

Gatekeepers of the Market? IMF Surveillance and Sovereign Access to Global Finance

Pablo Aguilar Perez

2025-32 Document de Travail/ Working Paper

EconomiX - UMR 7235 Bâtiment Maurice Allais
Université Paris Nanterre 200, Avenue de la République
92001 Nanterre Cedex

Site Web : economix.fr
Contact : secreteriat@economix.fr
Twitter : @EconomixU



Gatekeepers of the Market? IMF Surveillance and Sovereign Access to Global Finance

Pablo Aguilar Pérez^{a,*}

^a*EconomiX-CNRS, University Paris Nanterre*

Abstract

We study the effect of International Monetary Fund Article IV Public Information Notices on sovereign financing conditions. Using monthly data from 67 emerging market economies between 2004 and 2023, we estimate the causal impact of these surveillance disclosures through a dynamic panel matching framework that accounts for repeated observations. The release of Article IV statements leads to a statistically significant reduction in sovereign bond spreads and is associated with increased debt issuance and reserve accumulation, suggesting improved access to external finance. These effects are more pronounced when the surveillance message is optimistic and among countries more closely aligned geopolitically with major IMF shareholders, underscoring the importance of both tone and credibility context in shaping investor reactions. By isolating the influence of surveillance communications—independent of the Fund’s lending activities—this study contributes to the literature on the informational role of international organizations and the non-financial channels through which they affect sovereign risk and market behavior.

Keywords: IMF Surveillance; Dynamic Panel Matching; Geopolitical Alignment; Sentiment Analysis; Sovereign Risk

JEL: C23; F33; G15; H63

^{*}EconomiX-CNRS, University Paris Nanterre, 200 avenue de la République, 92001 Nanterre, France. ORCID ID: 0000-0002-2336-3870. Email: paguilar@parisnanterre.fr

1. Introduction

The International Monetary Fund (IMF) plays a key role in shaping global macroeconomic governance through its surveillance activities. Surveillance is widely regarded as a critical tool through which the IMF encourages compliance with policies that promote global stability and facilitate the macroeconomic coordination essential to sustaining it ([Mussa, 1997](#)). One of the core mechanisms the IMF uses to conduct surveillance is the Article IV consultation—a periodic, comprehensive assessment of a member country’s economic policies, institutional frameworks, and macroeconomic outlook ([Mussa, 1997](#); [Bird and Rowlands, 2017](#); [Momani and Hibben, 2024](#)).

These consultations typically culminate in the publication of Public Information Notices (PINs), which summarize IMF staff assessments, policy recommendations, and evaluations of macroeconomic risks ([Fratzscher and Reynaud, 2011](#)). While PINs are designed to enhance transparency and bolster policy credibility, these official communications may also have tangible financial consequences—particularly in sovereign bond markets, where country risk is shaped not only by fundamentals but also by assessments issued by international institutions.

Despite growing recognition of the IMF’s signaling role, empirical research has primarily focused on the effects of its lending programs, with relatively limited attention to surveillance ([Eichengreen et al., 2006](#); [Chapman et al., 2017](#); [Reinsberg et al., 2022](#); [Shim, 2022](#)). Among the few studies that do examine surveillance, methodological limitations persist. Most rely on panel regressions or event-study designs that fail to adequately address selection into surveillance or the dynamic and repeated nature of treatment timing ([Fratzscher and Reynaud, 2011](#); [Breen et al., 2020](#)). As a result, the causal impact of surveillance disclosures—such as PINs—on market outcomes remains insufficiently identified.

This article addresses these limitations by employing the generalized panel matching framework developed by [Imai et al. \(2023\)](#), which is specifically designed for settings with dynamic, repeated treatments. We adapt this approach to estimate the causal effect of IMF surveillance activities—measured by the release of PINs—on sovereign borrowing conditions. This framework is especially appropriate for our context, as it allows countries to be treated multiple times at irregular intervals, reflecting the institutional reality of recurring Article IV consultations. In addition to accommodating repeated treatments, the method facilitates identification of both pre-treatment trends and post-treatment dynamics. By implementing this flexible and credible identification strategy, we offer new evidence on how IMF surveillance communications influence financial markets—providing a more rigorous basis for assessing their signaling effects.

Using monthly data for 67 emerging market economies (EMEs) from 2004 to 2023, we find that Article IV PIN releases lead to a statistically significant reduction in sovereign bond spreads, consistent with improved investor confidence following public surveillance disclosures. These releases are also associated with increased sovereign debt issuance and reserve accumulation, indicating that the IMF’s signaling role can ease access to external finance. To validate these findings, we implement a series of robustness checks, including a future-leads falsification test, alternative treatment definitions using IMF Staff Reports, sensitivity analyses varying the time window and accounting for global shocks, such as the Global Financial Crisis and COVID-19. All results remain consistent, reinforcing the informational value of surveillance.

To examine the mechanisms behind market reactions, we test whether the effect of Article IV PINs varies systematically along two dimensions: the tone of the surveillance message and the geopolitical alignment of the recipient country. These factors may condition how investors interpret and respond to IMF communications. First, the tone—whether optimistic or critical—likely shapes expectations regarding a country’s macroeconomic trajectory. A favorable PIN may signal policy credibility and macroeconomic stability, lowering risk premiums. Conversely, a pessimistic tone may raise concerns over fiscal or external vulnerabilities. We quantify tone using a sentiment index derived from FinBERT, a language model trained on financial text (Yang et al., 2020).

Second, we explore how market reactions vary with geopolitical alignment, measured by United Nations General Assembly voting similarity (Fjelstul et al., 2025). Countries more closely aligned with key IMF shareholders—particularly the United States—may receive a credibility premium, whereby IMF endorsement is interpreted as a signal of broader political and financial backing. In contrast, countries with weaker geopolitical ties may face muted responses, as investors question the impartiality or weight of IMF surveillance.

Our findings suggest that IMF surveillance affects financial markets not only through the content of its assessments, but also through the credibility context in which they are received. Optimistic PINs and those issued in geopolitically aligned countries generate stronger market responses, indicating that both informational clarity and political proximity shape the interpretation and effectiveness of IMF signaling.

This article contributes to the literature in two important ways. First, it introduces a methodological innovation by applying the dynamic panel matching framework of Imai et al. (2023) to the study of IMF surveillance. This approach enables credible causal inference in a setting characterized by repeated and non-random treatment, addressing both selection and treatment timing heterogeneity. Second, it advances our understanding of the mechanisms

through which public IMF communications influence sovereign markets. By examining heterogeneity in market responses by tone and geopolitical alignment, the paper identifies when and for whom IMF surveillance has the greatest impact. This mechanism-based perspective highlights the non-financial influence of international organizations and their capacity to shape sovereign market access through surveillance-based communication—independent of lending or crisis interventions.

The paper proceeds as follows. Section 2 reviews the literature on IMF surveillance, policy signaling, and the financial market effects of public communication. Section 3 provides the institutional background on IMF Article IV consultations and PIN disclosures, and introduces the dataset used in the analysis. Section 4 details the empirical strategy. Section 5 presents baseline estimates of the impact of PIN releases on sovereign bond spreads, debt issuance, and reserve movements and conducts a series of robustness checks. Section 6 explores the underlying mechanisms, focusing on the informational content of sentiment and geopolitical alignment as moderators of investor response. Section 7 concludes.

2. Literature review and contribution

2.1. IMF surveillance and sovereign market access

A broader body of literature portrays the IMF as a multi-channel actor, shaping sovereign bond-market access through lending, surveillance, capacity development, and debt-restructuring advice. Together, these strands form a coherent narrative of how Fund interventions influence both the likelihood of tapping capital markets and the cost of borrowing.

Early work documented the catalytic effect of IMF programs on bond spreads, conditional on borrower characteristics (Mody and Saravia, 2006). Subsequent research emphasized that it is the credibility of conditionality (Chapman et al., 2017) and the borrower’s domestic political conditions (Shim, 2022) that drive market reactions. Building on these findings, Kogan et al. (2024) demonstrate that IMF program participation increases the probability of restoring market access for financially distressed sovereigns. Complementary evidence indicates that technical assistance related to debt-market reforms (IMF, 2022) and IMF-supported debt restructurings (Schlegl et al., 2019) both lower bond spreads and accelerate market reaccess.

This paper contributes to the literature on the relationship between IMF surveillance and sovereign access to international debt markets by providing a causal estimate of how the content and timing of surveillance communications influence borrowing conditions. We draw on two strands of existing research that form the foundation for our study.

First, Breen and Doak (2023) employ an event-study framework to examine how financial

markets respond to IMF surveillance publications. They find evidence that investor behavior reacts to the timing of Article IV releases, particularly in emerging markets, reinforcing the view that IMF surveillance acts as a significant informational signal. Second, [Choi and Hashimoto \(2018\)](#) evaluate the effects of data transparency reforms—such as subscription to the IMF’s Special Data Dissemination Standard (SDDS) and General Data Dissemination System (GDDS)—on sovereign borrowing costs. Their results indicate that greater transparency lowers bond spreads, consistent with markets rewarding improved information quality.

While this literature highlights the importance of IMF surveillance and communication in shaping investor expectations, the empirical strategies most often employed—notably event-study designs—face limitations related to selection bias and endogeneity. Specifically, such designs risk conflating IMF influence with coincident domestic developments. As a result, it remains difficult to isolate whether changes in borrowing conditions are driven by surveillance and communications or by concurrent shifts in macroeconomic or political fundamentals.

To determine an appropriate identification strategy, we draw on existing evidence regarding the non-randomness of PIN publication timing. As [Glennerster and Shin \(2008\)](#) note, although Article IV consultations are conducted on a regular schedule, PIN publication is not strictly exogenous. Delays or deviations from expected release dates often result from contemporaneous developments—such as missed program targets, political transitions, or disagreements over content. Each of these factors may correlate with changes in a country’s macroeconomic outlook, introducing potential endogeneity.

Given this potential non-randomness in publication timing, a conventional event-study approach may fail to eliminate pre-trends around the PIN release. In particular, if IMF and national authorities delay PIN publication in response to deteriorating market conditions—such as rising sovereign bond spreads or worsening macroeconomic indicators—then simply controlling for calendar time effects would be insufficient. Event-study estimates could thus capture both the true effect of the PIN release and pre-existing trends in market outcomes.

As a result, we adopt a difference-in-differences strategy that compares countries receiving a PIN release to those not treated in the same period. This setup allows us to control for both calendar time effects and event-specific dynamics, providing a more credible estimate of the causal impact of surveillance communications on sovereign borrowing conditions. Specifically, we adopt the non-parametric causal inference framework introduced by [Imai et al. \(2023\)](#), which accommodates repeated and staggered treatment assignments over time—a feature that closely reflects the institutional reality of IMF surveillance, where countries may receive consultations multiple times throughout the sample period. This framework enables

more credible causal identification by conditioning on a country’s full treatment history and a rich set of pre-treatment covariates, and by constructing matched counterfactuals that did not receive a PIN release in the same period.

In our setting, the treatment is defined as the publication of a PIN following an Article IV consultation. We then examine the causal impact of this surveillance communication on sovereign bond spreads and broader financial outcomes, isolating its signaling function while controlling for contemporaneous macroeconomic conditions and the non-random timing of PIN disclosures.

2.2. A sentiment-based mechanism framework for IMF surveillance effects

Recent work at the intersection of international political economy and computational linguistics has increasingly focused on the textual content of IMF communications. Similar to the rise of natural language processing (NLP) in analyzing financial disclosures (Du et al., 2025) and monetary policy statements (Pfeifer and Marohl, 2023), a parallel literature has emerged examining how the tone, framing, and thematic structure of IMF reports shape economic expectations, reform credibility, and market behavior.

The earliest illustration came from Glennerster and Shin (2003), who showed that sovereign spreads narrow once countries begin publishing Article IV reports and Reports on the Observance of Standards and Codes—evidence that investors reward transparency. Hand-coding 157 PINs, Fratzscher and Reynaud (2011) documented a country-specific relationship between more favorable IMF language and tighter EMBI spreads, pointing to a systematic signaling channel.

Advances in NLP have since moved the field from labor-intensive coding to scalable, automated text-as-data techniques. Mihalyi and Mate (2019) compiled a 2,500-plus-document corpus of Article IV and program reviews across 189 economies, using dictionary counts and regular-expression searches to track policy priorities over time and across countries. Their dataset underpins a new generation of automated topic models and sentiment indices. For example, Fayad et al. (2020) train a deep-learning sentiment model on surveillance texts to show how national authorities perceive IMF tone. Gehring and Lang (2018) analyze 117 rating-agency reports referencing the IMF and find that favorable references during a program’s tenure strengthen sovereign creditworthiness, revealing a “confidence” channel that complements formal surveillance signals.

Researchers have also begun to unpack the substantive focus of IMF communications. Ramos et al. (2022) track the rise of climate-risk language across reports, while Goes and Chapman (2024) show that resource-rich developing countries are more likely to pass extractive-sector reforms following Article IV consultations that emphasize natural resource governance.

We build on this literature by distinguishing not only the occurrence of surveillance events, but also the content and tone of the messages conveyed. Much of the existing work treats surveillance as a binary signal of engagement or transparency. In contrast, we examine both the occurrence and substance of IMF communications, with particular attention to the tone—that is, the degree of optimism or pessimism conveyed in official assessments.

When the IMF publishes a PIN, it transmits an institutional judgment on a country’s macroeconomic outlook, policy framework, and associated risks. Even when the underlying information is partly known, the official nature of the disclosure acts as a certification (if favorable) or warning (if critical) to external investors, rating agencies, and international institutions. A positive PIN may suggest policy credibility or debt sustainability, reassuring markets and reducing perceived risk. Conversely, a PIN that flags fiscal or institutional weaknesses may trigger investor caution and elevate borrowing costs.

Understanding this signaling mechanism is particularly important in settings characterized by high information asymmetries, such as emerging markets, or when domestic policy signals are ambiguous. In such contexts, the IMF’s assessments are not merely technical—they actively shape investor beliefs about future economic and policy trajectories, which in turn influence sovereign spreads, credit ratings, and access to capital markets.

To capture the tone embedded in PINs, we apply FinBERT (Bidirectional Encoder Representation from Transformers), a financial-domain language model developed by [Liu et al. \(2020\)](#) from Google AI, to the full text of IMF PINs. This allows us to generate sentiment scores that quantify optimism or pessimism in surveillance communications. By linking these scores to movements in sovereign financing conditions, we assess whether variation in how the IMF communicates—conditional on the occurrence of surveillance—affects investor perceptions of sovereign risk.

Additionally, we investigate how this relationship is moderated by geopolitical alignment, proxied by United Nations General Assembly (UNGA) voting similarity with key IMF shareholders ([Fjelstul et al., 2025](#)). This allows us to assess whether the credibility and market impact of IMF communications is shaped by the political context in which they are received.

In sum, this paper makes three novel contributions. First, it provides the first causal estimate of Article IV surveillance effects on sovereign spreads and issuance using a dynamic matched DiD design. Second, it develops a new monthly panel dataset integrating macro-financial indicators, political variables, and textual surveillance data. Third, it introduces a mechanism-based analysis of IMF surveillance tone and geopolitical context, advancing our understanding of how international institutions influence financial markets through non-lending channels.

3. Institutional context and data

3.1. Evolution of IMF surveillance and reporting

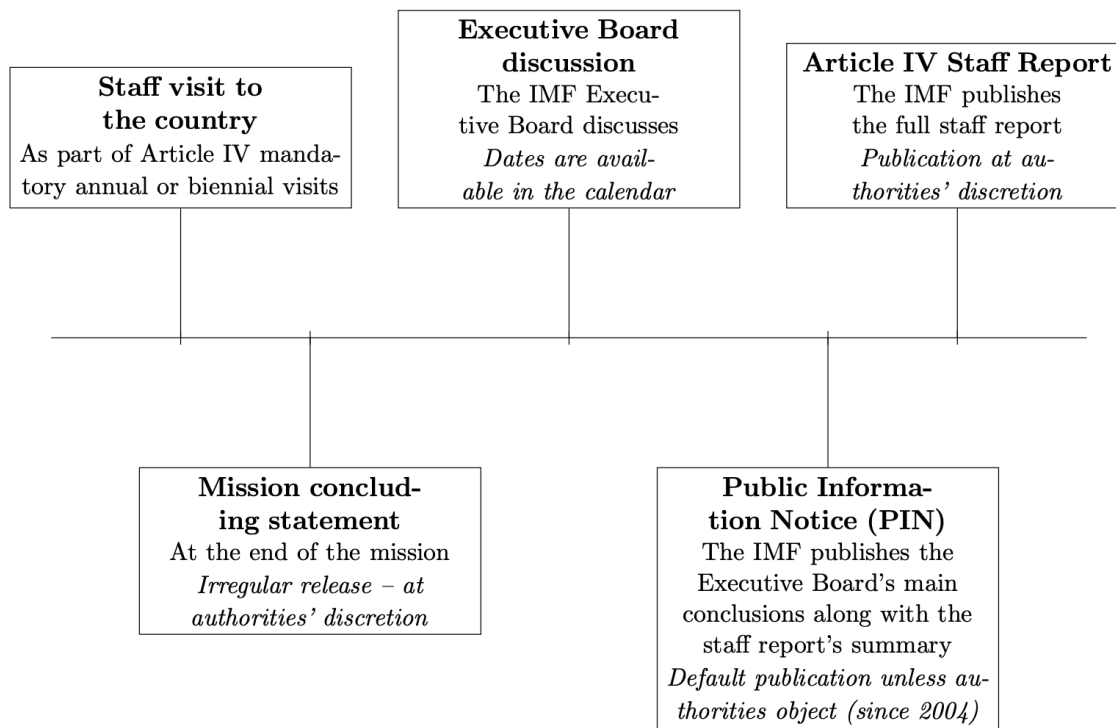
Following the collapse of the Bretton Woods system, the IMF amended its Articles of Agreement in 1977, redefining its surveillance responsibilities under Article IV. Section 3 specifies three key functions of surveillance: (i) monitoring the international monetary system; (ii) assessing members' adherence to obligations under Article IV, Section 1; and (iii) evaluating the consistency of each member's exchange rate policy with system stability (Mussa, 1997). As clarified in the IMF's 2003 Annual Report, surveillance also entails a comprehensive review of domestic and external policies, making it central to peer review and international policy signaling (IMF, 2003a).

IMF surveillance operates at two levels: multilateral and bilateral. Multilateral surveillance includes flagship reports such as the World Economic Outlook, Global Financial Stability Report, and Fiscal Monitor. Bilateral surveillance is carried out through Article IV consultations, which involve in-depth dialogue between IMF staff and individual member countries. These discussions encompass fiscal and monetary policy, financial supervision, exchange rate frameworks, and structural reform agendas (IMF, 2016a).

These consultations follow a structured, multi-stage process that combines policy dialogue with a formal reporting cycle aimed at enhancing transparency and external communication, as summarized in Figure 1. This entire cycle—from initial preparation to public disclosure—typically spans several months and reflects the IMF's dual mandate of technical surveillance and policy signaling. Typically conducted on an annual basis, the process begins with a two-week in-country mission, during which IMF staff meet with government officials, central bankers, legislators, business leaders, and civil society representatives. The goal is to assess macroeconomic developments and policy risks while fostering open dialogue. These consultations conclude with a final meeting convened among senior officials to review the findings. After returning to headquarters, the IMF team drafts a comprehensive Article IV staff report summarizing their findings, assessments, and recommendations. The report undergoes rigorous internal review, including clearance by IMF management, before being finalized and submitted to the Executive Board for formal review and discussion—typically during a one-day meeting in Washington, D.C. Following the Board's deliberations, the Chair (often the IMF's Managing Director) issues a confidential "Summing Up" of the Board's views, which is sent to national authorities. With the authorities' consent, the IMF proceeds to publicly disclose a set of key documents, including the full Article IV Staff Report, the "Summing Up," any accompanying statements by the authorities, and a Public Information Notice (PIN) summarizing the Board discussion in non-technical language. These materials

are published on the IMF website and disseminated via official press releases and media briefings (Boughton, 2012).

Figure 1: IMF visit timeline



Note: Author's representation of the timeline of IMF missions.

Over time, IMF surveillance has evolved from a confidential process to a highly transparent reporting regime. As Edwards et al. (2011) note, “A dramatic change has taken place in the manner in which the IMF conducts its annual surveillance of member state economies. These annual consultations (known as Article IV consultations) were originally viewed as private matters that were never intended to be made public.”

The first major shift occurred in April 1997, when the Executive Board authorized the release of PINs to summarize Article IV consultations. Although issuance required member consent, more than half (56%) of the 136 consultations conducted in 1997–98 resulted in a PIN (IMF, 1998, 2001). A second landmark occurred in June 1999, when the Board launched a pilot allowing full staff report publication. Between 1999 and 2001, roughly 47% of staff reports and 83% of consultations (via PINs) were published (IMF, 2001).

Encouraged by the pilot's success, the IMF adopted a standing voluntary publication policy in January 2001, introducing a presumption in favor of disclosing “policy-intention” documents. Publication rates rose significantly: by 2002-03, more than 90% of members had

issued at least one PIN (IMF, 2003b). Transparency became the norm following the February 2004 adoption of a “voluntary-but-presumed” policy, whereby staff reports are published unless authorities object. By 2008-09, over 95% of Article IV consultations generated a PIN, leading the IMF in 2013 to discontinue the separate “PIN” designation, issuing the identical format document as a standard press release (IMF, 2009).

These transparency gains were consolidated during the 2010s. The 2013 Transparency Policy Review reported a staff-report publication rate of 92% overall and virtually 100% for advanced economies (IMF, 2013). A 2016 follow-up showed that by 2015, 98% of consultations were accompanied by a press release, with emerging-market and developing countries achieving 80–90% publication rates (IMF, 2016b). Today, the regime is effectively universal, with press releases issued for nearly every consultation. Disclosure is no longer an optional add-on but an integral component of IMF surveillance (IMF, 2024).

In this study, we leverage this disclosure regime to assess how the release of PINs affects sovereign borrowing conditions. Because PINs contain direct and coded signals regarding macroeconomic risks, policy credibility, and reform priorities, they constitute a key informational channel through which the IMF can influence investor expectations, even in the absence of lending arrangements.

3.2. Data

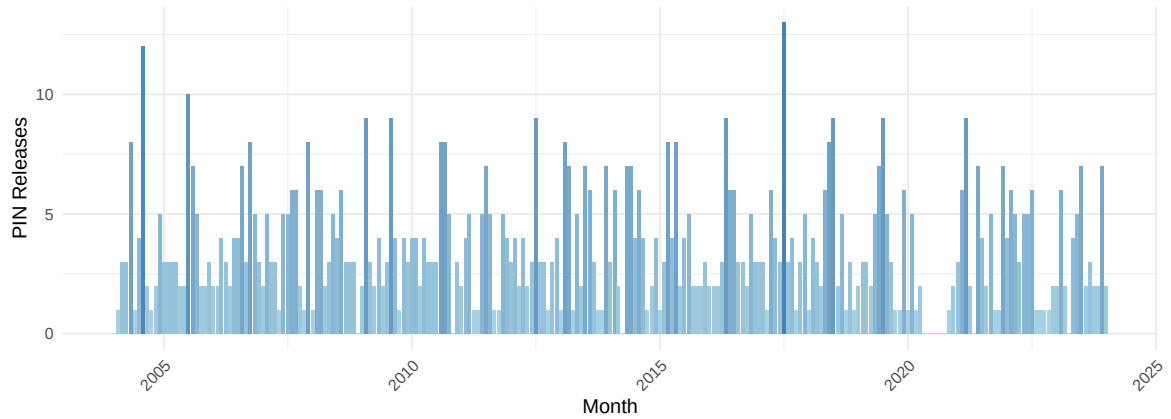
Our analysis uses monthly data for 67 emerging market economies (EMEs) covering the period from January 1, 2004, to January 1, 2024. This frequency is essential to capture short-term market responses and anticipation effects. The list of countries included in the sample is provided in Appendix A Table A.1.

To construct the dataset of IMF Article IV PIN release dates, we manually collected the publication timings from the official IMF Press Release Archive on the Fund’s website.¹ For each country between 2004 and 2023, we identified the month and year when the corresponding PIN was published. These PINs, which summarize the Article IV staff report and Executive Board discussions, serve as the key moment of public disclosure. We merged the release dates with a country-month panel to generate a binary treatment indicator, equal to 1 if a PIN was released in that month and 0 otherwise.

To describe the distribution of PIN disclosures, we present two complementary figures. Figure 2 displays the monthly frequency of PIN releases between 2004 and 2023, with each bar indicating the number of countries receiving a PIN in a given month.

¹<https://www.imf.org>

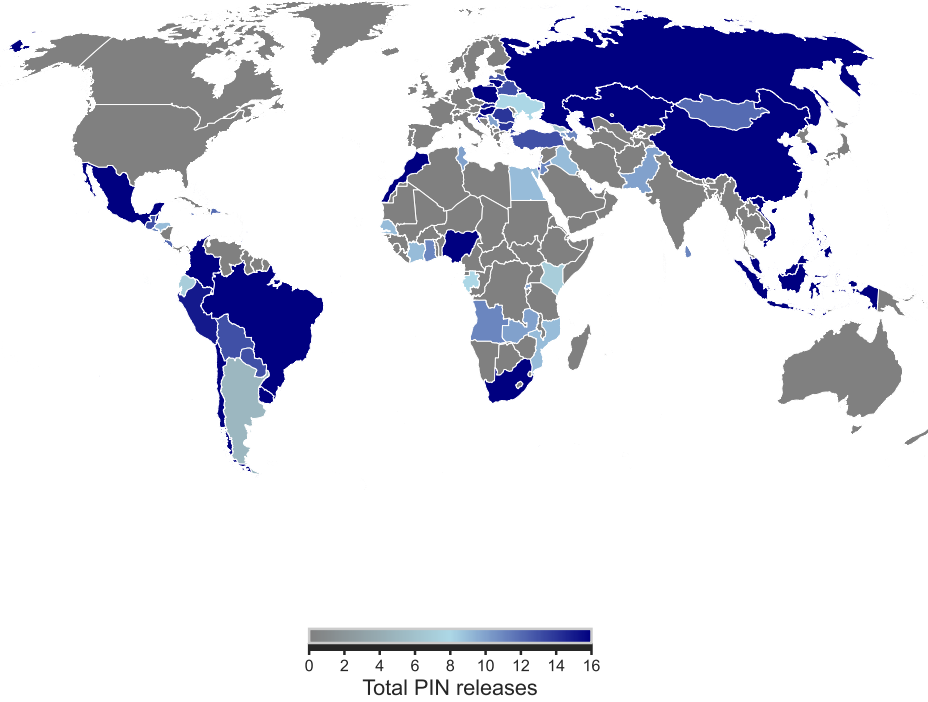
Figure 2: *Monthly Public Information Notice (PIN) releases, 2004-2023*



Note: Bar chart plots the number of countries receiving an IMF Article IV Public Information Notice (PIN) in each month between 2004 and 2023. Data compiled from IMF press release archives.

While the overall distribution reflects the IMF’s annual surveillance cycle, the figure reveals temporal clustering, including a sharp decline in 2020 due to COVID-19-related operational disruptions. This variation underscores the importance of conducting robustness checks that exclude or account for major global episodes—such as the Global Financial Crisis and the COVID-19 pandemic—to ensure that our results are not driven by these exceptional periods. Figure 3 illustrates the total number of PIN releases per country over the study period. It highlights substantial cross-country variation, confirming that surveillance frequency is not uniform and that many countries appear repeatedly in the panel.

Figure 3: *Global distribution of total IMF Article IV PIN releases, 2004-2023*



Note: The world map displays the total number of IMF Public Information Notices (PINs) released per country between 2004 and 2023, covering a sample of 63 emerging market economies (or countries formerly classified as such), for which JP Morgan EMBI spread data is available at the time of release. All data were collected from the IMF website.

Descriptive statistics presented in Table 1 further support this heterogeneity.² On average, each country received approximately 13 PINs between 2004 and 2023, with the number of disclosures ranging from a minimum of 5 to a maximum of 19. Across countries, the mean interval between PINs is about 20 months, with considerable dispersion: some countries experienced intervals as short as 12 months, while others had gaps extending up to 50 months. This variation suggests that, although IMF surveillance is broadly standardized in principle, its practical implementation differs considerably across countries and over time.

Together, these patterns justify our adoption of a dynamic, repeated-event design, recognizing that surveillance exposure is neither uniform nor one-time, but varies across countries and over time.

²Figure A.1 in Appendix A illustrates the distribution of the treatment variable used in our analysis of IMF Article IV consultations.

Table 1: *Descriptive statistics on IMF PIN disclosures (2004–2023)*

Statistic	Total PINs per country	Avg. months between PINs
Mean	12.9	19.8
Median	13.0	17.8
Min	5.0	12.2
Max	19.0	50.0
Standard Deviation	3.5	6.8

Note: This table summarizes the distribution of IMF Public Information Notice (PIN) disclosures across 63 countries over the period January 2004 to December 2023. The left column reports the total number of PINs received by each country, while the right column shows the average number of months between successive PINs per country. These statistics capture both the intensity and frequency of surveillance exposure across countries.

To capture sovereign access to international capital markets, we rely on a broad set of macro-financial variables: sovereign bond spreads (in basis points), the real effective exchange rate (REER index), foreign reserves (measured both in nominal USD and as a share of GDP), and hard-currency debt issuance. On the issuance side, we separately measure USD-denominated sovereign issuance (in nominal USD and relative to GDP), euro-denominated issuance (in nominal USD and relative to GDP), and total foreign-currency issuance (USD plus euro, both in nominal USD and as a share of GDP).

Our sovereign spread data come from two main sources: the J.P. Morgan Emerging Market Bond Index (EMBI) and the ICE BofA Option-Adjusted Spread (OAS). Both measure the premium countries pay to issue debt in US dollars relative to US Treasury securities of similar maturities. We also incorporate data from the ICE BofA US Emerging Markets External Sovereign Index, which tracks performance of sovereign debt issued in US dollars and euros by emerging markets. For our analysis, we focus on face value in US dollars at issuance, including all foreign-currency government debt with maturities over one year and no rating constraints. Additional macro-financial variables—including spreads, issuance volumes, reserves (including gold), and the real effective exchange rate (REER)—are sourced from Refinitiv and [Darvas \(2021\)](#).

Table 2 presents summary statistics for key financial variables. On average, sovereign issuance in US dollars exceeds that in euros, both in nominal terms (8,530 vs. 4,564 million USD) and relative to GDP (6.6% vs. 2.4%), with US dollar issuance observed more frequently.

Table 2: *Summary statistics of sovereign bond spreads and financial indicators*

Variable	Unit	N	Mean	25th	Median	75th
EMBI spread	Basis point spread	13647	501.02	160.46	281.88	468.99
USD sovereign face value issuance	Nominal MUSD	13202	8529.53	1200.00	3250.00	9630.48
EUR sovereign face value issuance	Nominal MUSD	5146	4563.96	780.00	2050.00	5500.00
Total FX sovereign face value issuance	Nominal MUSD	13689	9941.78	1525.00	3800.00	12200.00
FX Reserves	Nominal MUSD	16133	76754.65	2673.18	10364.7	37399.29
Real Effective Exchange Rate	Index	16017	108.08	93.63	102.34	115.45
Variables as shares of GDP						
USD sovereign face value issuance % GDP	Share of GDP	13202	6.59	1.76	3.90	8.24
EUR sovereign face value issuance % GDP	Share of GDP	5146	2.39	0.42	1.32	3.42
Total FX sovereign face value issuance % GDP	Share of GDP	13689	7.25	2.29	4.83	9.08
FX Reserves % GDP	Share of GDP	16133	16.25	7.88	12.22	18.59

Note: This table presents the number of observations (N), mean, 25th percentile, median, and 75th percentile for the variables of interest, covering both the nominal variables and the variables expressed as a share of GDP.

4. Empirical strategy

4.1. Identification strategy

Our empirical design leverages the staggered rollout of IMF PINs across countries and months from 2004 to 2023. While a natural starting point is the difference-in-differences (DiD) framework, the conventional two-way fixed effects (TWFE) estimator is known to produce biased or inconsistent estimates of the Average Treatment Effect on the Treated (ATT) when treatment effects are heterogeneous across units or over time (Goodman-Bacon, 2021; Callaway and Sant’Anna, 2021; Sun and Abraham, 2021; De Chaisemartin and d’Haultfoeuille, 2020). Recent methodological refinements—such as group-time DiD estimators, doubly robust procedures, and reweighted event-study designs—address some of these issues but typically rely on an “absorbing treatment” assumption (i.e., once treated, always treated) or require a never-treated control group (De Chaisemartin and d’Haultfoeuille, 2024).

However, in our context, every country receives multiple PINs over the sample period, violating both the absorption assumption and the never-treated condition. To address this, we implement the panel-matching approach proposed by Imai et al. (2023)—a non-parametric framework that (i) explicitly accommodates non-absorbing treatments, allowing countries to repeatedly enter and exit treatment, and (ii) permits treatment effect heterogeneity across countries and over time. By matching treated and control observations on recent pre-treatment outcomes, covariates, and treatment history, this method constructs valid counterfactual trajectories for each PIN event. It thus captures the complex, recurring nature of IMF surveillance and yields unbiased estimates of ATT for our key macro-financial

outcomes.

We define the treatment as the month in which a PIN is publicly released, aligning the intervention with the moment when IMF surveillance becomes visible to financial markets. To construct the counterfactual, we compare these treated country-months to a control group of country-months with no Article IV consultation or PIN release. While these control countries may eventually be treated later, they must remain untreated in both the current month and the preceding L months—the lookback window used for matching. This setup enables us to compare post-treatment outcomes of treated units to those of similar, recently untreated units, controlling for observed covariates and treatment history. By expanding the pool of control observations, this approach enhances statistical power and allows for more robust causal inference across the full sample period.

The approach proceeds as follows: for each treated country-month, we first identify a set of control country-months that share an identical IMF consultation history over a prespecified number of prior periods. We refer to this group as the matched set. Second, we refine these matches by adjusting for observable economic and political covariates using standard matching and weighting techniques to ensure that treated and control units are balanced on pre-treatment characteristics. We then apply a difference-in-differences estimator to account for underlying time trends in sovereign financial outcomes.

The treatment variable X_{im} equals 1 if country i experiences the release of a PIN in month m , and 0 otherwise. The average treatment effect of PIN disclosure—interpreted as the impact of transitioning from no recent IMF Article IV consultation to one resulting in a disclosed PIN—is defined in Equation (1).

$$\delta(F, L) = \mathbb{E}\left[Y_{i,m+F}(X_{im} = 1, X_{i,m-1} = 0, \{X_{i,m-\ell}\}_{\ell=2}^L) - Y_{i,m+F}(X_{im} = 0, X_{i,m-1} = 0, \{X_{i,m-\ell}\}_{\ell=2}^L) \mid X_{im} = 1, X_{i,m-1} = 0\right] \quad (1)$$

In this setup, treated observations are those where a PIN is publicly released in month m , but not in the preceding month $m - 1$ (i.e., $X_{i,m} = 1$ and $X_{i,m-1} = 0$). In our design, we exclude $X_{i,m-1}$ from the matching set because we explicitly condition on $X_{i,m-1} = 0$ when defining eligible treatment and control units. The treatment history used for matching therefore begins at $m - 2$ and extends back L months to account for prior patterns that may affect both treatment timing and potential outcomes. This estimand represents the average causal effect of initiating a PIN disclosure on outcomes such as sovereign bond spreads or debt issuance, measured F months after the release, and conditional on L months of prior treatment history.

The first step in our methodology is to select a non-negative integer F , which represents the

number of months after the public release of the IMF Article IV PIN at which the outcome is measured.³ Choosing $F > 0$ allows us to examine the cumulative effects of PIN disclosures on outcomes such as sovereign bond spreads or debt issuance.

In parallel, the method requires choosing a non-negative integer L , which defines the number of lagged months used in the matching procedure. Unlike the number of leads F , the choice of L is a core component of the identification strategy, as it governs the conditioning set for treatment history and covariates. Failing to adequately adjust for past IMF engagement may introduce bias if previous PIN releases influence both current treatment assignment and subsequent outcomes (Imai and Kim, 2019). As in fixed effects regressions, selecting L involves a bias–variance tradeoff: higher values of L improve the credibility of the unconfoundedness assumption but reduce statistical efficiency by narrowing the pool of valid matches.

We set the lag window $L = 4$, incorporating four months of treatment and covariate history. This four-month period is sufficient to absorb residual effects from prior consultations, control for seasonal patterns, and satisfy the carry-over assumption—that is, that older treatments no longer influence current outcomes. A longer lag would sharply reduce matchable observations, while a shorter lag could omit relevant pre-treatment dynamics. Empirically, we verify that $L = 4$ achieves a favorable balance between covariates, maintains adequate matched set sizes, and meets standard diagnostics, supporting its use in our baseline design.

For the post-treatment lead horizon, we examine outcomes over $F = 1$ to 9 months following each PIN release. This nine-month window aligns well with the typical policy—response cycle to Article IV recommendations, which usually unfolds over roughly three quarters. It also precedes the arrival of any subsequent PIN, given that the minimum observed interval between consultations is 12.2 months. At the same time, this window is short enough to limit contamination from other macroeconomic shocks or overlapping treatments. To assess the robustness of our parameter choices, we conduct sensitivity analyses by varying the lag window L and the lead horizon F . The results remain stable across alternative specifications, supporting the credibility and flexibility of our empirical design.

Given our choice of parameters L and F , and our goal of estimating the causal effect of IMF PIN releases on sovereign outcomes, we must rely on an additional assumption to identify the treatment effect. Specifically, we assume that—conditional on recent history—the timing of a PIN release behaves as if it were random with respect to future outcomes. Formally, this means that, conditional on the past values of the outcome variable, relevant covariates, and the history of PIN releases over the previous L months, the probability of receiving

³For instance, $F = 0$ captures the contemporaneous effect of the PIN release, while $F = 9$ reflects its impact nine months later.

treatment in month m is independent of the potential financial outcomes that would have occurred in the absence of treatment. This identifying condition is known as the sequential ignorability assumption (Robins et al., 2000), which generalizes the standard unconfoundedness assumption to settings with time-varying treatments in panel data. In our context, this implies that once we account for recent macroeconomic and political conditions, countries do not systematically time the release of PINs based on expectations about future market developments. If this assumption holds, the observed differences in outcomes between treated and matched control units can be interpreted as causally attributable to the PIN disclosure, rather than the result of selection or latent trends.

In addition to sequential ignorability, our difference-in-differences (DiD) framework relies on a conditional parallel trends assumption: in the absence of treatment, the outcome trajectories for treated and control units would have evolved similarly over time. This assumption allows us to control for time-invariant unobserved heterogeneity, provided that its effects are common across units. More precisely, we impose parallel trends conditional on past treatment status, lagged outcomes, and observed covariates. Under this assumption, any divergence in outcomes between the treated and control groups following a PIN release can be interpreted as the causal effect of the surveillance disclosure, not a reflection of pre-existing differences or unrelated shocks.

We formally define this identification strategy as:

$$\begin{aligned} & \mathbb{E}[Y_{i,t+F}(X_{it} = 0, X_{i,t-1} = 0, \{X_{i,t-\ell}\}_{\ell=2}^L, \{Z_{i,t-\ell}\}_{\ell=0}^L) - Y_{i,t-1} \mid X_{it} = 1, \\ & \quad X_{i,t-1} = 0, \{X_{i,t-\ell}, Y_{i,t-\ell}\}_{\ell=2}^L, \{Z_{i,t-\ell}\}_{\ell=0}^L] \\ &= \mathbb{E}[Y_{i,t+F}(X_{it} = 0, X_{i,t-1} = 0, \{X_{i,t-\ell}\}_{\ell=2}^L, \{Z_{i,t-\ell}\}_{\ell=0}^L) - Y_{i,t-1} \mid X_{it} = 0, \\ & \quad X_{i,t-1} = 0, \{X_{i,t-\ell}, Y_{i,t-\ell}\}_{\ell=2}^L, \{Z_{i,t-\ell}\}_{\ell=0}^L], \end{aligned} \tag{2}$$

where the conditioning set includes the treatment history, $\{X_{i,t-\ell}\}_{\ell=2}^L$, the lagged outcomes (except the immediate lag $Y_{i,t-1}$), and the covariate history, $\{Z_{i,t-\ell}\}_{\ell=0}^L$.

This conditioning strategy aims to make treated and control units as comparable as possible prior to treatment, thereby minimizing bias arising from systematic differences. By matching on observed covariates and treatment history, we reduce the risk of confounding due to observable characteristics. However, the possibility of unobserved, time-varying confounders remains—factors not captured in our dataset that could influence both the timing of PIN disclosures and subsequent financial outcomes.

To enhance the credibility of our causal interpretation, we conduct empirical diagnostics focused on pre-treatment outcome dynamics. In particular, we test whether treated and

matched control units followed similar trajectories in our key financial variables before the treatment occurred. If pre-treatment trends are statistically indistinguishable, this supports both the sequential ignorability assumption and the validity of the matching procedure. Evidence of parallel trends in the pre-treatment period increases confidence that any observed post-treatment divergence in outcomes is due to the PIN release rather than to pre-existing trends or omitted variables.

The next step in our methodology is to construct, for each treated observation i, m —that is, a country experiencing a public release of an IMF Article IV PIN at time m —a matched set of control country-months that share an identical treatment history over the L months leading up to the treatment. We perform exact matching on PIN release history to account for potential carryover effects, based on the assumption that prior consultations may influence both the likelihood of current treatment and subsequent outcomes.

To ensure that treatment and control groups are comparable on relevant covariates, we implement multiple matching refinement strategies widely used in the panel data causal inference literature. Specifically, we apply both Mahalanobis distance matching and propensity score-based methods to define matched sets. This dual approach allows us to evaluate the robustness of our results across alternative matching designs.

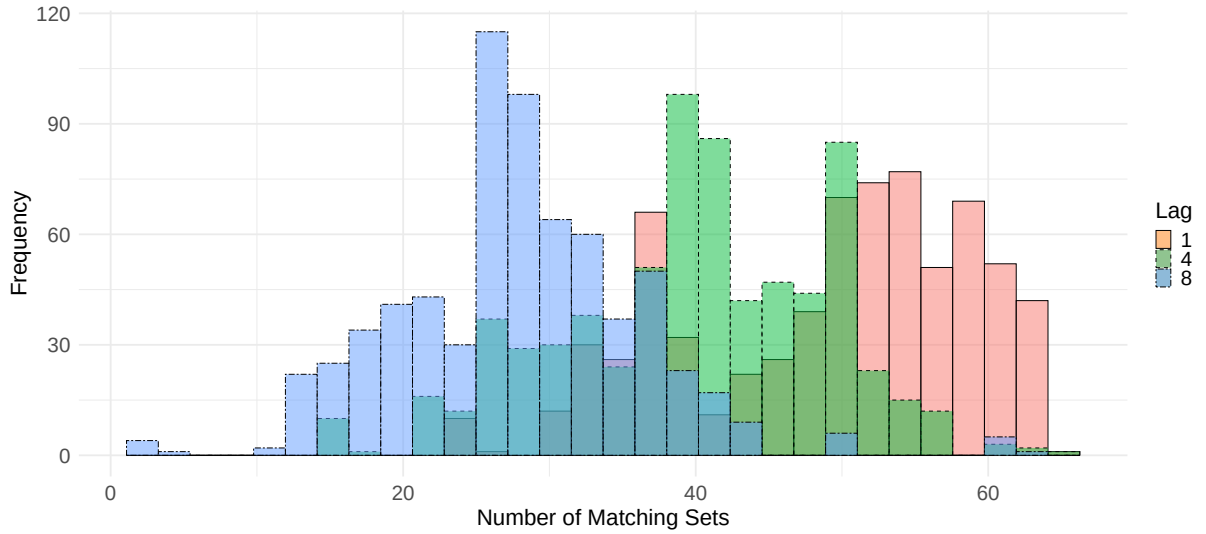
In the Mahalanobis matching approach, treated units are matched to control units based on the Mahalanobis distance, which adjusts for the variance and covariance structure of covariates. This method accounts for variable scaling and interdependence, yielding tighter control over covariate similarity than naïve distance metrics. In parallel, we implement propensity score matching, where we first estimate the probability of receiving treatment (i.e., a PIN release) using a logistic regression model that includes the treatment and covariate history over the past L months. Treated units are then matched to control units with similar estimated propensity scores. Additionally, we apply propensity score weighting, which retains the full sample while reweighting observations based on their proximity to the treated group in the propensity score distribution. This method places greater emphasis on control units that are more comparable to treated units, thereby improving internal validity without sacrificing sample size.

4.2. Constructing matched sets and assessing covariate balance

To illustrate our matching methodology, we begin by constructing the matched set for each treated observation based on its recent treatment and covariate history. Figure 4 presents the distribution of the number of matched control units per treated observation, comparing three specifications that vary the length of the lag history used in matching: a 1-month history (red solid bars), a 4-month history (green dashed bars), and an 8-month history

(blue double-dashed bars).

Figure 4: *Distribution of matched control units per treated observation*



Note: The figure plots the number of matched control units associated with each treated observation under three matching specifications: one with a 1-month treatment history (red solid bars), one with a 4-month treatment history (green dashed bars) and one with a 8-month treatment history (blue two-dashed bars).

As expected, increasing the conditioning window from 1 to 4- and 8-months results in a reduction in the number of matched control observations per treated unit, reflecting the more stringent requirements imposed by a longer treatment history. This pattern underscores the bias–variance tradeoff inherent in matching design: longer histories improve the credibility of identifying assumptions but reduce the number of comparable control observations. The comparison across specifications provides visual confirmation that our chosen lag length of $L = 4$ maintains a reasonable balance between covariate similarity and match availability.

We construct a pre-treatment covariate set designed to capture factors relevant to both treatment assignment—that is, the timing and likelihood of a PIN release—and the outcome variables. These confounders encompass macroeconomic fundamentals and political-institutional characteristics that are empirically linked to a country’s likelihood of IMF engagement and to investor assessments of sovereign creditworthiness. First, macroeconomic fundamentals include indicators such as the current-account balance, general government gross debt, net lending/borrowing, tax revenue, CPI inflation, real GDP growth, GDP per capita (PPP), and population size. These variables reflect economic performance and fiscal space—key dimensions that influence both IMF surveillance behavior and market perceptions of risk. Second, sovereign creditworthiness is proxied through long-term foreign-currency ratings from Fitch, Moody’s, and S&P, which provide a market-based assessment of default

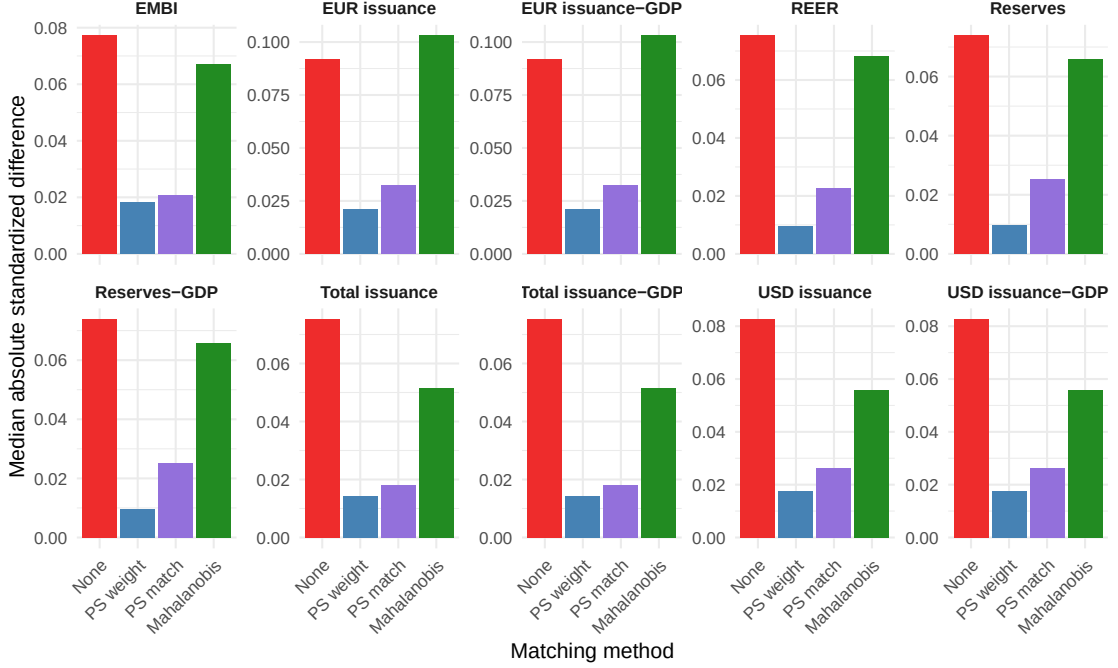
risk. Third, political-institutional attributes include the V-Dem liberal democracy score, electoral-calendar dummies, UNGA voting alignment with the United States, the Chinn-Ito capital-account openness index, and the de facto exchange rate regime. These variables account for geopolitical alignment, regime type, and external financial integration—factors that can shape both IMF engagement patterns and investor sentiment.

All covariates are measured using values from the most recent full year prior to the start of the 4-month treatment risk window, which helps avoid post-treatment bias and mitigates concerns about introducing “bad controls” into the identification strategy. A full description and discussion of the selected covariates is provided in Appendix B.

For each refinement strategy—propensity-score weighting (PSW), propensity-score matching (PSM), Mahalanobis matching, and no matching—we measure covariate balance via the standardized mean difference (SMD). Concretely, at each period in the four-lag window, we compute the difference between each treated unit’s covariate value and the weighted mean of its matched controls, average those differences across treated units, and then divide by the pooled standard deviation of that covariate. We report the absolute SMD (ASMDs) and adopt a threshold of 0.10 to signal acceptable balance, as recommended in the literature (Stuart et al., 2013; Zhao, 2019). To construct the final covariate set, we apply a two-step criterion (i) the ASMD must fall below 0.10 after weighting, and (ii) the matching or weighting procedure must improve balance relative to the unmatched sample, as indicated by a reduction in the absolute standardized difference.

Figure 5 presents the median ASMDs across all covariates for each outcome and matching method. Among all strategies, propensity score weighting (PS weighting) consistently produces the lowest ASMD values across outcome variables, indicating that it is the most effective method for covariate adjustment. While both propensity score matching (PS matching) and Mahalanobis distance matching also improve balance relative to the unmatched sample, they generally yield higher ASMDs than PS weighting, suggesting comparatively lower performance in aligning covariate distributions. As expected, the no matching specification results in the highest ASMDs, underscoring the importance of applying appropriate matching or weighting techniques to reduce pre-treatment differences and strengthen causal interpretation.

Figure 5: Median covariate balance before and after matching and weighting



Note: Each bar shows the median absolute standardized difference (ASD) across the full set of covariates for a given outcome variable and matching method –no matching, Mahalanobis matching, propensity-score matching, and propensity-score weighting–. ASDs are computed over four lags (as in our benchmark), and a lower median indicates better overall covariate balance.

These results indicate that PS weighting is the most reliable method for achieving covariate balance in our context. Its advantage stems from its ability to assign greater weight to control observations that closely resemble treated units in terms of propensity scores, thereby enhancing the comparability of the two groups and improving the internal validity of the estimated treatment effects.⁴

With the matched sets constructed, we proceed to estimate the Average Treatment Effect on the Treated (ATT) of an IMF PIN release, as defined in Equation (1). For each treated country-month (i, m) , we estimate the counterfactual outcome $Y_{i,m+F}(X_{im} = 0, X_{i,m-1} = 0, \dots, X_{i,m-L})$ by taking the weighted average of outcomes among control units in the matched set that share the same treatment history and are balanced on covariates. We then compute the difference-in-differences (DiD) estimate for each treated observation by comparing the actual post-treatment outcome to its estimated counterfactual. The overall ATT is obtained by averaging these DiD estimates across all treated units. Formally, our

⁴For a detailed, covariate-by-covariate breakdown under each method, see Appendix B.2 (Figure B.1 and Figure B.2). Appendix B.3 illustrates PSW’s use in constructing the counterfactual around Angola’s September 2014 PIN release.

estimator is expressed as:

$$\hat{\delta}(F, L) = \frac{1}{\sum_{i=1}^N \sum_{t=L+1}^{M-F} D_{it}} \sum_{i=1}^N \sum_{m=L+1}^{M-F} D_{im} \left\{ (Y_{i,m+F} - Y_{i,m-1}) - \sum_{i' \in \mathcal{M}_{im}} w_{i'm}^{i'} (Y_{i',m+F} - Y_{i',m-1}) \right\} \quad (3)$$

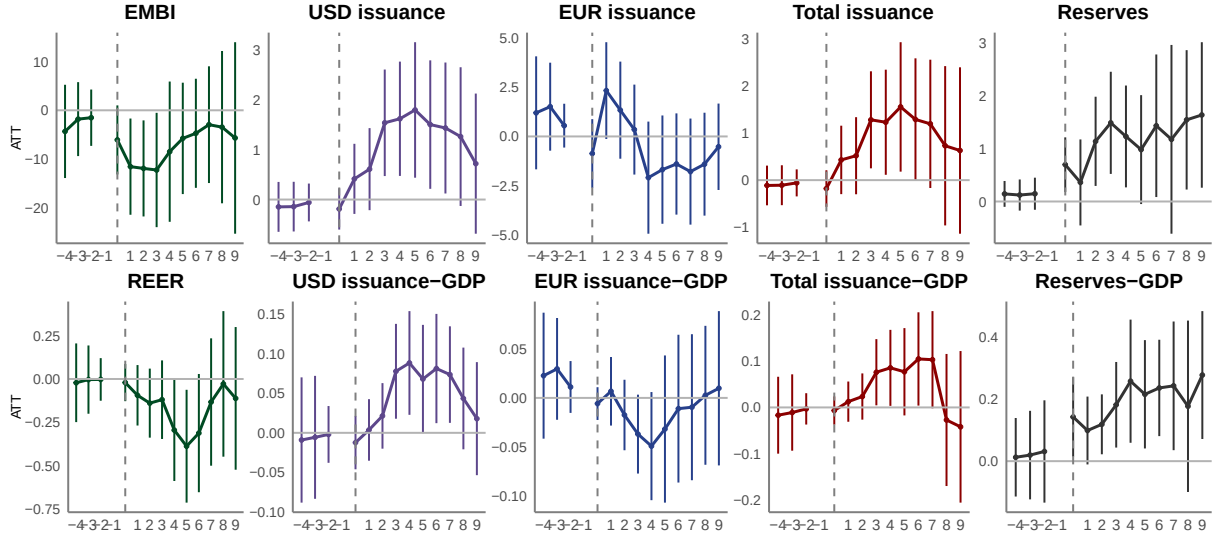
where $D_{im} = X_{im} (1 - X_{i,m-1}) \cdot \mathbb{1} \{ |\mathcal{M}_{im}| > 0 \}$, with $\mathcal{M}_{im}$, the matched set for the treated observations. $w_{i'm}^{i'}$ denotes the nonnegative normalized weight assigned to control country i' , such that $w_{im}^{i'} \geq 0$ and $\sum_{i' \in \mathcal{M}_{im}} w_{im}^{i'} = 1$.

5. Estimates of the economic impact of PIN releases

5.1. Baseline results

Figure 6 presents our benchmark estimates of the Average Treatment Effect on the Treated (ATT) following the release of an IMF Article IV PIN using dynamic treatment horizons up to nine months.

Figure 6: Estimated average effects of IMF PIN releases



Note: This figure displays the dynamic treatment effects of IMF Article IV PIN releases on our variables of interest. We report the effects using the propensity score weighting methodology showing the estimated Average Treatment Effect on the Treated (ATT) over the nine-month horizon following the PIN release and the four-month horizon before. Shaded areas represent 90% confidence intervals based on bootstrapped standard errors.

The top-left panel reports dynamic ATT estimates indicating that sovereign bond spreads (EMBI) decline by approximately 15 basis points beginning in the second month after a

positive IMF PIN release. This decline is statistically significant in months 2 through 4, before tapering off, with confidence intervals widening at longer horizons. These magnitudes are similar to the 17 bp abnormal spread returns reported by [Breen and Doak \(2023\)](#) for emerging markets, although our panel-matching DiD approach yields a more robust causal estimate and indicates stronger persistence.

Similarly, sovereign issuance indicators respond dynamically to the treatment. USD-denominated issuance increases significantly beginning in month four, with the ATT remaining statistically significant through month eight. This trajectory suggests that the observed decline in spreads enhances market access, enabling treated countries to increase issuance. Total foreign-currency sovereign issuance also rises, with the effect primarily driven by USD-denominated instruments, highlighting the dominance of USD markets in sovereign debt issuance.

The reserves panel shows a statistically significant and positive ATT across the post-release horizon, indicating that treated countries—those receiving a PIN release—accumulate more reserves compared to similar untreated counterparts. This causal effect is consistent with market reassurance following IMF communication, which may reduce capital flight risk or facilitate precautionary accumulation of reserves in response to improved sentiment.

Overall, our findings show that PIN releases produce a clear compression in sovereign borrowing costs, consistent with prior event-study estimates, and provide further evidence that this temporary easing strengthens investor confidence, broadens market access, and supports greater sovereign issuance capacity.

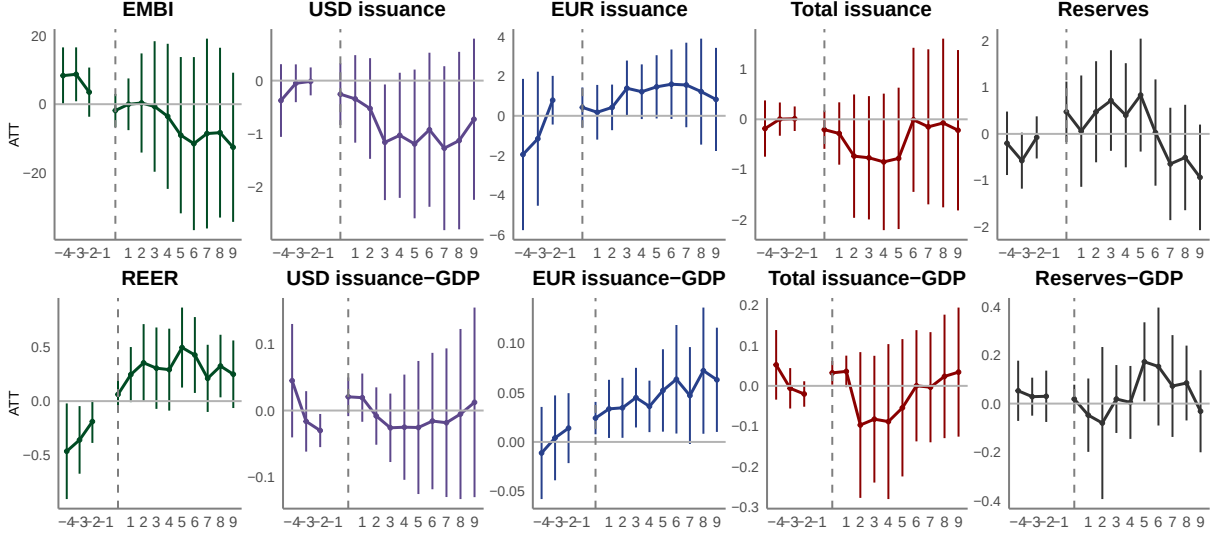
5.2. Robustness checks

To assess the robustness of our causal estimates and evaluate whether observed effects are attributable to the timing of IMF Article IV PIN disclosures—as opposed to spurious correlations or unobserved time trends—we implement a “future-leads” falsification test, in which each treated unit is artificially coded as treated not in month m , the actual month of PIN release, but at a forward-shifted date—specifically, at $m + 6$. This approach preserves all actual pre-treatment histories and covariates but introduces a pseudo-treatment indicator artificially assigned to a future date. The test directly evaluates whether treatment effects observed prior to the actual PIN publication may be the result of anticipation effects or dynamic selection into treatment, rather than causal responses to the PIN release itself.

We compute dynamic ATT estimates over the same forward-looking horizon $F = 1, \dots, 9$, using the same matching criteria (lag length L , treatment history, and covariates), standardization procedures, and estimation steps as in the baseline specification. [Figure 7](#) reports these placebo results. Consistent with our “information shock” hypothesis, the placebo esti-

mates are statistically insignificant across all event horizons.

Figure 7: *Future-leads placebo estimated average effects of IMF PIN releases*

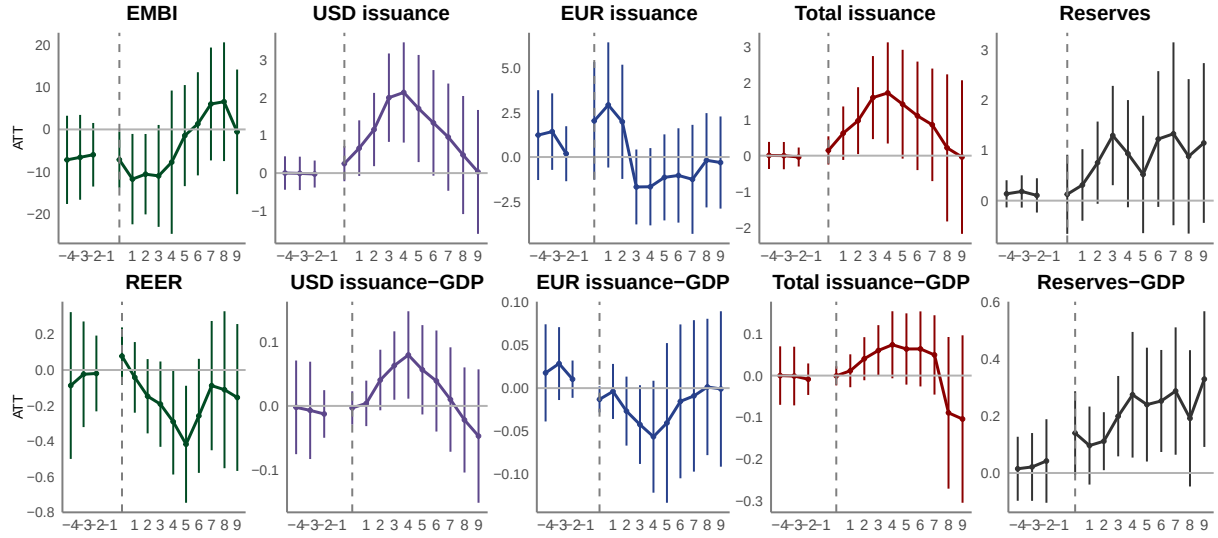


Note: This figure displays the dynamic treatment effects of IMF Article IV PIN placebo releases (coded at $m + 6$) on our variables of interest. We report the effects using the propensity score weighting methodology showing the estimated Average Treatment Effect on the Treated (ATT) over the nine-month horizon following the PIN release and the four-month horizon before. Shaded areas represent 90% confidence intervals based on bootstrapped standard errors.

As a further robustness check, we re-estimate our panel-matching model using the publication of the full IMF Article IV Staff Report—rather than the PIN—as the treatment indicator. Specifically, we define a binary indicator that equals one in any month when an Article IV Staff Report is published. These comprehensive documents contain detailed macroeconomic analysis, forecasts, vulnerability assessments, and stress tests, and are released less frequently than the shorter and more summary-focused PINs.

Figure 8 presents the resulting dynamic ATT estimates. The patterns closely resemble those in our baseline PIN-based results across most variables, with one notable exception: the real effective exchange rate shows no significant response to Staff Report releases. While overall statistical significance is somewhat lower than in the PIN specification, the effect on USD-denominated issuance remains both strong and highly significant.

Figure 8: Robustness - Estimated average effects of IMF Article IV Staff Report releases



Note: This figure displays the dynamic treatment effects of IMF Article IV Staff Report releases on our variables of interest. We report the effects using the propensity score weighting methodology showing the estimated Average Treatment Effect on the Treated (ATT) over the nine-month horizon following the PIN release and the four-month horizon before. Shaded areas represent 90% confidence intervals based on bootstrapped standard errors.

We conduct additional robustness checks to assess the stability of our findings. First, we perform sensitivity analyses by varying the lag structure (Figure C.1) and the number of forward months (Figure C.2), both reported in Appendix C. The results remain highly consistent across these alternative specifications, reinforcing the reliability of our dynamic ATT estimates. Second, although our nonparametric panel-matching framework accounts for global shocks—such as the Global Financial Crisis (GFC) and the COVID-19 pandemic—these extreme events may still interact with IMF PIN announcements under heightened financial uncertainty. To address this possibility, we re-estimate our model after excluding all observations during the GFC and COVID-19 periods. As shown in Figure C.3 (Appendix C), the core results remain robust, closely aligning with our benchmark estimates and confirming that the main findings are not driven by these global episodes.

6. Evidence on channels for PIN disclosures effects

Our benchmark estimates show that IMF Article IV PIN releases lead to lower sovereign bond spreads, increased external issuance, and higher reserve buffers. However, these aggregate effects do not fully explain how public disclosures influence investor behavior. In this section, we investigate two complementary transmission channels that may underlie the observed impacts.

We first examine the information channel. Under this interpretation, an optimistic PIN release serves as a public endorsement by the IMF of a country’s economic program. By reducing uncertainty regarding policy stability and signaling the potential for future financial support, a favorable PIN can directly enhance investor confidence. We test this channel by constructing measures of sentiment surprises within PIN reports and estimating how these surprises affect sovereign bond spreads and other macro-financial outcomes of interest.

Second, we explore a geopolitical alignment channel. PINs may carry greater weight in markets when they concern countries whose foreign policy positions are closely aligned with major IMF shareholders. To capture this dimension, we use voting similarity in the United Nations General Assembly as a proxy for alignment with the United States—led international order. We then assess whether PIN releases trigger larger spread compressions and issuance increases in countries with high alignment scores.

6.1. The informational channel of IMF Article IV PIN releases

To empirically assess the informational channel of IMF Article IV PIN releases, we investigate whether the observed effects on our outcome variables are driven by the signal content disclosed in these communications. We conduct a text-based sentiment analysis of PIN documents using natural language processing (NLP) techniques, classifying the tone of each release along a favorable-unfavorable spectrum. Specifically, we employ the FinBERT model, an adaptation of the Bidirectional Encoder Representations from Transformers (BERT) architecture, which is specifically designed for financial sentiment analysis (Yang et al., 2020).

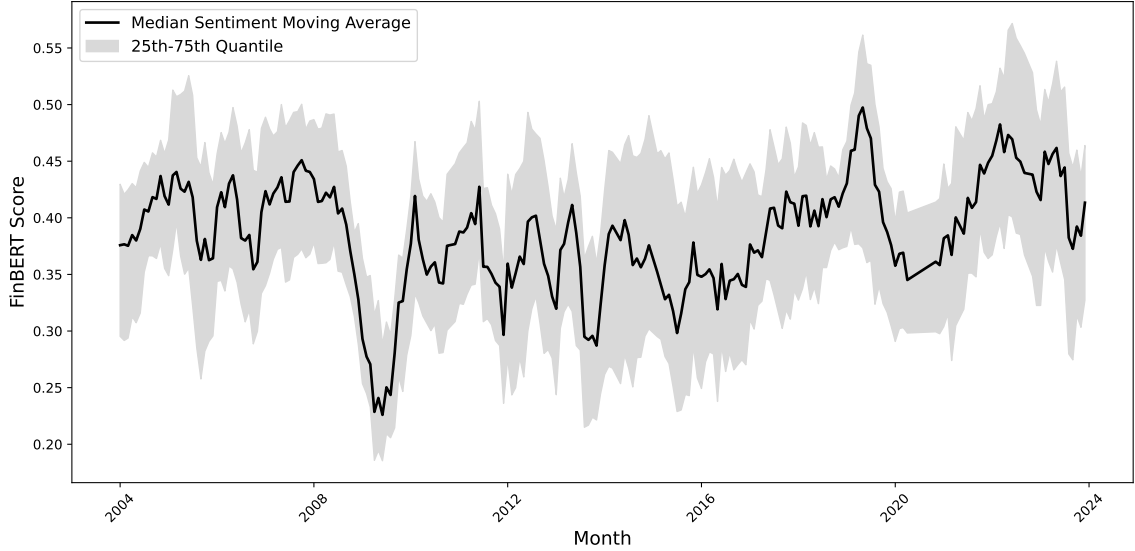
The BERT model is built on the Transformer architecture, which uses an attention mechanism capable of capturing contextual relationships across entire texts. Unlike traditional directional language models that analyze text sequentially (either left-to-right or right-to-left), the Transformer encoder processes entire sentences simultaneously, allowing it to incorporate bidirectional context and better understand each word in relation to its surroundings.

Due to the 512-token limit inherent in BERT-based models, we segment each PIN document into individual sentences prior to analysis. Each sentence is then evaluated using the FinBERT model, which assigns a sentiment score ranging from 0 (strongly negative) to 1 (strongly positive). We compute the average sentiment score across all sentences in a given PIN to derive a document-level measure of sentiment intensity. This approach enables us to quantify optimism or pessimism in IMF surveillance language, while capturing variation in tone both within and across documents.⁵

⁵To further ensure the robustness of our sentiment measurement, we benchmark our FinBERT-based approach against the foundational Loughran and McDonald dictionary method (Loughran and McDonald, 2011). Specifically, for the Loughran and McDonald method, we calculate sentiment as the difference between

Figure 9 illustrates the evolution of IMF sentiment toward countries in our sample, based on the textual content of Article IV PINs from 2004 to 2023. The black line represents the two-quarter moving average of the median FinBERT sentiment score, while the shaded area denotes the interquartile range, capturing the cross-country dispersion in sentiment.

Figure 9: Trends in PIN sentiment scores using FinBERT, 2004–2023



Note: The sentiment scores for each IMF Article IV Public Information Notice (PIN) are computed using the FinBERT model, which assigns a sentiment score to each sentence within the PIN text. These sentence-level scores are averaged to construct a document-level sentiment score per PIN. The figure presents a two-quarter moving average of median sentiment scores across all countries, along with the interquartile range (25th to 75th percentile), highlighting the evolution of IMF assessment tone over time.

The data reveal pronounced fluctuations in IMF tone, with marked declines during periods of global financial distress—such as the 2008–2009 financial crisis and the COVID-19 shock in 2020—followed by gradual recoveries. The widening dispersion in certain periods suggests increased heterogeneity in IMF assessments across countries, likely reflecting diverse country-specific vulnerabilities.

the number of positive and negative words, divided by the total number of positive and negative words present in each PIN: $\text{Sentiment} = ((\text{Positive Words} - \text{Negative Words}) / (\text{Positive Words} + \text{Negative Words}))$. This dictionary-based method imposes no tokenization limit, enabling analysis of the complete document text without segmentation. Figure D.1 in Appendix D compares both sentiment measures over time, highlighting a high correlation coefficient of 0.83 for the overall moving average series. Figure D.3 shows the distribution of the sentiment score for both the FinBERT and Loughran-McDonald (LM) methodologies along with the scatter of both measures. The FinBERT scores are more sharply centered around zero and exhibit lower variance, whereas the LM scores show a wider spread and a more left-skewed distribution, suggesting that LM may assign more extreme negative values. The correlation between the two sentiment measures, with a Pearson correlation coefficient of 0.69, indicate a relatively strong positive relationship.

In general, IMF PINs adopt cautious and measured language when discussing macroeconomic risks. Our FinBERT-based sentiment analysis indicates that the median sentiment score across all PINs is approximately 0.48, suggesting an overall neutral tone in the Fund’s communication. This neutrality likely reflects a structural tension in IMF surveillance: the need to deliver credible and transparent assessments while avoiding the risk of triggering adverse market reactions. Overly negative assessments may provoke self-fulfilling investor responses—such as widening bond spreads or capital outflows—that could exacerbate the vulnerabilities the Fund seeks to mitigate.

To navigate the complexity of interpreting IMF communications, we isolate the idiosyncratic—or “surprise”—component of each PIN’s tone; that is, the portion of sentiment that deviates from contemporaneous regional communication trends. Specifically, we construct a relative sentiment index by calculating the deviation of each PIN’s FinBERT sentiment score from the median sentiment of all PINs issued within the same IMF-defined region during a six-month window centered on the publication month. This approach controls for both global and regional shifts in tone, allowing us to extract the country-specific informational content of each release. The resulting deviation serves as a cleaner proxy for the “information shock” conveyed by the IMF—that is, the degree to which a given PIN is unexpectedly positive or negative relative to regional expectations.

Figure D.2 in Appendix D displays the evolution of FinBERT sentiment gap scores across the sample period, illustrating the de-trending effect of our methodology. By removing the common regional tone, the sentiment gap isolates idiosyncratic variation in IMF assessments. The resulting distribution fluctuates around zero, but exhibits pronounced positive and negative spikes. These sentiment “surprises” represent deviations from the regional average tone at the time of each PIN release, effectively capturing country-specific optimism or pessimism in IMF surveillance communications.

The presence of both upward and downward swings in the sentiment gap indicates that, at any given point in time, some countries received assessments notably more favorable or critical than their regional peers, even after conditioning on broader macroeconomic sentiment. This structure supports our interpretation of the FinBERT sentiment gap as a valid proxy for informational shocks, reflecting the unique tone embedded in each surveillance episode relative to prevailing regional narratives.

We follow the moderating-variable framework proposed by Imai et al. (2023), using each PIN’s FinBERT sentiment score as a moderator of the treatment effect. This approach allows us to estimate two separate dynamic ATT paths—one for “optimistic” PINs (positive sentiment surprises) and one for “pessimistic” PINs (negative sentiment surprises)—across

all of our outcome variables.

Figure 10 presents the dynamic ATT estimates, capturing the heterogeneous causal effects of IMF Article IV PIN releases by sentiment.

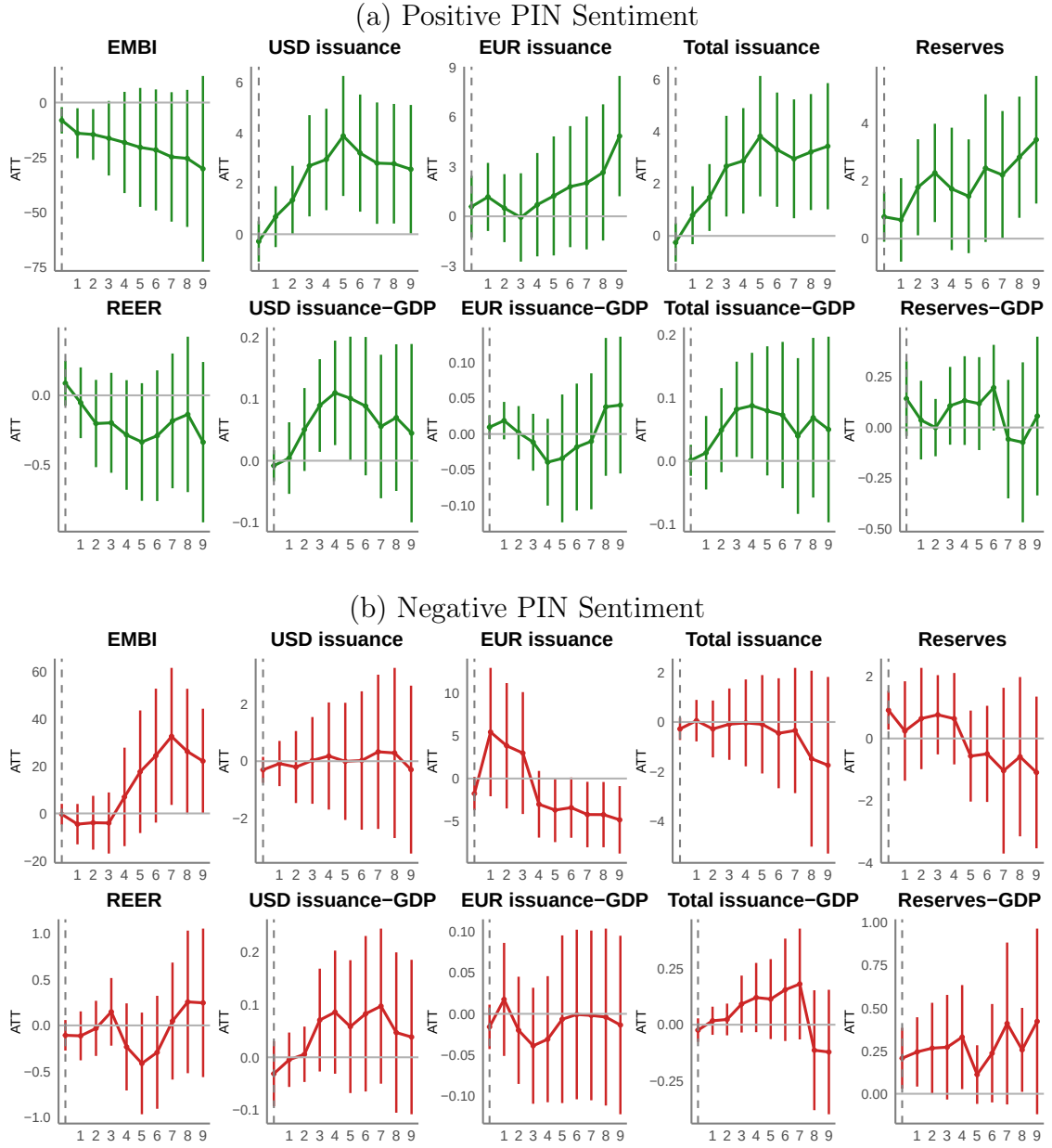
Panel (a) shows that optimistic sentiment surprises compress EMBI spreads by approximately 20–25 basis points—substantially larger than the 15 bp average effect in the baseline—and drive a 4 percent rise in USD-denominated sovereign issuance, compared to roughly 2 percent in the benchmark results. These effects are statistically significant between months 3 and 6, and the foreign-reserves-to-GDP ratio increases by 1–1.5 percentage points in months 4 through 6. Taken together, these findings suggest that positive PIN surprises not only lead to sharper reductions in borrowing costs but also translate into materially stronger improvements in market access.

Panel (b), by contrast, shows that negative sentiment surprises yield flat and statistically insignificant ATT estimates across most variables. EMBI spreads rise significantly only by month 7, indicating delayed investor concern. Similarly, EUR-denominated issuance declines significantly at this horizon, reflecting a contraction in market access following pessimistic IMF messaging. These findings support a “warning effect” interpretation of negative IMF assessments, where causal effects emerge more gradually and selectively.

Overall, the evidence points to a clear asymmetry: while positive IMF communications elicit immediate and favorable market responses, negative sentiment—when statistically significant—tends to generate delayed and adverse impacts.⁶

⁶As a robustness check, Figure D.4 in Appendix D replicates the dynamic treatment effect analysis using the Loughran–McDonald (LM) dictionary-based sentiment classification, in place of the FinBERT-based sentiment used in the main results. The core patterns remain broadly consistent: optimistic PINs are associated with improvements in market conditions and investor sentiment, whereas pessimistic PINs exhibit muted or negative effects.

Figure 10: *Heterogeneous treatment effects by PIN sentiment: optimistic vs. pessimistic releases*



Note: This figure displays the dynamic treatment effects of IMF Article IV PIN releases on sovereign bond spreads and related financial indicators, disaggregated by the sentiment of the PIN communication. Panel (a) reports results for PINs classified as optimistic, while Panel (b) reports results for PINs classified as pessimistic, based on the FinBERT sentiment score. Each subfigure plots the estimated Average Treatment Effect on the Treated (ATT) over a nine-month horizon following the PIN release. Shaded areas represent 90% confidence intervals based on bootstrapped standard errors. Estimates are derived using the same baseline specification as previous figures, with propensity score weighting (PSW) applied for covariate adjustment.

6.2. *The role of geopolitical alignment*

The second channel we investigate is geopolitical alignment: how investors’ reactions to IMF PINs vary depending on a country’s alignment with the United States–led international order. When markets perceive a country as closely aligned with major IMF shareholders or influential financial coalitions, favorable IMF surveillance reports may carry a “credibility premium,” signaling stronger institutional backing and greater policy stability. By contrast, countries viewed as less aligned—or even at odds diplomatically—may experience more muted or skeptical market responses to similar communications. We capture geopolitical alignment using voting similarity in the United Nations General Assembly (UNGA), which reflects long-run patterns of geopolitical affinity with key IMF stakeholders, particularly the United States.

Prior studies have shown that closer alignment with the United States predicts both higher IMF lending volumes (Dreher et al., 2009) and more favorable program conditionality (Stone, 2012). However, these findings are often confounded by the fact that lending programs typically respond to acute macroeconomic crises and involve significant resource commitments. In contrast, Article IV surveillance and PIN disclosures represent low-stakes, bilateral engagements, making them well-suited for a natural-experiment framework to isolate how geopolitical alignment influences market reactions—independent of crisis dynamics or financial bailouts.

To test this channel, we draw on UNGA ideal-point estimates from Fjelstul et al. (2025), which place countries along a latent dimension capturing their affinity with the United States–led liberal international order. We compute each country’s bilateral ideological distance from the United States using these ideal points.⁷

We then interact each country’s geopolitical distance from the United States in year t with the optimistic PIN surprises identified in the previous section—understood as the sentiment deviation of each PIN release compared to the regional sentiment over that period. This specification allows us to distinguish the political credibility channel from the pure informational channel, by testing whether positive IMF assessments elicit stronger market responses when issued for countries that are geopolitically aligned with the United States.

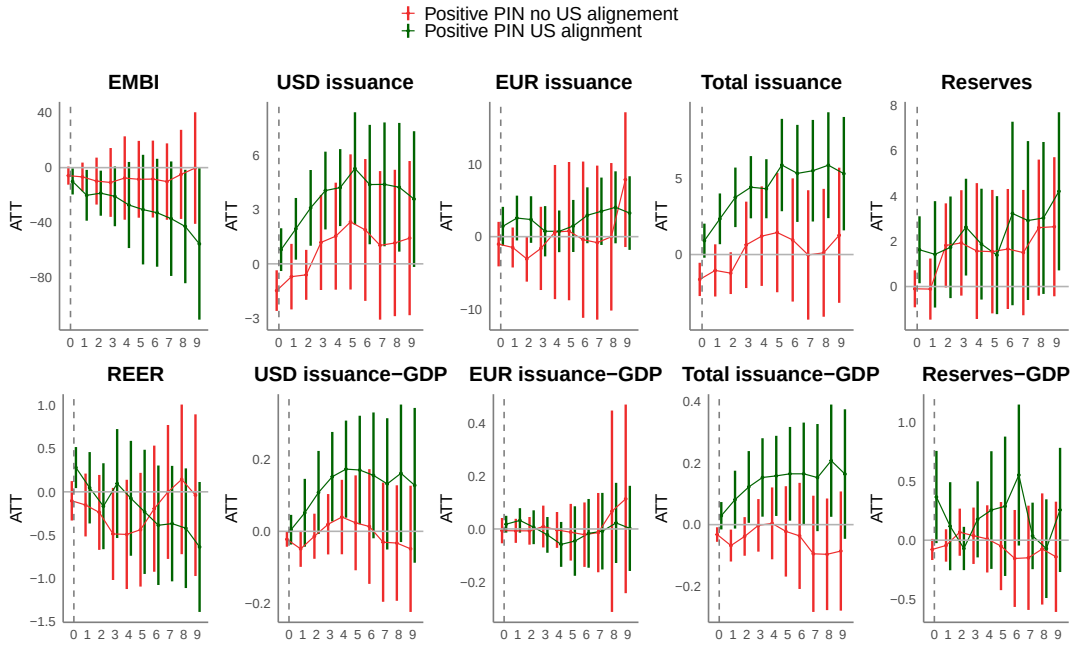
Figure 11 shows that the reduction in sovereign bond spreads and the corresponding increase in sovereign issuance—particularly in USD-denominated instruments—following optimistic PIN releases is concentrated among countries with above-median geopolitical alignment with

⁷The dynamic ordinal spatial model used to estimate ideal points treats each UNGA roll-call vote as an ordinal outcome, with probabilities determined by (a) the country’s latent ideal point, (b) vote-specific discrimination parameters, and (c) thresholds that separate “No,” “Abstain,” and “Yes” responses. See Bailey et al. (2017) for full methodological details.

the United States, as measured by UNGA voting similarity. In these geopolitically aligned states, USD-denominated issuance rises to approximately 6 percent, compared to a 4 percent peak in the overall optimistic-PIN sample. Similarly, the decline in sovereign spreads is sharply concentrated in this subgroup.

By contrast, for countries with below-median alignment, the effects of optimistic PINs on both spreads and issuance are statistically insignificant. This pattern suggests that positive IMF information surprises enhance access to international capital markets primarily for countries that are geopolitically proximate to the United States, reinforcing prior findings that geopolitical alignment conditions the credibility and market impact of IMF actions (Thacker, 1999; Barro and Lee, 2005).

Figure 11: Heterogeneity Results : Positive Sentiment and Geopolitical Alignment



Note: The figure presents heterogeneity analysis results comparing optimistic IMF PIN releases with additional conditioning on geopolitical alignment. Specifically, the new results highlight distinguishes countries that not only received optimistic PINs but also demonstrated higher geopolitical alignment with the United States, defined as above-median agreement in United Nations General Assembly (UNGA) voting patterns. The analysis uses the same baseline specification as previously shown, employing propensity score weighting (PSW) and bootstrapped 90 percent confidence intervals.

7. Conclusion

In this paper, we investigate how the International Monetary Fund's Article IV Public Information Notices (PINs) influence sovereign market access in 67 emerging market economies (EMEs) between January 2004 and January 2024. To address the limitations of standard

difference-in-differences (DiD) two-way fixed effects estimators in settings with repeated treatments, we apply the panel-matching DiD approach developed by [Imai et al. \(2023\)](#). This method allows us to construct control groups matched on observed covariates and accommodates non-absorbing treatment structures, thereby enhancing the credibility of our causal estimates.

Our core finding is that PIN releases lead to an immediate reduction in sovereign bond spreads—by approximately 15 basis points on average—accompanied by a contemporaneous surge in hard-currency debt issuance. Notably, this issuance effect is concentrated in US dollar-denominated borrowing, while euro-denominated issuance remains largely unaffected. This divergence suggests that financial markets place greater weight on IMF signals when evaluating dollar-denominated debt, consistent with the dollar’s role as the dominant global reserve currency.

Our results are robust across a series of validation checks, including: alternative treatment definitions based on staff report publication dates, placebo tests using future-dated PINs, and models that explicitly control for the Global Financial Crisis and COVID-19 pandemic. These findings also hold when outcomes are expressed in nominal terms or normalized as shares of GDP.

To examine the role of information content, we use natural language processing (NLP) to extract sentiment surprises from the textual content of PINs. We find a clear asymmetry: positive sentiment shocks significantly ease market conditions and stimulate greater issuance, while negative shocks elicit only muted responses, though they are followed by delayed upward pressure on spreads. This pattern suggests that optimistic IMF language sends a powerful confidence signal, whereas negative assessments are likely tempered to avoid triggering adverse market reactions.

Finally, we show that the strength of these effects is conditioned by a country’s geopolitical alignment, measured via voting similarity with the United States (US) in the United Nations General Assembly. Countries most closely aligned with the US experience the largest increases in dollar-denominated issuance following positive sentiment shocks. This mirrors earlier findings on IMF lending and conditionality, indicating that optimistic surveillance language only translates into improved market access for US geopolitical allies—revealing a structural bias in the operation of bilateral surveillance.

Together, these results underscore the powerful role of IMF surveillance in shaping sovereign access to global capital markets. By combining a causal panel-matching framework with textual sentiment analysis and measures of geopolitical proximity, we provide new evidence on the informational and political channels through which IMF PIN disclosures affect financial

market outcomes.

References

- Abeywickrama, K., Perera, N., Samarathunga, S., Pabasara, H., Jayathilaka, R., Wisenthige, K., 2024. Factors influencing IMF assistance in the Sub-Saharan African region. *PloS one* 19, e0307071.
- Bailey, M.A., Strezhnev, A., Voeten, E., 2017. Estimating Dynamic State Preferences from United Nations Voting Data. *Journal of Conflict Resolution* 61, 430–456.
- Baldacci, E., Gupta, S., Mati, A., 2008. Is it (Still) Mostly Fiscal? Determinants of Sovereign Spreads in Emerging Markets. IMF Working Paper WP/08/259. International Monetary Fund.
- Barro, R.J., Lee, J.W., 2005. IMF programs: Who is chosen and what are the effects? *Journal of Monetary Economics* 52, 1245–1269.
- Bird, G., Rowlands, D., 2017. The Effect of IMF Programmes on Economic Growth in Low Income Countries: An Empirical Analysis. *The Journal of Development Studies* 53, 2179–2196.
- Borgy, V., Laubach, T., Mésonnier, J.S., Renne, J.P., 2011. Fiscal Sustainability, Default Risk and Euro Area Sovereign Bond Spreads Markets. Technical Report 350. Banque de France.
- Boughton, J.M., 2012. Global Oversight: Strengthening and Broadening IMF Surveillance, in: *Tearing Down Walls: The International Monetary Fund 1990–1999*. International Monetary Fund, Washington, DC. chapter 4, pp. 109–185.
- Breen, M., Doak, E., 2023. The IMF as a global monitor: surveillance, information, and financial markets. *Review of International Political Economy* 30, 307–331.
- Breen, M., Hodson, D., Moschella, M., 2020. Incoherence in regime complexes: A sentiment analysis of eu-imf surveillance. *JCMS: Journal of common market studies* 58, 419–437.
- Brender, A., Drazen, A., 2005. Political budget cycles in new versus established democracies. *Journal of Monetary Economics* 52, 1271–1295.
- Brookhart, M.A., Schneeweiss, S., Rothman, K.J., Glynn, R.J., Avorn, J., Stürmer, T., 2006. Variable Selection for Propensity Score Models. *American Journal of Epidemiology* 163, 1149–1156.

- Callaway, B., Sant’Anna, P.H., 2021. Difference-in-differences with multiple time periods. *Journal of econometrics* 225, 200–230.
- Cantor, R., Packer, F., 1996. Determinants and Impact of Sovereign Credit Ratings. *Economic Policy Review* 2, 37–54.
- Chapman, T., Fang, S., Li, X., Stone, R.W., 2017. Mixed Signals: IMF Lending and Capital Markets. *British Journal of Political Science* 47, 329–349.
- Chinn, M.D., Ito, H., 2008. A New Measure of Financial Openness. *Journal of Comparative Policy Analysis: Research and Practice* 10, 309–322.
- Choi, S., Hashimoto, Y., 2018. Does transparency pay? evidence from IMF data transparency policy reforms and emerging market sovereign bond spreads. *Journal of International Money and Finance* 88, 171–190.
- Darvas, Z., 2021. Timely Measurement of Real Effective Exchange Rates. Working Paper 15/2021. Bruegel.
- De Chaisemartin, C., d’Haultfoeuille, X., 2024. Difference-in-Differences Estimators of Intertemporal Treatment Effects. *Review of Economics and Statistics* , 1–45.
- De Chaisemartin, C., d’Haultfoeuille, X., 2020. Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review* 110, 2964–2996.
- Dreher, A., Sturm, J.E., Vreeland, J.R., 2009. Global horse trading: IMF loans for votes in the United Nations Security Council. *European Economic Review* 53, 742–757.
- Du, K., Zhao, Y., Mao, R., Xing, F., Cambria, E., 2025. Natural language processing in finance: A survey. *Information Fusion* 115, 102755.
- Edwards, M.S., Coolidge, K.A., Preston, D.A., 2011. Who reveals? Transparency and the IMF’s article IV consultations. *SSRN Electronic Journal* .
- Eichengreen, B., Kletzer, K., Mody, A., 2006. The IMF in a world of private capital markets. *Journal of Banking & Finance* 30, 1335–1357.
- Eichengreen, B., Mussa, M., 1998. Capital Account Liberalization and the IMF. *Finance and Development* 35, 16.
- Fayad, G., Huang, C., Shibuya, Y., Zhao, P., 2020. How Do Member Countries Receive IMF Policy Advice: Results from a State-of-the-art Sentiment Index. IMF Working Paper WP/20/7. International Monetary Fund.

- Fjelstul, J., Hug, S., Kilby, C., 2025. Decision-making in the United Nations General Assembly: A comprehensive database of resolution-related decisions. *The Review of International Organizations* , 1–18.
- Fratzscher, M., Reynaud, J., 2011. IMF surveillance and financial markets—A political economy analysis. *European Journal of Political Economy* 27, 405–422.
- Gaillard, N., 2014. What Is the Value of Sovereign Ratings? *German Economic Review* 15, 208–224.
- Gehring, K., Lang, V.F., 2018. Stigma or Cushion? IMF Programs and Sovereign Creditworthiness. CIS Working Paper 98. Center for Comparative and International Studies.
- Glennerster, R., Shin, Y., 2003. Is Transparency Good For You, and Can the IMF Help? IMF Working Paper WP/03/132. International Monetary Fund.
- Glennerster, R., Shin, Y., 2008. Does transparency pay? *IMF Staff Papers* 55, 183–209.
- Goes, I., Chapman, T.L., 2024. Can “Soft” Advice from International Organizations Catalyze Natural Resource Sector Reform? *International Studies Quarterly* 68, sqae048.
- Goodman-Bacon, A., 2021. Difference-in-differences with variation in treatment timing. *Journal of Econometrics* 225, 254–277.
- Ilzetzki, E., Reinhart, C.M., Rogoff, K.S., 2022. Rethinking Exchange Rate Regimes, in: Gopinath, G., Helpman, E., Rogoff, K.S. (Eds.), *Handbook of International Economics*. Elsevier: North Holland. volume 6, pp. 91–145.
- Imai, K., Kim, I.S., 2019. When should we use unit fixed effects regression models for causal inference with longitudinal data? *American Journal of Political Science* 63, 467–490.
- Imai, K., Kim, I.S., Wang, E.H., 2023. Matching methods for causal inference with time-series cross-sectional data. *American Journal of Political Science* 67, 587–605.
- Imbens, G.W., Rubin, D.B., 2015. *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press, Cambridge.
- IMF, 1998. Annual Report 1998. International Monetary Fund, Washington, DC.
- IMF, 2001. Issues Brief: Transparency at the IMF.
- IMF, 2003a. Annual Report 2003: Making the Global Economy Work for All. International Monetary Fund, Washington, DC.

- IMF, 2003b. The Fund’s Transparency Policy–Review of Experience and Next Steps. Policy Paper. International Monetary Fund. Washington, DC.
- IMF, 2009. Review of the Fund’s Transparency Policy. Policy Paper. International Monetary Fund. Washington, DC.
- IMF, 2013. Press Release: IMF Executive Board Reviews the IMF’s Transparency Policy. Technical Report. International Monetary Fund. Washington, DC.
- IMF, 2016a. IMF Annual Report 2016: Finding Solutions Together. International Monetary Fund, Washington, DC.
- IMF, 2016b. Implementation Plan in Response to Board–Endorsed Recommendations from the IEO Evaluation. Technical Report. International Monetary Fund. Washington, DC.
- IMF, 2022. IMF Capacity Development Strategy 2022–25. International Monetary Fund, Washington, DC.
- IMF, 2024. Review of the Fund’s Transparency Policy and Open Archives Policy. Policy Paper. International Monetary Fund. Washington, DC.
- Jakobsen, J., De Soysa, I., 2006. Do Foreign Investors Punish Democracy? Theory and Empirics, 1984–2001. *Kyklos* 59, 383–410.
- Kogan, J., Kazandjian, R., Luo, S., Mbohou Mama, M., Miao, H., 2024. The Role of IMF Arrangements in Restoring Access to International Capital Markets. IMF Working Paper WP/24/173. International Monetary Fund.
- Liu, Z., Huang, D., Huang, K., Li, Z., Zhao, J., 2020. FinBERT: A Pre-trained Financial Language Representation Model for Financial Text Mining, in: Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20), Special Track on AI in FinTech, pp. 4513–4519.
- Loughran, T., McDonald, B., 2011. When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of finance* 66, 35–65.
- Mihalyi, D., Mate, A., 2019. Text-mining IMF country reports - an original dataset. MPRA Paper 100656. Munich Personal RePEc Archive.
- Mody, A., Saravia, D., 2006. Catalysing private capital flows: do IMF programmes work as commitment devices? *The Economic Journal* 116, 843–867.

- Momani, B., Hibben, M., 2024. Oxford Handbook of the International Monetary Fund. Oxford University Press.
- Müller, K., Xu, C., Lehibib, M., Chen, Z., 2025. The Global Macro Database: A New International Macroeconomic Dataset. Available at SSRN 5121271 .
- Mussa, M., 1997. IMF surveillance. *The American Economic Review* 87, 28–31.
- Oatley, T., Yackee, J., 2004. American Interests and IMF Lending. *International Politics* 41, 415–429.
- Pfeifer, M., Marohl, V.P., 2023. Centralbankroberta: A fine-tuned large language model for central bank communications. *The Journal of Finance and Data Science* 9, 100114.
- Ramos, L., Gallagher, K.P., Stephenson, C., Monasterolo, I., 2022. Climate risk and IMF surveillance policy: a baseline analysis. *Climate Policy* 22, 371–388.
- Reinhart, C.M., Rogoff, K.S., 2008. This Time Is Different: A Panoramic View of Eight Centuries of Financial Crises. NBER Working Paper 13882. National Bureau of Economic Research.
- Reinsberg, B., Stubbs, T., Kentikelenis, A., 2022. Compliance, defiance, and the dependency trap: International Monetary Fund program interruptions and their impact on capital markets. *Regulation & Governance* 16, 1022–1041.
- Robins, J.M., Hernán, M.A., Brumback, B., 2000. Marginal Structural Models and Causal Inference in Epidemiology. *Epidemiology* 11, 550–560.
- Schlegl, M., Trebesch, C., Wright, M.L.J., 2019. The seniority structure of sovereign debt. NBER Working Paper 25793. National Bureau of Economic Research.
- Shim, S., 2022. Who is credible? Government popularity and the catalytic effect of IMF lending. *Comparative Political Studies* 55, 2147–2177.
- Stone, R.W., 2012. Lending credibility: The International Monetary Fund and the post-communist transition. Princeton University Press.
- Stuart, E.A., Lee, B.K., Leacy, F.P., 2013. Prognostic score-based balance measures can be a useful diagnostic for propensity score methods in comparative effectiveness research. *Journal of clinical epidemiology* 66, S84–S90.
- Sun, L., Abraham, S., 2021. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics* 225, 175–199.

- Thacker, S.C., 1999. The High Politics of IMF Lending. *World politics* 52, 38–75.
- Yang, Y., Uy, M.C.S., Huang, A., 2020. Finbert: A pretrained language model for financial communications. *arXiv preprint arXiv:2006.08097* .
- Zhao, Q., 2019. Covariate balancing propensity score by tailored loss functions. *The Annals of Statistics* 47, 965.

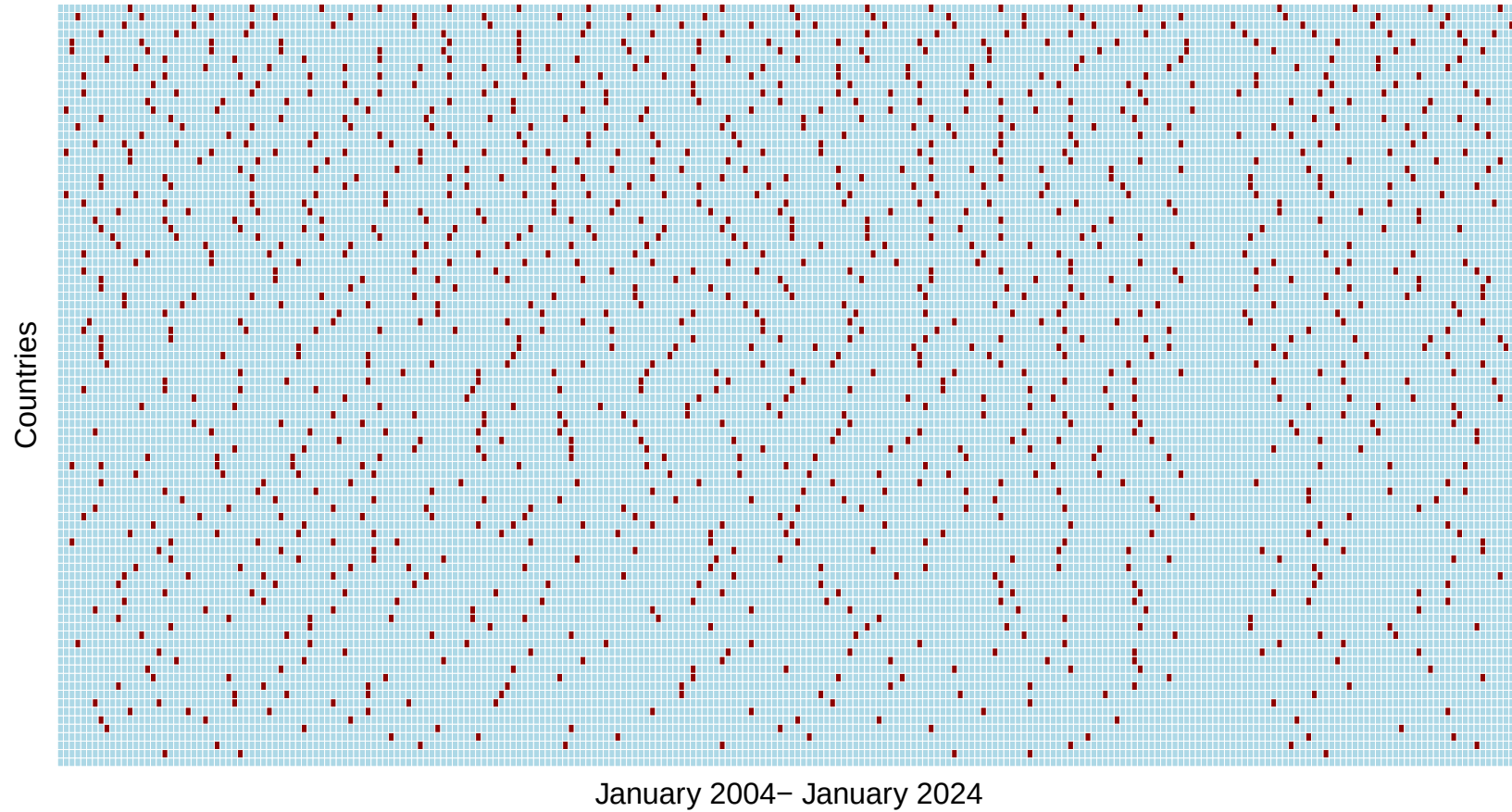
Appendix A - Data

Table A.1: List of countries included in the sample

Asia and Pacific	Europe	Middle East and Central Asia	Sub-Saharan Africa	Western Hemisphere
China	Bulgaria	Armenia	Angola	Aruba
Indonesia	Belarus	Azerbaijan	Côte d'Ivoire	Argentina
South Korea	Croatia	Bahrain	Gabon	Bahamas
Sri Lanka	Hungary	Egypt	Ghana	Belize
Mongolia	Lithuania	Georgia	Kenya	Bolivia
Malaysia	Latvia	Iraq	Mozambique	Brazil
Philippines	Poland	Jordan	Nigeria	Barbados
Vietnam	Romania	Kazakhstan	Rwanda	Chile
	Russia	Lebanon	Senegal	Colombia
	Serbia	Morocco	South Africa	Costa Rica
	Slovakia	Pakistan	Zambia	Dominican Republic
	Turkey	Qatar		Ecuador
	Ukraine	Tunisia		Guatemala
				Honduras
				Jamaica
				Mexico
				Peru
				Paraguay
				El Salvador
				Trinidad and Tobago
				Uruguay

Note: Countries are grouped according to the IMF regional classification. Africa includes Sub-Saharan African countries; Asia and Pacific covers countries in East Asia, South Asia, Southeast Asia, and the Pacific; Europe encompasses European countries, including both EU and non-EU members; Middle East and Central Asia comprises Middle Eastern countries, North Africa, and Central Asian republics; Western Hemisphere includes North America, Central America, the Caribbean, and South America.

Figure A.1: IMF Article IV PIN release timing across countries (2004–2023)



Note: The figure visualizes the distribution of IMF Public Information Notice (PIN) disclosures across countries and time. Each row corresponds to a country, while each column represents a month, spanning from January 2004 to January 2024. Red dots indicate treated country-month observations, defined as periods when a PIN was disclosed for a given country.

Appendix B - Covariates and covariate balance

B.1. Potential confounders

The choice of covariates should be guided by economic theory and prior empirical findings, while ensuring that the model remains parsimonious (Imbens and Rubin, 2015). Including extraneous variables can unnecessarily restrict common support, while omitting key confounders can introduce significant bias into estimated treatment effects (Brookhart et al., 2006).

We adjust for several types of covariates based on their theorized association with both treatment and outcome. First, we include macroeconomic variables that capture a country’s external and fiscal position, which directly influence market risk assessments and IMF surveillance decisions. The current account balance (as a percentage of GDP) reflects a country’s external financing needs and vulnerability to balance of payments crises. Countries with persistent current account deficits are more likely to attract IMF monitoring and face higher sovereign spreads (Reinhart and Rogoff, 2008). Similarly, general government gross debt and net lending/borrowing (both as a percentage of GDP) capture fiscal sustainability and potential financing gaps, making them key determinants of both IMF consultations and market pricing of default risk (Borgy et al., 2011). The tax revenue-to-GDP ratio reflects a government’s capacity to generate resources, which affects IMF fiscal assessments and market confidence (Baldacci et al., 2008).

We also include indicators of economic performance and stability. Consumer Price Index (CPI)-based inflation, expressed as year-on-year change, is a core macroeconomic indicator that signals monetary stability and may trigger IMF concern in the context of overheating or demand collapse. Real GDP growth (YoY) and GDP per capita (PPP) are standard metrics of output performance and development level, which both the IMF and financial markets consider when assessing creditworthiness. Population size is included to proxy for market size and economic scale, which may influence IMF strategic interest and market perceptions of liquidity and debt absorption capacity.

In terms of sovereign risk perception, we include the sovereign credit ratings from Fitch, Moody’s, and S&P. These ratings directly reflect market assessments of credit risk and often respond to IMF interactions, but also pre-determine market expectations of country risk (Cantor and Packer, 1996; Gaillard, 2014). They may also co-determine the likelihood of IMF consultations if declining ratings are viewed as a signal of increasing vulnerability.

We further include political and institutional factors that influence both the IMF’s surveillance intensity and investors’ trust in a country’s policy commitments. The liberal democracy index from V-Dem reflects the quality of governance and democratic accountability, which

shape IMF engagement strategies ([Abeywickrama et al., 2024](#)) and investor confidence in policy stability ([Jakobsen and De Soysa, 2006](#)). Political timing variables such as election years are included to control for political cycles that may influence the timing of Article IV consultations or PIN releases, as well as investor concerns over policy uncertainty or fiscal populism during elections ([Brender and Drazen, 2005](#)). Finally, we control for a country’s geopolitical alignment and global engagement using UN General Assembly voting patterns and ideal points ([Fjelstul et al., 2025](#)), which have been shown to influence IMF treatment and broader international financial relations [Thacker \(1999\)](#); [Oatley and Yackee \(2004\)](#).

We add the Chinn-Ito index ([Chinn and Ito, 2008](#)) to capture financial openness, since capital account liberalization increases vulnerability to external shocks and typically leads to greater IMF scrutiny ([Eichengreen and Mussa, 1998](#)). The exchange rate regime classification ([Ilzetzki et al., 2022](#)) is also critical, as exchange rate arrangements interact with monetary and fiscal policy credibility and IMF surveillance priorities.

The definitions and sources of data for confounding variables are presented in the following table.

Table B.1: *Covariates description and data sources*

Variable Name	Frequency	Source
Macroeconomic Data		
Current Account Balance (% of GDP)	Yearly data splined to monthly	S&P Global Market Intelligence
Consumer Price Index (CPI), Year-on-Year % Change	Monthly	S&P Global Market Intelligence
Real GDP, Growth Rate (YoY)	Yearly data splined to monthly	S&P Global Market Intelligence
Real GDP Per Capita (PPP 2019 USD)	Yearly data splined to monthly	S&P Global Market Intelligence
General government gross debt (% of GDP)	Yearly data splined to monthly	IMF World Economic Outlook
Sovereign rating (Fitch, Moody's and S&P)	Monthly	Trading Economics
General government net lending/borrowing (% of GDP)	Yearly data splined to monthly	IMF World Economic Outlook
General government tax revenues (% of GDP)	Yearly data splined to monthly	Global Macro Database (Müller et al., 2025)
Population: Total (millions)	Yearly data splined to monthly	S&P Global Market Intelligence
Political Data		
Liberal democracy index	Yearly data splined to monthly	V-Dem Dataset v15 ^a
Government stability index	Monthly	ICRG Historical Data
Chinn-Ito Index (KAOPEN Index)	Annual or latest available	Chinn and Ito (2008) ^b
Exchange rate regime classification	Monthly classification	Ilzetzi et al. (2022) ^c
Monitoring of Fund Arrangements (MONA)	Monthly	IMF ^d
United Nations GA Votes and Ideal Points	Annual Sessions	UN Votes (Fjelstul et al., 2025) ^e
Election year for parliament (if parliamentary) or president (if presidential)	Annual update	Database of Political Institutions 2020 ^f

^a <https://www.v-dem.net/data/the-v-dem-dataset/>

^b https://web.pdx.edu/~ito/Chinn-Ito_website.htm

^c <https://www.ilzetzi.com/irr-data>

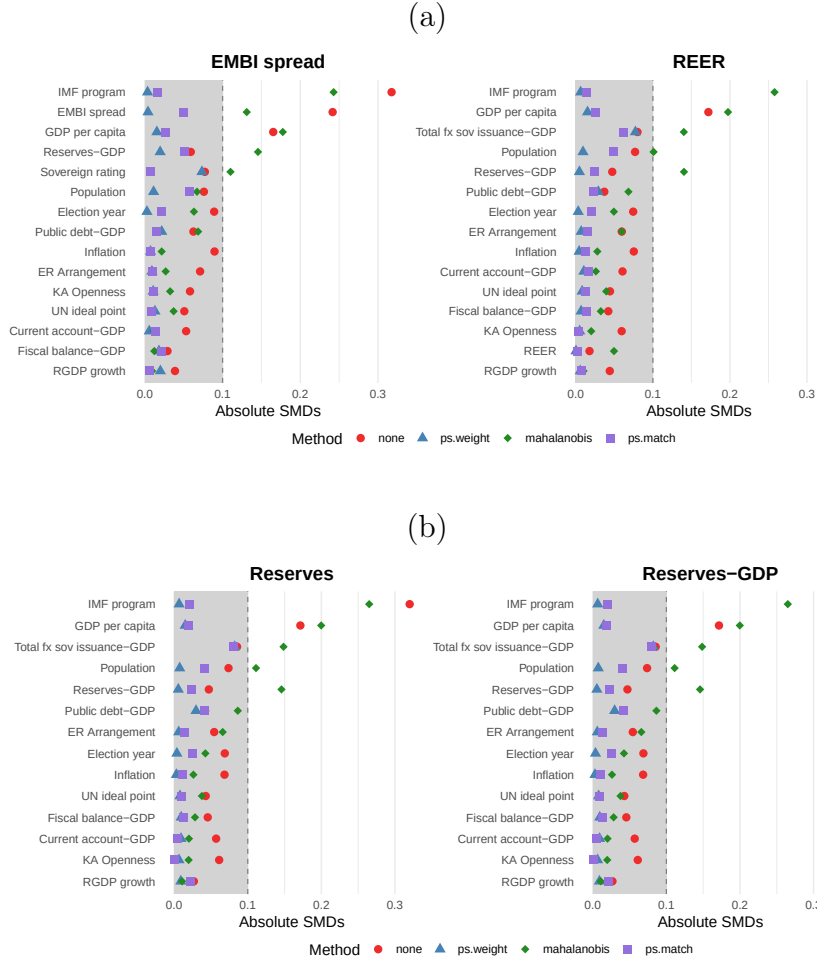
^d <https://www.imf.org/external/np/pdr/mona/index.aspx>

^e <https://github.com/evoeten/United-Nations-General-Assembly-Votes-and-Ideal-Points>

^f <https://publications.iadb.org/en/database-political-institutions-2020-dpi2020>

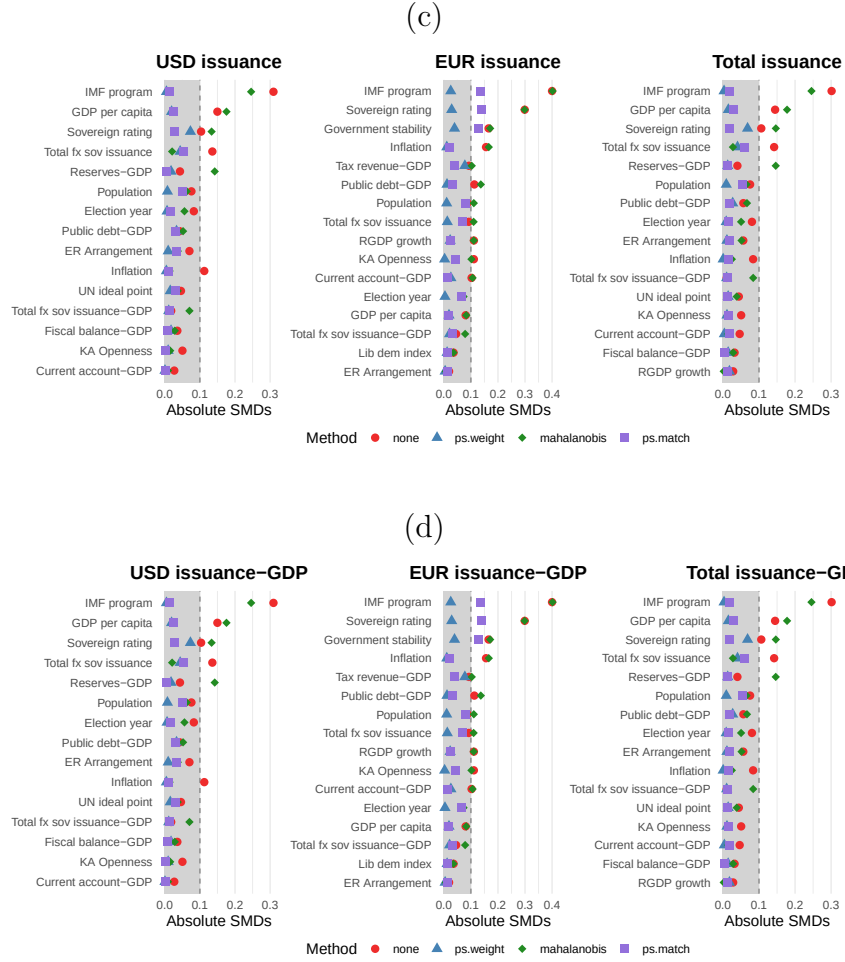
B.2. Covariate balance between treated and control groups

Figure B.1: Covariate balance panels (1)



Note: Panels (a) and (b) display the absolute standardized differences (ASDs) of selected covariates—each computed over four lags, in line with our benchmark specification—under four methods: no matching (red circles), Mahalanobis matching (green diamonds), propensity-score matching (purple squares), and propensity-score weighting (blue triangles). Covariates are chosen based on the two rules for each given outcome, (1) their ASD under propensity-score weighting is below 0.1 and (2) that ASD is lower than under no matching. Panel (a) reports these ASDs for the EMBI spread and the real effective exchange rate (REER), while Panel (b) shows them for foreign reserves and foreign reserves-to-GDP.

Figure B.2: Covariate balance panels (2)



Note: Panels (c) and (d) plot the absolute standardized differences (ASDs) of selected covariates—each measured over four lags, as in our benchmark specification—under four approaches: no matching (red circles), Mahalanobis matching (green diamonds), propensity-score matching (purple squares), and propensity-score weighting (blue triangles). Covariates are chosen based on the two rules for each given outcome, (1) their ASD under propensity-score weighting is below 0.1 and (2) that ASD is lower than under no matching. Panel (c) reports these ASDs for the nominal sovereign issuance outcomes (USD, EUR, and total issuance in USD), while Panel (d) presents the ASDs for those same issuance measures expressed as percentages of GDP.

B.3. Example: September 2014 Angola PIN release

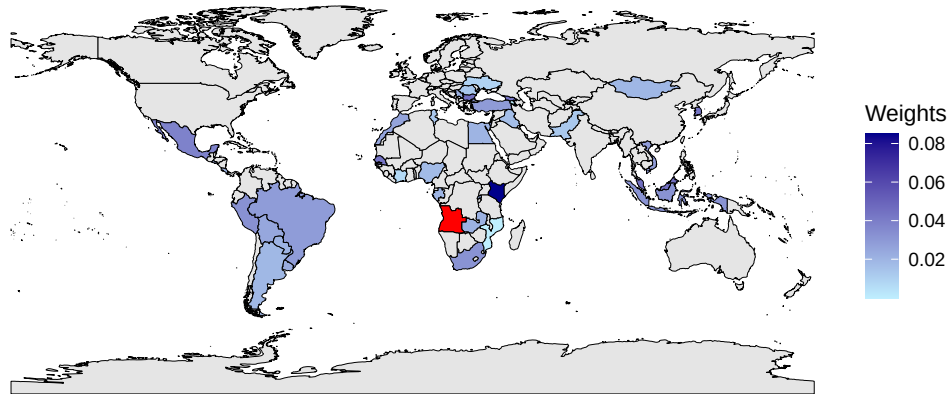
Figure B.3 illustrates the application of Propensity Score Weighting (PSW) in building a counterfactual for Angola's economic performance around the September 2014 Angola PIN release.

The figure is divided into two parts: the world map displays the weights attributed to various countries based on the Propensity Score Weighting methodology. Angola is highlighted in red, signifying that it is assigned the highest weight, reflecting its key role in the counterfactual analysis. Countries are color-coded based on the strength of the weight assigned, with shades of blue indicating countries with lower weights. The color gradient from dark blue to red emphasizes how the weights differ across countries, with Angola being the most influential country in the analysis due to its role in the study. This map showcases how countries are weighted to create a counterfactual comparison for Angola's economic trajectory post-2014.

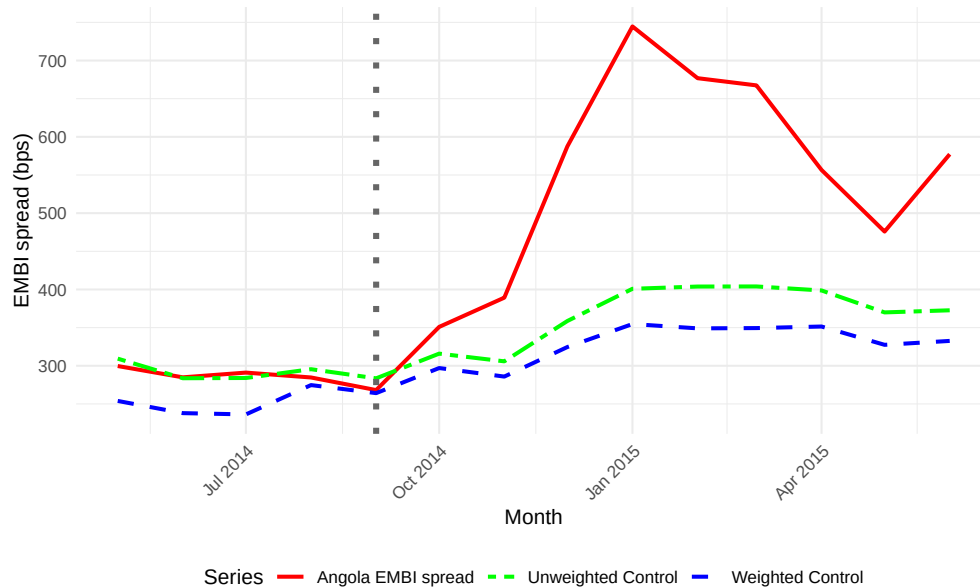
The plot illustrates the construction of the counterfactual series for Angola's EMBI spread. The figure compares Angola's observed EMBI spread (red line) with the weighted average EMBI spread of control countries (blue and green lines). The solid green line represents the "Weighted Control" series, which is the weighted average of control countries using Propensity Score Weights. The dashed blue line represents the "Unweighted Control" series, which is simply the unweighted average of control countries. The graph provides a clear visual comparison, showing how the post-release series for Angola differs from the counterfactual, constructed using the weighted control series.

Figure B.3: Propensity score weighting at work : September 2014 Angola PIN release

(a) World map propensity score weights



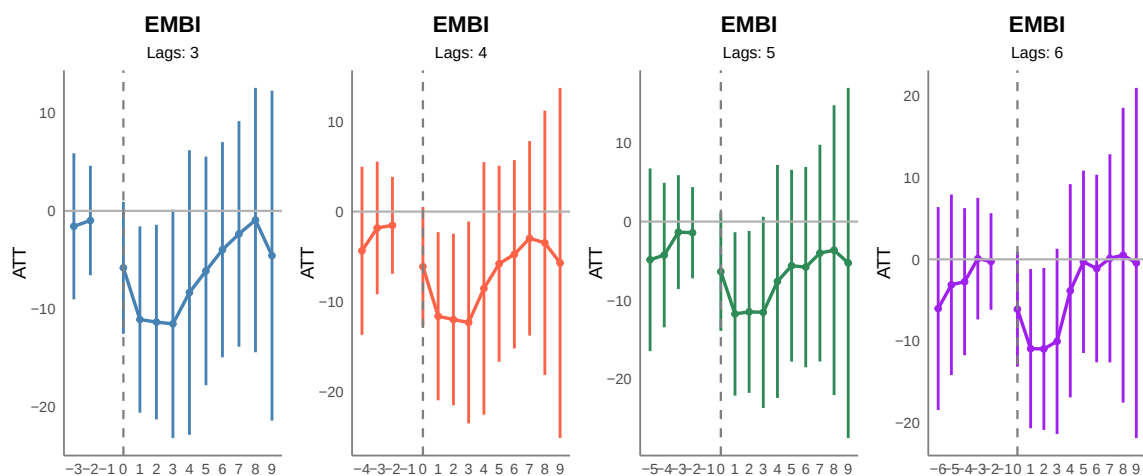
(b) Example on EMBI spread counterfactual



Note: The world map in Graph (a) shows the weights attributed to the other countries to build a counterfactual for the Angola 2014 PIN release. Angola is represented with a red color while the blue shade is stronger based on the weight given to each of the countries in the propensity score weighting methodology. The graph (b) illustrates the construction of the counterfactual scenario by comparing the observed post-event series for Angola with the weighted average of control countries, calculated using propensity score weights. The treatment effect is then measured as the difference between the observed series and this counterfactual series following the event.

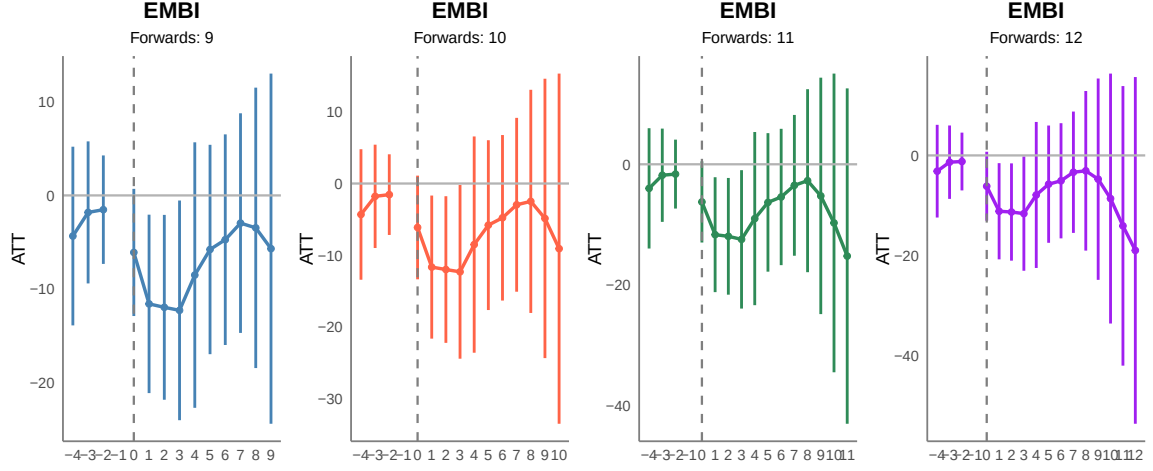
Appendix C - Additional robustness checks

Figure C.1: Lag length robustness : Effect of Public Information Notice (PIN) Releases



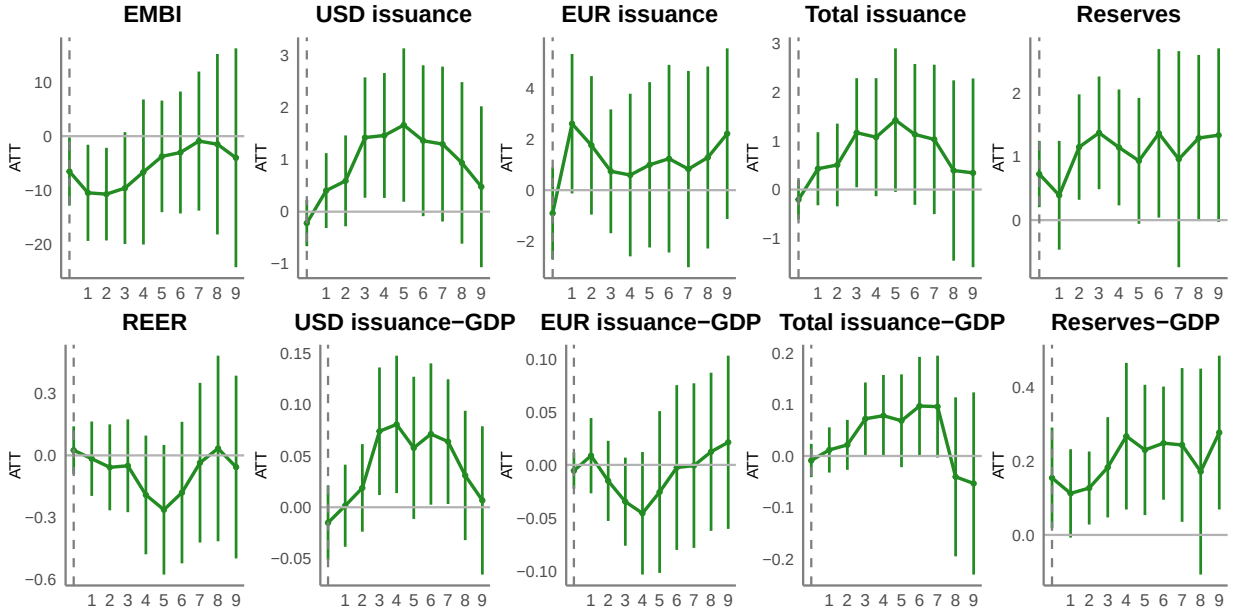
Note: This figure displays the dynamic treatment effects of IMF Article IV PIN releases on the EMBI spread in basis points. We report the effects using four different lag specifications, from 3 months, to 6 months. Each subfigure plots the estimated Average Treatment Effect on the Treated (ATT) over a nine-month horizon following the PIN release. Vertical lines represent 90% confidence intervals based on bootstrapped standard errors.

Figure C.2: Forward horizon robustness : Effect of Public Information Notice (PIN) Releases



Note: This figure displays the dynamic treatment effects of IMF Article IV PIN releases on the EMBI spread in basis points. We report the effects using four different forward horizon specifications, from 9, 10, 11 and 12 months. Each subfigure plots the estimated Average Treatment Effect on the Treated (ATT) over the chosen horizon following the PIN release. Vertical lines represent 90% confidence intervals based on bootstrapped standard errors.

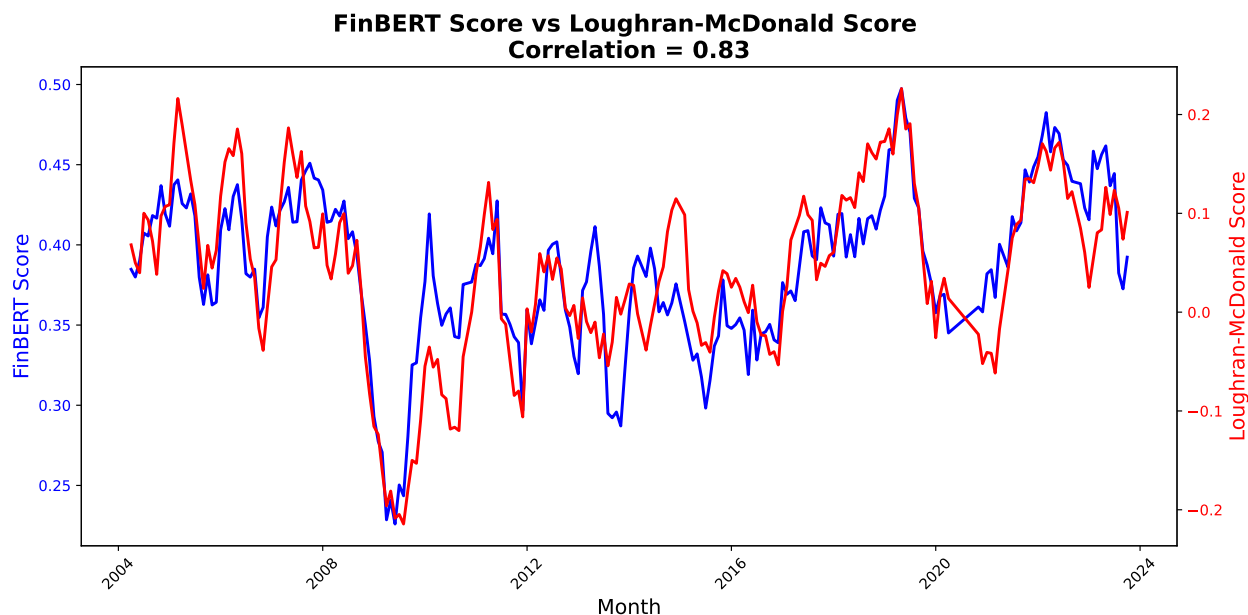
Figure C.3: Excluding crises periods: Effect of Public Information Notice (PIN) Releases



Note: This figure displays the dynamic treatment effects of IMF Article IV PIN releases on our variables of interest. We report the effects using the propensity score weighting methodology showing the estimated Average Treatment Effect on the Treated (ATT) over the nine-month horizon following the PIN release and the four-month horizon before. Shaded areas represent 90% confidence intervals based on bootstrapped standard errors. We exclude from the sample the August 2008 to March 2009 for the GFC and the March 2020 to May 2020 for the pandemic based on spikes on the VIX.

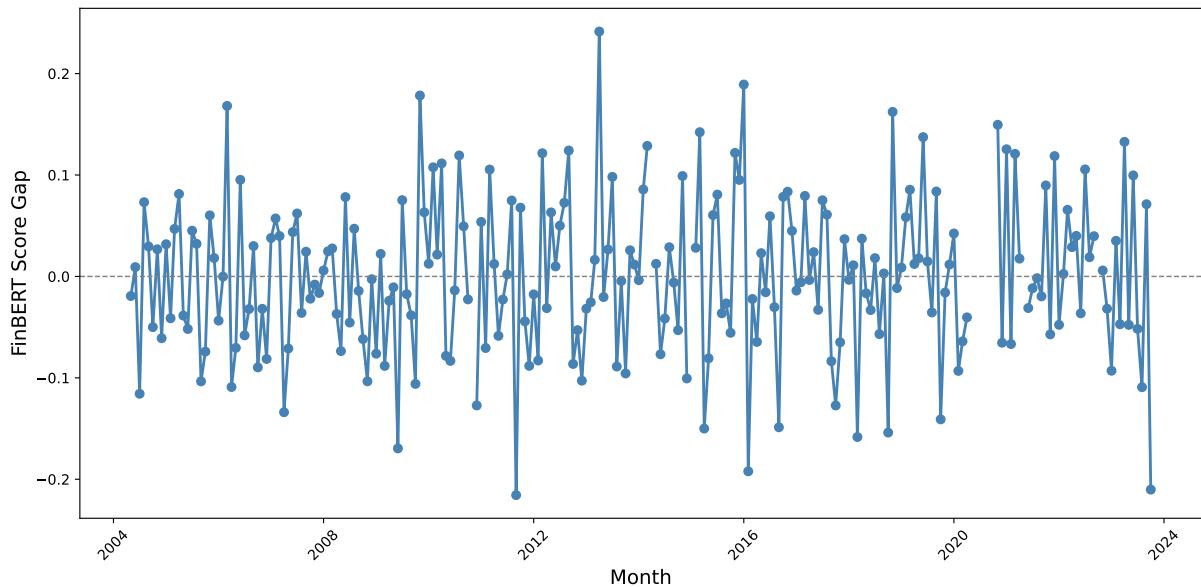
Appendix D - Sentiment analysis

Figure D.1: Monthly sentiment scores from FinBERT and Loughran–McDonald Dictionary methods, 2004–2023



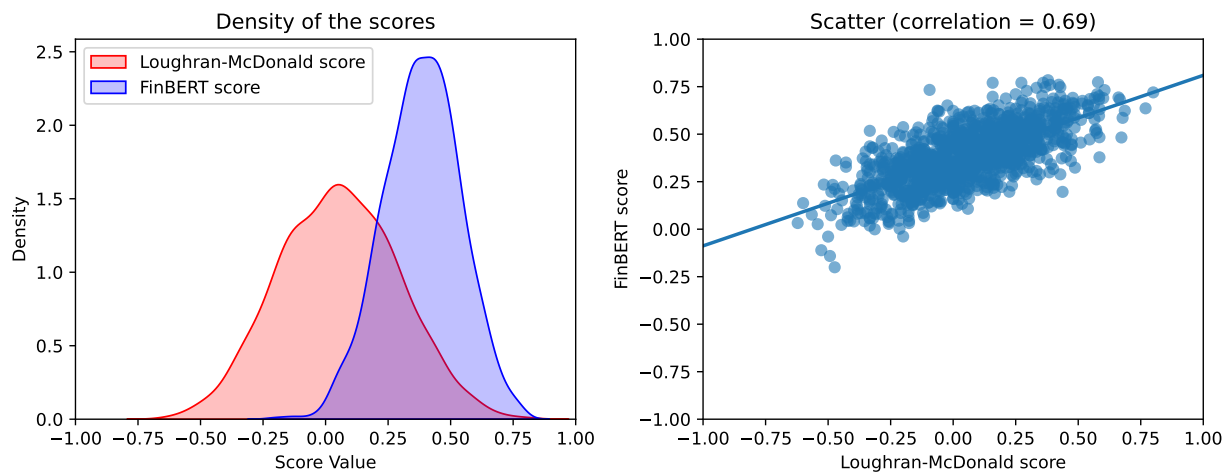
Note: This figure compares sentiment scores computed using the FinBERT model with those generated from the Loughran–McDonald financial dictionary approach. Each series represents the monthly average sentiment score across IMF Article IV PINs. The observed correlation of 0.83 suggests a strong alignment between the two sentiment measurement techniques, validating the consistency of FinBERT-based sentiment extraction over the analyzed period.

Figure D.2: Country-specific sentiment gaps in IMF Article IV PINs, 2004–2023



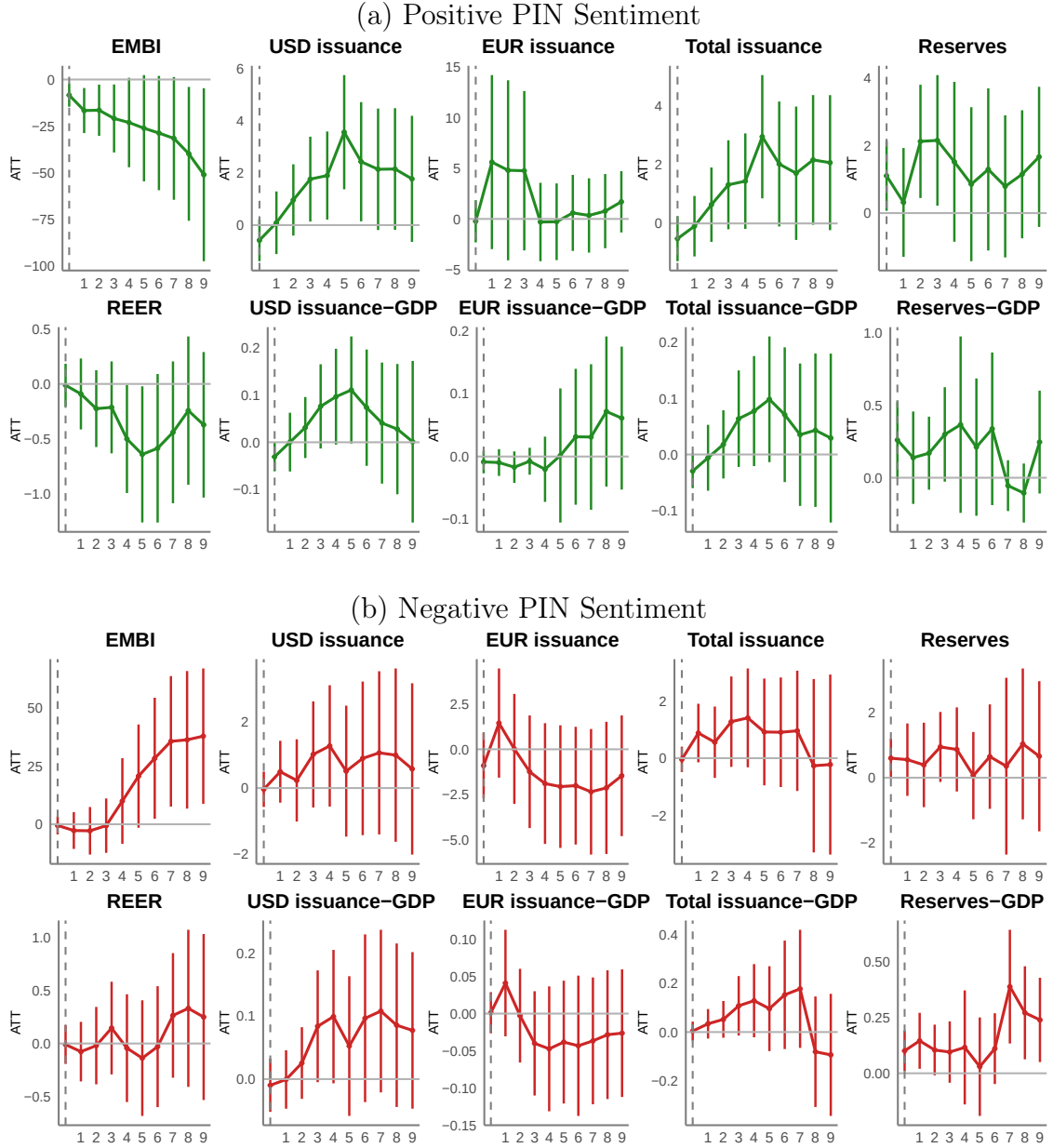
Note: This figure displays the monthly distribution of FinBERT sentiment gap scores, defined as the difference between a country’s PIN sentiment score and the median sentiment of all PINs published in the same IMF-defined region during the same six-month window. This gap captures country-specific deviations from the regional communication tone, isolating the “surprise” or idiosyncratic component of IMF messaging. Positive values indicate sentiment more favorable than regional peers; negative values suggest more pessimistic or critical assessments.

Figure D.3: FinBERT vs Loughran-McDonald Sentiment Scores



Note: Comparing the sentiment scores on IMF Article IV Public Information Notice (PIN) computed with the FinBERT model and the Loughran-McDonald model. The left-hand graph shows the density plot of both sentiment scores, while the right-hand graph shows the correlation between the two scores in a scatter plot.

Figure D.4: Loughran-McDonald Indicator - Heterogeneous treatment effects by PIN sentiment: optimistic vs. pessimistic releases



Note: This figure displays the dynamic treatment effects of IMF Article IV PIN releases on sovereign bond spreads and related financial indicators, disaggregated by the sentiment of the PIN communication. Panel (a) reports results for PINs classified as optimistic, while Panel (b) reports results for PINs classified as pessimistic, based on the FinBERT sentiment score. Each subfigure plots the estimated Average Treatment Effect on the Treated (ATT) over a nine-month horizon following the PIN release. Shaded areas represent 90% confidence intervals based on bootstrapped standard errors. Estimates are derived using the same baseline specification as previous figures, with propensity score weighting (PSW) applied for covariate adjustment.