Institut de Recerca en Economia Aplicada Regional i Pública
Research Institute of Applied Economics

Document de Treball 2023/15 1/38 pág.
Working Paper 2023/15 1/38 pag.

"Sovereign Risk and Economic Complexity: Machine Learning Insights on Causality and Prediction"

Jose E. Gomez-Gonzalez, Jorge M. Uribe and Oscar M. Valencia

UBIREA | Institut de Recerca en Economia Aplicada Regional i Pública

Abstract

We investigate how a country's economic complexity influences its sovereign yield spread with respect to the US. We analyze various maturities across 28 countries, consisting of 16 emerging and 12 advanced economies. Notably, a one-unit increase in the economic complexity index is associated to a reduction of about 87 basis points in the 10-year yield spread ($p<0.01$). However, this effect is largely non-significant for maturities under 3 years and, when significant ($p<0.1$), the reduction is around 54 bps. This suggests that economic complexity affects not only the level of the sovereign yield spreads but also the curve slope. Our first set of models utilizes advanced causal machine learning tools, allowing us to control for a large set of potential confounders. This is crucial given our relatively small dataset of countries and roughly 15 years of data, as well as the low frequency of annual variables. In the second part of our analysis, we shift our focus to economic complexity's predictive power. Our findings reveal that economic complexity is a robust predictor of sovereign spreads at 5-year and 10-year maturities, ranking among the top three predictors, alongside inflation and institutional factors like the rule of law. We also discuss the potential mechanisms through which economic complexity reduces sovereign risk and emphasize its role as a long-run determinant of productivity, output and income stability, and the likelihood of fiscal crises.

JEL classification: F34, G12, G15, H63, O40.

Keywords: Sovereign Credit Risk, Convenience Yields, Yield Curve, Government Debt, Double-Machine-Learning, XGBoost.

Jose E. Gomez-Gonzalez: City University of New York-Lehman College (USA). Visiting Professor - Universidad de la Sabana. Email: jose.gomezgonzalez@lehman.cuny.edu

Jorge M. Uribe: Universitat Oberta de Catalunya, Barcelona (Spain). Email: juribeg@uoc.edu

Oscar M. Valencia: Fiscal Management Division, Inter-American Development Bank, Washington (USA). Email: oscarva@iadb.org

1. Introduction

In the current landscape of globalization shaped by limited fiscal space, rising interest rates, and a pressing need for financing, both from developed and non-developed nations, particularly in response to the urgent need for an ecological transition to address climate change, fiscal considerations have regained paramount importance. In particular, the exploration of factors influencing a country government's capacity to secure funds in international debt markets at favorable terms has garnered significant attention in academic and policy circles. This study makes a significant contribution in this respect. We investigate the role of a country's level of economic complexity as a determinant of sovereign credit risk or, in the case of developed countries, their convenience yield.

Economic complexity has recently gained prominence as a new paradigm for economic development. According to Balland and others (2022) various institutions, including the World Bank, the European Commission, the World Economic Forum, the OECD, and numerous national and regional organizations, have increasingly adopted the principles of economic complexity and incorporated its analytical framework. This concept revolves around a nation's ability to produce complex products that are not easily substitutable in global markets and are highly valued by trade partners, such as specialized machinery, as opposed to basic commodities.

In what follows, by leveraging recent advancements in causal machine learning, known as double machine learning (Chernozhukov and others 2018), we assess the impact of economic complexity on the sovereign credit spreads (with respect to the US) of a diverse panel of 28 countries, encompassing both emerging and developed nations. Notably, our analysis incorporates an extensive array of control variables, including relevant factors previously identified in the literature, comprising macroeconomic, market, debt-related and institutional variables (see section 2 for the motivation of our control variables).

In fact, we are the first to comprehensively consider this range of control variables while examining the direct impact of economic complexity across various maturities on the yield spread curve, spanning from 3 months to 10 years. This approach enables us to effectively isolate the influence of multiple confounding factors when estimating the effects of interest across countries and maturities. This sort of analysis would be unattainable through conventional panel econometrics and factor models as employed in the extant literature. This is primarily because of the substantial number of confounding variables, around 30, that require consideration when estimating the direct causal effects of complexity on spreads, given the relatively limited dataset available for both countries and over the time, and especially the low frequency of the variables (annual).

Crucially, our approach openly acknowledges the potential researcher-induced bias when employing machine learning or related techniques to reduce the dimensionality of variables that could impact sovereign debt. This awareness is crucial for accurately estimating both direct and indirect causal effects. In contrast, prior studies that have relied on a wide array of variables to investigate the determinants of sovereign debt, contributing significantly to our understanding in this field, as exemplified by Maltritz and Molchanov (2013) who employed Bayesian Moving Averaging, have overlooked this vital aspect essential for extracting causal insights from machine learning and large-dimensional factor analysis.

Our findings clearly highlight the influence of economic complexity on sovereign credit risk, particularly in the longer maturities. According to our baseline calculations, an increase of one standard deviation in the economic complexity index (ECI) of Hidalgo and Hausmann (2009) leads to an approximate reduction of 87 basis points (bp) in the 10-year spread ($p<0.01$) and 54 bp in the 3-month spread ($p<0.10$). This highlights how economic complexity not only impacts the *level* of the spread, thereby affecting a country's ability to secure international funding at a lower cost but also shapes the *slope* of the yield spread curve, a critical factor in a country's capacity to mitigate rollover risks without incurring the typically greater expenses of funding with longer maturity debt.

In the second part of our results, we evaluate the relative importance of economic complexity as a predictor of sovereign risk spreads. This analysis, employing a different machine-learning algorithm known as Extreme Gradient Boosting (XGBoost), complements our initial findings. We demonstrate that economic complexity, aside from its statistical and economic significance in determining sovereign risk, exhibits considerable predictive power. It ranks third among more than 30 variables in explaining sovereign spreads across both short and long maturities (i.e., 5 years and 10 years). Our assessment of relative performance is achieved by constructing Shap values for the XGBoost model. Interestingly, only inflation and institutional variables appear to exert a stronger influence than economic complexity, which is more relevant than traditional determinants in the literature such as real growth or the debt-to-GDP ratio.

Our contribution extends to two distinct branches of the existing literature. Firstly, we align with a body of research that scrutinizes the long-run factors influencing sovereign yields and spreads (e.g., Bellas, Papaioannou, and Petrova 2010; Poghosyan 2014; Wang, Xue, and Zheng 2021), by introducing economic complexity as a crucial determinant. Secondly, we contribute to a strand of studies that delve into the varied dynamics across different maturities of yield and spread curves (Eichler and Maltritz 2013). These studies underscore the significance of shifts in curve slope dynamics and changes in debt maturity in the face of different economic and political shocks (e.g., Afonso and Martins 2012; Wellmann and Trück 2018; Augustin 2018; Sánchez, Sapriza, and Yurdagul 2018).

The remainder of this document is structured as follows: In section two we position our study in the literature. In section three we revisit the expected theoretical relationship between economic complexity and sovereign credit risk, with a particular focus on recent literature that stresses economic complexity as a significant determinant of economic development and fiscal performance. Section four provides an overview of our methodology with an emphasis on the description of our credit risk spread measure taken from Du and Schreger (2016) and, the causal and non-causal machine learning tools that we use to answer our research questions. Section five describes our data and sources. Section six presents our main findings and the concluding section is section seven.

2. Related Literature

We contribute to two distinct areas of international finance. Firstly, our study adds to the body of research that examines the factors influencing sovereign risk as measured by sovereign yields. This body of literature has emphasized the importance of fiscal discipline and potential output growth in reducing risk spreads, particularly in longer terms. For example, when distinguishing between long-term and short-term determinants, Poghosyan (2014) found that in the long run, a 1-percentage point (pp) increase in the government debt-to-GDP ratio corresponds to an approximate 2 bp increase in government bond yields[1]. Additionally, a 1-pp

[1] See also Wang, Xue, and Zheng (2021) for a recent assessment of the relationship between debt and growth.

increase in the potential growth rate is associated with an approximate 45 bp increase in yields. In the short term, sovereign bond yields may deviate temporarily from their long-term fundamental levels, but approximately half of these deviations correct themselves within a year. Similarly, following the same distinction, Bellas, Papaioannou, and Petrova (2010) proposed that fundamental factors have a substantial influence on shaping emerging market sovereign bond spreads in the long term, while, conversely, in the short term, financial volatility emerges as a more dominant determinant. We add economic complexity to set of long-run factors previously investigated in the field.

Other authors have explored a different set of factors influencing sovereign yields, such as the local and foreign monetary policy conditions (Arora and Cerisola 2001; Dailami, Masson, and Padou 2008); unconventional monetary policy interventions (Krishnamurthy, Nagel, and Vissing-Jorgensen 2018; De Santis 2020) and the zero lower bound of interest rates (Coroneo and Pastorello 2020). Local inflation rates and deficit-to-GDP ratios (Liu and Spencer 2013; Gill 2018) have also been studied, as well as the terms-of-trade and their volatility (Hilscher and Nosbusch 2010; Maltritz 2012) and market uncertainty indicators, in particular the VIX (Matsumura and Machado 2010; Afonso and Jalles 2019). Since numerous studies have pinpointed external factors as the main influencers of sovereign risk, a subset of research has delved into the impact of financial and trade openness on sovereign spreads (e.g. Maltritz 2012; Maltritz and Molchanov 2014).

Some studies highlight the convergence of fiscal and/or political factors in determining sovereign yields in emerging and advanced economies (Sanjeev, Mati, and Baldacci 2008; Caggiano and Greco 2012; Afonso and Jalles 2019; Beqiraj, Patella, and Tancioni 2021), including the impact of political factors (Eichler 2014; Chatterjee and Eyigungor 2019; Brooks, Cunha, and Mosley 2022). Additionally, there is extensive research demonstrating the global factors that influence sovereign credit risk commonality worldwide, particularly involving the US stock and bond market dynamics (Longstaff and others 2011; Liu and Spencer 2013), and global financial risk (Gilchrist and others 2022). These prior studies provide the rationale for our comprehensive set of control variables and the use of double machine learning to conduct our main analyses.

Our study also aligns with a set of research efforts that investigate how different maturities of sovereign yields and spreads respond to economic shocks. Theoretically, long-term interest rates are closely intertwined with market expectations concerning a government's future solvency and financing requirements, whereas short-term interest rates reflect concerns related to liquidity and short-term performance outlooks (Freixas and Rochet 2008; Eichler and Maltritz 2013). Consequently, it is reasonable to anticipate that the determinants of short- and long-term yield spreads may be different. The composition of long-term and short-term debt plays a fundamental role in emerging market economies, as highlighted by Arellano and Ramanarayanan (2012). Long-term debt serves as a safeguard against fluctuations in interest-rate spreads, whereas short-term debt effectively incentivizes prompt repayment. In a related study, Sánchez, Sapriza, and Yurdagul (2018) introduce a framework for an endogenous determination of sovereign debt maturity, which highlights that sovereign debt tends to have durations and maturities that commonly exceed one year and tend to move in harmony with the economic cycle. Secondly, it observes that sovereign yield spread curves often exhibit non-linear, upward-sloping patterns. Finally, factors like output volatility, individual impatience, risk aversion, and particularly abrupt cessation of capital inflows, are identified as fundamental determinants of debt maturity.

Remarkably, Eichler and Maltritz (2013) investigate the factors influencing government bond yield spreads in EMU countries. These authors emphasize the evaluation of default risk across varying timeframes as indicated by spreads of different maturities. Their findings indicate that low economic growth and greater economic openness amplify default risk across all maturity levels. However, heightened indebtedness exclusively heightens short-term risk, while factors like net lending, trade balance, and interest rate costs predominantly impact long-term default risk.

Most prior research in this second branch has focused on extracting common factors through principal component analysis (PCA) of yields (or spreads) across a broad set of countries, aiming to uncover the global factors that shape the yield (spread) curves. These studies typically identify three latent factors known as the level, slope, and curvature, which suffice to describe the time series variations in interest rates across countries (examples of this literature can be found in recent works such as Afonso and Martins 2012; Wellmann and Trück 2018; Augustin 2018; and references therein). Our approach is different. Given that economic complexity is a relatively slow-moving variable primarily associated with long-term investments in productivity and knowledge diffusion (Hidalgo 2021), our primary focus is on understanding cross-sectional variations among countries in economic complexity that contribute to explaining sovereign risk. We do not emphasize the high-frequency time series movements, which are often the focus of more financially oriented research in this area. Nonetheless, as we examine different maturities ranging from 3 months to 10 years, our results also provide insights into this line of research by highlighting the anticipated relationship between the slope of the yield spread curve and the novel long-term determinant that we explore.

3. Complexity and sovereign yields, some preliminary facts and theory

The level of economic complexity of a country can be assessed using the economic complexity index constructed by -and publicly available from- the Harvard's Growth Lab (Hausmann and others 2014), which assigns a numerical value within -3 to 3, allowing for a quantitative assessment of the matter. In short, the ECI presents a comprehensive approach for simultaneously measuring economic development and resilience to shocks. It surpasses other broad indicators such as the Human Development Index or even the GDP, which tend to focus on specific aspects of the economy and overlook the relative position of a country in the global trade network.

A country is more complex if it produces goods that are relatively rare and combine a highly diversified set of knowledge and capabilities (Hidalgo 2021). Previous research has established significant relationships between a country's ECI and various economic and social outcomes. Countries with a high ECI have demonstrated an ability to effectively optimize their production inputs for enhanced output value (Hidalgo 2021), exhibit resilience in the face of macroeconomic shocks (Hausmann and others 2014), tend to experience reduced income inequality (Hartmann and others 2017), and show a positive association with gender equality (Nguyen 2021). Additionally, societies characterized by economic complexity are more inclined towards technological innovation (Gala and others 2018), enjoy greater macroeconomic stability, particularly in fiscal matters (Gomez, Uribe, and Valencia 2023), and tend to adopt more environmentally sustainable production practices (Romero and Gramkow 2021).

Given these established relationships, it is reasonable to expect a negative association between economic complexity and sovereign risk. At the national level, high economic complexity consistently correlates with long-term economic growth, as evidenced by numerous studies (e.g., Hidalgo and Haussmann 2009; Haussmann and others 2014; Ferrarini and Scaramozzino 2016; Tachella, Mazzilli, and Pietronero 2018; Nepelski and De Prato 2020). In economic literature, it is well established that more sophisticated exports are linked to higher future economic growth (Hallak 2006; Hausmann, Hwang, and Rodrik 2007). Furthermore, countries boasting greater economic complexity tend to exhibit more stable growth patterns due to reduced output volatility (Gu̇̈neri and Yalta 2021) and enhanced total factor productivity (Sweet and Eterovic 2019). Economic complexity plays a pivotal role in achieving export stabilization, as evidenced by the findings of Zou and others (2023). Their research underscores the significance of product sophistication as a key factor in both initiating and maintaining stable export partnerships. These factors are crucial drivers of sustained economic expansion and contribute to fiscal budget stability, enabling nations to navigate turbulent economic periods without succumbing to fiscal crises, as demonstrated by Gomez, Uribe, and Valencia (2023). At the microeconomic level, there is compelling evidence suggesting that companies with a more complex product portfolio experience reduced fluctuations in their output (Maggioni, Turco, and Gallegati 2016).

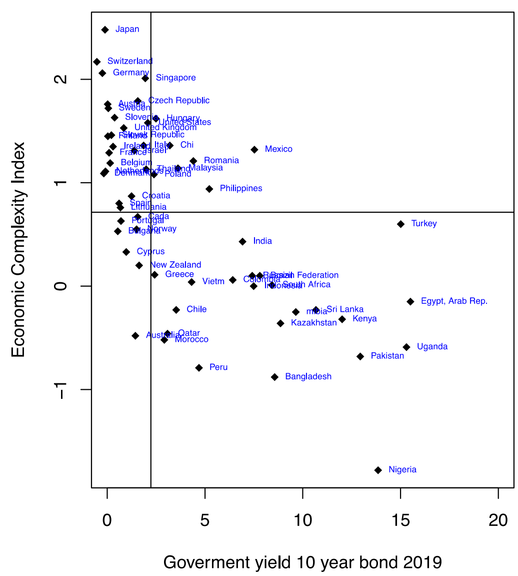
Achieving productive diversification and sophistication is critical for maintaining macroeconomic and fiscal stability. Countries heavily reliant on the production of basic and ubiquitous goods are vulnerable to fluctuations in international market prices, which can adversely impact their overall income (Deaton 1999). Similarly, nations heavily dependent on tourism are susceptible to global economic cycles, resulting in a sharp reduction in tourism demand during periods of low global economic activity (Aronica, Pizzuto, and Sciortino 2021). In contrast, economies equipped with the capability to produce complex goods through complex networks involving various forms of expertise and capabilities tend to exhibit greater resilience to external shocks. Consequently, they can be expected to face a lower risk of experiencing fiscal crises and enjoy a lower credit risk prospect, which is priced by the market.

Drawing a parallel to financial asset investments, a diversified portfolio, especially consisting of low risk assets, helps mitigate risk and generate more stable income for investors over time. Similarly, countries with more complex production structures benefit from more stable income streams, leading to less fluctuation in tax revenues for governments. As a result, it can be expected that complexity is associated with lower costs of sovereign debt and reduced risk premiums.

Figure 1 shows the ECI of 60 countries for the year 2019 plotted against the sovereign yields for 10-year bonds in the same year. Our analysis excludes the years 2020 and 2021, for which ECI data is readily available, due to the extraordinary disruptions caused by the Covid-19 pandemic, which significantly influenced international debt market dynamics in a way orthogonal to our interests (see, for instance, Candelon and Moura 2023).

As evident from the figure, there is a distinct negative correlation (i.e. -0.65, p< 0.001) between the ECI and the yields paid. Furthermore, the stability of this relationship is evident in Figure 2, which presents the same variables as in the previous figure but focuses on 27 countries with available data for the year 2000. Once again, a significant correlation (p<0.001) of -0.65 is observed between the two variables.

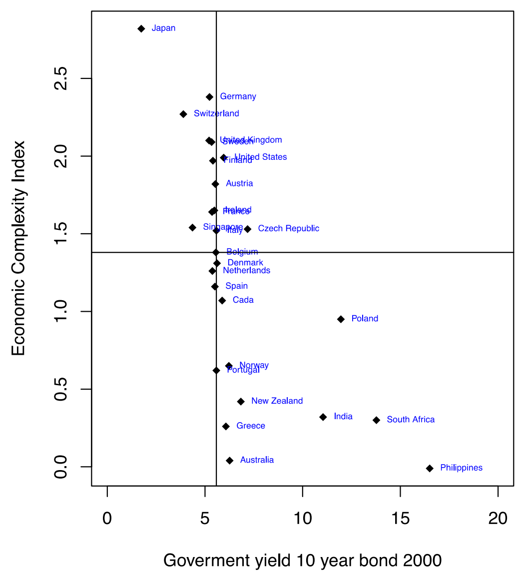
Figure 1. Plot of ECI against Sovereign yield 10 years, 2019



Note: The figure shows the relationship between a country's economic complexity level and the yields of 10-year maturity sovereign bonds in the year 2019 for a sample of 60 countries. Correlation -0.65.

While the simple correlation between the ECI and yields serves as a useful starting point, it falls short of fully quantifying the direct causal impact of economic complexity on sovereign credit risk. To address this issue, two additional key steps are required explained in the methodology, an adequate measure of risk and methodological tools from the recent causal machine learning literature.

Figure 2. Plot of ECI against Sovereign yield 10 years, 2000



Note: The figure shows the relationship between a country's economic complexity level and the yields of 10-year maturity sovereign bonds in the year 2000 for a sample of 27 countries. Correlation -0.65.

4. Methodology

4.1. Sovereign credit risk

Since our primary concern is a country's sovereign credit risk, we employ the spread with respect to the US for a given maturity at each year, rather than the raw yields in Figures 1 and 2. Specifically, we utilize the local currency sovereign risk indicators developed by Du and Schreger (2016) and Du, Im, and Schreger (2018).

These indicators are constructed as deviations from covered interest rate parity (CIP) between government bond yields in the United States and other countries, denoted as $\Phi_{i,n,t}$:

$$\Phi_{i,n,t} = y_{i,n,t}^{Govt} - \rho_{i,n,t} - y_{USD,n,t}^{Govt} , \quad (1)$$

here, $y_{i,n,t}^{Govt}$ represents the n-year local currency government bond yield in country i, $\rho_{i,n,t}$ represents the n-year market-implied forward premium for hedging currency i against the US dollar, and $y_{USD,n,t}^{Govt}$ is the n-year US Treasury bond yield.

The Treasury CIP deviation measures the distinction between the synthetic dollar borrowing cost of country i and the direct dollar borrowing cost of the United States. This allows comparing sovereign borrowing costs after converting the promised cash flows of local currency sovereign bonds into US dollars. The primary factors influencing CIP deviations for government bond yields include differences in default risk between US and foreign government bonds, variations in convenience yields between US and foreign government bonds, and financial frictions. The relative significance of these factors depends on the specific country and maturity being studied (Du and Schreger 2016; Du, Im, and Schreger 2018).

Du and Schreger (2016) attribute most of the spread variation to credit risk, particularly in the case of emerging markets. In contrast, for developed markets with negligible sovereign default risk and open capital accounts, Du, Im, and Schreger (2018) attribute the spread to convenience yields stemming from factors such as liquidity and other potential non-pecuniary benefits of US bonds compared to others. This approach to constructing spreads effectively mitigates currency risk factors reflected on traditional spreads, enabling us to concentrate on the analysis of sovereign credit risk.

The second step we undertake pertains to the challenge of identifying direct causal effects amid the presence of numerous potential confounding factors within a dataset that typically contains relatively few data points (compared to typical machine learning tasks). Building on prior research, we are aware that various macroeconomic factors, spanning both the supply and demand sides of the economy, play significant roles. These factors include real growth, investment and consumption growth, institutional variables such as the rule of law and regulatory quality, global uncertainty, levels of capital account openness, terms of trade, export commodity rents, population size, and debt-related metrics like the debt-to-GDP ratio, fiscal balances, primary balances, and government revenue. All these factors are expected to exert an influence on the spread of sovereign bonds.

However, the data on spreads available from the original authors' website typically commences in the mid-2000s for emerging market countries. This is primarily due to limitations in the original data sources, such as Bloomberg. In certain cases, like Chile, there are only three years with enough observations (2005, 2011, 2017) between 1995 and 2019 when information for the economic complexity index is available (excluding the period of the Covid-19 pandemic). Given the combination of relatively limited data points, the low frequency of economic complexity measurements (annual), and the substantial number of potential confounding variables, it becomes essential to employ non-traditional machine learning models that have been recently developed in the econometrics literature to specifically address causal inquiries in the presence of numerous confounding variables.

4.2. Double Machine Learning

We adopt the methodology developed by Chernozhukov and others (2018)[2]. We follow closely the presentation by Bach and others (2023) for our exposition of the methods, adapting the notation to our case.

In general lines, when investigating causal relationships, it is often necessary to control for other variables, which we refer to as confounders. This becomes particularly crucial in observational studies, like the one at hand, where randomization is impossible to perform, making the consideration of confounders essential to estimating both direct and indirect causal effects (Pearl 2009).

In the context of our study, we have a multitude of potential control variables, encompassing macroeconomic, institutional, and fiscal factors. Therefore, it is imperative to carefully select the most pertinent variables before proceeding with the analysis, particularly when our primary focus is on examining the impact of economic complexity on sovereign credit risk. Furthermore, the interplay between these variables and their relationship with both spreads and economic complexity can be intricate, potentially involving nonlinearities and interactions.

In such scenarios, machine-learning algorithms, such as tree-based methods, and regularization and shrinkage techniques, are well suited for the task of variable selection. However, it's important to acknowledge that utilizing these methods to choose from our initially extensive set of control variables introduces a form of bias known as regularization bias or pre-selection bias, which can affect subsequent estimations of causal effects. Double Debiased Machine

[2] The methodology has been implemented by Bach and others (2023) in the R package DoubleML.

Learning or, simply, Double Machine Learning (DML) is a method designed to estimate causal effects in the presence of a high number of confounders.

In our case, we can represent our problem as a partially linear regression model-PLR (Robinson 1988) through the following equations:

$$\Phi_{i,n,t} = \alpha ECI_{i,t} + g_0\left(X_{i,t}\right) + u_{i,n,t}, \tag{2}$$

$$ECI_{i,t} = m_0\left(X_{i,t}\right) + v_{i,n,t}, \tag{3}$$

with $E(u|ECI, X) = 0$, $E(v|X) = 0$ and where, $\Phi_{i,n,t}$ is defined as before, $ECI_{i,t}$ is the economic complexity index for country *i* at year *t*, and $X_{i,t}$ is a high dimensional vector of confounder variables that influence both $\Phi_{i,n,t}$ and $ECI_{i,t}$, including country and time fixed effects. DML developed by Chernozhukov and others (2018) allows us to estimate very accurately the functions $g_0(\cdot)$ and $m_0(\cdot)$, which can be linear or not. In addition, it allows us to correct for pre-selection bias by a procedure that is called post-double-selection (Belloni, Chernozhukov, and Hansen 2014).

The model above can be rewritten in residual form as follows, where we omit year, maturity, and country indexes, to ease notation. In this way, it becomes transparent that we run a single regression that uses DML for each maturity, separately, always including time and country fixed effects in the set of controls X:

$$v = ECI - m_0(X), \tag{4}$$

$$w = \left(\Phi - l_o(X)\right), \tag{5}$$

$$w = v\alpha + u, \tag{6}$$

with $l_o(X) = E(\Phi|X) = \alpha m_0(X) + g_0(X)$, $E(u|ECI, X) = 0$, $E(v|X) = 0$, $m_0(X) = E(ECI|X)$. The variables w and v are just the original variables after taking out the effect of X. This is called partialling out the effect of X. In this equation α is identified as long as $var(v) \neq 0$.

Estimation algorithm of the PLR model reads as follows:

i. Estimate l_o and m_0 by $\hat{l}_o$ and $\widehat{m}_0$, which can be done by solving the two problems of predicting Φ and ECI using a generic ML method. In our case, we use random forest[3]. In this case, the estimated residuals are given by:

$$\hat{v} = ECI - \widehat{m}_0(X), \tag{7}$$

$$\widehat{w} = \left(\Phi - \hat{l}_o(X)\right), \tag{8}$$

Notice that these residuals should be obtained by cross-validation, to avoid biases and over-fitting (Belloni, Chernozhukov, and Hansen 2014).

ii. Estimate α by regressing the residual $\widehat{w}$ on $\hat{v}$. This can be done using conventional inference tools as shown by Chernozhukov and others (2018).

In terms of inference, to construct point and interval estimator with ML we use the method-of-moment estimator for α based on the empirical moment condition given by:

$$E[\psi(w; \alpha, \eta_0)] = 0, \tag{9}$$

where ψ is know as the score function, $w = (\Phi, \alpha, X)$, α is our parameter of interest which corresponds to the effect of economic complexity on the yields spread at a given maturity. η denotes nuisance functions equal to η_0 in population (i.e. functions g_0 and m_0 in equations 2 and 3). Inference relies on choosing a score function that satisfies the so-called Neyman orthogonally condition (Neyman 1979) given by:

$$\partial_\eta E[\psi(w; \alpha, \eta_0)]|_{\eta=\eta_0} = 0. \tag{10}$$

Employing a Neyman-orthogonal score makes estimation of the parameter α robust against first order bias that arises from regularization. In the PLR model two alternatives for the score function are available, from which we select the partialling-out score given by:

$$\psi(w; \alpha, \eta_0) \coloneqq \left(\Phi - l(X) - \alpha\big(ECI - m(X)\big)\right)\big(ECI - m(X)\big), \tag{11}$$

where $\eta = (l, m), \eta_o = (l_0, m_0)$, $w = (\Phi, \mathrm{ECI}, \mathrm{X})$ and l, m are P-square-integrable functions mapping X on $\mathbb{R}$ (see Bach and others 2023 and Chernozhukov and others 2018, for additional details).

[3] See Giraldo and others (2023) and Gu, Kelly, and Xiu (2020), which describe the advantages of tree-based models in the case of relatively small datasets like ours.

Two methods exist to carry out the estimation with DML considering the cross-fitting nature of the problem, and which employ a form of sample splitting to eliminate the over-fitting. For this, let us assume that we have a sample $(w_i)_{i_1}^{N}$, which is i.i.d. To simplify notation, we also assume that N is divisible by *K*. Then,

$$E_N[g(w)] \coloneqq \frac{1}{N}\sum_{i=1}^{N} g(w_i). \tag{12}$$

Method 1: The sample $(w_i)_{I=1}^{N}$ is split into K fragments, and indexed with $(I_k)_{k=1}^{K}$, for $[N] = \{1, \ldots, N\}$, such that the size of each fragment, I_k, is $n = N/K$. For each part, $k \in [K] = \{1, \ldots, K\}$, we construct a random forest estimator $\hat{\eta}_{0,k} = \eta_{0,k}\left((w_i)_{i \notin I_k}\right)$ of $\eta_{0,k}$. Notice that $x \rightarrow \hat{\eta}_{0,k}(x)$ depends only on the subset of data $(w_i)_{i \notin I_k}$. Then, for each $k \in [K]$, we construct the estimator $\hat{\alpha}_k$ as to solve the following equation:

$$\frac{1}{n}\sum_{i \in I_k} \psi\left(w_i; \hat{\alpha}_k, \hat{\eta}_{0,k}(x)\right) = 0\ . \tag{13}$$

And the causal effect is obtained via aggregation as follows:

$$\tilde{\alpha}_k = \frac{1}{K}\sum_{k=1}^{K} \hat{\alpha}_k. \tag{14}$$

Method 2: The sample is split into *K* fragments and indexed with $(I_k)_{k=1}^{K}$, in the sample $[N] = \{1, \ldots, N\}$ such that the size of each fragment is $n = N/K$. A random forest is constructed for each part, $k \in [K] = \{1, \ldots, K\}$, $\hat{\eta}_{0,k}$. This time the estimator of the causal parameter $\tilde{\alpha}_k$ is constructed by solving the following equation:

$$\frac{1}{n}\sum_{k=1}^{K}\sum_{i \in I_k} \psi\left(w_i; \hat{\alpha}, \hat{\eta}_{0,k}\right) = 0. \tag{15}$$

We present our main estimations in the results section using both methods.

4.3. Extreme Gradient Boosting for assessing the relative role of economic complexity as a predictor

Finally, in the third part of our results section we use non-causal machine learning algorithms to assess the relative importance of economic complexity as a predictor of sovereign risk. In this case, the focus is not on the estimation of a causal effect, but rather on comparing the prediction power of economic complexity in relation with a large set of macroeconomic, institutional and debt related variables.

Specifically, we use XGBoost, developed by Chen and Guestrin (2016), as an efficient implementation of Gradient Tree Boosting (GTB). GTB, a variant of Gradient Boosting introduced by Friedman (2001), employs decision trees as base learners. The key idea behind GTB is the iterative fitting of regression trees to the residuals of the preceding trees, aiming to minimize the loss function of the model. In a similar vein, XGBoost operates iteratively, creating an ensemble of decision trees. Each new tree is trained to correct the prediction errors of the previous models. Notably, XGBoost is well suited for handling datasets with numerous features in relation with the number of observations. Models are fitted using any differentiable loss function, and in our case, we employ a standard square loss, optimized through gradient descent.

To interpret the results of our models, we employ SHAP (SHapley Additive exPlanations) values, a methodology introduced by Lundberg and Lee (2017). SHAP values are designed to quantify the contribution of each feature to the final prediction, considering the interactions between features (such as covariates) and the value ranges of each feature. This approach yields an accurate and intuitive explanation of how the model arrives at its predictions.

The use of SHAP values in conjunction with XGBoost is particularly valuable when it is crucial to comprehend the factors influencing the model's predictions. This is especially relevant in scenarios where we seek to understand the determinants of sovereign risk. By examining the SHAP values associated with each feature, we gain insights into which features exert the most significant influence on spread over time and how they relate to one another in our longitudinal data comprising a variety of countries over the years. This facilitates a more in-depth understanding of the dynamics underlying our model's predictions and complements the results on the effects obtained via DML.

5. Data

We rely on a comprehensive dataset comprising fiscal variables and associated macroeconomic, financial, and institutional factors, encompassing both emerging economies and advanced economies. The dataset covers a total of 28 countries, consisting of 16 emerging markets and 12 advanced economies. Developed markets: Australia (AUD), Canada (CAD), Switzerland (CHF), Denmark (DKK), Germany (EUR), United Kingdom (GBP), Japan (JPY), Norway (NOK), New Zealand (NZD), Sweden (SEK), Israel (ILS) and South Korea (KRW). Emerging markets: Brazil (BRL), Chile (CLP), China (CNY), Colombia (COP), Hungary (HUF), Indonesia (IDR), India (INR), Mexico (MXN), Malaysia (MYR), Peru (PEN), Philippines (PHP), Poland (PLN), Russia (RUB), Thailand (THB), Turkey (TRY) and South Africa (ZAR).

Most of the dataset features annual data points spanning from 1995 to 2019, ensuring comprehensive coverage of fiscal distress and related macroeconomic and financial upheavals in the global economy. Our dataset also includes various maturities of sovereign yield spreads relative to US government yields of the same maturity. Specifically, it encompasses spreads for 3 months, 1 year, 2 years, 3 years, 5 years, 7 years, and 10 years. These spreads are computed as either an average of monthly data within a year or as observations taken on the last day of the calendar year. Spreads are only available for the 28 countries listed above.

It is important to note that we have excluded the years 2020 and 2021, for which ECI indicators are already available, from our calculations. This decision is based on our primary focus on the long-term determinants of sovereign risk, such as economic complexity. The years 2020 and 2021, were marked by the onset of the Covid-19 pandemic, experienced abrupt changes in spreads that may not reflect the structural transformations that are the primary interest of this study.

In Table 1 we present the summary statistics of the data set including the spreads, macroeconomic and institutional factors. We gathered an extensive array of variables that vary over time, drawing inspiration from existing literature on fiscal crises and determinants of sovereign debt. The first three columns of Table 1 provide the complete list of variables and their definitions, while the last four columns display their means, medians, standard deviations, maximum, and minimum values. Our model's baseline specification includes indicators like real growth, the debt-to-GDP ratio, interest payments, revenue from fuels, real exchange rates, and institutional quality, among others.

Our selection of variables considers well-known sources of sovereign risk identified in existing literature, and each variable is theoretically justified. To maintain theoretical coherence, we refrain from including any transformations of the original variables, such as differences, squares, or interactions, in our dataset.

As for proxy of economic diversification, we have included the Economic Complexity Index from Harvard's Growth Lab, regularly updated on their Atlas of Economic Complexity webpage (Hidalgo and Haussmann 2009). It's worth noting that the ECI data are available for the subset of 28 countries, as indicated before, which also have good information regarding both the sovereign yields and macroeconomic and institutional variables, alongside fiscal information.

Table 1. Summary statistics of the data set and source

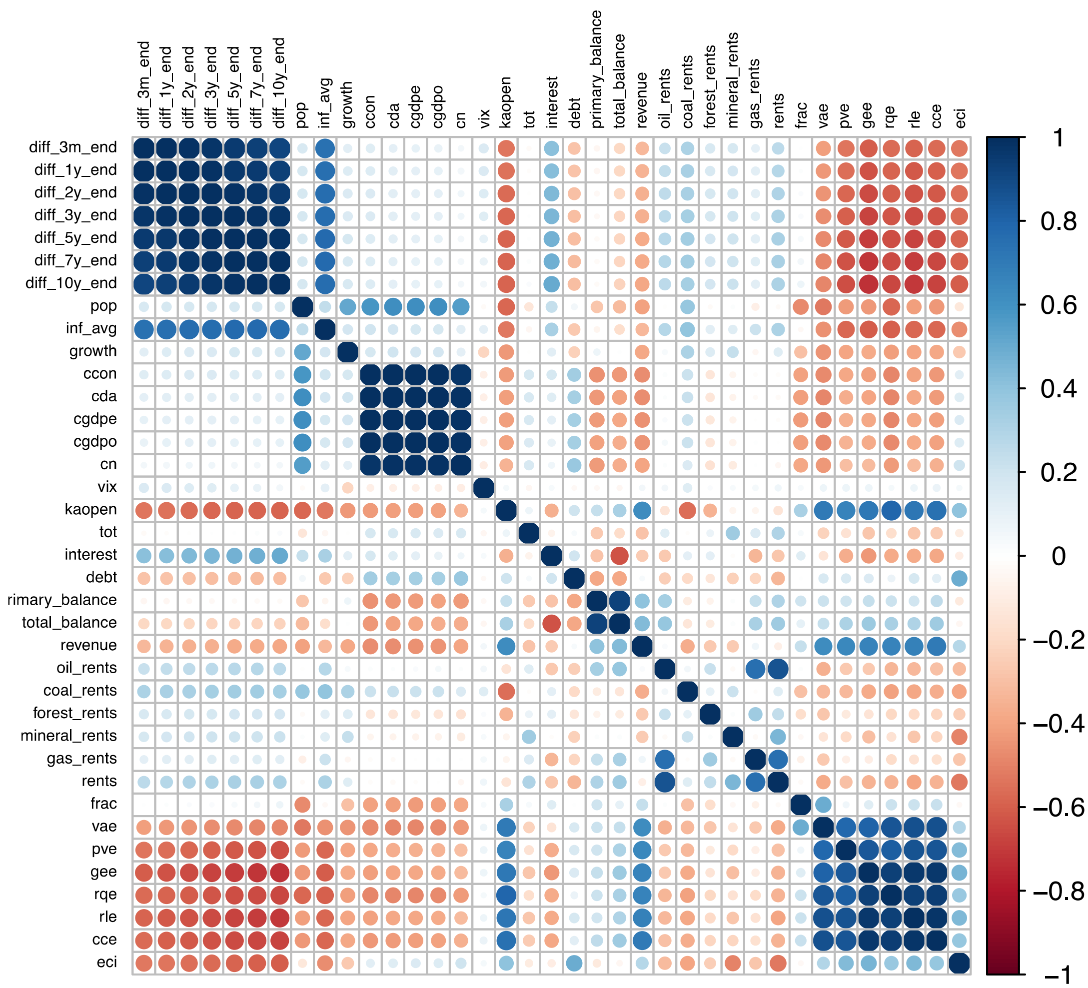
Indicator	*Abreviation*	*Source*	*Mean*	*Median*	*Std.Dev*	*Max.*	*Min.*
3 month yield spread	diff_3m_end	Du's CIP web page	2.12	1.65	3.42	19.72	-5.64
1 year yield spread	diff_1y_end	Du's CIP web page	2.09	1.56	3.41	18.43	-6.02
2 year yield spread	diff_2y_end	Du's CIP web page	2.07	1.43	3.39	16.7	-5.98
3 year yield spread	diff_3y_end	Du's CIP web page	2.03	1.36	3.35	15.77	-5.89
5 year yield spread	diff_5y_end	Du's CIP web page	1.88	1.13	3.22	15.69	-5.29
7 year yield spread	diff_7y_end	Du's CIP web page	1.77	1	3.15	15.3	-5.15
10 year yield spread	diff_10y_end	Du's CIP web page	1.69	0.88	3.06	15	-4.6
Population in millions	pop	WEO	155.62	38.47	342.3	1433.78	3.72
Inflation rate, average of the year	inf_avg	WEO	3	2.34	2.67	16.33	-1.33
Real GDP growth	growth	WEO	3.21	2.97	2.71	14.25	-7.82
Log of per capita real consumption	ccon	Penn World Tables	13.49	13.47	1.18	16.19	11.21
Log of per capita domestic absorption	cda	Penn World Tables	13.8	13.78	1.2	16.8	11.48
Log of expenditure-side real GDP at current PPPs in mil. 2017US$	cgdpe	WEO	13.82	13.78	1.18	16.81	11.49
Log of output-side real GDP at current PPPs in mil. 2017US$	cgdpo	WEO	13.83	13.76	1.19	16.82	11.48
Log of capital stock at current PPPs in mil. 2017US$	cn	Penn World Tables	15.2	15.05	1.26	18.44	12.71
Chicago Board Options Exchange Volatility Index	vix	Bloomberg	19.39	16.67	6.15	32.7	11.09
Financial openness, Chinn-	kaopen	Ito's web	1.19	2.32	1.37	2.32	-1.23

Ito index							
Terms of trade change in %	tot	WEO	102.67	100	15.9	159.88	49.3
Interest expenses as % GDP	interest	WEO (estimate)	1.6	1.36	1.62	8.37	-3.09
Gross debt as % of GDP, general government	debt	WEO	53.37	43.67	36.78	236.14	6.86
Primary balance as % GDP	primary_balance	WEO	0.16	0.07	3.24	15.83	-8.73
Fiscal balance as % GDP	total_balance	WEO	-1.44	-1.53	4.02	18.64	-11.23
Fiscal revenue as % GDP	revenue	WEO	34.08	33.65	11.22	58.63	14.05
Oil rents as % of GDP	oil_rents	World Bank	1.36	0.39	2.25	11.6	0
Coal rents as % of GDP	coal_rents	World Bank	0.39	0.03	0.79	7.25	0
Forest rents as % of GDP	forest_rents	World Bank	0.27	0.13	0.5	4.5	0
Mineral rents as % of GDP	mineral_rents	World Bank	0.7	0.15	1.57	12.63	0
Gas rents as % of GDP (gas_rents)	gas_rents	World Bank	0.4	0.11	0.78	4.83	0
Natural resources rents as % of GDP	rents	World Bank	2.46	1.12	3.25	17.1	0
Historical ethnic fractionalization	frac	HIEF-Harvard	0.66	0.67	0.18	0.95	0
Voice and accountability	vae	World Bank	0.76	0.99	0.84	1.8	-1.75
Political stability and absence of violence	pve	World Bank	0.31	0.62	0.95	1.61	-2.06
Government Effectiveness	gee	World Bank	1.03	1.21	0.82	2.35	-0.52
Regulatory quality	rqe	World Bank	0.96	1.11	0.76	2.09	-0.63
Rule of law	rle	World Bank	0.89	1.06	0.96	2.11	-0.97
Control of corruption	cce	World Bank	0.92	0.84	1.11	2.47	-1.13
Economic Complexity Index	eci	Harvard's Growth Lab	0.99	1	0.86	2.86	-0.84

Note: The table shows summary statistics of the variables in our sample along with their respective data sources.

Figure 3 presents the Pearson's correlation among the variables described in Table 2. Notably, the proxy variables for institutional quality in each country exhibit strong correlations with each other, including government effectiveness, regulatory quality, rule of law, and control of corruption. Likewise, there are nearly perfect correlations between the aggregate demand proxies, such as domestic absorption and real consumption. Finally, there are notably high correlations between all spreads in our data set.

Figure 3. Plot of the correlation among the variables in Table 1



Note: the figure shows the correlation among the continuous variables in the study sample.

Notably, sovereign yield spreads show a significant negative correlation with institutional variables and with the economic complexity index. This finding is particularly noteworthy as it sheds light on the importance of properly considering a large set of institutional quality proxies if one wants to assess the true impact of complexity on sovereign risk.

6. Results

Our results are organized into three sections: In the first section, we provide our baseline estimates, examining the impact of economic complexity on sovereign risk spreads at various maturities, ranging from 3 months to 10 years. The second section explores alternative model specifications to assess the robustness of our main claims. We consistently find statistically significant effects for maturities exceeding 3 years, even when they are reduced by up to 33% in the most extreme cases. This reaffirms the documented economic significance of these effects. Shorter maturities, in certain specifications, may exhibit statistical non-significance, reinforcing our argument about the distinct impact of economic complexity along the yield-spread curve, particularly its slope. In the third section, we present the outcomes of a purely statistical exercise using extreme gradient boosting. Here, our focus shifts from estimating the effect of complexity to evaluating its predictive power compared to a broad set of traditional determinants of yield spreads. This new set of results emphasizes the high predictive capability of economic complexity, emphasizing its significance as a long-term determinant of sovereign risk alongside factors such as inflation and institutional quality.

6.1. Baseline results: Effect of Economic Complexity on Sovereign Risk

Table 2 contains our primary findings, which include point and interval estimates of the impact of economic complexity on sovereign spreads of various maturities. The intervals were calculated at a 99% confidence level. Additionally, columns 3 to 5 present the standard errors, p-values, and t-statistics associated with these estimates.

In Panel A, we implement Method 1 as described in the methodology, while Panel B corresponds to Method 2. The table reveals that, in Panel A, the effects are statistically significant at a 90% confidence level for all maturities. At a 95% confidence level, significance holds for maturities longer than one year, and at a 99% confidence level, significance is observed for maturities equal to or exceeding two years. Regarding the magnitude of the effects, they exhibit a positive correlation with maturity. The smallest effects are observed for 3-month maturities, where an increase of one point (equivalent to one standard deviation) in the Economic Complexity Index leads to a reduction of 54 basis points in the spread. Conversely, the most substantial effect is observed in the 10-year spread, corresponding to an 87 basis point reduction. Remarkably, the effect shows a steady increase in between.

In Panel B, the disparity between the 3-month and 10-year maturities becomes even more pronounced, ranging from 39 bps in the former case to 83 bps in the latter. Notably, for maturities less than 2 years, this effect does not attain statistical significance at a 95% confidence level. Furthermore, it remains statistically insignificant for the 3-month maturity at any traditional level of confidence.

Table 2. Base line results: with both year and country fixed effects

Panel A. Method 1

	Effect	*S.E.*	*P.Value*	*t.Statistic*	*Lower.CI*	*Upper.CI*
3 month yield spread	-0.54	0.29	0.06	-1.87	-1.28	0.20
1 year yield spread	-0.67	0.29	0.02	-2.34	-1.41	0.07
2 year yield spread	-0.73	0.26	0.01	-2.77	-1.41	-0.05
3 year yield spread	-0.72	0.25	0.00	-2.85	-1.37	-0.07
5 year yield spread	-0.84	0.23	0.00	-3.69	-1.42	-0.25
7 year yield spread	-0.84	0.21	0.00	-4.02	-1.38	-0.30
10 year yield spread	-0.87	0.20	0.00	-4.42	-1.37	-0.36

Panel B. Method 2

	Effect	*S.E.*	*P.Value*	*t.Statistic*	*Lower.CI*	*Upper.CI*
3 month yield spread	-0.39	0.28	0.17	-1.38	-1.13	0.34
1 year yield spread	-0.54	0.28	0.06	-1.89	-1.27	0.19
2 year yield spread	-0.62	0.26	0.02	-2.36	-1.29	0.06
3 year yield spread	-0.61	0.25	0.01	-2.46	-1.26	0.03
5 year yield spread	-0.76	0.22	0.00	-3.40	-1.34	-0.19
7 year yield spread	-0.80	0.21	0.00	-3.81	-1.33	-0.26
10 year yield spread	-0.83	0.20	0.00	-4.23	-1.33	-0.32

Note: The table shows the impact of a unitary variation in the economic complexity index on various sovereign spread maturities, ranging from 3 months to 10 years. We utilized a random forest approach with 15 trees, a minimum node size of 2, and a maximum depth limit of 5 to estimate the nuisance functions. All variables outlined in Table 1 were included as control variables, along with dummy variables for each country and year. In Panel A Cross-fitting was conducted using Method 1. In Panel B we show the results using Method 2 as explained in the methodology. In both cases we applied a Neyman-orthogonality condition of partialling out, as detailed in section 3.

Our findings in this respect introduce a novel perspective to the field, as prior literature has not explored the influence of economic complexity on sovereign risk and convenience yields. However, they align with certain aspects of earlier research, such as Sánchez, Sapriza, and Yurdagul (2018), who present a model for endogenously determining sovereign debt maturity and emphasize the pro-cyclical nature of sovereign debt maturity. Likewise, Eichler and Maltritz (2013) investigate the determinants of government bond yield spreads at varying maturities. Their conclusions highlight that increased indebtedness primarily affects short-term maturities, while factors like net lending, trade balance, and interest rate costs predominantly impact long-term default risk. Our results complement these prior studies and others by demonstrating the notable influence of economic complexity on longer maturities which, indeed, is associated with different effects along the spread curve, supporting previous arguments advanced by this literature.

6.2. Alternative Specifications

Tables 3 and 4 present models identical to those in Table 1, with the exception that Table 3 excludes the inclusion of dummy variables for countries and years in the pool of controls, and Table 4 exclusively incorporates country dummy variables. These variations aim to assess the sensitivity of the results to different model specifications. Broadly speaking, the primary findings remain consistent. Economic complexity exhibits statistical and economic significance in impacting sovereign spreads across all specifications for maturities exceeding 3 years. However, in most instances, the magnitude of the effects is diminished by approximately 30%.

We favor the results presented in our baseline specification in Table 2, as the inclusion of country and year dummy variables serves to account for potential confounding factors. The exclusion of these variables, as seen in traditional panel data specifications, could introduce biases, particularly in the context of an unbalanced panel as we have in this study. In this case, these biases appear to mitigate the impact of economic complexity on sovereign risk at all maturities.

Table 3. Effects without year and country fixed effects

Panel A. Method 1

	Effect	*S.E.*	*P.Value*	*t.Statistic*	*Lower.CI*	*Upper.CI*
3 month yield spread	-0.58	0.31	0.06	-1.90	-1.38	0.21
1 year yield spread	-0.64	0.29	0.03	-2.19	-1.38	0.11
2 year yield spread	-0.59	0.27	0.03	-2.21	-1.28	0.10
3 year yield spread	-0.66	0.25	0.01	-2.60	-1.30	-0.01
5 year yield spread	-0.73	0.24	0.00	-2.98	-1.36	-0.10
7 year yield spread	-0.69	0.23	0.00	-2.98	-1.29	-0.09
10 year yield spread	-0.73	0.22	0.00	-3.29	-1.30	-0.16

Panel B. Method 2

	Effect	*S.E.*	*P.Value*	*t.Statistic*	*Lower.CI*	*Upper.CI*
3 month yield spread	-0.40	0.30	0.19	-1.30	-1.18	0.39
1 year yield spread	-0.45	0.29	0.11	-1.58	-1.19	0.29
2 year yield spread	-0.43	0.27	0.11	-1.60	-1.11	0.26
3 year yield spread	-0.50	0.25	0.04	-2.01	-1.14	0.14
5 year yield spread	-0.59	0.24	0.01	-2.43	-1.21	0.03
7 year yield spread	-0.60	0.23	0.01	-2.59	-1.19	-0.00
10 year yield spread	-0.64	0.22	0.00	-2.91	-1.20	-0.07

Note: The table shows the impact of a unitary variation in the economic complexity index on various sovereign spread maturities, ranging from 3 months to 10 years. We utilized a random forest approach with 15 trees, a minimum node size of 2, and a maximum depth limit of 5 to estimate the nuisance functions. All variables outlined in Table 1 were included as control variables. In Panel A Cross-fitting was conducted using Method 1. In Panel B we show the results using Method 2 as explained in the methodology. In both cases we applied a Neyman-orthogonality condition of partialling out, as detailed in section 3.

Table 4. Results with only country effects.
Panel A. Method 1

	Effect	*S.E.*	*P.Value*	*t.Statistic*	*Lower.CI*	*Upper.CI*
3 month yield spread	-0.63	0.29	0.03	-2.14	-1.38	0.13
1 year yield spread	-0.59	0.29	0.04	-2.02	-1.33	0.16
2 year yield spread	-0.64	0.28	0.02	-2.27	-1.36	0.08
3 year yield spread	-0.57	0.27	0.03	-2.13	-1.25	0.12
5 year yield spread	-0.55	0.23	0.02	-2.39	-1.13	0.04
7 year yield spread	-0.67	0.22	0.00	-3.06	-1.23	-0.11
10 year yield spread	-0.62	0.20	0.00	-3.10	-1.13	-0.11

Panel B. Method 2

	Effect	*S.E.*	*P.Value*	*t.Statistic*	*Lower.CI*	*Upper.CI*
3 month yield spread	-0.51	0.29	0.08	-1.76	-1.26	0.24
1 year yield spread	-0.44	0.29	0.12	-1.55	-1.18	0.30
2 year yield spread	-0.49	0.28	0.08	-1.76	-1.20	0.23
3 year yield spread	-0.44	0.26	0.10	-1.67	-1.12	0.24
5 year yield spread	-0.46	0.23	0.04	-2.05	-1.05	0.12
7 year yield spread	-0.59	0.22	0.01	-2.70	-1.15	-0.03
10 year yield spread	-0.57	0.20	0.00	-2.86	-1.08	-0.06

Note: The table shows the impact of a unitary variation in the economic complexity index on various sovereign spread maturities, ranging from 3 months to 10 years. We utilized a random forest approach with 15 trees, a minimum node size of 2, and a maximum depth limit of 5 to estimate the nuisance functions. All variables outlined in Table 1 were included as control variables, along with dummy variables for each country in the sample. In Panel A Cross-fitting was conducted using Method 1. In Panel B we show the results using Method 2 as explained in the methodology. In both cases we applied a Neyman-orthogonality condition of partialling out, as detailed in section 3.

The influence of economic complexity becomes particularly prominent in the case of longer debt maturities. This aspect holds significant importance, especially concerning debt restructuring during distress episodes in emerging market economies. Such episodes are often linked to increased borrowing costs, as countries are compelled to secure financing through longer-term contracts. This situation arises due to the generally positive slope of the yield curve during the debt restructuring process, which naturally results in higher borrowing costs for these countries. However, our findings indicate that this mechanism does not apply uniformly to countries with higher levels of economic complexity. All other factors being equal, these countries experience lower yields for longer maturities compared to other nations. Consequently, during times of crises, debt restructuring for more economically complex economies proves to be a more cost-effective option, thus alleviating pressure on these countries' government budget. In essence, economic complexity emerges as an attractive feature for risk mitigation within sovereign debt markets.

In essence, greater economic complexity enables countries to achieve a dual objective. It allows them to reduce roll-over risk during crisis episodes by issuing long-term debt instruments to replace short-term maturities, all while avoiding a substantial increase in their borrowing costs associated with this strategic shift from short to long debt. For a comprehensive exploration of the underlying mechanisms governing the choice of maturity in sovereign debt issuance, refer to Beetsma and others (2021).

6.3. Relative importance of economic complexity when explaining sovereign credit risk

Figures 4 and 5 show the SHAP values for all the predictors in our dataset, including the Economic Complexity Index, when making predictions for sovereign spreads with 5 and 10-year maturities. The abbreviations used in these figures are defined in Table 1.

The numbers next to each variable in both figures represent the SHAP values, which quantify the average impact of each predictor variable on the predictions. The colored points within the figures represent the individualized predictive influence of each variable on sovereign spreads at 5 and 10 years. Darker violet points correspond to larger values of the predictor variable, while lighter yellow points correspond to smaller values. It's important to note that the order of the variables is the focus here, as SHAP values are normalized.

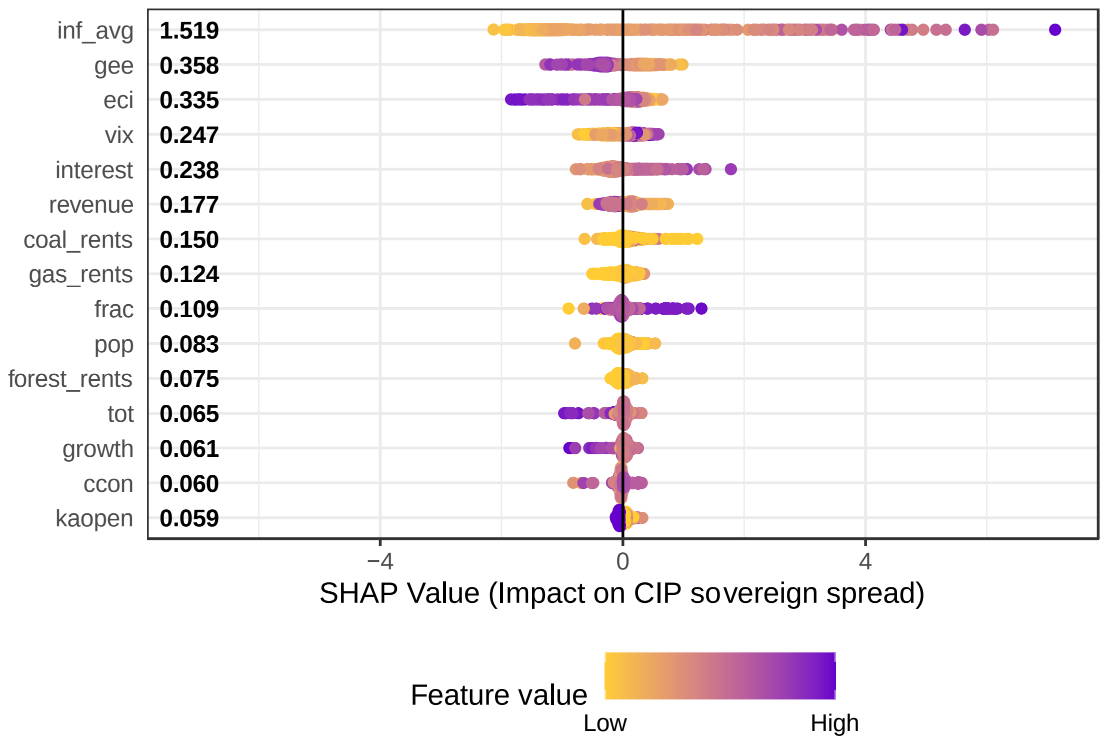
For both maturities, inflation emerges as the most influential predictor of yield spreads. Looking at Figure 4, it's evident that countries with low inflation (depicted in yellow) typically experience lower spreads. In other words, lower inflation has a negative impact on spreads, reducing the sovereign risk. Conversely, countries with high inflation (represented by darker yellow and violet shades) tend to face higher risk, as indicated by wider spreads. Remarkably, exceptionally high inflation substantially amplifies the spread, and this effect is asymmetrical, as demonstrated by the pronounced dark point on the far-right side of the figure.

The significant role of inflation as a key predictor of sovereign spreads is expected and well documented. Inflation is known to exert a substantial impact on a country's sovereign bond yields. Notably, inflation erodes the real value of bonds, particularly affecting longer-term debt instruments. As a result, it is anticipated that nations with higher inflation rates would be forced to offer greater risk compensation to investors holding their government bonds. High inflation tends to drive up a nation's nominal GDP, leading to immediate improvements in debt-to-GDP ratios—a phenomenon often referred to as "inflating debt away". This process introduces additional sources of risk. Notable studies in this field include the works of Buraschi and Jiltsov (2005), Gürkaynak, Sack, and Wright (2010), Hördahl and Tristani (2012), Camba-Méndez and Werner (2017), D'Amico, Kim, and Wei (2018), and Camba-Méndez (2020). Our findings underscore the paramount role of inflation in international debt markets from a novel methodological perspective.

The second most relevant predictor is institutional quality, a factor closely associated with a country's level of development. In Figure 4, this second predictor is denoted as "gee", corresponding to the Government Effectiveness indicator as defined in Table 1. In Figure 5, the second most influential predictor is "rle", representing the Rule of Law indicator, both of which are World Bank estimates.

Political institutions and the quality of governance are natural determinants of sovereign credit risk. Nations with fragile institutions and governance structures often face higher sovereign yields, reflecting the perceived greater risk of default, as noted by Eichler (2014). In broad terms, institutional risk encompasses the overall quality of a country's institutions, encompassing its legal and political framework. Increased institutional risk typically translates into higher sovereign yields. Moreover, in line with the insights presented by Butler and Fauver (2006), the institutional environment can significantly influence sovereign credit ratings, thereby impacting a country's sovereign spreads. Our findings underline the crucial role of institutions in shaping sovereign yield spreads.

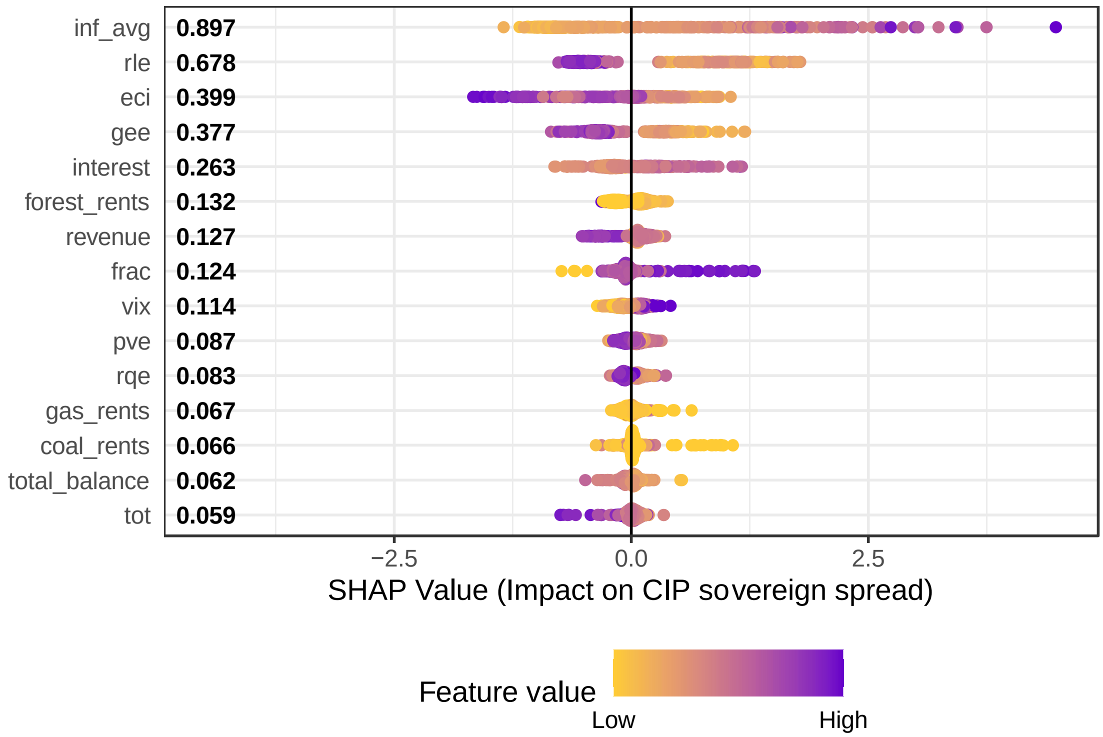
Figure 4. SHAP value of the top 15 variables in table 1 when predicting sovereign spreads



Note. The figure shows the SHAP values of the top-15 predictor variables of sovereign spreads at 5 years.

Taking the third position among the 30 variables is the economic complexity indicator. The magnitude of the SHAP value linked to the ECI is strikingly comparable to that of Government Effectiveness in Figure 4 for the 5-year maturity. In Figure 5, the ECI's effect is approximately two-thirds of the impact of the Rule of Law indicator. This highlights the substantial relative influence of economic complexity in shaping sovereign spreads. These results align with findings in Gomez, Uribe, and Valencia (2023), which identify economic complexity as a pivotal factor in determining the likelihood of fiscal crises. In this context, complexity risk becomes relevant for nations characterized by limited productive diversification and lower resilience to economic shocks. Our study demonstrates that this type of risk is indeed factored into international debt markets, thus establishing economic complexity as a significant predictor of sovereign yield spreads.

Figure 5. SHAP value of the top 15 variables in table 1 when predicting sovereign spreads



Note. The figure shows the SHAP values of the top-15 predictor variables of sovereign spreads at 10 years.

7. Conclusions and policy implications

In the current global context, fiscal considerations carry first-order importance due to limited fiscal resources, rising interest rates, and the urgency of financing for various needs, including the ecological transition. We contribute by postulating economic complexity as one of the main factors influencing a country's ability to secure favorable international debt financing, especially at maturities between 5 and 10 years.

Particularly, this study investigates the role of a country's economic complexity in determining sovereign credit risk, employing double machine learning for causal inference. The analysis covers a diverse panel of 28 countries, including both emerging and developed economies, considering a large array of control variables. Unique to the study is the simultaneous examination of the direct impact of economic complexity on yield spreads at various maturities, from 3 months to 10 years while considering a large set of controls.

In our baseline specifications the effect of economic complexity is shown to be significant for all maturities at 10% of significance, and only for maturities greater than 3 years at 99%. Our robustness checks include constructing two different DML estimators and including only country fixed effects and non-effects in the pool of controls. In all cases the results hold. Nonetheless, in some cases they are attenuated. For instance, the effect of complexity on sovereign spreads at 10 years range between – 57 bps in a model that only includes country fixed effects and uses the second method for cross fitting, and -0.87 bps in our baseline results. In all cases this effect is significant. All in all, our findings reveal economic complexity significantly influences sovereign credit risk, particularly in longer maturities, impacting both spread level and slope.

In the second part of the study, XGBoost machine learning shows economic complexity's substantial predictive power, ranking third among over 30 variables, with only inflation and institutional variables exerting a stronger influence.

This study contributes to international finance by highlighting the importance of economic complexity as a determinant of sovereign risk and exploring how different maturities of sovereign yields respond to economic shocks. Our research topic becomes particularly pertinent amidst recent global crises, encompassing financial crises, pandemics, wars, along with disruptions in value chains and political fragmentation.

By highlighting the importance of economic complexity in securing more favorable financing terms in international debt markets for countries, we indirectly emphasize the need for diversifying their range of export products, especially for economies at low and intermediate levels of development. The effectiveness of diversification and industrial policies, whether currently in place or in the process of implementation, can be evaluated by consistently tracking a country's complexity metrics over time.

This topic has taken center stage in both academic and economic policy discussions, fueled by the pressing need for economies to enhance their resilience and adaptability. In fiscal matters, the urgency is even more pronounced, given the increasing levels of public and private debt that render economies more vulnerable to external shocks, resulting in prolonged financial pressures.

References

Afonso, António, and João Tovar Jalles, 2019, "Quantitative Easing and Sovereign Yield Spreads: Euro-area Time-varying Evidence," Journal of International Financial Markets, Institutions and Money, Vol. 58, pp. 208-224.

Afonso, António, and Manuel M.F. Martins, 2012, "Level, Slope, Curvature of the Sovereign Yield Curve, and Fiscal Behaviour," Journal of Banking & Finance, Vol. 36, No. 6, pp. 1789-1807.

Arellano, Cristina, and Ananth Ramanarayanan, 2012, "Default and the Maturity Structure in Sovereign Bonds," Journal of Political Economy, Vol. 120, No. 2, pp. 187-232.

Aronica, Martina, Pietro Pizzuto, and Caterina Sciortino, 2022, "COVID‐19 and Tourism: What Can We Learn from the Past?," The World Economy, Vol. 45, No. 2, pp. 430-444.

Arora, Vivek, and Martin Cerisola, 2001, "How Does US Monetary Policy Influence Sovereign Spreads in Emerging Markets?," IMF Staff Papers, Vol. 48, No. 3, pp. 474-498.

Augustin, Patrick, 2018, "The Term Structure of CDS Spreads and Sovereign Credit Risk," Journal of Monetary Economics, Vol. 96, pp. 53-76.

Bach, Philipp, Victor Chernozhukov, Malte S. Kurz, and Martin Spindler, 2023, "DoubleML--An Object-Oriented Implementation of Double Machine Learning in R," arXiv preprint arXiv:2103.09603v4.

Balland, Pierre-Alexandre, Tom Broekel, Dario Diodato, Elisa Giuliani, Ricardo Hausmann, Neave O'Clery, and David Rigby, 2022, "Reprint of The New Paradigm of Economic Complexity," Research Policy, Vol. 51, No. 8, p. 104568.

Banerji, Sanjay, Alexia Ventouri, and Zilong Wang, 2014, "The Sovereign Spread in Asian Emerging Economies: The Significance of External Versus Internal Factors," Economic Modelling, Vol. 36, pp. 566-576.

Beetsma, Roel, Oana Furtuna, Massimo Giuliodori, and Haroon Mumtaz, 2021, "Revenue- versus Spending-based Fiscal Consolidation Announcements: Multipliers and Follow-up," Journal of International Economics, Vol. 131, p. 103455.

Bellas, Dimitri, Michael Papaioannou, and Iva Petrova, 2010, "Determinants of Emerging Market Sovereign Bond Spreads: Fundamentals vs. Financial Stress," International Monetary Fund.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, 2014, "High-dimensional Methods and Inference on Structural and Treatment Effects," Journal of Economic Perspectives, Vol. 28, No. 2, pp. 29-50.

Beqiraj, Elton, Valeria Patella, and Massimiliano Tancioni, 2021, "Fiscal Stance and the Sovereign Risk Pass-Through," Economic Modelling, Vol. 102, p. 105573.

Brooks, Sarah M., Raphael Cunha, and Layna Mosley, 2022, "Sovereign Risk and Government Change: Elections, Ideology and Experience," Comparative Political Studies, Vol. 55, No. 9, pp. 1501-1538.

Buraschi, Andrea, and Alexei Jiltsov, 2005, "Inflation Risk Premia and the Expectations Hypothesis," Journal of Financial Economics, Vol. 75, No. 2, pp. 429-490.

Butler, Alexander W., and Larry Fauver, 2006, "Institutional Environment and Sovereign Credit Ratings," Financial Management, Vol. 35, No. 3, pp. 53-79.

Caggiano, Giovanni, and Luciano Greco, 2012, "Fiscal and Financial Determinants of Eurozone Sovereign Spreads," Economics Letters, Vol. 117, No. 3, pp. 774-776.

Camba-Méndez, Gonzalo, 2020, "On the Inflation Risks Embedded in Sovereign Bond Yields" (No. 2423), ECB Working Paper.

Camba-Méndez, Gonzalo, and Thomas Werner, 2017, "The Inflation Risk Premium in the Post-Lehman Period" (No. 2033), ECB Working Paper.

Candelon, Bertrand, and Rubens Moura, 2023, "Sovereign Yield Curves and the COVID-19 in Emerging Markets," Economic Modelling, Vol. 127, p. 106453.

Canzoneri, Matthew B., Robert E. Cumby, and Behzad T. Diba, 2002, "Should the European Central Bank and the Federal Reserve Be Concerned about Fiscal Policy?," Rethinking Stabilization Policy, pp. 29-31.

Chatterjee, Satyajit, and Burcu Eyigungor, 2019, "Endogenous Political Turnover and Fluctuations in Sovereign Default Risk," Journal of International Economics, Vol. 117, pp. 37-50.

Chen, Tianqi, and Carlos Guestrin, 2016, "Xgboost: A Scalable Tree Boosting System," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785-794.

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins, 2018, "Double/Debiased Machine Learning for Treatment and Structural Parameters," The Econometrics Journal, Vol. 21, No. 1, pp. C1-C68.

Coroneo, Laura, and Sergio Pastorello, 2020, "European Spreads at the Interest Rate Lower Bound," Journal of Economic Dynamics and Control, Vol. 119, p. 103979.

D'Amico, Stefania, Don H. Kim, and Min Wei, 2018, "Tips from TIPS: The Informational Content of Treasury Inflation-Protected Security Prices," Journal of Financial and Quantitative Analysis, Vol. 53, No. 1, pp. 395-436.

Dailami, Mansoor, Paul R. Masson, and Jean-Jose Padou, 2008, "Global Monetary Conditions Versus Country-Specific Factors in the Determination of Emerging Market Debt Spreads," Journal of International Money and Finance, Vol. 27, No. 8, pp. 1325-1336.

De Santis, Roberto A., 2020, "Impact of the Asset Purchase Programme on Euro Area Government Bond Yields Using Market News," Economic Modelling, Vol. 86, pp. 192-209.

Deaton, Angus, 1999, "Commodity Prices and Growth in Africa," Journal of Economic Perspectives, Vol. 13, No. 3, pp. 23-40.

Du, Wenxin, and Jesse Schreger, 2016, "Local Currency Sovereign Risk," The Journal of Finance, Vol. 71, No. 3, pp. 1027-1070.

Du, Wenxin, Joanne Im, and Jesse Schreger, 2018, "The US Treasury Premium," Journal of International Economics, Vol. 112, pp. 167-181.

Eichler, Stefan, 2014, "The Political Determinants of Sovereign Bond Yield Spreads," Journal of International Money and Finance, Vol. 46, pp. 82-103.

Eichler, Stefan, and Dominik Maltritz, 2013, "The Term Structure of Sovereign Default Risk in EMU Member Countries and Its Determinants," Journal of Banking & Finance, Vol. 37, No. 6, pp. 1810-1816.

Ferrarini, Benno, and Pasquale Scaramozzino, 2016, “Production complexity, adaptability and economic growth,” Structural Change and Economic Dynamics, Vol. 37, pp. 52-61.

Freixas, Xavier, and Jean-Charles Rochet, 2008, "Microeconomics of Banking," MIT Press.

Friedman, Jerome H., 2001, "Greedy Function Approximation: A Gradient Boosting Machine," Annals of Statistics, pp. 1189-1232.

Gala, Paulo, Jhean Camargo, Guilherme Magacho, and Igor Rocha, 2018, "Sophisticated Jobs Matter for Economic Complexity: An Empirical Analysis Based on Input-Output Matrices and Employment Data," Structural Change and Economic Dynamics, Vol. 45, pp. 1-8.

Gilchrist, Simon, Bin Wei, Vivian Z. Yue, and Egon Zakrajšek, 2022, "Sovereign Risk and Financial Risk," Journal of International Economics, Vol. 136, p. 103603.

Gill, Navjeet, 2018, "A Research on Eurozone Bond Market and Determinants of Sovereign Bond Yields," Journal of Financial Risk Management, Vol. 7, No. 2, pp. 174-190.

Giraldo, Carlos, Iader Giraldo, Jose E. Gomez-Gonzalez, and Jorge M. Uribe, 2023, "An Explained Extreme Gradient Boosting Approach for Identifying the Time-Varying Determinants of Sovereign Risk," Finance Research Letters, Vol. 57, p. 104273.

Gomez-Gonzalez, Jose E., Jorge M. Uribe, and Oscar M. Valencia, 2023, "Does Economic Complexity Reduce the Probability of a Fiscal Crisis?," World Development, Vol. 168, p. 106250.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu, 2020, "Empirical Asset Pricing via Machine Learning," The Review of Financial Studies, Vol. 33, No. 5, pp. 2223-2273.

Gumus, Inci, 2011, "Exchange Rate Policy and Sovereign Spreads in Emerging Market Economies," Review of International Economics, Vol. 19, No. 4, pp. 649-663.

Güneri, Barbaros, and Ahmet Y. Yalta, 2021, "Does Economic Complexity Reduce Output Volatility in Developing Countries?," Bulletin of Economic Research, Vol. 73, No. 3, pp. 411-431.

Gupta, Sanjeev, Amine Mati, Emanuele Baldacci, 2008, "Is It (Still) Mostly Fiscal? Determinants of Sovereign Spreads in Emerging Markets," International Monetary Fund.

Gürkaynak, Refet S., Brian Sack, and Jonathan H. Wright, 2010, "The TIPS Yield Curve and Inflation Compensation," American Economic Journal: Macroeconomics, Vol. 2, No. 1, pp. 70-92.

Hallak, Juan Carlos, 2006, "Product Quality and the Direction of Trade," Journal of International Economics, Vol. 68, No. 1, pp. 238-265.

Hartmann, Dominik, Miguel R. Guevara, Crsitian Jara-Figueroa, Manuel Aristarán and César A. Hidalgo, 2017, "Linking Economic Complexity, Institutions, and Income Inequality," World Development, Vol. 93, pp. 75-93.

Hausmann, Ricardo, César A. Hidalgo, Sebastián Bustos, Michele Coscia, and Alexander Simoes, 2014, "The Atlas of Economic Complexity: Mapping Paths to Prosperity," MIT Press.

Hausmann, Ricardo, Jason Hwang, and Dani Rodrik, 2007, "What You Export Matters," Journal of Economic Growth, Vol. 12, pp. 1-25.

Hidalgo, César A., 2021, "Economic Complexity Theory and Applications," Nature Reviews Physics, Vol. 3, No. 2, pp. 92-113.

Hidalgo, César A., and Ricardo Hausmann, 2009, "The Building Blocks of Economic Complexity," Proceedings of the National Academy of Sciences, Vol. 106, No. 26, pp. 10570-10575.

Hilscher, Jens, and Yves Nosbusch, 2010, "Determinants of Sovereign Risk: Macroeconomic Fundamentals and the Pricing of Sovereign Debt," Review of Finance, Vol. 14, No. 2, pp. 235-262.

Hördahl, Peter, and Oreste Tristani, 2012, "Inflation Risk Premia in the Term Structure of Interest Rates," Journal of the European Economic Association, Vol. 10, No. 3, pp. 634-657.

Krishnamurthy, Arvind, Stefan Nagel, and Annette Vissing-Jorgensen, 2018, "ECB Policies Involving Government Bond Purchases: Impact and Channels," Review of Finance, Vol. 22, No. 1, pp. 1-44.

Liu, Zhuoshi, and Peter Spencer, 2013, "Modelling Sovereign Credit Spreads with International Macro-Factors: The Case of Brazil 1998–2009," Journal of Banking & Finance, Vol. 37, No. 2, pp. 241-256.

Longstaff, Francis A., Jun Pan, Lasse H. Pedersen, and Kenneth J. Singleton, 2011, "How Sovereign Is Sovereign Credit Risk?," American Economic Journal: Macroeconomics, Vol. 3, No. 2, pp. 75-103.

Lundberg, Scott, and Sun-In Lee, 2017, "A Unified Approach to Interpreting Model Predictions," Advances in Neural Information Processing Systems, Vol. 30.

Maggioni, Daniela, Alessia L. Turco, and Mauro Gallegati, 2016, "Does Product Complexity Matter for Firms' Output Volatility?," Journal of Development Economics, Vol. 121, pp. 94-109.

Maltritz, Dominik, 2012, "Determinants of Sovereign Yield Spreads in the Eurozone: A Bayesian Approach," Journal of International Money and Finance, Vol. 31, No. 3, pp. 657-672.

Maltritz, Dominik, and Alexander Molchanov, 2013, "Analyzing Determinants of Bond Yield Spreads with Bayesian Model Averaging," Journal of Banking & Finance, Vol. 37, No. 12, pp. 5275-5284.

Maltritz, Dominik, and Alexander Molchanov, 2014, "Country Credit Risk Determinants with Model Uncertainty," International Review of Economics & Finance, Vol. 29, pp. 224-234.

Matsumura, Marco S., and José V. Machado V., 2010, "The Role of Macroeconomic Variables in Sovereign Risk," Emerging Markets Review, Vol. 11, No. 3, pp. 229-249.

Mealy, Penny, and Alexander Teytelboym, 2022, "Economic Complexity and the Green Economy," Research Policy, Vol. 51, No. 8, p. 103948.

Medas, Paulo, Tigran Poghosyan, Yizhi Xu, Juan Farah-Yacoub, and Kerstin Gerling, 2018, "Fiscal Crises," Journal of International Money and Finance, Vol. 88, pp. 191-207.

Moreno-Badia, Marialuz M., Paulo Medas, Pranav Gupta, and Yuan Xiang, 2022, "Debt Is Not Free," Journal of International Money and Finance, Vol. 127, p. 102654.

Nepelski, Daniel, and Giuditta De Prato, 2020, "Technological Complexity and Economic Development," Review of Development Economics, Vol. 24, No. 2, pp. 448-470.

Neyman, Jerzy, 1979, "C (α) Tests and Their Use," Sankhyā: The Indian Journal of Statistics, Series A, 1-21.

Nguyen, Canh Phuc, 2021, "Gender Equality and Economic Complexity," Economic Systems, Vol. 45, No. 4, p. 100921.

Poghosyan, Tigran, 2014, "Long-Run and Short-Run Determinants of Sovereign Bond Yields in Advanced Economies," Economic Systems, Vol. 38, No. 1, pp. 100-114.

Robinson, Peter M., 1988, "Root-N-Consistent Semiparametric Regression," Econometrica: Journal of the Econometric Society, pp. 931-954.

Romero, João P., and Camila Gramkow, 2021, "Economic Complexity and Greenhouse Gas Emissions," World Development, Vol. 139, p. 105317.

Sánchez, Juan M., Horacio Sapriza, and Emircan Yurdagul, 2018, "Sovereign Default and Maturity Choice," Journal of Monetary Economics, Vol. 95, pp. 72-85.

Sweet, Cassandra, and Dalibor Eterovic, 2019, "Do Patent Rights Matter? 40 Years of Innovation, Complexity, and Productivity," World Development, Vol. 115, pp. 78-93.

Tacchella, Andrea, Dario Mazzilli, and Luciano Pietronero, 2018, "A Dynamical Systems Approach to Gross Domestic Product Forecasting," Nature Physics, Vol. 14, No. 8, pp. 861-865.

Wang, Ruohan, Yi Xue, and Wenping Zheng, 2021, "Does High External Debt Predict Lower Economic Growth? Role of Sovereign Spreads and Institutional Quality," Economic Modelling, Vol. 103, p. 105591.

Wellmann, Dennis, and Stefan Trück, 2018, "Factors of the Term Structure of Sovereign Yield Spreads," Journal of International Money and Finance, Vol. 81, pp. 56-75.

Zou, Zongsen, Yu Zhang, Xin Li, and Xiuling Wang, 2023. "The nexus between product sophistication and export survival: Evidence from China," Structural Change and Economic Dynamics, Vol. 67, pp. 480-502.

UBIREA
Institut de Recerca en Economia Aplicada Regional i Pública
Research Institute of Applied Economics
Universitat de Barcelona
Av. Diagonal, 690 • 08034 Barcelona

WEBSITE: www.ub.edu/irea • **CONTACT:** irea@ub.edu