
Eco. recession	large economic slowdown
Eco. recession	severe recession
Eco. recession	profond recession
Eco. recession	contraction in output

Appendix C (Continued)

Category	Keyword
Eco. recession	deep recession
Eco. recession	severe contraction
Eco. recession	deep contraction
Eco. recession	profond contraction
Eco. recession	large decline in income per capita
Eco. recession	deep economic downturn
Eco. recession	severe economic downturn
Eco. recession	deep economic downturn
Eco. slowdown	slowdown in the economic activity
Eco. slowdown	slowdown in economic growth
Eco. slowdown	slowdown of the economy
Eco. slowdown	slowdown of output
Eco. slowdown	economic decline
Eco. slowdown	activity remains weak
Eco. slowdown	the economy slowed down
Eco. slowdown	declining trend in economic activity
Eco. slowdown	decline in economic activity
Eco. slowdown	slowing down of business activity

Appendix C (Continued)

Category	Keyword
Eco. slowdown	slow down
Eco. slowdown	low rates of economic growth
Eco. slowdown	low rate of economic growth
Eco. slowdown	economic activity on a downward trend
Eco. slowdown	depressed level of economic activity
Eco. slowdown	the economic situation worsen
Eco. slowdown	slowing the pace of economic recovery
Eco. slowdown	decline in economic activity
Eco. slowdown	weakening of economic fundamental
Eco. slowdown	recession
Eco. slowdown	contraction of output
Eco. slowdown	sluggish recovery
Eco. slowdown	contraction of economic activity
Eco. slowdown	economic downturn
Eco. slowdown	output is estimated to have contracted
Eco. slowdown	slowdown in the economic activity
Eco. slowdown	slowdown of output
Eco. slowdown	slow economic activity

Appendix C (Continued)

Category	Keyword
Epidemics	epidemic
Epidemics	epidemia
Epidemics	pandemia
Epidemics	pandemic
Epidemics	virus
Epidemics	infection
Epidemics	\sflu\s
Epidemics	relapsing fever
Epidemics	typhoid fever
Epidemics	leishmaniasis
Epidemics	dengue
Epidemics	mumps
Epidemics	meningitis
Epidemics	poliomyelitis
Epidemics	measles
Epidemics	zika
Epidemics	encephalitis
Epidemics	\ssars\s

Appendix C (Continued)

Category	Keyword
Epidemics	\smers\s
Epidemics	nipah
Epidemics	vcjd
Epidemics	\shiv\s
Epidemics	hiv/aids
Epidemics	typhus
Epidemics	hepatitis
Epidemics	h1n1
Epidemics	h5n1
Epidemics	ebola
Epidemics	\ssida\s
Epidemics	rotavirus
Epidemics	\slyme\s
Epidemics	hepatite
Epidemics	chikungunya
Epidemics	dysentaria
Epidemics	dysentery
Epidemics	smallpox

Appendix C (Continued)

Category	Keyword
Epidemics	yellow fever
Epidemics	cholera
Epidemics	malaria
Epidemics	coronavirus
Epidemics	covid 19
Epidemics	\splague\s
Expectations	crisis risks
Expectations	market reversal
Expectations	economic sentiment remains poor
Expectations	market sentiment has collapsed
Expectations	increase uncertainty in the international environment
Expectations	heightened risk aversion
Expectations	high level of risk
Expectations	general uncertainty
Expectations	crisis of confidence
Expectations	risk of crisis
Expectations	confidence crisis
Expectations	panic

Appendix C (Continued)

Category	Keyword
Expectations	potential risks
Expectations	upward risk
Expectations	market confidence
Expectations	high risk
Expectations	downside risks
Expectations	increase the risks
Expectations	self fulfilling crises
Expectations	potential risks
Expectations	restoring market confidence
Expectations	major risks
Expectations	heightening risks
Expectations	deterioration in market sentiment
Expectations	increase uncertainty in the international environment
Expectations	deterioration in market sentiment
Expectations	weakening of investor confidence
Expectations	market confidence
Expectations	uncertainty in international capital markets
Expectations	uncertainty among market participant

Appendix C (Continued)

Category	Keyword
Expectations	change in expectations
Expectations	speculative capital movements
Expectations	speculative attack
Expectations	uncertainty among market participant
Expectations	a time of heightened global uncertainty
Expectations	change in investors sentiment
Expectations	reassure the markets
Expectations	extreme global risk aversion
Expectations	provide assurances to financial markets
Expectations	restore market confidence
Expectations	reduce market uncertainty
Expectations	bolster confidence
Expectations	economic credibility
Expectations	slump in confidence
Expectations	undermining confidence
Expectations	confidence crisis
Expectations	signals to markets
Expectations	market confidence sagged

Appendix C (Continued)

Category	Keyword
Expectations	vulnerable to abrupt swings in market sentiment
Expectations	heightened risk aversion
Expectations	increase in global risk aversion
Expectations	weakening of market confidence
Expectations	vulnerable to changes in the international investment climate
Expectations	confidence in the liquidity of the foreign exchange market
Expectations	increase uncertainty in the international environment
Expectations	pressures on confidence
Expectations	self-fulfilling
Expectations	shifts in investor sentiment
Expectations	bolstering market confidence
Expectations	confidence crisis
Financial	financial stability crisis
Financial	international monetary crisis
Financial	crisis in financial market
Financial	financial risks
Financial	turmoil in financial markets
Financial	turmoil in international financial markets

Appendix C (Continued)

Category	Keyword
Financial	volatility in financial markets
Financial	restore the strength of the financial sector
Financial	unfolding financial crisis
Financial	global market sell-off
Financial	global financial shock
Financial	financial shock
Financial	financial contagion
Financial	financial crisis
Financial	collapse of financial markets
Financial	fire sells
Financial	collapse of equity prices
Financial	financial market panic
Financial	global financial turbulence
Financial	viability and health of the financial sector
Housing	home prices have been declining
Housing	drops in real estate prices
Housing	house price trends
Housing	home-price overvaluation

Appendix C (Continued)

Category	Keyword
Housing	real house prices declining
Housing	foreclosures
Housing	house price inflation
Housing	house-price inflation
Housing	foreclosures
Housing	bust in housing
Housing	home-price declines
Housing	house-price declines
Housing	house prices fall
Housing	stalling house prices
Housing	slower house price
Housing	slowing housing wealth
Housing	declines in house prices
Housing	headwinds from housing
Housing	problems in housing
Housing	housing downturn
Housing	cooling housing market
Housing	cooling in the housing market

Appendix C (Continued)

Category	Keyword
Housing	change in housing wealth
Housing	deceleration in house prices
Housing	slowdown in the housing market
Housing	housing slowdown
Housing	house prices seemed overvalued
Housing	housing boom
Housing	falling house prices
Housing	spillovers from the housing market
Housing	spillovers from housing
Housing	housing market weakness
Housing	slowdown in the housing market
Housing	subprime
Housing	residential investment has declined rapidly
Inflation	inflation pressure
Inflation	inflationary pressure
Inflation	high. {0,10}inflation
Inflation	high rate of inflation
Inflation	severe. {0,10}inflation

Appendix C (Continued)

Category	Keyword
Inflation	large.{0,10}inflation
Inflation	virulence.{0,10}inflation
Inflation	unprecedented.{10}inflation
Inflation	sharp.{0,2}increase in domestic prices
Inflation	large increase in.{0,10}prices
Inflation	high pressure on.{0,10}prices
Inflation	inflation.*critical
Inflation	inflation.*unprecedented levels
Inflation	despite the acceleration of inflation
Inflation	the rate of inflation accelerated sharply
Inflation	inflation crisis
Inflation	hyperinflation
Inflation	large monetary creation
Inflation	combat inflation
Inflation	halting inflation
Inflation	halt to inflation
Inflation	efforts against inflation
Inflation	quick reduction.*inflation

Appendix C (Continued)

Category	Keyword
Inflation	inflation down quickly
Inflation	lowering the rate of inflation
Inflation	entrenchment of inflationary behavior
Migration	refugee
Migration	migrant
Migration	inward migration
Migration	population inflow
Migration	asylum
Migration	immigrant
Migration	immigration
Nat. disaster	flood
Nat. disaster	drought
Nat. disaster	rainfall
Nat. disaster	torrential rains
Nat. disaster	natural calamities
Nat. disaster	power shortage
Nat. disaster	natural disaster
Nat. disaster	earthquake

Appendix C (Continued)

Category	Keyword
Nat. disaster	hurricane
Nat. disaster	typhoon
Nat. disaster	cyclone
Nat. disaster	calamity
Nat. disaster	adverse weather conditions
Nat. disaster	tsunami
Political	political turmoil
Political	internal security situation
Political	political atmosphere
Political	political crisis
Political	political uncertainty
Political	political instability
Political	political transition spillovers
Political	political turn-over
Political	policies risks
Political	political turmoil
Political	political risk
Political	unstable political

Appendix C (Continued)

Category	Keyword
Political	political instability
Political	poor governance
Political	disturbed political conditions
Political	political and economic developments
Political	political and security situation
Political	economic and political situation
Political	political crisis
Political	unsettled political situation
Political	political tensions
Political	geopolitical events
Political	policy-related uncertainty
Political	policy related uncertainty
Political	geopolitical risk
Political	election related uncertainty
Political	election related uncertainties
Political	governance issues
Political	complex geopolitical situation
Political	geopolitical tensions

Appendix C (Continued)

Category	Keyword
Political	geopolitical turmoil
Political	weak governance
Political	adverse geopolitical events
Political	adverse geopolitical
Political	unexpected political events
Political	revolution
Political	uncertain policies
Political	uncertainty about policy
Political	political contagion
Political	euro exit
Political	exit of the eurozone
Political	uncertain national election
Political	political transition
Political	political pressures
Political	change of administration
Political	risks linked to the electoral calendar
Political	uncertainty surrounding the outcome of the presidential election
Political	uncertainty regarding the political transition

Appendix C (Continued)

Category	Keyword
Political	domestic political developments
Political	political risk
Political	facilitate an orderly transition to a new administration
Political	uncertainty about the continuity of policies
Political	uncertainty regarding future policies
Political	unstable political environment
Political	military coup
Political	coup d'etat
Political	annulment of the election
Political	parliamentary upheavals
Political	critical political.*juncture
Political	lack of an approved government
Social	social risk
Social	social strain
Social	social.*turmoil
Social	social disruption
Social	social climate as deteriorate
Social	social tension

Appendix C (Continued)

Category	Keyword
Social	protest
Social	railroad-transport strike
Social	deteriorating social climate
Social	blockade
Social	social unrest
Social	walkouts
Social	events of may-june 1968
Sovereign	rescheduled debt
Sovereign	external payments crisis
Sovereign	difficulties in servicing its external debt
Sovereign	difficult time in rolling over its debt
Sovereign	rescheduling of external debt
Sovereign	rescheduling agreement
Sovereign	suspend service payments
Sovereign	fiscal crisis
Sovereign	debt relief
Sovereign	failure to roll over debt
Sovereign	government bonds crisis

Appendix C (Continued)

Category	Keyword
Sovereign	government bonds crisis
Sovereign	bond crisis
Sovereign	debt reprofiling
Sovereign	sovereign debt crisis
Sovereign	public debt crisis
Sovereign	default risks
Sovereign	self fulfilling crises
Sovereign	debt restructuring program
Sovereign	gouvernement default
Sovereign	restructuring of debt
Sovereign	suspension of payments
Sovereign	debt swap
Sovereign	debt restructuring
Sovereign	debt rescheduling
Sovereign	debt service reduction
Sovereign	debt restructuring program
Sovereign	rescheduling of the debt
Sovereign	arrears

Appendix C (Continued)

Category	Keyword
Sovereign	rescheduling of arrears
Sovereign	arrears in the payment
Sovereign	restructuring of its external debt
Sovereign	restructuring agreements
Sovereign	external payment arrears
Sovereign	debt service reduction
Sovereign	no debt service payments
Sovereign	relation with external creditors
Sovereign	paris club
Sovereign	club of paris
Sovereign	debt relief
Sovereign	debt exchange
Trade	trade war
Trade	trade policy tension
Trade	trade tension
Trade	trade conflict
Trade	escalation of trade restrictions
Trade	disruption of trade

Appendix C (Continued)

Category	Keyword
Trade	trade crisis
Trade	trade restrictions
Trade	trade volatility
Trade	weak trade
Trade	disruption to trade
Trade	slowdown in trade
Trade	trade restricting measure
Trade	decline in fdi
Trade	decline in foreign direct investment
Trade	fdi flows declined
Trade	fdi have declined
Trade	trade issues
Trade	slowdown in trade
Trade	trade slowdown
Trade	uncertainty about future trade policy
Trade	uncertainty about trade policy
Trade	trade restraints
Trade	trade policy unpredictability

Appendix C (Continued)

Category	Keyword
Trade	slowdown in global trade
Trade	stronger competition of countries
Violence	war damage
Violence	insurgency crisis
Violence	security crisis
Violence	civil conflict
Violence	civil war
Violence	ensuing conflict
Violence	armed conflict
Violence	armed internal conflict
Violence	armed domestic conflict
Violence	ongoing conflict
Violence	violent conflict
Violence	atlantic conflict
Violence	internal conflict
Violence	regional conflict
Violence	conflicts in the region
Violence	conflict zone

Appendix C (Continued)

Category	Keyword
Violence	conflict regions
Violence	military coup
Violence	military take-over
Violence	coup d'etat
Violence	escalated attacks
Violence	breakdown of cease-fire
Violence	ethnic rivalries
Violence	terrorist attacks
Violence	terrorism
Violence	guerilla offensive
Violence	continuing external aggression
Violence	incidence de la guerre
World	world-wide recession
World	global economic crisis
World	global crisis
World	world recession
World	worldwide recession
World	international crisis

Appendix C (Continued)

Category	Keyword
World	global financial crisis
World	deep international recession
World	international downturn
World	international recession
World	ongoing global downturn
World	deterioration of external environment
World	weakening of international economic activity
World	turbulence in international markets
World	external conditions deteriorated markedly
World	unfavorable developments in the international economic environment
World	deterioration of external environment
World	recession in the world economy
World	international monetary crisis
World	worsening international environment
World	difficult external environment
World	downside risks in the international environment
World	further deterioration in the international environment
World	uncertain external environment

Appendix C (Continued)

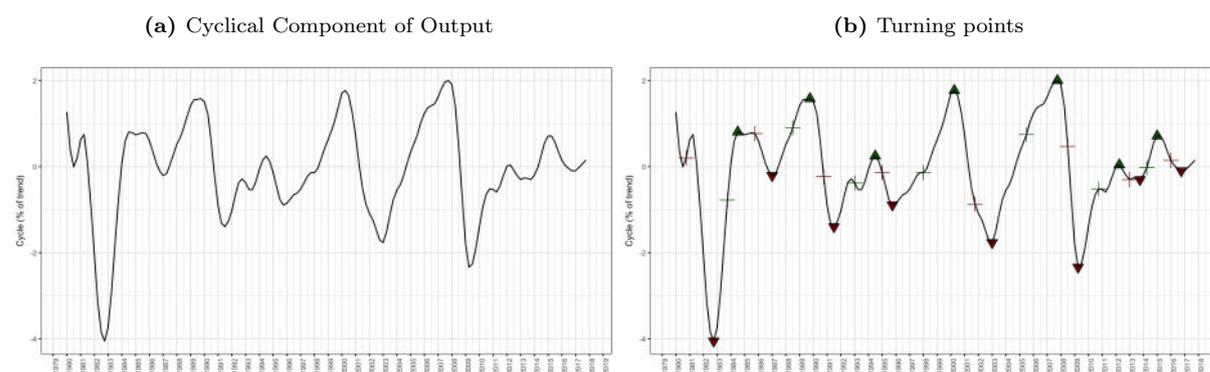
Category	Keyword
World	slowdown in international economy
World	fragile global outlook
World	international financial turmoil
World	sharply deteriorating external conditions

Note: Authors' own elaboration.

C.2 A Database On Crises Discussion

C.2.1 Comparison with Benchmarks

Figure 43: Locating Turning Points in Economic Activity: United States



Source: Analytical Database, OECD and Authors' Computations

Appendix C (Continued)

TABLE XLI: Comparison with Benchmark: Soft Recession

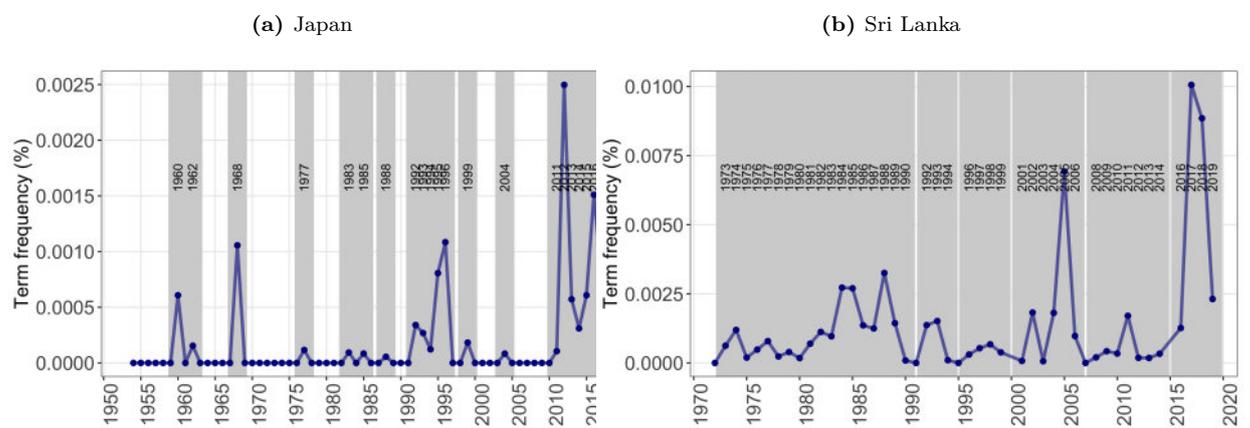
	<i>Dependent variable:</i>						
	g	g	g<0	g<0	Phase B	Large B	Phase B2
Y soft	-0.73*** (0.14)	-0.31*** (0.11)					
Y soft>0			0.03*** (0.01)	0.08*** (0.01)	0.02 (0.02)	0.003 (0.01)	0.04** (0.01)
Constant	3.89*** (0.31)			0.05*** (0.01)			
Country FE	No	Yes	Yes	No	No	No	No
Time FE	No	Yes	Yes	No	No	No	No
Controls	No	Yes	Yes	No	No	No	No
Robust se	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F	25.83***	3.09***	8.33***	45.65***	2.97***	2.45***	4.06***
Observations	2,061	2,060	2,061	2,061	2,098	2,021	2,098

Note: Y soft is the soft recession term-frequency. G corresponds to real GDP growth rate. Y soft>0 and g<0 are dummies equal to 1 when the condition is satisfied. Phase B, Large B and Phase B2 are dummies indicating the cyclical component of real GDP obtained following [Harding and Pagan \(2002\)](#). Phase B is equal to 1 for all the years in between the peak and the trough of the cycle. Large B indicates the downturn phases with the largest amplitude. Phase B2 refers to the second half of the downturn. GDP data are from the Analytical Database of the OECD. ***: significant at 1% level, **: significant at 5% level, *: significant at 10% level.

Appendix C (Continued)

C.2.2 Zeroing In Non-Economic Indicator

Figure 44: Examples of Natural Disasters Indicator



Note: The blue line corresponds to the natural disaster term-frequency for, respectively, Japan and Sri Lanka. Shaded gray areas are years of strictly positive term-frequency.

Japan has a long history of natural disasters: the geographic position of the archipelago predispose it to different kinds of natural calamities. Four peaks stand out in the last 70 years: 1960, 1968, 1995, 2011: the 1960 peaks correspond to the Typhoon Vera, the 1995 peak to the Kobe earthquake and the 2011 one to the Tōhoku earthquake and the related meltdown of the Fukushima nuclear power plant. All these events constituted major setbacks in the Japanese economy. Sri Lanka displays a startling pattern: a constant vulnerability to natural disaster, with discussions happening almost every year

Appendix C (Continued)

since 1970. Among the largest events peaks, we identify the 2004 Indian Ocean earthquake and the 2017 extensive floods that wrought havoc the country .