ISSN 2392-1641
e-ISSN 2450-0097

Economics and Business Review

Volume 12 (1) 2026

CONTENTS

Poznań University of Economics and Business Press

Economics and Business Review

Volume 12 (1) 2026

Paper based publication

https://doi.org/10.18559/ebr.2026.1

ISSN 2392-1641
e-ISSN 2450-0097

POZNAŃ UNIVERSITY OF ECONOMICS AND BUSINESS PRESS
ul. Powstańców Wielkopolskich 16, 61-895 Poznań, Poland
phone +48 61 854 31 54, +48 61 854 31 55
https://wydawnictwo.ue.poznan.pl, e-mail: wydawnictwo@ue.poznan.pl
postal address: Al. Niepodległości 10, 61-875 Poznań, Poland

Printed and bound in Poland by:
Perfekt – Gaul i wspólnicy sp. k.

Circulation: 80 copies

Economics and Business Review

Volume 12 (1) 2026

CONTENTS

PUEB PRESS
Economics and Business Review Vol. 12 (1), 2026: 5–8
https://www.ebr.edu.pl

Editorial introduction

One of the defining characteristics of recent history is the tremendous growth of many economies that, not long ago, were described as 'third-world'. As of 2025 (according to the IMF), developing economies constitute 41.5% of world GDP–more than twice the share in 1995. That growth fuelled many scientific inquiries that reveal distinctive characteristics of economic processes in developing countries, thereby broadening economic knowledge beyond what is presented in typical economics courses in Western countries. *Economics and Business Review*, due to its focus on emerging market economies, is one of the beneficiaries of that growth, because many of its publications show how well-known economic processes operate in local, often non-Western circumstances. In this context, the editors are glad to welcome Dani Rodrik, who, along with John Cantwell and Ari Van Assche, has recently agreed to join the journal's Editorial Advisory Board. Professor Rodrik has devoted much of his writing to the problems of international trade, globalisation, and the role of emerging markets in that process, and repeatedly reminded readers that the local context must be taken into account when analysing the problems and policy challenges in particular economies. He stems from that view in the opening piece of the current issue of *Economics and Business Review*, where he notes that advanced economies have started to experience problems typical of developing economies and offers tailored, original policy recommendations to address them.

Apart from the invited piece by Professor Rodrik, the current issue comprises six research articles by fifteen scholars based in China, India, Nigeria, Pakistan, Poland, Saudi Arabia, Türkiye, and the United Arab Emirates. In other words, this issue, to a large extent, offers a perspective from scholars in developing economies and shows that local contexts inspire a rethink of theories developed in advanced Western economies. The editors *of Economics and Business Review* hope that readers of this issue will experience such inspiration while reading it.

As noted, the issue begins with an invited perspective by Dani Rodrik titled **On productivism**, which formed the basis of his lecture at Poznań University of Economics and Business during the ceremony granting him the degree of *doctor honoris causa*. Professor Rodrik addresses the problem of unequal development in advanced economies, observed in the lagging growth of peripheral regions, the decline in the number of stable, adequately paid jobs, the shrinking middle class, and, as a consequence, the rise in political populism. The essay argues that existing policies are inadequate to address these challenges and proposes a new policy paradigm called productivism. It aims to achieve, as Professor Rodrik puts it, an equal distribution of productive economic opportunities across local communities and socio-economic groups. Thus, the approach aims not at redistributing income or stimulating aggregate demand but at providing equal chances to pursue individual economic prosperity. The direct aim is to create good jobs that generate relatively stable, decent income, safe working conditions, and some prospects for career progression. Subsequently, the essay outlines how the existing policy instruments can be modified to achieve that aim.

The first regular article, titled **Economic complexity and the shadow economy in Africa: An assessment of nonlinearity and asymmetry**, by James Temitope Dada et al., examines the relationship between economic complexity and the shadow economy, an area that has been addressed to date with mixed results. Economic complexity is understood as a synonym for economic development and is reflected in a country's productive capacity and diversity, as observed in its export structure. The authors assess the relationship between this measure and the shadow economy across 28 African countries from 1995 to 2020 using a battery of econometric techniques. They establish that the analysed relationship is asymmetric: when economic complexity is low and begins to grow, it initially increases the shadow economy; however, once a certain level of economic complexity is reached, further growth reduces informal economic activity.

The subsequent article, **Demystifying Foreign Direct Investment dynamics in emerging economies: An ISM–MICMAC analysis**, by Srabani Paul Grover et al., examines the structural interdependencies among key determinants of Foreign Direct Investment (FDI) in emerging markets. Moving beyond conventional multiple regression frameworks commonly employed in the FDI literature, the study applies Interpretive Structural Modelling (ISM) combined with MICMAC analysis to identify hierarchical linkages and driving-dependence relationships among FDI determinants. The results indicate that political stability, corruption, and infrastructure quality are the most influential driving factors, shaping other economic and institutional variables within the system. By clustering determinants according to their driving and dependence power, the paper offers policy-relevant insights for enhancing investment attractiveness in emerging economies.

Political connection and corporate ESG performance: Evidence from China by Congming Ding and Xuezhenzi Hu investigates the impact of executives' political connections on firms' ESG performance in China's capital market. Using panel data for A-share listed companies over the period 2009–2022, the authors examine whether politically connected executives affect ESG ratings. The results reveal a significant positive association between political ties and ESG scores. Mechanistic analysis indicates that political connections enhance ESG performance by increasing media scrutiny, reducing financing constraints, and improving access to government subsidies. The robustness of the findings is confirmed using Two-Stage Least Squares (2SLS) estimation, highlighting the role of political capital in fostering sustainable corporate practices.

The next contribution, entitled **Liquidity risk and liquidity timing in the cross-section of Indian equity mutual fund returns** and written by Suresh Kumar and Hyder Ali, investigates the dual role of liquidity exposure and liquidity timing in shaping the performance of Indian equity mutual funds over the period 2007–2024, encompassing the Global Financial Crisis, the COVID-19 shock, and the post-pandemic recovery. Sorting funds by liquidity betas and constructing a high-minus-low liquidity spread (HMLiq), the authors find a positive liquidity premium in tranquil and recovery periods that weakens during crises. The results indicate that timing ability is concentrated among high-liquidity-beta funds and offers limited protection under systemic stress. The findings are particularly relevant for institutional investors, fund managers, and financial regulators concerned with portfolio resilience and liquidity risk management in emerging markets.

Forecasting cryptocurrencies in turbulent times: Evidence on parsimony versus model complexity, by Anna Tatarczak and Oleksandra Humeniuk, investigates short-term return predictability for Bitcoin, Ether, and Litecoin over the period 2020–2024. The authors compare autoregressive benchmarks with Kitchen Sink and VARX-type models using both point and density forecast accuracy measures, supported by Diebold–Mariano tests and Model Confidence Set inference. The results show that a first-order autoregressive specification and other parsimonious models incorporating cryptocurrency-specific variables consistently outperform more complex linear frameworks, while macro-financial predictors add little incremental value. The findings are particularly relevant for quantitative investors, portfolio managers, and risk analysts seeking forecasting strategies in highly volatile cryptocurrency markets.

The final paper of the issue, **From digital mining to market prices: An empirical analysis of the relationship between energy consumption and price dynamics of Bitcoin and Ether**, authored by Levent Sezal, also situated within the cryptocurrency research stream, comparatively examines the relationship between energy consumption and price dynamics for Bitcoin and Ether. Using daily data and applying Augmented Dickey–Fuller and Phillips–Perron unit root tests, ARDL cointegration analysis, and Toda–Yamamoto causality tests, the study evaluates both short- and long-run linkages. The findings reveal a long-term cointegration relationship for Bitcoin, with unidirectional causality running from prices to energy consumption, whereas no such relationship is identified for Ether following its transition to a Proof-of-

-Stake mechanism. The study contributes to the literature at the intersection of digital finance and energy economics and offers insights for policymakers and energy market analysts assessing the environmental implications of cryptocurrency markets.

Michał Pilc
Konrad Sobański
Lead Editors

PUEB PRESS
Economics and Business Review
https://www.ebr.edu.pl
Vol. 12 (1), 2026: 9–28
https://doi.org/10.18559/ebr.2026.1.3121

On productivism[1]

Dani Rodrik[2]

Abstract

'Productivism' refers to an approach that prioritises the dissemination of productive economic opportunities throughout the entire economy and segments of the labour force. It differs from what has come to be called 'neoliberalism' by assigning governments and civil society significant roles in achieving this goal. Productivism puts less faith in markets and is suspicious of large corporations. It emphasises production and investment over finance and the revitalisation of local communities over globalisation. It also departs from the Keynesian welfare state by focusing less on redistribution, social transfers, and macroeconomic management, and more on creating economic opportunity by working on the supply side of the economy to create good, productive jobs for everyone. This article relates the contemporary labour market problems of advanced economies to the dualism literature in economic development, which focuses on the divergence between 'modern' and 'traditional' segments within poor economies. It then highlights the nature of the new challenges and why established models of economic growth and Keynesian social welfare need to be updated. It describes new modes of industrial policy required to deal with these challenges and questions whether

Keywords

- productivism
- labour market
- globalisation
- industrial policy
- economic growth
- social welfare

Suggested citation: Rodrik, D. (2026). On productivism. *Economics and Business Review*, *12*(1), 9–28. https://doi.org/10.18559/ebr.2026.1.3121

[1] Invited perspective. The text covers the key points of the lecture given by Dani Rodrik during the ceremony in which he was awarded Doctor Honoris Causa from Poznań University of Economics and Business. Original place of publication: Besley, Tim, Bucelli, Irene and Velasco, Andrés (eds) (2025) The London Consensus: Economic Principles for the 21st Century, London: LSE Press, pp. 77–96 https://doi.org/10.31389/lsepress.tlc.c

[2] Harvard Kennedy School, 79 John F. Kennedy Street, Cambridge, MA 02138, USA.

our governments are up to it. It also discusses how the elements of this new strategy are drawing support from both sides of the political spectrum.

JEL codes: J21, L52, O10, O14, O25.

Introduction: An old problem in a new setting

How to overcome 'productive dualism' is our central economic challenge. Dualism is an old idea that lies at the core of development economics and has become increasingly relevant to advanced economies as well. The economists who founded the field of development economics, such as the Nobel Prize winning W. Arthur Lewis, noted that the economies of poor nations are split between a narrow 'modern' sector that uses advanced technologies and a much larger 'traditional' sector characterised by extremely low productivity. For a long time, scholars considered dualism as the defining feature of developing countries, in contrast to advanced nations where they assumed that frontier technologies and high productivity prevailed across the entire economy. This marked development economics as a distinct sub-branch of economics, separate from conventional neoclassical economics. Correspondingly, the task of development policy became the establishment of new institutional arrangements to overcome the disparities in incomes, education, health, and life chances more broadly created by productive dualism. While the developed-developing country distinction may have made some sense in the 1950s and 1960s, in the 2020s it no longer appears to be relevant.

Industrialisation has been the traditional vehicle for overcoming dualism; as workers get absorbed into more productive manufacturing activities, wages rise, and the economy's overall productivity increases. But this old and powerful remedy no longer works. As a result of automation and other innovations that have been labour-saving, manufacturing has lost its ability to create plentiful jobs in both developing and advanced economies (Diao et al., 2021). Globalisation has accelerated the process as a small number of countries with strong comparative advantages in manufacturing have squeezed production in middle- and high-income economies.

Employment de-industrialisation has been a common feature of all advanced economies. Manufacturing employment has declined (as a share of total employment), even in countries like South Korea or Germany that have maintained strong industrial sectors. Increasingly, developing countries have also struggled to create significant employment in formal manufacturing firms.

Changes in manufacturing technologies have made it difficult for low-income countries to successfully compete in manufacturing without using skill- and capital-intensive technologies that absorb limited labour and are inappropriate in light of these countries' underlying factor endowments, since they are abundant in low-skilled labour and scarce in capital.

Hence, productive dualism is becoming an entrenched feature of developing and advanced economies alike, requiring remedies that come straight out of the development policy toolbox. In a 2017 book called *The vanishing middle class*, the MIT economic historian Peter Temin pointed out that the Lewis model of a dual economy had become increasingly relevant to the US (Temin, 2017). De-industrialisation, globalisation, new technologies that favour professionals and capitalists, and declining protections for labour have widened the gap between the winners from these developments and those who are left behind. Convergence between poor and rich parts of the economy has been arrested, labour markets became increasingly polarised between high- and low-educated workers, and regional disparities widened. In Europe, the increase in inequality has not been as marked thanks to stronger welfare states, but the same forces have operated there too. The gaps between the most productive firms and regions, and those lagging behind, grew while the middle class shrank (Vacas-Soriano & Fernandez-Macias, 2017).

Consequently, policymakers in advanced economies are now grappling with the same questions that have preoccupied development policymakers for a long time: how to attract investment, create jobs, increase skills, spur entrepreneurship, enhance access to credit and technology—in short, how to close the gap with the more advanced, productive parts of the national economy. The starting points may be different, but the problems of a region where good jobs have disappeared, productive employment has become scarce, social problems (such as crime and addiction) have mushroomed, and there is low trust between government officials and various social groups, and the business community looks distressingly familiar to a development economist. The obstacles that racial or ethnic minorities, recent immigrants, or low-educated workers must surmount in such settings are the bread-and- butter of development economics.

Localities that are left behind in advanced economies may have access to greater amounts of financial resources. In the United States, state and local governments spend tens of billions of dollars, not very effectively, on tax incentives and other subsidies to attract large firms (Slattery & Zidar, 2020). But their officials typically operate under structural and bureaucratic constraints that would be familiar to their counterparts in poor nations. They lack the requisite information on where the most important opportunities and bottlenecks are, they are subject to political pressure and lobbying from parochial private interests, and the capabilities they need to mobilise, even when they exist, are spread across a wide range of public and private organisations that they do not directly control. The new realities of labour markets require updates to established models of growth and the Keynesian social welfare state.

In this article, I describe the 'productivism' approach, which is a remedy that targets productive dualism at its source. I first outline this approach and then compare it to other policy frameworks with the help of a taxonomy of public policies. I discuss the relationship between productivist policies and what are commonly called industrial policies, providing an example of how they can be deployed in service sectors. Since economists and many others tend to be sceptical of the capacity of governments to undertake transformational policies, I will address some of the traditional objections to government interference in the productive sphere. I also suggest that productivism carries appeal for many elements of both the right and left side of the political spectrum. I end the article with some cautions about the dangers of taking economic paradigms too seriously.

1. A new approach

Productivism is an approach that prioritises the dissemination of productive economic opportunities throughout all parts of the economy and segments of the labour force. Our core economic and social problems—poverty, inequality, exclusion, and insecurity—have many roots. But they are reproduced and reinforced on a daily basis as immediate by-products of firms' employment, investment, and innovation decisions. In the language of economists, these decisions are rife with externalities for society, i.e., they have consequences that spill over to many people, firms, and other parts of the economy. Some of these externalities are well recognised in economics. Learning and innovation spillovers from research and development (R&D) form the rationale for tax credits and other public subsidies. Environmental externalities and the effects of greenhouse gas emissions on climate change form the basis for environmental regulation.

But today, these externalities are broader and include what we can call 'good jobs' externalities. Good jobs are a pathway to the middle class. They pay well enough to allow for a reasonable living standard with some security and savings, are relatively stable, have safe working conditions, and offer some career progression. Firms that generate good jobs contribute to the vitality of their communities. Conversely, a shortage of good jobs comes at social, political, and economic costs. Social consequences can take the form of exclusion, broken families, drug abuse, addiction, and crime. Political ills can follow, such as polarisation, the rise of populism, backlashes against globalisation and immigration, decline in trust in government, experts, and institutions. The prevalence of 'bad jobs' is also symptomatic of economic dualism, which creates its own inefficiency: productive technologies remain bottled up in a few firms and do not disseminate throughout the rest of the economy and the labour force.

Firms' decisions on how many workers to employ, how much to pay, what kind of technologies to deploy and how to organise work affect not just the bottom line, but the life opportunities of prospective employees and their communities. When a company decides to automate its production line or outsource part of its production to another country, society may suffer long-term damage that is not internalised by its managers or shareholders. Framing the problem as an 'externality'—or as a 'coordination failure' that prevents firms and governments from undertaking complementary actions (in training, technology adoption, investment decisions) for broad-based prosperity—clarifies that productivism is about productivity, and not about redistribution or social/labour standards. But it does not presume productivity trickles down. It aims to enhance wellbeing across all sectors of society by directly broadening access to productive employment opportunities.

Productivism differs from what has come to be called 'neoliberalism' in that it gives governments and civil society significant roles in achieving productive employment goals. It puts less faith in markets and is suspicious of large corporations. It emphasises production and investment over finance, and the revitalisation of local communities over globalisation. It also departs from the Keynesian welfare state—the paradigm that neoliberalism replaced—in that it focuses less on redistribution, social transfers, and macroeconomic management, and more on creating economic opportunity by working on the supply side of the economy to create good, productive jobs for everyone. And productivism diverges from both of its antecedents by exhibiting greater scepticism towards technocrats and being less instinctively hostile to populism in the economic sphere (Rodrik, 2018).

2. Where conventional models fall short

To see how productivism differs from alternative approaches, it is useful to consider our policy options through a matrix that categorises different approaches to prosperity and inequality (Figure 1). First, I divide policies into pre-production, production, and post-production stage interventions. To understand fully the range of options for creating inclusive prosperity, this is a better categorisation of policies than the conventional pre-distribution/re-distribution distinction. Within the pre-distribution category, my framework makes a further distinction between policies that affect endowments people

		At what stage of the economy does policy intervene?		
		pre-production	production	post-production
Which segment of the economy do we care about?	low productivity			
	middle productivity			
	high productivity			

Figure 1. Remedies for prosperity and inequality

Source: own work.

bring to markets (such as education) and policies that directly influence production, employment, and investment decisions (such as industrial policies or labour market regulations). Second, I divide interventions into those that intend to redress inequities at the bottom, middle, or top of the income distribution. Minimum wages, e.g., target the incomes of the working poor while wealth taxes target incomes at the very top.

The traditional welfare state model operates largely within the first and third columns: it targets the educational and other endowments of workers before they join labour markets and *ex-post* redistribution through taxes and social insurance policies (see Figure 2). The government's role is to finance education, engage in progressive taxation, and provide social insurance against idiosyncratic risks, such as unemployment, illness, and disability. The presumption is that good/middle-class jobs will be available to everyone with adequate education and skills.

Traditional growth strategies, on the other hand, focus on the most productive segments of the economy and encompass interventions within the bottom row (see Figure 3). These may include innovation systems, intellectual property rules, appropriate regulatory structures, and export and innovation incentives. The presumption is that high growth eventually pulls everyone up and leaves few regions or pockets of the labour market behind.

When the inadequacy of good/middle-class jobs is driven by secular trends, such as technology and globalisation, neither of these strategies work well. Economic insecurity, inequality, and poor productivity (except for those at the very top) are important structural problems today. Secular trends in technology and globalisation are hollowing out the middle of the employment distri-

		At what stage of the economy does policy intervene?		
		pre-production	production	post-production
Which segment of the economy do we care about?	low productivity	investments in eduction and training		transfers; full-employment macro policies
	middle productivity			social insurance, pensions, safety nets
	high productivity			

Figure 2. Traditional welfare state model

Source: own work.

		At what stage of the economy does policy intervene?		
		pre-production	production	post-production
Which segment of the economy do we care about?	low productivity			
	middle productivity			
	high productivity	innovation systems, IPR rules, trade agreements	market-friendly regulations, R&D and export incentives	corporate tax incentives

Figure 3. Traditional growth model

Source: own work.

bution. These trends exhibit themselves in the form of bad jobs that do not offer stability, sufficient pay, and career progression, and in permanently depressed labour markets outside major metropolitan centres. These problems need a different strategy that tackles the creation of good jobs directly. The focus necessarily turns to firms; to help them internalise the economic and social spillovers that they generate. Hence, the productive sector must be at the heart of such a strategy. This calls for targeting the middle cell of the matrix, focusing on direct interventions in the productive sphere with the goal of expanding the supply of middle-skill jobs (Figure 4). Altogether, we must change what we produce, how we produce it and who gets a say in production decisions. This requires not just new policies, but also a reconfiguration of existing ones.

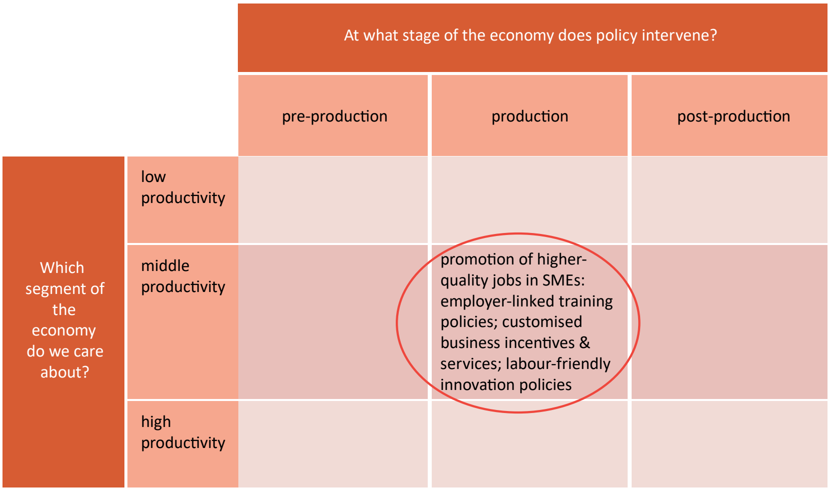

Figure 4. The productivist 'good-jobs' model

Source: own work.

Advanced and developing nations alike will need a new breed of coordinated policies aimed at the supply and demand sides of labour markets, combining skill training programmes with support for firms (Rodrik & Stantcheva, 2021a). Good jobs require good firms and vice versa. Active labour market policies designed to increase skills and employability should broaden into partnerships with firms explicitly targeting the creation of good jobs (Rodrik & Sabel, 2019). Industrial and regional policies that currently centre on tax incentives and investment subsidies should be replaced by customised business services and amenities to facilitate maximum employment creation (Bartik, 2019). National innovation systems should be redesigned to orient investments in new technologies in a more employment- friendly direction (Acemoglu & Restrepo, 2019). Policies that tackle climate change, such as the European Green Deal, should be explicitly linked to programmes of job creation in lagging communities (European Commission, 2019). Recognising that in the future prosperity will have to rely much more on services and smaller and medium-sized enterprises, the focus of industrial policy should be reoriented away from manufacturers and 'national champions', large private corporations that receive priority in government policies.

A new economic order requires an explicit quid pro quo between private firms and public authorities. To prosper, firms need a reliable and skilled workforce, good infrastructure, an ecosystem of suppliers and collaborators, easy access to technology, and a sound regime of contracts and property rights. Most of these are provided through public and collective action, which is the government's side of the bargain. Governments, in turn, need firms to internalise the various externalities they produce for their communities and societies when they make their labour, investment and innovation decisions. So, firms must live up to their side of the bargain too, not as corporate social responsibility, but as part of an explicit regulatory and governance framework.

Looking at our policy challenge in these terms makes it clear that the conventional separation between growth policies and social policies no longer makes sense. Faster economic growth requires that new technologies and productive opportunities are disseminated among smaller firms and wider segments of the labour force, and that their use is not confined to narrow segments of the elite. Reducing inequality and economic insecurity is more effective when it happens through better employment prospects than through fiscal redistribution only. The economic growth and the social agenda are increasingly one and the same.

3. New types of industrial policies

If productivism is to be successful it will have to internalise the lessons learned from the failures of past policies and adapt to fundamentally new challenges. State interventions aimed at reshaping the structure of an economy—so-called 'industrial policies'—have been traditionally faulted for being ineffective and getting captured by special interests. 'Governments cannot pick winners', as the old adage goes. In reality, much of this criticism is overdone. While there have been notable failures (Lincicome, 2021), systematic studies in the 2010s and early 2020s find that industrial policies incentivising investment and job creation in disadvantaged regions have done surprisingly well (Criscuolo et al., 2019).

Public initiatives have been behind some of the most startling high-tech successes of our time, including the internet and GPS. For every Solyndra, a solar cell manufacturer that failed spectacularly after half a billion dollars in government loan guarantees (Stephens & Leonnig, 2011), there is often a Tesla, the phenomenally successful electric battery and vehicle manufactures that also received government support at a critical phase of its development (Overly, 2017).

Nevertheless, there is much room for improvement. The most effective industrial policies entail close, collaborative interactions between government agencies and private firms, whereby firms receive critical public inputs—financial support, skilled workers or technological assistance—in return for meeting soft and evolving targets on investment and employment. This kind of industrial policy is likely to work much better, whether in promoting local economic development or in directing major national technological efforts, than open-ended subsidies or tax incentives.

Productivism focuses on enhancing the productive capabilities of all segments and regions of a society. While traditional forms of social assistance and especially better access to education and healthcare can help in this regard, connecting people with productive employment opportunities requires further intervention. It requires improvements both on the demand and the supply side of the labour market (Rodrik, 2021b). Policies must encourage an increase in the quantity and quality of jobs that are available for the less educated and less skilled members of the workforce, where they choose (or can afford to) live.

In the future, the bulk of these jobs will not come from manufacturing, but from services, such as health and long-term care and retail. In the United States, less than one in ten workers are currently employed in manufacturing. Virtually all new net job creation in the private sector since the late 1970s has taken place in services. Even if policy succeeds in reshoring manufacturing and supply chains, the impact on employment is likely to remain limited. The experiences of East Asian manufacturing superstars, such as South Korea and Taiwan, provide sobering examples. These two countries have managed to rapidly increase the share of manufacturing value added in gross domestic product (GDP) (at constant prices), yet, they have experienced steady declines in manufacturing employment ratios.

This is important since so much of the policy effort in the United States is focused on promoting high-tech manufacturing. The most recent example is the CHIPS and Science Act that the US Congress has passed, providing $52 billion in funding for semiconductors and related manufacturing (Moore, 2022). The initiative aims at enhancing national security vis-a-vis China and creating good jobs. Unfortunately, even if the first objective is met, the second objective is likely to remain elusive. A strategy fixated on geopolitical competition with China will not be effective on the jobs front. A similar point can be made about the subsidies to green technologies that are a core component of the so-called Inflation Reduction Act that US President Joe Biden signed in 2022. The green transition is undoubtedly an urgent priority that the new paradigm needs to tackle. But here, governments also cannot achieve multiple objectives with a single instrument. Policies that target climate change are not a substitute for good-job policies and vice versa. Shoring up the middle class and disseminating the benefits of technology broadly through society requires an explicit good-jobs strategy.

4. A good-jobs strategy for services

But is an industrial policy for services possible? I have discussed elsewhere what such a strategy might look like in the contexts of the US, French, and British economies. Here, I will briefly summarise the US proposals (Doshi et al., 2023; Rodrik, 2022; Rodrik & Stancheva, 2021b).

My proposed programme has both local and national components. The local approach would build on existing development and business assistance programmes that are already loosely structured along the lines advocated here. These are collaborative partnerships between local development agencies, firms, and other partners aiming to revitalise local communities and create good jobs. They are organised around an implicit (and evolving) quid pro quo: the provision of public services (such as business extension services, infrastructure, or customised training) in return for soft commitments by firms on investment and employment creation. Such partnerships align with a new, more flexible, and contextual model of industrial policy that is better suited to the challenge of creating good jobs.

The federal initiative would be the establishment of an Advanced Research Projects Agency (ARPA) focused on the promotion of employment-friendly technologies: ARPA-W(orkers). Starting from the premise that innovations that *complement* rather than *displace* workers are feasible, yet, currently undersupplied, ARPA-W would promote early-stage investments in digital and other technologies that enhance prevailing worker skills and create good jobs.

Consider what is perhaps the toughest test case for these ideas: longterm care. Employment in this sector will increase rapidly in future years as the population continues to age and, consequently, demand for in-home or assisted living arrangements increases. Much of long-term care work is done in homes (through agencies that provide the caregivers or through self-employed caregivers) or in assisted living or retirement communities where, unlike hospitals or nursing homes, regulations are weak. In such settings, remuneration and work conditions have traditionally been very poor—characteristics that epitomise bad jobs. Employees are mostly women and disproportionately are people of colour. Long-term care workers are typically regarded as performing low-skill jobs and are often not viewed as real professionals.

As Paul Osterman has noted, there are three ways in which jobs in longterm care can be improved (Osterman, 2020). First, the government can regulate and impose standards (such as high minimum wages). Second, the government can increase reimbursement rates from Medicaid and Medicare in the hope that higher rates translate into increased wages. Third, the productivity of direct-care workers can be raised, allowing the long-term care system to serve patients' needs better and to reduce costs, generating room for better compensation. While the first two strategies might be useful, greater productivity is ultimately the most reliable source of better jobs.

Osterman suggests that it could be useful to increase productivity in long-term care through a strategy that is analogous to the deployment of innovations in manufacturing pioneered by Japanese car producers. This entails a combination of investing in worker skills, providing workers with greater voice, discretion and autonomy, and giving them more responsibility for the quality of the service. Care workers that are empowered with greater autonomy and decision-making can use their knowledge of residents and patients to customise their services and provide more flexibility (e.g., in schedules, food, and treatment). An important component of the strategy could be the introduction of new technologies that complement caregivers' skills, such as digital tools that enable caregivers to collect real-time information and to respond quickly and efficiently to the needs of individual residents.

These changes would require a willingness to experiment with novel work practices and a continuum of efforts, from R&D and the introduction of new technologies for long-term care, on the one hand, and to their local adoption, adaptation, and contextualisation in specific communities, on the other. If long-term care is managed better in these ways, productivity benefits would show up in lower turnover among care workers, reduced hospitalisation rates, better management of chronic conditions, and quicker and smoother transitions out of acute care facilities. None of this is easy. Enhancing productivity in services is notoriously difficult and often impeded by a myriad of well-meaning licensing, safety, and other regulations. But if we cannot find ways of increasing productivity in jobs that our workers are destined for, we will end up with economies that are both worse performing and less inclusive.

5. Are governments up to it?

Scepticism about the ability of governments to lead and achieve positive change is near universal. To many, 'effective government action' is an oxymoron. Given the state of our contemporary politics, such doubts may be well-placed. Authoritarian populism and polarisation—which interact with and reinforce each other—have infested our public sphere to the detriment of our capacity to mount collective action against common problems.

But there is a longer-standing concern about government action that relates to administrative capabilities. Governments do not have the information and capabilities, the argument goes, needed to achieve positive structural change in the economy. Give governments too much power and they will direct resources towards the wrong places and turn into captive tools of special interests. That was the argument at the heart of neoliberalism and a key source of its appeal. It is the argument that must be overcome by any successor narrative on economic policy, and productivism especially, if it is to become successful.

In reality, government capabilities are not inherited or static. They are built over time, once appropriate priorities are set and as a result of experience, learning and building trust with private entities. For public officials, the relevant questions should not be 'do we have the capacity?' but rather, 'do we have in place the right priorities and the correct mode of governance?'

The sceptic might say this all sounds good in theory, but it is not achievable in practice. Look around and public governance seems to be failing throughout, from the local and national to the global level. In fact, as Charles Sabel and David Victor point out in their book, effective models of governance already exist and have made a big difference (Sabel & Victor, 2022). The practice is there, but so far, theory has been lacking. Sabel and Victor focus on climate change, which is the greatest policy challenge of our time, and it is also an area where governance is doubly difficult: regulations have to be not only effective at the national level, but they also have to be negotiated globally among states with different interests and circumstances. They build their argument on the example of the Montreal Protocol on ozone (UN Environment Programme, 2018). First negotiated in 1987, the protocol has been successful at curbing ozone depleting substances (ODS), to the point where the ozone layer is now on course to full recovery.

The ozone layer and climate change challenges looked similar at the outset, with significant scientific and technological uncertainty and considerable differences among the positions of advanced and developing nations. The United Nations Framework Convention on Climate Change of 1992, the first global climate agreement, in fact took the Montreal Protocol as its model. Both global regimes started out as 'thin' regimes, with broad commitments to cut emissions—ozone depleting substances in the first case and greenhouse gases in the second—by a certain date, but otherwise it had little operational content.

But the agreements evolved very differently. The Montreal Protocol made steady progress by bringing firms and governments into collaboration in solving concrete technological problems, while climate change agreements got stalled in global negotiations. Sabel and Victor show that a key difference was the creation of sectoral committees under Montreal, in which ODS-emitting firms joined national regulators and scientists in search for technological alternatives. The groups started small and were few in number, but expanded as knowledge accumulated, actors acquired new capabilities, and parties built trust between each other. The virtue of the sectoral committees was that actual problem solving was devolved to local actors, the firms with the requisite technological know-how. When innovation stalled, targets were reset.

The result was a virtuous loop of on-the-ground innovation and top-level goal setting. In the climate regime, by contrast, firms were kept at arms' length from regulators, for fear that they would control the process. Instead, these entrenched conflicts of interest and resulted in inadequate innovation.

The Montreal Protocol is not the only successful case of what the authors call 'experimental governance'. They discuss in detail a wide range of national and sub-national programmes, ranging from the Advanced Research Projects Agency—Energy (ARPA-E) in the United States to the control of agricultural pollution in Ireland. In each of these cases, governance revolves around ground-level experimentation married to higher-level goal setting. Successful practices that emerge from these collaborations are routinised subsequently through dissemination and standard setting.

These examples are not limited to environmental policy. The operation of ARPA-E is modelled after the Defense Advanced Research Projects Agency (DARPA), a US agency that is responsible for some of the landmark innovations of our time, such as the internet and GPS. At the local level, the most successful initiatives to revitalise communities and create jobs take the form of private-public collaborations that bring training programmes, businesses, non-profit groups, and public officials together to create new pathways to economic opportunity (Fallows & Fallows, 2019). Effective national industrial policies take a similar collaborative, cross-sectoral approach (Ghezzi, 2017). The important point is that there are enough concrete, real-world examples of these collaborative approaches to give us hope that these ideas are not utopian.

As Sabel and Victor explain, the general strategy in all these domains is to start out with ambitious, somewhat ill-defined goals. Programme leaders must acknowledge the deep uncertainty and, hence, the likelihood of mistakes and false starts. There must be incentives for the actors with the most detailed and accurate information—typically firms—to look for solutions, which means public agencies must contribute some combination of sticks (the threat of regulation) and incentives (public inputs). Milestones and monitoring are key to permit reassessment and revision. Solutions are generalised, as they emerge, in the form of standards or regulations for all. Innovation is key, since higher standards (cleaner environment, better jobs) are possible only through productivity-enhancing innovations.

This kind of policymaking differs significantly from the conventional approaches that dominate today's thinking. From my perspective, the state versus market dichotomy no longer makes sense. States and markets are complements, not substitutes. Economists' standard top-down, principal-agent model of regulation (with it top-down, principal-agent framing) becomes unhelpful.

6. A paradigm beyond right and left?

If productivism is to be successful, it will have to transcend the stale ideologies of the past. A new economic paradigm becomes truly established when even its purported opponents start to see the world through its lens. At its height, the Keynesian welfare state received as much support from conservative politicians as it did from left-wing liberals. In the United States, e.g., Republican presidents Dwight Eisenhower and Richard Nixon bought fully into its essential tenets—regulated markets, redistribution, social insurance, and counter-cyclical macroeconomic policies—and worked to expand social welfare programmes and strengthen workplace and environmental regulation (Gerstle, 2022).

It was similar with the neoliberal approach. The impetus for it came from economists and politicians—such as Milton Friedman, Ronald Reagan, and Margaret Thatcher who were all market enthusiasts. But if the paradigm eventually became dominant, it was in no small part thanks to centre-left leaders, such as Bill Clinton and Tony Blair, who had internalised much of its pro-market agenda (Rodrik, 2016). These leaders pushed for deregulation, financialisation, and hyper-globalisation, while paying lip service to ameliorate the consequent rise in inequality and economic insecurity.

As with previous paradigms, productivism will have to find support eventually from both ends of the political spectrum. The polarisation that prevails in our political life makes such an outcome seem outlandish. Yet, there are in fact signs of convergence.

We saw many of these elements in the Biden administration's narrative and in at least some of its policies. The wholesale embrace of industrial policies to facilitate the green transition, rebuild domestic supply chains, and stimulate good jobs, the finger-pointing at corporate profits as a partial culprit behind inflation and the refusal to revoke Trump's tariffs against China are some examples. When the administration's most senior economist, Secretary of Treasury Janet Yellen, extols the virtues of 'friend-shoring'—sourcing supplies from US allies—over the World Trade Organization (WTO), we know we are in a different world (US Department of the Treasury, 2022).

But similar strands exist on the political right as well. Alarmed by China's rise, Republicans have made common cause with Democrats in pushing for active investment and innovation policies to bolster US manufacturing (Franck, 2021). Past (and likely future) Republican presidential candidate Senator Marco Rubio has made impassioned pleas for industrial policy—promoting financial, marketing, and technological assistance to small businesses as well as manufacturing and high-tech sectors (Rubio, 2019, 2021). "In those instances in which the market's most efficient outcome is one that's bad for our people," says Rubio, "what we need is targeted industrial policy to further the common good". Progressives on the left could not agree more. The architect of Trump's China trade policy, Robert Lighthizer, similarly has won many fans on the left for his hard-ball tactics vis-a-vis the WTO. Robert Kuttner, a leading voice among the progressives, has argued that Lighthizer's views on trade, industrial policy, and economic nationalism 'were more those of a progressive Democrat' (Kuttner, 2022).

The Niskanen Center, named after the libertarian economist William Niskanen who was a principal advisor to Reagan, has made 'state capacity', the ability of governments to provide public goods, one of its main planks, emphasising its importance for a healthy economy (Lindsey, 2021). Oren Cass, advisor to 2008 Republican presidential candidate Mitt Romney and a former senior fellow at market-promoting Manhattan Institute, is a critic of financialised capitalism and supports reshoring supply chains and investment in local communities. Patrick Deneen, one of the leading intellectuals of the 'populist right' talks about the importance of 'pro-worker policies' and 'the encouragement, through government policy, of domestic production'. Listening recently to Deneen discussing these and other economic policies, the *New York Times* writer Ezra Klein was moved to say: "What's funny about that to me is that they seem to me to resemble what the current Democratic Party is" (Klein, 2022).

Pragmatism can override political partisanship when it comes to the real work of fostering local businesses and job creation and the public-private partnerships necessary to achieve that end. That was the revelation of the husband-and-wife team of James and Deborah Fallows when they travelled around America on their single-engine plane to study experiences with local economic development (Fallows and Fallows, 2019). Confronted by the challenges of economic decline and joblessness, local politicians were engaged along with community groups, entrepreneurs, and other stakeholders in extensive policy experimentation—and in many of those cases whether they were Democrats, Republicans, or Independents made little difference to what they did.

However, deep divides between the two parties on social and cultural issues, such as abortion rights, race and gender, remain. Many in the Republican Party, including key figures such as Marco Rubio, have yet to give up their allegiance to Donald Trump, who continues to be a threat to US democracy. And there is always the danger that the 'new' industrial policies that conservatives and progressives alike favour will fizzle out or turn into the policies of the past.

Whether it goes astray or not, there are signs of a major reorientation in economic policy—one that is rooted in production, work and localism instead of finance, consumerism and globalism. And it might turn into a new paradigm that captures the imagination of both sides of the political spectrum.

7. Beware economists bearing paradigms

At present, we are in the midst of a transition away from neoliberalism, with much uncertainty about what will replace it. We might approach the absence of a solidified new paradigm with mixed feelings. On the one hand, we certainly do not need yet another orthodoxy offering cookie-cutter solutions and ready-made blueprints for nations and regions with very different circumstances and needs (Rodrik, 2021a). On the other hand, economic policy needs to be guided by an overall animating vision. If history is a guide, the vacuum left by the waning of neoliberal ideas will soon be filled by a new paradigm—and the more appropriate and adaptable that paradigm, the better.

All our previous policy paradigms—whether mercantilist, classical liberal, Keynesian, social-democratic, ordo-liberal or neoliberal—had important blind spots because they were conceived of as universal programmes to be applied at all times and everywhere. Inevitably, the innovations they brought to how we think about economic governance were overshadowed by those blind spots. The result was over-reach and a back-and-forth swing in the pendulum between excessive optimism and pessimism about the role of the government in the economy.

The answer to any policy question in economics is 'it depends'. It may seem this would render economics useless and irrelevant. But in fact, the opposite is true. We need economic analysis and evidence to fill out the details of *what it depends upon*. The keywords of a truly useful economics paradigm are contingency, contextuality, and non-universality. Economics teaches us that there is a time for fiscal profligacy and a time for fiscal conservatism. A time when government should intervene in supply chains and a time when it should leave markets to their own devices. Taxes should be sometimes high, sometimes low. Trade should be freer in some areas and regulated in others. Mapping the links between varying real-world circumstances and the desirability of different types of interventions is what good economics is about.

Our societies are confronted today with vital challenges that require new economic approaches and significant policy experimentation. But those who are looking for a new economic paradigm—or actively trying to develop one—should be careful what they are wishing for. Our goal should be not to create tomorrow's ossified vision, but to learn how to adapt our policies and institutions to changing exigencies. Ultimately, what our economy demands is sound ideas, and not necessarily a new paradigm (Rodrik, 2021b).

By the time any set of ideas becomes conventional wisdom, it is riddled with one-size-fits-all generalisations and truisms that are bound to be unhelpful and misleading as a general orientation to policy. As such, what I have described here as productivism must be understood as a contingent set of policies—a set of policies that at best meets the demand of our time. The more successful it is, the less relevant it will become to future challenges.

References

Acemoglu, D., & Restrepo, P. (2019). The wrong kind of AI? Artificial Intelligence and the future of labor demand. *Cambridge Journal of Regions, Economy and Society*, *13*(1), 25–35. https://www.iza.org/publications/dp/12292/the-wrong-kind-of-ai-artificial-intelligence-and-the-future-of-labor-demand

Bartik, T. J. (2019). *Making sense of incentives: Taming business incentives to promote prosperity*. W. E. Upjohn Institute for Employment Research. https://doi.org/10.17848/9780880996693

Criscuolo, C., Martin, R., Overman, H. G., & Van Reenen, J. (2019). Some causal effects of an industrial policy. *American Economic Review*, *109*(1), 48–85. https://doi.org/10.1257/aer.20160034

Diao, X., Ellis, M., McMillan, M., & Rodrik, D. (2021). *Africa's manufacturing puzzle: Evidence from Tanzanian and Ethiopian firms*. NBER Working Paper, 28344. https://www.nber.org/system/files/working_papers/w28344/w28344.pdf

Doshi, V., Spencer, H., & Rodrik, D. (2023, July). Creating a good-jobs economy in the UK. *The Economy Inquiry 2030*. https://economy2030.resolutionfoundation.org/wp-content/uploads/2023/07/Creating-a-good-jobs-economy-1.pdf

European Commission. (2019). *The European Green Deal*. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en

Fallows, J., & Fallows, D. (2019). *Our towns: A 100,000-mile journey into the heart of America*. Penguin Random House, Pantheon Books.

Franck, T. (2021, June 8). The Senate is set to pass a massive bipartisan tech and manufacturing bill that takes on China. *CNBC*. https://www.cnbc.com/2021/06/08/senate-takes-on-china-with-bipartisan-manufacturing-bill.html

Gerstle, G. (2022). *The rise and fall of the neoliberal order: America and the world in the free market era*. Oxford University Press.

Ghezzi, P. (2017). Mesas ejecutivas in Peru: Lessons for productive development policies. *Global Policy*, *8*(3), 369–380. https://doi.org/10.1111/1758-5899.12457

Klein, E. (2022, May 13). Transcript: Ezra Klein interviews Patrick Deneen. *New York Times*. https://www.nytimes.com/2022/05/13/podcasts/transcript-ezra-klein-interviews-patrick-deneen.html

Kuttner, R. (2022, May 31). After hyper-globalization. *The American Prospect*. https://prospect.org/economy/after-hyper-globalization/

Lincicome, S. (2021). *Industrial policy: A bad idea is back*. CATO Institute. https://www.cato.org/policy-report/july/august-2021/industrial-policy-bad-idea-back

Lindsey, B. (2021). State capacity: What is it, how we lost it, and how to get it back. *Niskanen Center*. https://www.niskanencenter.org/state-capacity-what-is-it-how-we-lost-it-and-how-to-get-it-back/

Moore, S. K. (2022). U.S. passes Landmark Act to fund semiconductor manufacturing. *IEEE Spectrum*. https://spectrum.ieee.org/chips-act-of-2022

Osterman, P. (2020). *Creating good jobs*. MIT Press.

Overly, S. (2017, March 16). This government loan program helped Tesla at a critical time. Trump wants to cut it. *The Washington Post*. https://www.washingtonpost.com/news/innovations/wp/2017/03/16/this-government-loan-program-helped-tesla-at-a-critical-time-trump-wants-to-cut-it/

Rodrik, D. (2016). The abdication of the left. *Project Syndicate*. https://www.project-syndicate.org/commentary/anti-globalization-backlash-from-right-by-dani-rodrik-2016-07

Rodrik, D. (2018). Is populism necessarily bad economics? *AEA Papers and Proceedings*, *108*, 196–199. https://doi.org/10.1257/pandp.20181122

Rodrik, D. (2021a). A better deal for the world's workers. *Project Syndicate*. https://www.project-syndicate.org/commentary/new-deal-for-workers-bargaining-power-good-jobs-by-dani-rodrik-2021-12?barrier=accesspaylog

Rodrik, D. (2021b). Beware economists bearing policy paradigms. *Project Syndicate*. https://www.project-syndicate.org/commentary/economic-policy-must-abandon-universal-paradigms-by-dani-rodrik-2021-05

Rodrik, D. (2022). *An industrial policy for good jobs*. The Hamilton Project Brookings Institution. https://www.brookings.edu/wp-content/uploads/2022/09/20220928_THP_Proposal_Rodrik_GoodJobs.pdf

Rodrik, D., & Sabel, C. (2019). *Building a good jobs economy*. https://drodrik.scholar.harvard.edu/publications/building-good-jobs-economy

Rodrik, D., & Stantcheva, S. (2021a). Fixing capitalism's good jobs problem. *Oxford Review of Economic Policy*, *37*(4), 824–837. https://doi.org/10.1093/oxrep/grab024

Rodrik, D., & Stantcheva, S. (2021b). *Economic inequality and insecurity: Policies for an inclusive economy*. https://drodrik.scholar.harvard.edu/publications/economic-inequality-and-insecurity-policies-inclusive-economy

Rubio, M. (2019). *Rubio releases Chairman's mark to reauthorize the Small Business Act*. U.S. Senate Committee on Small Business Entrepreneurship. https://www.sbc.senate.gov/public/index.cfm/2019/7/rubio-releases-chairman-s-mark-to-reauthorize-the-small-business-act

Rubio, M. (2021). Exclusive: American industrial policy and the rise of China. *The American Mind*. https://americanmind.org/memo/american-industrial-policy-and-the-rise-of-china/

Sabel, C. F., & Victor, D. G. (2022). *Fixing the climate: Strategies for an uncertain world*. Princeton University Press.

Slattery, C., & Zidar, O. (2020). Evaluating state and local business incentives. *Journal of Economic Perspectives*, *34*(2), 90–118. https://doi.org/10.1257/jep.34.2.90

Stephens, J., & Leonnig, C. D. (2011, August 31). Solyndra solar company fails after getting federal loan guarantees. *The Washington Post*. https://www.washingtonpost.com/politics/solyndra-solar-company-fails-after-getting-controversial-federal-loan-guarantees/2011/08/31/gIQAB8IRsJ_story.html

Temin, P. (2017). *The vanishing middle class: Prejudice and power in a dual economy*. MIT Press.

UN Environment Programme. (2018). *About Montreal Protocol*. http://www.unep.org/ozonaction/who-we-are/about-montreal-protocol

US Department of the Treasury. (2022, April 13). *Remarks by secretary of the treasury Janet L. Yellen on way forward for the global economy*. https://home.treasury.gov/news/press-releases/jy0714

Vacas-Soriano, C., & Fernandez-Macias, E. (2017). *Europe's shrinking middle class*. Eurofound. https://www.eurofound.europa.eu/en/commentary-and-analysis/all-content/europes-shrinking-middle-class

PUEB PRESS

Economics and Business Review
https://www.ebr.edu.pl

Vol. 12 (1), 2026: 29-54
https://doi.org/10.18559/ebr.2026.1.2610

Economic complexity and the shadow economy in Africa: An assessment of nonlinearity and asymmetry

James Temitope Dada[1] Folorunsho M. Ajide[2]
Mosab I. Tabash[3] Mamdouh Abdulaziz Saleh Al-Faryan[4]

Abstract

Recent literature reveals that economic complexity (EC) has important implications for shadow economy and has generated mixed empirical conclusions. This paper contributes to this debate by investigating the nonlinear and asymmetric effects of EC on shadow economy in 28 Africa countries between 1995 and 2020. Granger and Yoon's (2002) approach is used to decompose the EC into positive and negative components, while the dynamic panel threshold regression, two-step system generalised method of moments, pooled mean group, augmented mean group, and common correlated mean group are employed as the estimation techniques. The results indicate that positive EC shocks reduce the size of the shadow economy,

Keywords

- economic complexity
- shadow economy
- Africa
- nonlinearity

Suggested citation: Dada, J. T., Ajide, F. M., Tabash, M. I., & Al-Faryan, M. A. S. (2026). Economic complexity and the shadow economy in Africa: An assessment of nonlinearity and asymmetry. *Economics and Business Review*, *12*(1), 29-54. https://doi.org/10.18559/ebr.2026.1.2610

[1] Department of Economics, Obafemi Awolowo University, P.M.B 13 Ile-Ife, Osun State, 220282, Nigeria, corresponding author: jamesdada@oauife.edu.ng, https://orcid.org/0000-0002-8434-6660.

[2] Department of Economics, University of Ilorin, P.M.B 1515, Ilorin, Kwara State, 234031, Nigeria, ajide.fm@unilorin.edu.ng, https://orcid.org/0000-0003-3231-2423.

[3] College of Business, Al Ain University, P.O. Box 64141, Al Ain, United Arab Emirates, mosab.tabash@aau.ac.ae, https://orcid.org/0000-0003-3688-7224.

[4] Saudi Economic Association, P.O. Box 71115, Riyadh-11587, Saudi Arabia, Al-Faryan@hotmail.com, https://orcid.org/0000-0003-1665-807X.

whereas negative shocks contribute to the growth of informality, thus suggesting the presence of asymmetry. The threshold of economic complexity was found to be 0.41.

JEL codes: O3, O17, O55.

Article received 29 September 2025, accepted 15 January 2026.

Introduction

The proponents of economic complexity argue that a nation's wealth is based on the amount of productive knowledge and the cumulative know-how utilised in product diversity. They explain that the country's wealth is not just based on its resources alone, but productive knowledge and how it is transformed in the production process. Their framework by focuses on a country's productive capabilities rather than the size of GDP (Hausmann et al., 2014), thus offering a nuanced understanding of economic growth compared to the conventional models). Economic complexity is a measure of a country's overall productive capabilities that emerges from the diversity and complexity reflected in its export product basket. An economy with more diverse skills, knowledge and innovative infrastructural facilities is capable of producing more complex goods and services (Ajide, 2022). In the literature, economic complexity has been linked to the shadow economy due to the high level of innovativeness and skills required in the production process for increasing the export product basket (Canh & Thanh, 2020).

Economic complexity is expected to reduce the shadow economy by bringing numerous opportunities into the official economy, including job openings, entrepreneurial development, and other benefits that can be leveraged from a nation's international competitiveness (Canh & Thanh, 2020; Nguyen, 2022). However, it may intensify the shadow economy due to undesirable economic shocks resulting from an over-reliance on external economies and financial crises (Nguyen, 2022). In this case, economic actors view the informal sector as a safety net. Also, the nexus between shadow economy and economic complexity may be an inverted N-shaped or a U-shaped pattern, due to the interaction of the structural, institutional, and economic transformation as well as the stages of economic development (Bolarinwa & Simatele, 2025; Fotié & Mbratana, 2024; Lewis, 1954). Drawing on the dualistic view, it can be hypothesised that at the early stage of materialised knowledge in the national productive structure (economic complexity), the shadow economy tends to decline as labour moves from subsistence economic activities to a higher level of economic sophistication. Increased economic complexity provides workers with good wages and job security. Industrialisation and a knowledge-based economy therefore boost opportunities in the formal economy.

Moreover, higher economic complexity is associated with better institutional quality (Dada et al., 2025), which can eventually reduce the size of the shadow economy (Ketu et al., 2024). Higher economic complexity can be accompanied by increased taxes, labour law rigidity, and other regulatory burdens. The excessive costs of formalisation over the benefits of operating in the official economy may lead economic agents to consider the shadow sector as the best alternative. These scenarios may in turn generate nonlinear reactions in the shadow economy, suggesting context-dependent and threshold sensitivity (Bolarinwa & Simatele, 2025). Interestingly, the nonlinear impact of economic factors on the shadow economy has been explored in the recent literature (Ajide & Dada, 2024; Bolarinwa & Simatele, 2023), providing insights into appropriate policy responses to this phenomenon in developing nations (Wu & Schneider, 2019; Yu & Ohnsorge, 2019). Different studies have argued that the shadow economy distorts formal economic operations, thereby hindering the ability of government to achieve her objectives (Dell'Anno, 2022). Other authors suggest that the shadow economy protects marginalised individuals who experience income disparities, uneducated agents, and women entrepreneurs, who find it challenging to access justice in formal employment and entrepreneurial opportunities (Ajide et al., 2025; Medina & Schneider, 2019). This issue was also predicted theoretically in Lewis's (1954) dualistic ideology; Lewis argues that informal and formal economic systems may coexist in developing regions due to unbalanced features in traditional and modern economic structures, lapses in institutional settings, inequalities, and stringent regulations in the official economy (Ajide et al., 2025).

Given that many theoretical and empirical studies support the coexistence of formal and informal sectors in developing nations, the new empirical focus has shifted to ascertaining the appropriate balance or threshold for the two economic systems (see Ajide & Dada, 2024; Bolarinwa & Simatele, 2023; Canh & Thanh, 2020; Wu & Schneider, 2019; Yu & Ohnsorge, 2019). In this paper, we present the case of the nonlinear and asymmetric effect of economic complexity on the shadow economy in selected Africa countries. This constitutes our primary contribution to the literature. The recent empirical debate on the influence of economic complexity on the shadow-economy activities has yielded mixed results (Ajide & Dada, 2024; Ketu et al., 2024; Nguyen, 2022). Specifically, Nguyen (2022) finds that economic complexity reduces the shadow economy in a global sample; this is consistent with the empirical findings of Ketu et al. (2024), who also employed the generalised method of moments. However, the mean group panel estimators employed by Ajide and Dada (2024) indicate that in the African setting, the impact of economic complexity on the shadow economy is heterogeneous, showcasing a mix of positive and negative effects, as well as nonsignificant results in the country-specific analysis. One of the main unresolved issues is the nonlinear effect of economic complexity, as observed in previous studies. The present study bridges this gap in the literature. Our objective is to examine the nonlinear impact of economic complexity on the shadow economy in Africa.

Africa serves as an interesting region to investigate the nonlinearity between economic complexity and the shadow economy. The magnitude of Africa's shadow economy has been estimated to exceed the world average of 31% of GDP (Elgin et al., 2021; Medina & Schneider, 2019). Likewise, the economic complexity index suggests that African economies operate at lower levels (Olaniyi & Odhiambo, 2023). This is due to a technology gap, over-reliance on unrefined natural resources, for foreign exchange, and production based on less sophisticated products (Dada et al., 2024b; Mealy et al., 2019). Based on this, our study offers new insights into the relationship between economic complexity and the shadow economy, and proposes that the effect of economic complexity on this economy follows a threshold-based pattern. This suggests that the shadow economy can be curbed if economic complexity reaches a particular threshold, but may harm the system if the institutional structure is not effectively tightened.

To examine this, the study employs the Granger and Yoon (2002) decomposition framework, splitting the changes in economic complexity index into positive and negative shocks to capture the asymmetric relationship as expanded by Habib et al. (2017). It also employs a two-step system generalised method of moments, pooled mean group, augmented mean group (AMG), and common correlated effects mean group (CCEMG) within the asymmetric framework over panel data of 28 African countries between 1995 and 2020. The study found the threshold value of economic complexity to be 0.41. Below the threshold level, economic complexity has a significantly positive impact on the shadow economy, while above the threshold, its impact is significantly negative. In addition, positive economic complexity shocks mitigate the size of the shadow economy. However, a negative economic complexity shock contributes to the growth of the shadow economy. This result establishes the presence of asymmetries: the shadow economy responds differently to positive and negative shocks in terms of economic complexity.

The paper is structured as follows: Section 1 reports the theories and relevant empirical studies; Section 2 discusses the materials and methods used to conduct the study; in Section 3, we present the empirical results and interpretation; in the final section, we discuss the conclusion and policy implications of the findings.

1. Literature review

The study draws on the neoclassical approach and the dualistic theory of development to explain the interrelationship between the shadow economy and economic complexity. The neoclassicists clarify that the choice to participate in the shadow economy is guided by welfare-optimisation principles as perceived by economic agents (Dell'Anno, 2022). This view seems to rely heavily on how economic agents perceive costs and the benefits of informality. In a more refined approach, authors like Lewis (1954) and Tokman (1979) view the shadow economy as performing a positive transitional role by absorbing marginalised groups in society and, at the same time, reducing income inequality, since this sector provides additional income for survival. This view is consistent with the early stage of economic progress but may become ineffective as higher levels of growth and economic sophistication are experienced. Furthermore, the theory of institutional economics explains that the legal system and strong governance system can shape the shadow economy's roles as the economy matures, implying that the relationship between the shadow economy and economic complexity may be threshold-sensitive (Bolarinwa & Simatele, 2025; Dada et al., 2025; La Porta & Shleifer, 2008; North, 1990).

Institutional and socioeconomic factors are the principal drivers of the decision to participate in the informal sector. For instance, income inequality promotes informality, particularly in regions with significant income disparities (Ajide et al., 2024). This is because the informal sector may offer opportunities to bridge income gaps in the official economy. Economic agents may participate in the informal economy to escape restrictive institutional factors (Dell'Anno, 2022). The shadow economy may increase due to the prevalence of corruption. Weak control over corruption may allow businesses to bribe their way through informal channels, thereby circumventing stringent regulatory policies (Schneider & Buehn, 2018).

Recent studies have revealed that economic complexity downgrades the magnitude of the shadow economy (Ajide & Dada, 2024; Ketu et al., 2024; Nguyen, 2022; Nguyen & Thanh, 2020). Economic complexity captures many factors that are left out in the computation of GDP, including the productive knowledge that materialises and is employed in the production process (Hidalgo, 2021). Enhanced economic complexity is expected to reduce informality (Nguyen, 2022). Economic complexity offers further benefits to economic agents in the formal economic system, bolstering the production structure and discouraging operations in the informal sector (Nguyen, 2022). Informality may be discouraged because economic complexity enhances the functioning of the financial sector, thereby increasing access to funding for advanced technologies within the official economic system (Ajide & Osinubi, 2024; Canh & Thanh, 2020).

For instance, Nguyen (2022) examines the influence of economic complexity across a panel of 115 countries from 1995 to 2017, finding that economic complexity reduces the shadow economy in both high- and low-income countries. Similarly, Canh et al. (2021) find that economic integration, institutional factors, and economic complexity are key drivers of the shadow economy in 112 economies. Most importantly, economic complexity reduces informality. Canh and Thanh (2020) reveal a nonlinear relationship between export diversification and the informal economy in 116 economies.

Similarly, Ketu et al. (2024) found that within 24 countries from 1996 to 2016, economic complexity was associated with a reduction in the shadow economy. However, this finding is inconsistent with the study by Ajide and Dada (2024), which found no correlation between the variables using different mean-group estimators, such as the augmented mean-group and the dynamic common correlated effect mean-group, over a panel data set of 27 countries from 1995 to 2017. In a country-specific estimation, economic complexity and shadow economy were found to be negatively related in Congo, Uganda, and Ghana, while the opposite was documented for Tunisia, Madagascar, and Botswana. No significant effects were reported for other countries. This study points to the heterogeneous nature of African economies.

Furthermore, in panel data for 99 countries, Pham et al. (2024) established that economic complexity is non-linearly linked to income inequality and the informal economy: there is a U-shaped association between economic complexity and income disparities. Therefore, income disparities and economic complexity must be taken into account when addressing the shadow economy. This aligns with the study by Ajide et al. (2025), which suggests that the informal sector helps address income inequality by enabling marginalised individuals to bridge the income gaps created in the official economy.

In conclusion, only a few studies have investigated the influence of economic complexity on shadow economy, emphasising linearity among the variables. However, no study has examined the nonlinearity between shadow economy and economic complexity. We aim to address this gap with the use of a panel of African countries. Recent studies have highlighted the importance of examining the nonlinearity between the shadow economy and socioeconomic factors to inform effective policy responses (Ajide & Dada, 2024; Bolarinwa & Simatele, 2023; Wu & Schneider, 2019; Yu & Ohnsorge, 2019).

The literature suggests that economic complexity influences shadow economy through direct and indirect channels (Ajide & Dada, 2024; Canh et al., 2021; Canh & Thanh, 2020). Hidalgo and Hausmann (2009) show that income and economic complexity are correlated. Specifically, while economic complexity stimulates growth in the official economy, a deviation from this connection may lead to an expansion of the informal sector. Such a deviation may indicate a potentially positive impact on long-term growth. However, a short-term impact may lead households to move into the shadow economy to compensate for income shortfalls (Ajide & Dada, 2024; Ketu et al., 2024). Research has shown that output volatility reduces income per capita and factor productivity, negatively impacting economic growth (Ramey & Ramey, 1991). During this period, the shadow economy may serve as a survival channel for households to cope with the challenges of such volatility, especially in developing economies, where financial and governance systems are very weak (Ajide et al., 2024). Therefore, a complex economy is believed to have more paths to growth stability, as studies confirm that economic complexity reduces volatility in national output and increases labour productivity (Sweet & Eterovic, 2019). In this case, economic complexity is expected to prevent a potential shift in the shadow economy (Ketu et al., 2024).

Various studies indicate that a weak institutional system is a primary source of the shadow economy's growth (Dell'Anno, 2022; Elgin et al., 2021). In addition, formal and informal economies share some key determinants; therefore, economic complexity is expected to influence the shadow economy by attracting economic actors to the official economy in the presence of innovations and new knowledge being disseminated and reflected in national export baskets (Ajide, 2022; Canh & Thanh, 2020; Nguyen, 2022; Schneider & Buehn, 2018). However, entrepreneurial firms and informal workers may not fully benefit from the opportunities presented by economic complexity because of the semi- or lower-level skills they possess; thus, they may still prefer informal-sector operations (Nguyen, 2022).

Djeunankan et al. (2023) explain that economic complexity enhances governance and institutional reforms by influencing foreign direct investors, who demand greater protection of property rights. Economic complexity promotes quality institutions through a skilled workforce and human capital development, thereby reducing income inequality and enhancing the lives of poor households while also reducing the participation of shadow economic actors (Dada et al., 2024b). Economic complexity improves transparency and leaders' accountability, an ingredient for discouraging participation in shadow economy. However, a country with a skewed industrial structure may have higher cohesion among the elites, and an overreliance on elites as a means of survival may reduce accountability (Olaniyi & Odhiambo, 2023). With a highly skewed industrial and productive structure, citizens may lose economic privileges, leading to increased income inequality and expanding the shadow economy.

Furthermore, knowledge is an essential input in the production process and is embedded in the economy's productive structures. Economic complexity is closely tied to knowledge accumulation, as reflected in the diversity and ubiquity of the economy's export baskets (Hartmann, 2014). Therefore, progress in product complexity is closely tied to knowledge and technological transformation, as well as their application in the economy. This process creates more economic opportunities for average citizens and reduces the attractiveness of the shadow economy (Ajide, 2022; Canh et al., 2021; Hartmann et al., 2017). Moreover, in the formal economy, workers enjoy higher wages, greater job security, and greater specialisation due to the economic transformation enabled by economic complexity and its associated higher returns to scale (Lee & Vu, 2020). Consequently, enhanced job quality and security help reduce marginalisation and income inequality, thereby diminishing reliance on the informal sector for survival.

However, based on the view of the routine-biased technological hypothesis and skill-based technical change theory, economic transformation resulted from product sophistication, and higher-skilled labour demand may increase with the use of sophisticated machines. Therefore, the demand for low-skilled labour will be reduced (Acemoglu & Autor, 2011). In turn, this could increase the level of marginalisation and deepen the shadow economy. These positive and negative forces in the relationship between economic complexity and the shadow economy suggest the possibility of nonlinearity between the two. While studies such as Bandeira Morais et al. (2021) and Le et al. (2020) provide empirical evidence of an inverted U-shaped relationship between income gaps and economic complexity, to our knowledge, the asymmetric effect of economic complexity on the shadow economy has yet to be explored empirically. The focus of our study is to shed light on this issue in the context of African economies.

2. Data, model, and estimation strategies

2.1. Data

The dataset for this study spans from 1995 to 2020 for 28 African countries. Table A1 in the Appendix lists the countries. The study aims to investigate the asymmetric and nonlinear impact of economic complexity on the shadow economy. The shadow economy data are sourced from the dataset of Elgin et al. (2021). This data comprises all economic activities (legal or illegal), regardless of official status. Economic complexity data are extracted from the Atlas of Economic Complexity dataset, which assesses the nation's export competitiveness and the sophistication of its production system (Ajide & Dada, 2024; Hausmann et al., 2014).

To avoid the omitted variable problem, additional control variables are added, including per capita income as a proxy for the official economy, population, and domestic credit to the private sector as a proxy for financial development. Data on per capita income, financial development, and population growth are obtained from the World Bank's World Development Indicators. Per capita income is measured in 2015 constant USD and is incorporated because the formal economy plays a role in determining the magnitude of the shadow economy (Ajide et al., 2024, 2025; Dada et al., 2024a,b).

2.2. Model specification

This study follows the model specification used in existing studies to examine the relationship between economic complexity and the shadow economy in African economies. The choice of control variables is based on the previous literature. The dualist approach suggests that the shadow economy exists alongside the official economy (Lewis, 1954), providing an alternative for marginalised households (Loayza, 2018). A formal economy, typically measured by economic growth or GDP, is expected to have an inverse relationship with the shadow economy (Berdiev et al., 2018; Dada & Al-Faryan, 2024). Studies by Awoleye et al. (2025), Njangang et al. (2020), and Capasso and Jappelli (2013) show that financial services influence the size of the shadow economy; therefore, financial development is included as a control variable. Similarly, Kpognon (2022) and Njangang et al. (2020) suggest that population growth drives the expansion of the shadow economy, especially in developing countries where formal-sector unemployment is high. In specific terms, and adding the control variables to the model, equation 1 can be expressed as:

$$SHE_{it} = \alpha + \beta\, EC_{it} + \gamma\, GDP_{it} + \delta\, POP_{it} + \theta\, FD_{it} + \varepsilon_{it} \tag{1}$$

SHE represents the shadow economy, EC represents economic complexity, and GDP, POP, and FD are per capita income, population, and financial development, respectively. i and t are the cross-section and time indicators, respectively.

To capture the asymmetric relationship, we apply the Granger–Yoon (2002) approach to decompose the economic complexity index into positive and negative shocks. This approach was later expanded by Habib et al. (2017). The process of decomposing the variable is as follows.

$$EC_{it} = EC_{it-1} + \varepsilon_{it} = EC_0 + \sum_{j=1}^{p} \varepsilon_{itj} \tag{2}$$

where $t = 1, 2, 3, \ldots T$, $i = 1, 2, 3, \ldots I$, the term EC_0 is the initial value of the series, whereas ε_{it} is the noise. The negative and positive components (asymmetries) are detailed as:

$$\varepsilon^{+}_{it} = \max(\varepsilon, 0) \tag{3}$$

$$\varepsilon^{-}_{it} = \min(\varepsilon, 0) \tag{4}$$

Hence,
$$\varepsilon_{it} = \varepsilon^{+}_{it} + \varepsilon^{-}_{it} \tag{5}$$

Putting equations (3) and (4) in (2), then:

$$EC_{it} = EC_{it-1} + \varepsilon_{it} = EC_{0} + \sum_{j=1}^{p} \varepsilon^{+}_{itj} + \sum_{j=1}^{p} \varepsilon^{-}_{itj} \tag{6}$$

Finally, the positive economic complexity shocks and negative economic complexity shocks is described as a cumulative:

$$EC^{+}_{it} = \sum_{j=1}^{p} \varepsilon^{+}_{itj} \tag{7}$$

$$EC^{-}_{it} = \sum_{j=1}^{p} \varepsilon^{-}_{itj} \tag{8}$$

Incorporating equations (7) and (8) into equation (1), the testable model becomes:

$$SHE_{it} = \alpha + \rho\, EC^{+}_{it} + \pi\, EC^{-}_{it} + \gamma\, GDP_{it} + \delta\, POP_{it} + \theta\, FD_{it} + \varepsilon_{it} \tag{9}$$

2.3. Estimation procedures

Firstly, the study tests the econometric characteristics of the variables. The cross-sectional dependency (CD) in the data is investigated. The financialisation and globalisation of markets have rendered the assumption of country independence unrealistic. Furthermore, African countries are not immune to global events and are thus susceptible to shocks emanating from both within and outside the region. Thus, it is essential to account for the likelihood of CD in the data to guide the estimation methods. The CD in the data is tested using the Pesaran-scaled Lagrange Multiplier (Ps-LM), the Bias-corrected scaled Lagrange Multiplier (Bcs-LM), and the Pesaran CD (P-CD).

Secondly, the study employs first- and second-generation estimators to examine the asymmetric impact of economic complexity on the SHE. Specifically, the study uses the two-step system generalised method of moments (2SGMM), pooled mean group (PMG), augmented mean group (AMG), and common correlated effects mean group (CCEMG) within the asymmetric framework. The 2SGMM addresses the endogeneity and serial correlation commonly found in the literature on the shadow economy (Ajide et al., 2025). Additionally, the

2SGMM addresses the omitted variable bias and simultaneity issues inherent in the shadow economy-economic complexity model.

To estimate the 2SGMM, the following linear restriction is imposed on equation (9) and used as the instruments.

$$E(SHE_{i,t-s}, \Delta\varepsilon_t) = E(EC^{+}_{i,t-s}, \Delta\varepsilon_t) = E(EC^{-}_{i,t-s}, \Delta\varepsilon_t) = \\ = E\left(GDP_{i,t-s}, \Delta\varepsilon_t\right) = E\left(POP_{i,t-s}, \Delta\varepsilon_t\right) = E\left(FD_{i,t-s}, \Delta\varepsilon_t\right) = 0 \qquad (10)$$

For $s \geq 2$, and $t = 3, 4, \ldots, T$.

On the other hand, the PMG, AMG, and CCEMG account for cross-sectional dependence, spatial correlation, and slope heterogeneity, which are commonly found in heterogeneous panels (Dada et al., 2024b). Furthermore, we employed standard errors that are robust to heteroscedasticity, autocorrelation, and spatial or cross-sectional dependence. In addition, the PMG, AMG, and CCEMG do not require stationarity and cointegration tests before being utilised (Pesaran, 2006). To save space, the models for these estimation techniques are not specified.

Thirdly, the dynamic panel threshold regression of Seo et al. (2019) is used to estimate the threshold of economic complexity in determining the size of the shadow economy. The threshold model is as follows:

$$y_{it} = (1, x'_{it})\,\beta_1\,1(q_{it} \leq \gamma) + (1, x'_{it})\,\beta_2\,1(q_{it} > \gamma) + \mu_i + \varepsilon_{it} \\ i = 1, 2, \ldots, n; \qquad t = 1, 2, \ldots, T \qquad (11)$$

where y_{it} is the scalar of the dependent variable (shadow economy), x'_{it} is the vector of time-varying independent variables that also include the lag value of the dependent variable, 1( ) is the indicator function, q is the transition variable (economic complexity), γ is the threshold parameter, β_1 and β_2 are the slope parameters linked with different periods, μ_i is a time-invariant unobserved country-specific effect term, and ε_{it} is a zero-mean idiosyncratic random disturbance.

3. Results and discussion

Table 1 presents descriptive statistics and a correlation matrix. Applying four different estimation techniques, as reported in Table A2 in the Appendix, the null hypothesis of cross-sectional independence in the variable is rejected at the conventional level. These results point to the presence of CD across all variables used. This result reveals that any factors that cause a change in one nation will also influence other countries in the group. Hence, there are spill-over and contagion effects within the panel.

Table 1. Descriptive statistics and correlation matrix

	SHE	EC^{+}	EC^{-}	GDP	POP	FD
Mean	40.230	1.006	–0.978	2246.103	2.786	25.728
Median	39.795	0.796	–0.738	1262.221	2.878	15.565
Maximum	64.586	3.696	0.000	11937.640	5.842	160.125
Minimum	26.250	0.000	–3.373	43.231	0.595	0.429
Standard deviation	7.953	0.844	0.829	2367.782	8.924	28.580
SHE	1	0.055[b] (0.045)	0.015[c] (0.061)	–0.213[a] (0.000)	0.043[a] (0.0257)	–0.414[a] (0.000)
EC^{+}		1	–0.914[a] (0.000)	0.186[a] (0.000)	0.011[a] (0.775)	–0.121[a] (0.001)
EC^{-}			1	–0.156[a] (0.000)	0.093[a] (0.013)	0.123[a] (0.001)
GDP				1	–0.297[a] (0.000)	0.453[a] (0.000)
POP					1	0.013[a] (0.723)
FD						1

Note: p-values in parentheses ( ). a, b and c suggest 1%, 5% and 10% levels of significance.

Source: own work.

Due to the presence of the CD, the unit root is examined using CADF and CIPS tests. The result of the stationarity test is presented in Table A3 in the Appendix. All variables are stationary at the first difference (I(1)) at the 1% significance level. The long-run cointegration among the variables is investigated using the Westerlund panel cointegration test, and the result is reported in Table A4 in the Appendix. The result suggests that two of the four test statistics are significant (Gt and Pt), suggesting that long-run cointegration exists among the variables. Furthermore, the result implies that series tend to cointegrate and move to a long-run steady state, even after an initial short-run disequilibrium. Based on this outcome, the study accounted for the long-run relationship using the PMG, AMG, and CCEMG estimation techniques.

3.1. Baseline results

This paper investigates the asymmetric effect of economic complexity on the size of shadow economy in selected African countries. Four different estimation techniques (2SGMM, PMG, AMG, and CCEMG) are employed to establish the relationship, and the results are presented in Table 2. The results of the 2SGMM confirm that the instrument used, as indicated by the Hansen-J statistic test, is valid and free from serial correlation. Furthermore, the 2SGMM reveals the presence of first-order autocorrelation (AR1) and the absence of second-order autocorrelation (AR2). The lagged value of the dependent variable is significantly positive at the conventional level. This reveals that the past value of the shadow economy increases the magnitude of the present value.

Regarding the asymmetric components, this study builds upon the works of Ahmed et al. (2021), Altintas and Kassouri (2020), and Arnaut et al. (2023) in interpreting the coefficients. The positive shock (EC^+) is interpreted directly; conversely, the negative shock (EC^-) is interpreted with the coefficients reversed. Considering this, Table 4 suggests that positive economic complexity (EC^+) has a significantly negative effect on shadow economy, whereas the effect is not significant in the AMG model. Specifically, the results indicate that a unit increase in positive economic complexity shocks reduces the shadow economy by 0.116, 0.054, and 0.120 in the 2SGMM, PMG, and CCEMG models, respectively. This indicates that the growth in the sophistication and complexity of the economic structure helps curb the extent of the shadow economy. Furthermore, producing high-quality products makes activities in the shadow economy less attractive and reduces its size.

Meanwhile, the negative shock of economic complexity (EC^-) has a significant influence on the size of the shadow economy in the region in the 2SGMM, AMG, and CCEMG models. Specifically, the coefficients indicate that a unit increase in the negative economic complexity shock (EC^-) increases the size of the shadow economy on average by 0.205, 0.985, and 0.201 in the 2SGMM, AMG, and CCEMG models, respectively. The presence of asymmetry in economic complexity is further examined using the Wald coefficient test. The test results in Table 2 indicate asymmetry in the relationships across all models. Specifically, this indicates that the relationship's assumptions of linearity and symmetry may be too restrictive and do not accurately reflect the realities of the nexus between economic complexity and the shadow economy.

Correspondingly, economic growth, a measure of the official or formal economy, has a significant negative effect on the shadow economy in all models. In contrast population growth rate and financial development have a significantly positive influence on the size of the shadow economy in the long run in selected African countries. This result shows that the shadow economy provides alternative to the growing population in terms of employment opportunities. Furthermore, the result reveals that the financial architecture of African countries supports the growth of informal economy through the diversion of fund. Furthermore, the institutional lapses allow credit to be channelled from the official economy to the unofficial economy. This subsequently spurs the growth of the shadow economy.

Table 2. Baseline regression results

	2SGMM	PMG	AMG	CCEMG
SHE(–1)	0.852[a] (0.000)			
EC^{+}	–0.111[b] (0.031)	–0.054[a] (0.000)	–0.097 (0.620)	–0.120[b] (0.014)
EC^{-}	–0.205[a] (0.002)	–0.976 (0.255)	–0.985[a] (0.000)	–0.201[a] (0.002)
GDP	–0.002[b] (0.046)	–0.001[a] (0.000)	–0.001[a] (0.000)	–0.004[b] (0.023)
POP	0.016[b] (0.037)	0.053 (0.000)	0.065[a] (0.000)	–0.139 (0.618)
FD	0.017[a] (0.000)	–0.168 (0.000)	0.021[a] (0.000)	–0.114 (0.248)
Constant		–3.556[a] (0.000)	8.295[a] (0.000)	0.038[a] (0.000)
Wald test (asymmetric components)	–4.065[a] (0.000)	38.71[a] (0.000)	7.31[a] (0.007)	14.88[a] (0.000)
AR(1)	–2.402[b] (0.016)			
AR(2)	–0.173 (0.862)			
J-statistic test	23.649 (0.482)			
RMSE			1.813	6.41

Note: a, b, and c suggest 1%, 5% and 10% levels of significance. First-differenced variables in used in 2SGMM.

Source: own work.

3.2. Robustness check: Alternative measures of shadow economy

The robustness check verifies the estimates from Table 2 using additional proxies for the shadow economy. The study uses two alternative sources of data. Firstly, the shadow economy data from Medina and Schneider (2019), based on the Multiple Indicators Multiple Causes framework, are reported in Table 3. Second, shadow economy data based on a dynamic general equilibrium model by Elgin et al. (2021) are also applied, and the results are presented in Table 4. The findings from Table 3 reveal that positive shocks in economic complexity (EC^{+}) significantly reduce the size of the shadow economy, whereas negative shocks in economic complexity (EC^-) have the opposite impact. Similarly, the result in Table 4 (DGE proxy) also aligns with the main findings presented in Table 2. These results from the alternative proxies establish the presence of an asymmetric effect of economic complexity on the size of the shadow economy in Africa. For other control variables, economic growth substantially reduces the size of the informal economy, while financial development positively influences the growth of the shadow economy. However, population growth has a mixed impact on the size of shadow economy. In summary, the results from the alternative long-run estimates are consistent with those from Table 2 (baseline results), with only slight variations in the size and significance of the coefficient.

Table 3. Regression results for alternative proxy of shadow economy MIMIC from Medina and Schneider (2019)

	2SGMM	PMG	AMG	CCEMG
SHE(–1)	0.863[a] (0.000)			
EC^+	–0.828[a] (0.000)	–0.771[c] (0.064)	–0.132 (0.634)	–0.574[b] (0.026)
EC^-	0.472 (0.113)	–1.464 (0.220)	–0.847[a] (0.000)	0.091 (0.055)
GDP	–0.002[c] (0.072)	–0.002[a] (0.000)	–0.001[a] (0.000)	–0.044[b] (0.033)
POP	–0.002 (0.945)	0.074[a] (0.000)	–0.091[a] (0.000)	0.313 (0.304)
FD	0.085[a] (0.000)	0.146[a] (0.000)	0.041[a] (0.000)	–0.159 (0.461)
Constant		–4.440[a] (0.000)	7.519[a] (0.000)	6.992[a] (0.000)
Wald test (asymmetric components)	–4.961[a] (0.000)	0.92 (0.339)	5.04[b] (0.025)	0.929 (0.761)
AR(1)	–3.498[a] (0.000)			
AR(2)	–1.633 (0.102)			
J-statistic test	25.577 (0.482)			
RMSE			2.449	7.56

Note: a, b, and c suggest 1%, 5% and 10% levels of significance.

Source: own work.

Table 4. Regression results for alternative proxy of shadow economy DGE from Elgin et al. (2021)

	2SGMM	PMG	AMG	CCEMG
SHE(–1)	0.942[a] (0.000)			
EC^+	–0.100[c] (0.094)	–0.681[a] (0.000)	–0.271[c] (0.098)	–0.077 (0.872)
EC^-	–0.255[a] (0.000)	–0.596[b] (0.027)	–0.024[b] (0.031)	0.354 (0.975)
GDP	–0.002[a] (0.005)	–0.001[c] (0.067)	–0.003[b] (0.042)	–0.001[b] (0.028)
POP	–0.001 (0.992)	–0.056[b] (0.034)	–0.206 (0.625)	0.287 (0.977)
FD	–0.001 (0.201)	0.201[a] (0.000)	0.010 (0.585)	–0.011 (0.955)
Constant		1.757 (0.155)	7.349[a] (0.000)	4.641 (0.796)
Wald test (asymmetric components)	–1.906[c] (0.057)	4.329[b] (0.031)	6.502[b] (0.025)	1.02 (0.783)
AR(1)	–3.964[a] (0.000)			
AR(2)	1.898 (0.102)			
J-statistic test	19.048 (0.518)			
RMSE			0.288	0.30

Note: a, b, and c suggest 1%, 5% and 10% levels of significance.

Source: own work.

3.3. Threshold effect

This section examines the threshold effect of economic complexity on shadow economy using the dynamic panel threshold approach proposed by Seo et al. (2019). The result of the threshold analysis is presented in Table 5. The results reveal that when economic complexity is below the threshold level of 0.41, it contributes significantly to the growth of the informal economy. Conversely, at the upper threshold of 0.41, economic complexity reduces the size of the shadow economy. This result provides evidence of nonlinearity in the relationship between economic complexity and the shadow economy.

Regarding other variables, the lagged shadow economy has a significant positive impact on itself at the lower regime, indicating that the historical value of the shadow economy has a substantial influence on its current value. Furthermore, the non-significant impact of the lagged variable in the upper regime highlights the importance of economic sophistication in mitigating the shadow economy. Economic growth, as a proxy for the official economy, has a significantly negative impact on the shadow economy in both regimes. It shows that the shadow economy and the official economy are inversely related. On the other hand, financial development has a significant positive effect on the shadow economy at the upper regime, whereas population growth has no significant impact in any of the regimes.

Table 5. Threshold estimate of economic complexity using dynamic panel threshold (dependent variable: SHE)

r	**0.41[a] (0.000)**	
	Low (<0.41)	**High (>0.41)**
SHE(–1)	0.460[a] (0.000)	0.022 (0.959)
EC	0.479[a] (0.043)	–0.925[a] (0.009)
GDP	–0.001[b] (0.032)	–0.002[b] (0.025)
POP	0.318 (0.124)	–0.046 (0.109)
FD	–2.903 (0.172)	5.282[a] (0.002)
Constant	5.953 (0.49)	

Note: a and b are 1% and 5% levels of significance.

Source: own work.

3.4. Discussion

This study found evidence of asymmetric and nonlinear relationships between the shadow economy and economic complexity in 28 African countries. The results consistently indicate that positive shocks to economic complexity decrease the size of the shadow economy. Conversely, negative shocks in economic complexity contribute to the growth of shadow economy. Furthermore, technological sophistication and knowledge-based productive capacities create opportunities and expand exports, thereby affecting the degree of trade openness and output in the official market, and consequently reducing the motivation to engage in the shadow economy (Hidalgo, 2021).

Studies have revealed that increased economic complexity facilitates foreign direct investment flows, expanding the official economy's opportunities (Boleti et al., 2021; Osinubi & Ajide, 2022). Highly complex economies foster skilled, knowledge-based job places, allowing those in the shadow economy to migrate into the main economy. To be more precise, this result from the asymmetric relationship indicates that increasing economic complexity is a crucial tool for achieving sustainable development, thereby strengthening human capital, which in turn leads to a reduction in income inequality and the presence of the shadow economy (Dada et al., 2024b; Djeunankan et al., 2023; Lee & Vu, 2020). This result does not stand alone, however, as it is comparable to the empirical studies by Nguyen (2022), Ajide (2022), and Ketu et al. (2024), who found that economic complexity serves as an important policy tool for limiting the growth of the shadow economy. Studies have also documented the adverse effect of a fall in economic complexity on the growth of the informal economy. For instance, Ajide and Dada (2024) found that the effect of economic complexity on the shadow economy in Africa is heterogeneous, with economic complexity positively influencing the shadow economy in countries such as Botswana, Madagascar, and Tunisia.

The threshold regression buttresses the result obtained from the asymmetric analysis. The results show that economic complexity reduces the shadow economy at the upper threshold regime but exacerbates it at the lower threshold regime. Studies have documented the threshold level of economic complexity in influencing socioeconomic variables. For example, Azami et al. (2024) found a threshold of economic complexity of 1.12 in the relationship between economic complexity and ecological footprint across selected Asian countries. Similarly, Azizi (2020) identified a threshold of –1.15 in the relationship between economic complexity and energy consumption in Iran. Additionally, Elazhary et al. (2024) found the threshold of economic complexity to be -0.301 in the relationship between foreign direct investment and industry value-added in Egypt. This finding corresponds with a few studies that have noted a nonlinear relationship between the shadow economy and other economic factors (Ajide & Dada, 2024; Wu & Schneider, 2019).

Economic growth, as a proxy for the formal economy, has an inverse relationship with the shadow economy. This result aligns with the neoclassical, dualist, and voluntarist theories. Such an outcome supports various empirical findings on the relationship between the shadow economy and the official economy. For example, the study by Chen (2012) found that economic growth reduces the extent of the informal economy.

Population growth significantly increases the size of the shadow economy. This economy provides an escape route for those who cannot work in the official economy due to unemployment or limited opportunities. This result aligns with the findings of Dada et al. (2024a), Kpognon (2022), and Njangang et al. (2020). Lastly, the finding shows that financial development contributes to the growth of the shadow economy. This reveals that loopholes and lapses in the financial sector culminate in opportunistic behaviour among economic agents, leading to credit being diverted from the official to the shadow economy. It suggests that the financial market diverts funds from the real sector of the economy, thereby thwarting entrepreneurial activities in the official economy. This result is consistent with the research conducted by Ajide and Dada (2024) and Nguyen and Thanh (2020).

Conclusions

The impact of economic complexity on socioeconomic activities, especially the shadow economy, remains a matter of debate in the research community. While some studies acknowledged the positive role of economic complexity in reducing the size of the shadow economy, others found a negative influence. This study therefore takes a step forward by scrutinising the asymmetric impact of economic complexity on the size of the shadow economy using a dataset of 28 African nations from 1995 to 2020. To this end, estimation techniques such as 2SGMM, PMG, AMG, and CCEMG are employed.

The study's results reveal that positive economic complexity shocks mitigate the size of shadow economy, while negative economic complexity shocks contribute to its growth. This result establishes the presence of asymmetries, as the shadow economy reacts differently to positive and negative shocks in terms of economic complexity. Furthermore, the threshold level of economic complexity in the region is 0.41 on a scale of –2.5 to +2.5. While economic complexity contributes to the shadow economy below the threshold, it reduces its size above the threshold. Economic growth has a significant negative impact on the shadow economy. However, population growth and financial development positively affect the size of this economy. These findings are robust to alternative proxies of the shadow economy.

Based on the conclusions of this study, the following policy recommendations are proposed. The study found that positive shocks in economic complexity has a beneficial effect in reducing the size of the shadow economy in the region. Hence, policies aimed to increase the complexity of the economy should be pursued effectively. These policies include massive investment in human capital, infrastructure, research and development, and technological innovation, which could boost the productive knowledge in the economy.

Economic complexity is found to improve in the presence of advanced technology and automation. Therefore, efforts should be directed to technological advancements and a transition to a knowledge-driven economic system. All these can further reduce the size of Africa's shadow economy. The complexity of the economic system can only emerge from high-quality human education, highly skilled training, and an empowerment programme. Therefore, policymakers should prioritise this area to discourage potential participation in the shadow economy.

Although this study has limitations, it can serve as a guide for future research. Firstly, the asymmetry in the relationship between economic complexity and shadow economy is assessed within African countries. Future studies can extend this asymmetry to Asian and Latin American countries. Secondly, future studies might investigate the interaction between economic complexity and other economic characteristics. Thirdly, the study does not focus on the dual nature of informality (informal employment vs. informal production), which could have differential relationships with economic complexity. Moreover, institutional indicators that capture labour market dynamics, rigid infrastructure, and institutional inefficiency are not included in the model due to the unavailability of reliable data. Thus, future studies could consider this important aspect when examining informality as reliable data for African economies becomes available. Lastly, the study focuses on a heterogeneous panel of countries. Hence, a country-specific study that considers each country's specificities could be conducted on the relationship in an asymmetric framework.

Appendix

Table A1. List of countries

Algeria	Malawi
Angola	Mali
Botswana	Morocco
Burkina Faso	Mozambique
Cameroon	Namibia
Congo Republic	Nigeria
Congo Democratic	Senegal
Egypt	South Africa
Gabon	Sudan
Ghana	Togo
Guinea	Tunisia
Ivory Coast	Uganda
Kenya	Zambia
Madagascar	Zimbabwe

Source: own work.

Table A2. CD test

	PS LM	BCS LM.	P CD
SHE	132.799[a]	132.195[a]	44.207[a]
EC^{+}	304.187[a]	303.583[a]	95.355[a]
EC^{-}	301.808[a]	301.204[a]	94.994[a]
GDP	216.437[a]	215.833[a]	57.725[a]
POP	313.885[a]	313.281[a]	86.559[a]
FD	195.349[a]	194.744[a]	60.271[a]

Note: a suggests 1% level of significance.

Source: own work.

Table A3. Stationarity test

Variables	CADF		CIPS	
SHE	–1.871	–3.641[a]	–2.013	–4.647[a]
EC^{+}	–1.831	–3.041[a]	–1.954	–5.091[a]
EC^{-}	–1.364	–3.445[a]	–1.689	–4.769[a]
GDP	–1.362	–2.618[a]	–1.176	–3.835[a]
POP	–1.881	–4.428[a]	–1.516	–3.284[a]
FD	–1.365	–2.905[a]	–2.071[c]	–5.092[a]

Note: a and c suggest 1% and 10% levels of significance.

Source: own work.

Appendix 4. Westerlund cointegration test

	Value	z-value	Robust p-value
G_t	–6.192[a]	–20.146	0.000
G_a	–1.841	5.874	1.000
P_t	–8.132[b]	–6.252	0.022
P_a	–1.952	1.643	0.847

Note: a and b suggest 1% and 5% levels of significance.

Source: own work.

References

Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In D. Card & O. Ashenfelter (Eds.), *Handbook of labor economics* (vol. 4, pp. 1043–1171). https://doi.org/10.1016/S0169-7218(11)02410-5

Ahmed, Z., Cary, M., Shahbaz, M., & Vo, X.V. (2021). Asymmetric nexus between economic policy uncertainty, renewable energy technology budgets, and environmental sustainability: Evidence from the United States. *Journal of Cleaner Production*, *313*, 127723. https://doi.org/10.1016/j.jclepro.2021.127723

Ajide, F. M. (2022). economic complexity and entrepreneurship: Insights from Africa. *International Journal of Development Issues*, *21*(3), 367–388. https://doi.org/10.1108/IJDI-03-2022-0047

Ajide, F. M., & Dada, J. T. (2024). Economic complexity and shadow economy: Insights from Africa. *Iranian Economic Review*, *28*(3), 788–813.

Ajide, F. M., Dada, J. T., Al-Faryan, M. A. S., & Tabash, M. I. (2024). Shadow economy-income inequality nexus: A panel analysis of ECOWAS countries. *Journal of Economic Policy Reforms*, *28*(2), 119–141. https://doi.org/10.1080/17487870.2024.2392571

Ajide, F. M., Dada, J. T., Tabash, M. I., & Al-Faryan, M. A. S. (2025). Financial inclusion women empowerment and informal economy in Africa. *Poverty and Public Policy*, 17(4), e70029. https://doi.org/10.1002/pop4.70029

Ajide, F. M., & Osinubi, T. T. (2024). Upgrading economic complexity in Africa: The role of remittances and financial development. *Global Business and Economics Review*, *30*(3), 359–382. https://doi.org/10.1504/GBER.2024.137620

Altintas, H., & Kassouri, Y. (2020). The impact of energy technology innovations on cleaner energy supply and carbon footprints in Europe: A linear versus nonlinear approach. *Journal of Cleaner Production*, *276*, 124140. https://doi.org/10.1016/j.jclepro.2020.124140

Arnaut, M., Dada, J. T., Sharimakin, A., & Al-Faryan, M. A. S. (2023). Does environmental quality respond (A)symmetrically to (in)formal economies? Evidence from Nigeria. *Society and Business Review*, *18*(4), 646–667. https://doi.org/10.1108/SBR-07-2022-0181

Awoleye, E. O, Dada, J. T. & Al-Faryan, M. A. S. (2025). Analyzing the role of remittances and financial development in the size of the shadow economy: Insight from Africa. *Man and the Economy*, *12*(1) 21–44. https://doi.org/10.1515/me-2024-0002

Azami, S., Pour Akbari, F., & Najafi, S. M. B. (2024). The asymmetric impact of economic complexity on ecological footprints in large Asian emitters: New evidence from threshold and non-threshold panel models. *Quarterly Journal of Quantitative Economics*. https://doi.org/10.22055/jqe.2024.47028.2632

Azizi, Z. (2020). The threshold effect of economic complexity on energy consumption in Iran using smooth transition regression model. *Journal of Applied Economics Studies in Iran*, *8*(32), 103–127.

Bandeira Morais, M., Swart, J., & Jordaan, J. A. (2021). Economic complexity and inequality: Does regional productive structure affect income inequality in Brazilian states? *Sustainability*, *13*(2), 1006. https://doi.org/10.3390/su13021006

Berdiev, A. N., Saunoris, J. W., & Schneider, F. (2018). Give me liberty, or I will produce underground: Effects of economic freedom on the shadow economy. *Southern Economic Journal*, *85*(2), 537–562. https://doi.org/10.1002/soej.12303

Bolarinwa, S. T., & Simatele, M. (2023). Informality and poverty in Africa: Which comes first? *Sustainable Development*, *31*(3), 1581–1592. https://doi.org/10.1002/sd.2468

Bolarinwa, S. T., & Simatele, M. (2025). Informality and development: Revisiting the nonlinear dynamics with threshold evidence. *Economics Letters*, *256*, 112588. https://doi.org/10.1016/j.econlet.2025.112588

Boleti, E., Garas, A., Kyriakou, A. & Lapatinas, A. (2021). Economic complexity and environmental performance: Evidence from a world sample. *Environmental Modeling and Assessment*, *26*(3), 251–270 https://doi.org/10.1007/s10666-021-09750-0

Canh, P. N., Schinckus, C. & Thanh, D. S. (2021). What are the drivers of shadow economy? A further evidence of economic integration and institutional quality. *Journal of International Trade & Development, 30*(1), 47–67. https://doi.org/10.1080/09638199.2020.1799428

Canh, P. N., & Thanh D. S. (2020). Exports and the shadow economy: Nonlinear effects. *Journal of International Trade & Economic Development*, *29*(7), 865–890. https://doi.org/10.1080/09638199.2020.1759676

Capasso, S., & Jappelli, T. (2013). Financial development and the underground economy. *Journal of Development Economics*, *101*, 167–178. https://doi.org/10.1016/j.jdeveco.2012.10.005

Chen, M. A. (2012). *The informal economy: Definitions, theories and policies*. Women in Informal Employment: Globalizing and Organizing.

Dada, J. T., Ajide, F. M., Arnaut, M., & Al-Faryan, M. A. S. (2024a). On the contributing factors to shadow economy in Africa: Do natural resources, ethnicity and religious diversity make any difference? *Resources Policy*, *88*, 104478. https://doi.org/10.1016/j.resourpol.2023.104478

Dada, J. T., Ajide, F. M., Al-Faryan, M. A. S., & Tabash, M. I. (2024b). The moderating effect of economic complexity in the shadow economy-renewable energy transition nexus: Evidence from African economies. *Economic Change and Restructuring*, *57*(6), 180. https://doi.org/10.1007/s10644-024-09741-7

Dada J. T., & Al-Faryan, M. A. S. (2024). Analyzing the role of shadow and official economies on energy demand in the United Arab Emirates: An (a)symmetric approach. *Social Sciences and Humanities Open*, *9*, 100767 https://doi.org/10.1016/j.ssaho.2023.100767

Dada, J. T., Awoleye, E. O., Tabash, M. I., Ray, S., & Al-Faryan, M. A. S. (2025). The criticality of institutions in the link between natural resources rents and informal economy: Insight from a panel of African countries. *Mineral Economics*. https://doi.org/10.1007/s13563-025-00515-9

Dell'Anno, R. (2022). Theories and definitions of the informal economy: A survey. *Journal of Economic Surveys*, *36*(5), 1610–1643 https://doi.org/10.1111/joes.12487

Djeunankan, R., Njangang, H., & Tékam, H. (2023). How does economic complexity improve energy efficiency? Mechanism discussion and empirical test. *Environmental Science and Pollution Research*, *30*(43), 96906–96925. https://doi.org/10.1007/s11356-023-28920-z

Elazhary, H. R., Tohamy Eltayb, A., & Hamdy Hashem, A. (2024). *Foreign Direct Investment and industry value-added in Egypt: Economic complexity threshold analysis*. https://www.researchgate.net/journal/almjlt-almsryt-lldrasat-altjaryt-2785-9703?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InB1YmxpY2F0aW9uIn19

Elgin, C., Kose, M. A., Ohnsorge, F., & Yu, S. (2021). *Understanding informality*. CAMA Working Paper 76/2021. https://doi.org/10.2139/ssrn.3916568

Fotié, A.K., & Mbratana, T. (2024). Informality and development: The nonlinear effect. *Economics Letters*, *234*, 111470. https://doi.org/10.1016/j.econlet.2023.111470

Granger, C. W., & Yoon, G. (2002). *Hidden cointegration*. Economics Working Paper, 2. https://escholarship.org/uc/item/9qn5f61j

Habib, M. M., Mileva, E., & Stracca, L. (2017). The real exchange rate and economic growth: Revisiting the case using external instruments. *Journal of International Money and Finance*, *73*, 386–398. https://doi.org/10.1016/j.jimonfin.2017.02.014

Hartmann, D. (2014). *Economic complexity and human development: How economic diversification and social networks affect human agency and welfare*. Taylor & Francis.

Hartmann, D., Guevara, M. R., Jara-Figueroa, C., Aristarán, M., & Hidalgo, C. A. (2017). Linking economic complexity, institutions, and income inequality. *World Development*, *93*, 75–93. https://doi.org/10.1016/j.worlddev.2016.12.020

Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., & Simoes, A. (2014). *The atlas of economic complexity: Mapping paths to prosperity*. MIT Press.

Hidalgo, C. A. (2021). Economic complexity theory and applications. *Nature Reviews Physics*, *3*(2), 92–113. https://doi.org/10.1038/s42254-020-00275-1

Hidalgo, C. A., & Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, *106*(26), 10570–10575. https://doi.org/10.1073/pnas.0900943106

Ketu, I., Kelly, A. M, & Tchouto. J. E. (2024). Does economic complexity reduce the size of the shadow economy in African countries? *SN Business & Economics*, *4*(1), 1–27 https://doi.org/10.1007/s43546-023-00610-5

Kpognon, K. D. (2022). Effect of natural resources on the size of informal economy in sub-Saharan Africa: An empirical investigation. *Structural Change and Economic Dynamics*, *63*, 1–14. https://doi.org/10.1016/j.strueco.2022.08.005

La Porta, R., & Shleifer, A. (2008). *The unofficial economy and economic development*. NBER Working Paper, 14520. https://www.nber.org/papers/w14520

Le Caous, E., & Huarng, F. (2020). Economic complexity and the mediating effects of income inequality: Reaching sustainable development in developing countries. *Sustainability*, *12*(5), 2089. https://doi.org/10.3390/su12052089

Lee, K. K., & Vu, T. V. (2020). Economic complexity, human capital and income inequality: A cross-country analysis. *The Japanese Economic Review*, *71*(4), 695–718. https://doi.org/10.1007/s42973-019-00026-7

Lewis, A. (1954). Economic development with unlimited supplies of labour. *Manchester School of Economic and Social Studies*, *23*, 1–30.

Loayza, N. (2018). *Informality: Why is it so widespread and how can it be reduced?* World Bank Research and Policy Briefs, 133110.

Mealy, P., Farmer, J. D., & Teytelboym, A. (2019). Interpreting economic complexity. *Science Advances*, *5*(1). https://doi.org/10.1126/sciadv.aau1705

Medina, L., & Schneider, F. (2019). *Shedding light on the shadow economy: A global database and the interaction with the official one*. CESifo Working Paper, 7981. https://papers.ssrn.com/sol3/papers.cfm?Abstract_id=3502028

Nguyen, C. P. (2022). Does economic complexity matter for the shadow economy? *Economic Analysis and Policy*, *73*, 210–227. https://doi.org/10.1016/j.eap.2021.12.001

Nguyen, C. P., & Thanh, S. D. (2020). Financial development and the shadow economy: A multi-dimensional analysis. *Economic Analysis and Policy*, *67*, 37–54. https://doi.org/10.1016/j.eap.2020.05.002

Njangang, H., Nembot, L. N., & Ngameni, J. P. (2020). Does financial development reduce the size of the informal economy in sub-Saharan African countries? *African Development Review*, *32*(3), 375–391. https://doi.org/10.1111/1467-8268.12446

North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge University Press.

Olaniyi, C.O., & Odhiambo, N. M. (2023). Does institutional quality matter in the financial development-economic complexity nexus? Empirical insights from Africa. *Research in Globalization*, *7*, 100173. https://doi.org/10.1016/j.resglo.2023.100173

Osinubi, T. T., & Ajide, F. M. (2022). Foreign direct investment and economic complexity in emerging economies. *Economic Journal of Emerging Markets*, 259–270. https://doi.org/10.20885/ejem.vol14.iss2.art9

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, *74*(4), 967–1012 https://doi.org/10.1111/j.1468-0262.2006.00692.x

Pham, M. H., Truong, H. D. H., & Hoang, D. P. (2024). Economic complexity, shadow economy, and income inequality: Fresh evidence from panel data. *International Economic Journal*, *38*(2), 270–292. https://doi.org/10.1080/10168737.2024.2311704

Ramey, G., & Ramey, V. A. (1991). *Technology commitment and the cost of economic fluctuations*. NBER Working Papers, 3755. https://doi.org/10.3386/w3755

Seo, M. H., Kim, S., & Kim, Y. J. (2019). Estimation of dynamic panel threshold model using Stata. *Stata Journal*, *19*(3), 685–697. https://doi.org/10.1177/1536867X19874243

Schneider, F., & Buehn, A. (2018). Shadow economy: Estimation methods, problems, results and open questions. *Open Economics*, *1*(1), 1–29. https://doi.org/10.1515/openec-2017-0001

Sweet, C., & Eterovic, D. (2019). Do patent rights matter? 40 years of innovation, complexity and productivity. *World Development*, *115*, 78–93. https://doi.org/10.1016/j.worlddev.2018.10.009

Tokman, V. E. (1979). An exploration into the nature of informal-formal sector relationships. *The Urban Informal Sector*, 1065–1075. https://doi.org/10.1016/B978-0-08-024270-5.50007-7

Wu, F. D., & Schneider, F. (2019). *Nonlinearity between the shadow economy and level of development*. IMF Working Paper, WP/19/48. Institute for Capacity Development. http://hdl.handle.net/10419/202731

Yu, S., & Ohnsorge, F. (2019). The challenges of informality. *World Bank Blogs*. https://blogs.worldbank.org/developmenttalk/challenges-informality

PUEB PRESS

Economics and Business Review
https://www.ebr.edu.pl

Vol. 12 (1), 2026: 55-79
https://doi.org/10.18559/ebr.2026.1.2558

Demystifying Foreign Direct Investment dynamics in emerging economies: An ISM–MICMAC analysis

Srabani Paul Grover[1] **Varun Chotia**[2]
Satyaki Datta[3] **Sham Ranjan Shetty**[4]

Abstract

Foreign Direct Investment (FDI) has emerged as a pivotal force in the economic development trajectory of emerging economies, catalysing growth, technological transfer, and global integration. The World Investment Report of 2025 states that, as of 2024, 57.5% ($867.2 billion) of all global FDI inflows went to emerging economies, up from just 16.4% ($222.7 billion) in 2000. The existing literature on FDI determinants mostly uses the multiple regression approach. However, we employ a novel approach using Interpretive Structural Modelling (ISM) coupled with MICMAC analysis. This qualitative methodology provides a holistic understanding of the hierarchical relationships and interdependencies

Keywords

- Foreign Direct Investment (FDI)
- FDI determinants
- FDI drivers
- emerging economies
- Interpretive Structure Modelling (ISM)
- MICMAC

Suggested citation: Grover, S. P., Chotia, V., Datta, S., & Shetty, S. R. (2026). Demystifying Foreign Direct Investment dynamics in emerging economies: An ISM–MICMAC analysis. *Economics and Business Review*, *12*(1), 55-79. https://doi.org/10.18559/ebr.2026.1.2558

[1] Jaipuria Institute of Management, Jaipur, 1, Bambala Institutional Area, Pratap Nagar, Sanganer, Jaipur, Rajasthan, India, srabani.paul.fpm22j@jaipuria.ac.in, https://orcid.org/0009-0005-9463-0939.

[2] Jaipuria Institute of Management, Jaipur, 1, Bambala Institutional Area, Pratap Nagar, Sanganer, Jaipur, Rajasthan, India, varun.chotia@jaipuria.ac.in, https://orcid.org/0000-0003-1844-9060.

[3] T A Pai Management Institute, Manipal Academy of Higher Education, Manipal, India, 80 Badagabettu, Manipal, Udupi, Karnataka, India, corresponding author: satyaki.datta@manipal.edu, https://orcid.org/0000-0003-2410-7171.

[4] T A Pai Management Institute, Manipal Academy of Higher Education, Manipal, India, 80 Badagabettu, Manipal, Udupi, Karnataka, India, shamranjan.shetty@manipal.edu, https://orcid.org/0000-0003-1488-5877.

among FDI drivers. A key finding of this study is that the three most significant factors influencing FDI in emerging economies are political stability, corruption, and the state of infrastructure. These factors significantly impact all other drivers in the system. The drivers were categorised into clusters using MICMAC based on their dependence and driving power.

JEL codes: C65, F21, O10.

Article received 20 October 2025, accepted 4 February 2026.

Introduction

In the recent economic history of most emerging countries, Foreign Direct Investment (FDI) has been a growing trend (Elfakhani & Mackie, 2015) and is a vital component of these countries' economic development. One of the most prominent indicators of the world economy's globalization over the past two decades has been the significant rise in FDI flows among nations, particularly to emerging economies, as documented by recent studies analysing global FDI trends from the 1990s through the 2010s and beyond (Al-Kasasbeh et al., 2022; Nazzal et al., 2025; World Investment Report, 2006). This is depicted in Figure 1.

The percentage of global FDI inflow to emerging economies has increased from a mere 16.4% (222.7 billion dollars) in 2000 to 57.5% (867.2 billion dollars) as of 2024 (World Investment Report, 2025), increasing by almost 289% in 2024 as compared to 2000. FDI impacts the host nation's economy in a myriad of ways, such as the recipient country's general welfare, employment, economic growth, development, and output. According to Blomstrom and Kokko (2001), FDI is the primary means of disseminating modern technologies. Since the swift and effective adoption of "best practices" across national boundaries is a fundamental component of economic development, emerging economies rely substantially on FDI for growth. Previous research presented in the *Economics and Business Review* has shown that FDI inflows have substantial implications for emerging economies' growth. For instance, a study on the Polish economy found that increases in FDI were associated with higher economic growth, illustrating the macroeconomic significance of foreign investment in transitional contexts (Soylu, 2019).

As demonstrated in Figure 1, Asia remains the leading recipient of FDI among emerging regions, capturing 40% of the global FDI inflows in 2024, well ahead of Latin America at 11%. Africa's share, however, remains comparatively modest at just 6%. Notably, FDI inflows into Asian economies have risen sharply, from USD 132.4 billion in 2000 to USD 604.5 billion in 2024, marking an impressive growth of nearly 357%.

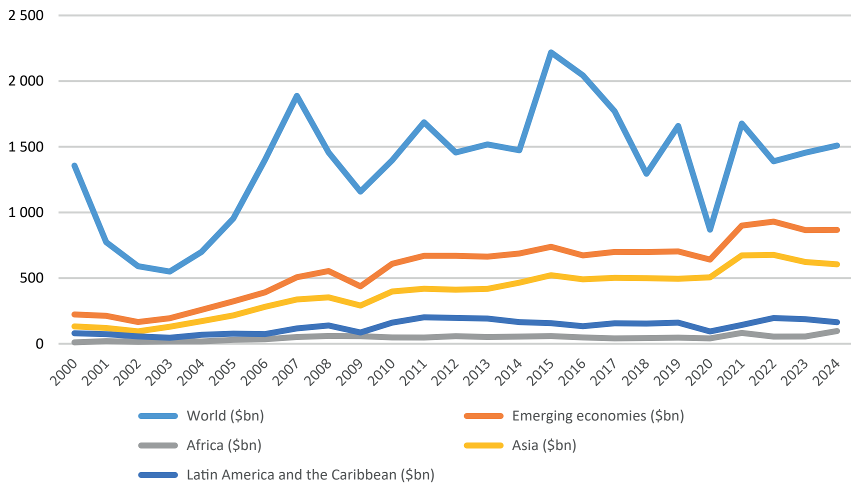

Figure 1. Trends in global and regional FDI inflows (world, emerging economies, Asia, Africa, and Latin America), 2000–2024 (US$ billion)

Note: Figure 1 illustrates the trend of FDI inflows from 2000 to 2024, comparing global inflows with those directed towards emerging economies. The breakdown highlights regional distribution across Asia, Africa, and Latin America & the Caribbean, showing Asia as the dominant recipient of FDI among emerging regions. Please note that FDI inflows are expressed in current (nominal) US dollars, as reported by UNCTAD, and are not adjusted for inflation.

Source: UNCTAD website, World Investment Report (2025), compiled by the authors.

The key question is: What drives FDI inflows into emerging economies? Several studies (Adhikary, 2017; Anyanwu, 2012; Elfakhani & Mackie, 2015) have explored this, highlighting various determinants. Kumari and Sharma (2017), analysing 20 Asian emerging nations (1990–2012), found market size to be the most influential factor. Saini and Singhania (2018) used panel data from 11 developed and 9 emerging countries (2004–2013) and observed diverse FDI drivers across nations. In particular, the authors found that policy-related determinants of FDI, such as GDP growth, the freedom index, and trade openness, are more significant in developed economies, but economic determinants of FDI, including trade openness, and efficiency-related variables, play a bigger role in developing ones. Maryam and Mittal (2020) examined BRICS nations (1994–2018), identifying GDP, trade openness, exchange rates, and infrastructure as long-term influencers. Wagner and Delios (2023) focused on India (2000–2017), concluding that its digital strength, R&D, and skilled professionals attract long-term FDI, positioning the country as a rising knowledge economy.

This study addresses three key research gaps in the literature on FDI determinants. Firstly, most prior studies rely on time-series data, vulnerable to economic shocks like recessions, often distorting results and producing inconsistent or unclear relationships. Secondly, quantitative studies frequently suffer from limited data on crucial variables such as corruption, political stability, and legal system efficiency, weakening their analytical strength. Thirdly, a methodological gap exists, as past research largely depends on multiple regression analysis, which, while helpful in showing direct relationships, fails to capture the complex interlinkages among FDI drivers that influence investment decisions in emerging economies.

The study employs the ISM-MICMAC approach to address the identified research gaps. ISM is a qualitative method that relies on expert opinions to map complex relationships among components, thereby providing a structured framework for analysing challenging problems (Hughes et al., 2020). MICMAC, on the other hand, classifies elements by examining their driving and dependence power through matrix analysis, helping to reveal their relative significance in the system (Bashir & Ojiako, 2020). When combined, ISM-MICMAC offers a robust framework that not only visualises hierarchical relationships but also highlights the degree of interdependency between variables (Chowdhury et al., 2020). This integrated approach enhances the understanding of how determinants interact and supports policymakers in prioritising the most critical areas for attracting FDI into an economy.

The study looks for answers to the following research questions: (1) What factors influence the inflow of FDI in emerging economies? (2) How do the identified determinants link to one another? (3) What is the precise driving and dependence power of every FDI determinant that has been identified? (4) Which FDI determinants, in terms of their driving and dependent powers, are the most important and prominent? A key finding of the analysis is that political stability, corruption and infrastructure emerge as the three most critical determinants, as they strongly influence the behaviour of the other drivers in the system.

The paper is structured as follows: Section 1 presents the research methodology, including an overview of the ISM–MICMAC approach, the characteristics of the expert panel, and the identification of FDI drivers from the literature. Section 2 reports the empirical results of the analysis. Section 3 discusses the findings and outlines their implications for policy and practice. Last section concludes the study and highlights its limitations as well as directions for future research.

1. Research methodology

This study used a mix of approaches to achieve its objectives (see Figure 2). A thorough literature analysis was conducted to determine emerging economies' primary FDI drivers. Subsequently, an expert survey was carried out to gather information about the relevant interactions between the drivers. The application of the ISM technique resulted in the development of the Reachability Matrix and Structural Self-Interaction Matrix. This was followed by the emergence of a hierarchical structure, with levels assigned to each driver. These drivers were then classified into different clusters using the MICMAC approach.

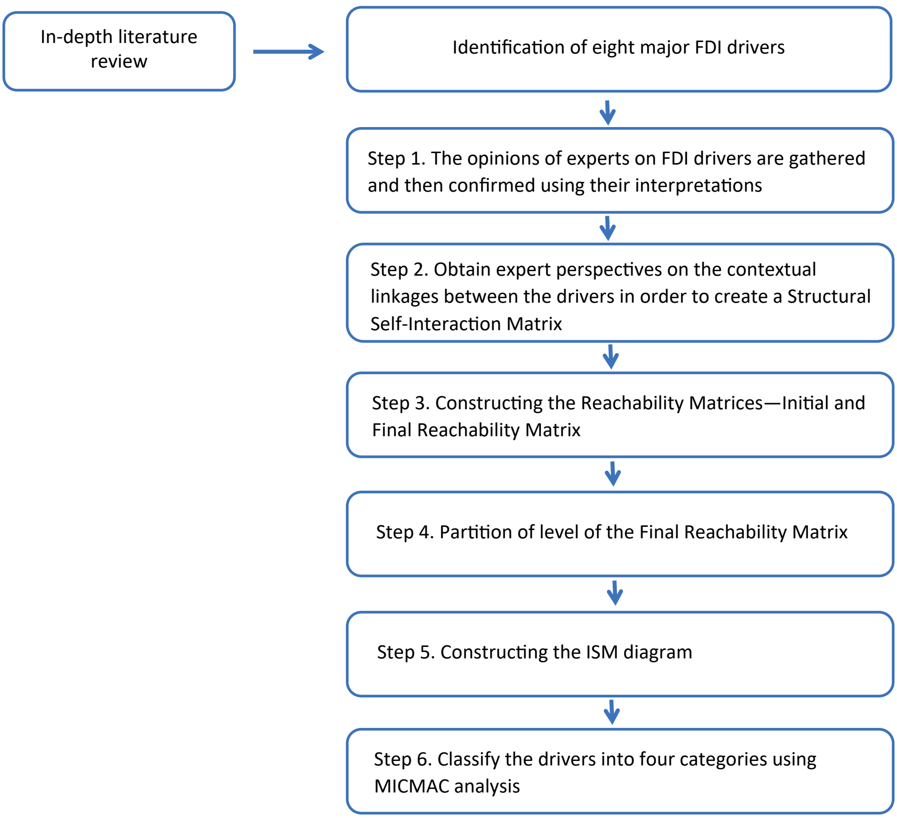

Figure 2. Step-by-step research framework based on the ISM–MICMAC methodology for identifying and classifying FDI drivers

Note: Figure 2 illustrates the six key steps of the ISM–MICMAC methodology, beginning with identifying FDI drivers and expert inputs, followed by the construction of the Structural Self-Interaction Matrix and the reachability matrix. The process then develops a hierarchical ISM model, which is finalised through MICMAC analysis to classify drivers based on their driving and dependence power.

Source: own work.

The integration of ISM-MICMAC provides a powerful methodological framework for analysing complex, multi-variable problems such as FDI determinants. ISM helps in establishing a clear hierarchical structure by identifying which factors act as the most influential drivers and which remain dependent within the system. This systematic layering ensures that relationships are not only captured but also organized in a logical sequence, facilitating interpretation. MICMAC, on the other hand, complements ISM by classifying variables based on their driving and dependence power, thereby offering a nuanced understanding of their roles within the system. Together, these techniques enable the study to move beyond a simple identification of factors, allowing a more comprehensive mapping of interdependencies and strategic insights for policymakers. Figure 3 outlines the six major steps involved in the process.

1.1. Overview of ISM methodology and MICMAC analysis

ISM is an approach that creates multi-level models by breaking down complicated systems into several subsystem components (Watson, 1978). The direction and strength of the direct and indirect interactions between variables are determined using the ISM approach by aggregating the viewpoints of experts and writers (Saxena et al., 1990). It uses terminology, graphs and differential mathematics to help researchers determine when and how the variables are related (Ansari et al., 2013). The following is an outline of the six crucial steps of the ISM approach, as depicted in Figure 3:

Step I: The FDI drivers are first identified through a thorough literature analysis.

Step II: Semi-structured interviews were undertaken with experts from academia and industry based on the drivers verified in Step I to determine the contextual interrelationships between the drivers. It was suggested that the experts (seven from the academic community and 10 from industry, with an experience of over ten years) compare the drivers in pairs. Four letters are used to represent the contextual interactions between the drivers *i* and *j*: (1) V signifies that "driver *i* led to driver *j*". (2) A signifies that "driver *j* led to driver *i*". (3) *X* signifies that "drivers *i* and *j* led to each other". (4) *O* signifies that "drivers *i* and *j* are not related to each other". To create the contextual interrelationships between the drivers, the acquired data were loaded into the Structural Self-Interaction Matrix .

Step III: The third step is to create a reachability matrix. The following guidelines (see Table 1) are used to develop the initial reachability matrix from the Structural Self-Interaction Matrix created in Step II (Peeters et al., 2019).

Table 1. Transformation rules for converting the Structural Self-Interaction Matrix into the Reachability Matrix (RM) in the ISM methodology

Structural Self-Interaction Matrix	Reachability Matrix (*i*, *j*)	Reachability Matrix (*j*, *i*)
V	1	0
A	0	1
X	1	1
O	0	0

Note: Table 1 demonstrates how Structural Self-Interaction Matrix symbols are systematically converted into binary values for the Reachability Matrix. For example, *V* indicates that factor *i* influences factor *j*, resulting in (*i*, *j*) = 1 and (*j*, *i*) = 0. Similarly, *A* shows the reverse influence, *X* reflects mutual influence, and *O* denotes no relation between the two factors.

Source: own work.

The initial reachability matrix does not depict the indirect linkages between the drivers but only the direct interactions. Therefore, it is transformed into the final reachability matrix using the rule of transitivity (i.e., if $A = B$ and $B = C$, then $A = C$).

Step IV: Creating the divisions between the levels. Based on reachability and antecedent sets, the levels are partitioned. Each driver and any other drivers might lead to making up the reachability set. Every driver and any other driver could form a part of the antecedent set. The intersection set is then determined based on the reachability and the antecedent set. Following this, the reachability set is compared with the intersection set to determine each driver's level. The drivers fall into level I, the highest level of the ISM hierarchy, when the intersection sets and reachability are identical. The higher-level drivers suggest that other drivers have a more significant potential to influence them. The drivers at level I should not be included in the reachability sets of any other drivers. After that, the subsequent iteration will start to detect further levels and keep on until all drivers have been classified. Ultimately, a hierarchical system is formed.

Step V: The structural model can be represented as a digraph after determining the levels.

Step VI: Grouping the drivers according to the MICMAC technique. The MICMAC (cross-impact matrix multiplication applied to classification) technique is employed to group the drivers based on their dependent and driving power coming from the final version of the reachability matrix. Generally, a driver's dependence power suggests that other drivers should be investigated before them. A higher driving power indicates that a driver can outperform many different drivers (Ansari et al., 2013). Therefore, based on the methodology employed in earlier research, the barriers can be divided into four categories: autonomous clusters are the ones with low driving and dependence power; linkage clusters have high driving and dependence power; dependent clusters consist of elements whose driving power is low and dependence power is high; and independent clusters consist of elements whose driving power is high and dependence power is low (Nandal et al., 2019; Zhou et al., 2019).

1.2. Characteristics of the expert panel

To ensure the reliability and validity of expert inputs in this ISM-MICMAC-based study, a purposive sampling approach was adopted to recruit participants with significant domain expertise. A total of 17 experts were selected, comprising senior academics and industry professionals, each with over a decade of experience in fields related to FDI, international business, and policy advisory roles in emerging economies. The academic professionals, who had research and teaching experience directly applicable to foreign direct investment and emerging market dynamics, represented the fields of economics, international business, finance, development studies, and public policy. The industry professionals were chosen from different fields, such as financial services, infrastructure development, investment advisory, multinational corporations, and policy-oriented research organisations. They were in charge of strategic planning in new economies, project assessment, and cross-border investment decision-making. Experts were identified through academic networks, industry associations, and published work in relevant domains. They were contacted via email and provided with an overview of the study, along with a consent form outlining their expected role and the voluntary nature of their participation. The final group was selected to ensure a balanced representation of perspectives from both academia and industry, in line with best practices in ISM studies. Table 2 presents the profiles of seventeen academic and industry professionals with over a decade of expertise in the field.

The eight determinants of FDI considered in this study were exclusively derived from an extensive review of the existing literature. Prior empirical and conceptual studies on FDI inflows into emerging economies were carefully examined, and recurrent factors were identified, enabling us to narrow down to eight widely acknowledged determinants: market size, infrastructure, trade openness, human capital, political stability, exchange rate, inflation, and corruption. Once these drivers were established from the literature, inputs were sought from a panel of experts from academia and industry. The role of experts was not to identify or propose the determinants

Table 2. Professional profile of experts participating in the ISM–MICMAC analysis

Participant No.	Designation	Experience (in years)	Field
1	Associate Professor	11	Academics
2	Professor	15	Academics
3	Professor	13	Academics
4	Research Specialist	12	Industry
5	Senior Manager	15	Industry
6	Associate Professor	13	Academics
7	Senior Manager	15	Industry
8	Associate Director	17	Industry
9	Associate Professor	13	Academics
10	Professor	14	Academics
11	Associate Director	17	Industry
12	Senior Manager	14	Industry
13	Associate Director	18	Industry
14	Assistant Professor	11	Academics
15	Senior Vice-President	19	Industry
16	Research Specialist-Fund Manager	16	Industry
17	Vice-President	15	Industry

Note: Table 2 outlines the profile of 17 experts consulted for the ISM-MICMAC analysis, comprising senior academics and industry professionals with extensive experience in FDI, international business, and policy advisory in emerging economies.

Source: own work.

but to provide their opinions on the contextual interrelationships among these drivers. This was done through semi-structured interviews and brainstorming sessions with the chosen experts. Experts then evaluated the relationship between these drivers by answering whether one directly influences the other. In the split of opinions, the majority rule principle was applied, where the minority is subordinate to the majority (Gan et al., 2018). If nine or more experts agreed, a relationship was confirmed. These contextual links were finalised through discussions and recorded in the Structural Self-Interaction Matrix. Their assessments formed the basis for developing the Structural Self-Interaction Matrix, which was subsequently converted into the Reachability Matrix and used for constructing the ISM model and conducting the MICMAC analysis.

1.3. Identification of FDI drivers from literature

The key determinants of FDI consistently feature in empirical analyses as critical influencers of investment flows. These are market size (Ayomitunde et al., 2020; Ullah & Khan, 2017), infrastructure (Kingori, 2022; Rehman et al., 2023), trade openness (Albahouth & Tahir, 2024; Moraghen et al., 2023), human capital (Abbas et al., 2021; Ibbaro-Olivo et al., 2024), exchange rate dynamics (Muhammad et al., 2018; Sasana & Fathoni, 2019), inflation (Agudze & Ibhagui, 2021; Imran & Rashid, 2023), political stability (Bhujabal et al., 2024; Kechagia & Metaxas, 2022), and corruption (Kechagia & Metaxas, 2022; Qureshi et al., 2021). These factors shape the investment climate by affecting both the potential returns and the risks perceived by the multinational enterprises. While economic determinants signal growth opportunities and efficiency gains, institutional and policy-related variables play an equally important

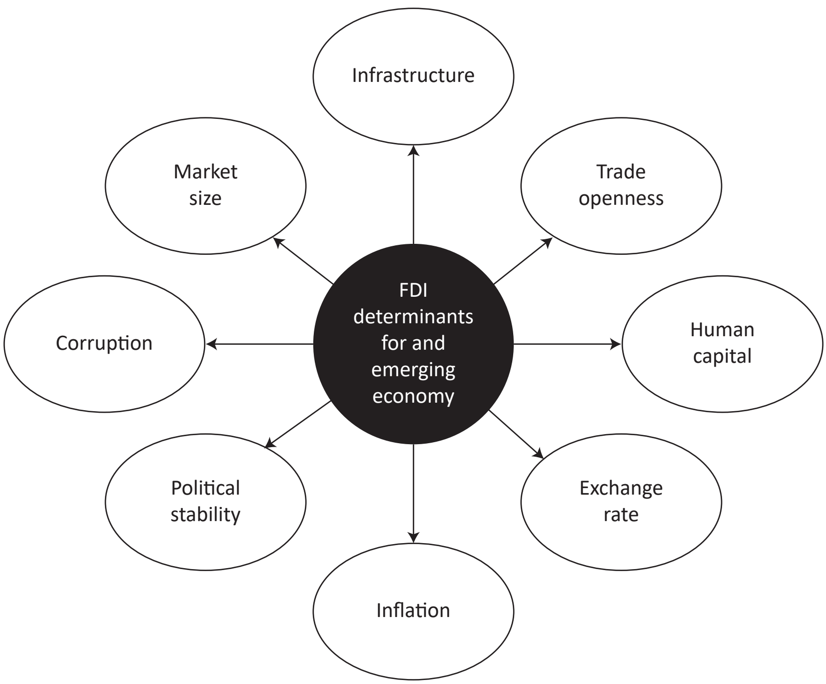

Figure 3. Key determinants of Foreign Direct Investment in emerging economies identified in the study

Note: Figure 3 illustrates the eight major determinants of FDI in emerging economies, as identified in the study. These factors span macroeconomic conditions (e.g., inflation, exchange rate), institutional aspects (e.g., political stability, corruption), market and structural factors (e.g., market size, infrastructure), and trade and human capital (e.g., trade openness and human capital).

Source: own elaboration.

role in ensuring stability and predictability. Together, they form a multidimensional framework, within which investors make strategic decisions, making it essential to examine how these determinants interact and influence FDI inflows in emerging economies. Several studies published in the *Economics and Business Review* also contribute to understanding FDI dynamics. For example, Trąpczyński (2013) critically examines how theoretical models of FDI explain subsidiary performance and the heterogeneity of empirical findings across contexts, underscoring the complexity of FDI drivers and the need for nuanced frameworks beyond standard econometric analysis.

An overview of the main FDI determinants identified in the literature to date is provided below, and Figure 3 illustrates these determinants. Table 4 displays the initial RM, while Table 5 shows the final RM.

1.3.1. Market size and trade openness

Market size continues to be one of the strongest predictors of FDI, as larger economies signal greater consumer demand and scale economies. Recent studies confirm this linkage: Ullah and Khan (2017) and Ayomitunde et al. (2020) highlight the importance of economy size in South Asia and Nigeria, while Sasana and Fathoni (2019) report similar evidence from ASEAN. Other empirical contributions published in the *Economics and Business Review* have examined the behaviour and motives of foreign direct investment across contexts. For example, Gorynia et al. (2015) use quantitative survey data to investigate the outward FDI motives of Polish multinational firms, identifying market-seeking incentives as the most prominent drivers.

Openness to trade further complements market size by reducing transaction costs and integrating economies into global value chains. Albahouth and Tahir (2024) show that trade openness significantly drives FDI in ASEAN, while Sabir et al. (2019) and Moraghen et al. (2023) demonstrate its positive role in emerging and small island economies. Together, these studies reaffirm that expanding markets and liberal trade policies are major attractions for investors.

1.3.2. Infrastructure and human capital

Well-developed infrastructure enhances efficiency-seeking FDI by reducing operational costs and improving competitiveness. Rehman et al. (2023) found that transport, energy, and telecommunication infrastructure significantly increased FDI inflows into BRICS, while Kingori (2022) noted a bidirectional link between infrastructure and FDI in Kenya. Similarly, human capital is crucial for absorbing the technology and knowledge that accompanies FDI. Dobrota et al. (2021) and Osei and Kim (2020) highlight the role of skilled labour in enhancing productivity, and Abbas et al. (2021) showed how FDI fosters skill development. Ibbaro-Olivo et al. (2024) further confirm that FDI inflows in Southeast Asia have been positively associated with the growth of technical and vocational education.

1.3.3. Exchange rate and inflation

Exchange rate conditions shape investor decisions by affecting cost structures. Muhammad et al. (2018) found that depreciation can attract FDI by lowering local costs, although excessive volatility discourages inflows, a finding echoed by Sasana and Fathoni (2019). Inflation, another indicator of macroeconomic stability, shows mixed effects. While Imran and Rashid (2023) observed a positive relationship between inflation and FDI in emerging countries, Agudze and Ibhagui (2021) demonstrated that high inflation thresholds negatively affect inflows. Dewi and Septriani (2023) noted that, in the ASEAN context, inflation, alongside growth and interest rates, contributes positively to FDI, suggesting that the effect remains context-dependent.

1.3.4. Political stability and corruption

Institutional quality strongly influences investment attractiveness. Political stability builds investor confidence, as highlighted by Bhujabal et al. (2024) for South and Southeast Asia, and by Kechagia and Metaxas (2022) for BRIC and CIVET nations. Kiptoo (2024) further demonstrate that contract enforcement and rule of law in stable political environments increase investor trust, while instability deters inflows. Similarly, corruption control is essential in sustaining FDI. Qureshi et al. (2021) found that a lower level of corruption significantly enhances FDI across 54 economies, and Kechagia and Metaxas (2022) reported a similar effect in CIVETS countries. These findings underscore that good governance remains central to attracting and retaining foreign investment.

2. Empirical results

The ISM technique addresses the eight identified factors that influence the inflow of FDI to an emerging economy. This section presents the investigation's results and provides a thorough discussion of them.

2.1. Creation of structural self-interaction matrix

The Structural Self-Interaction Matrix provides a pairwise comparison between the drivers. In this step, the experts determine which elements have the power to impact one another and how different components in the matrix may affect the same factor. The Structural Self-Interaction Matrix that was obtained from the study is shown in Table 3.

Table 3. Structural Self-Interaction Matrix of FDI drivers in emerging economies

FDI drivers	Corruption	Inflation	Exchange rate	Political stability	Human capital	Trade openness	Infrastructure	Market size
Market size	A	A	A	A	X	X	X	X
Infrastructure	A	X	O	A	X	X	X	
Trade openness	A	V	X	A	X	X		
Human capital	X	A	O	A	X			
Political stability	X	V	V	X				
Exchange rate	O	X	X					
Inflation	A	X						
Corruption	X							

Note: Table 3 presents the Structural Self-Interaction Matrix, which captures the contextual relationships among the eight FDI drivers. The symbols denote the direction of influence: *V* = driver *i* influences driver *j*; *A* = driver *j* influences driver *I*; *X* = drivers *i* and *j* influence each other; *O* no direct influence between *i* and *j*. The Structural Self-Interaction Matrix forms the foundation for deriving the Reachability Matrix, which is then subsequently used in ISM to develop the hierarchical structure of interrelationships.

Source: own work.

2.2. Development of initial Reachability Matrix and final Reachability Matrix

The Structural Self-Interaction Matrix is used to construct the Initial reachability matrix. Structural Self-Interaction Matrix transforms the letters *V*, *A*, *X*, and *O* into binary input. The final reachability matrix is generated from the initial reachability matrix using the transitivity rule.

Table 4. Initial Reachability Matrix of FDI drivers in emerging economies

FDI drivers	Market size	Infrastructure	Trade openness	Human capital	Political stability	Exchange rate	Inflation	Corruption
Market size	1	1	1	1	0	0	0	0
Infrastructure	1	1	1	1	0	0	1	0
Trade openness	1	1	1	1	0	1	1	0
Human capital	1	1	1	1	0	0	0	1
Political stability	1	1	1	1	1	1	1	1
Exchange rate	1	0	1	0	0	1	1	0
Inflation	1	1	0	1	0	1	1	0
Corruption	1	1	1	1	1	0	1	1

Note: Tables 4 and 5 present the Initial and Final Reachability Matrices derived from the Structural Self-Interaction Matrix. The contextual symbols (*V*, *A*, *X*, *O*) are converted into binary inputs (1 and O), where 1 indicates the presence of influence between two drivers and 0 its absence. The Final Reachability Matrix further incorporates the principle of transitivity, making it the basis for level partitioning in the ISM process.

Source: own work.

Table 5. Final Reachability Matrix of FDI drivers in emerging economies

FDI drivers	Market size	Infrastructure	Trade openness	Human capital	Political stability	Exchange rate	Inflation	Corruption	Driving power
1. Market size	1	1	1	1	0	1	1	1	7
2. Infrastructure	1	1	1	1	0	1	1	0	6
3. Trade openness	1	1	1	1	0	1	1	0	6
4. Human capital	1	1	1	1	1	0	1	1	7
5. Political stability	1	1	1	1	1	1	1	1	8
6. Exchange rate	1	0	1	0	0	1	1	0	4
7. Inflation	1	1	1	1	0	1	1	0	6
8. Corruption	1	1	1	1	1	1	1	1	8
Dependence power	8	7	8	7	3	7	8	4	

Source: own work.

2.3. Sectioning the final Reachability Matrix at various levels

The reachability and the antecedent sets are used for each matrix's elements, and the final Reachability Matrix for level partitioning is assessed (Warfield, 1974). The element itself and those elements it would aid in accomplishing make up the reachability set. The element itself and every other element that could be useful in getting it to make up the antecedent set. The drivers belong to level 1, the highest level in the hierarchical structure, based on ISM, provided their intersection sets and reachability are the same. The iteration process continues till all the drivers are assigned a level. In the end, a hierarchy is established. Tables 6, 7, and 8 show the three iterations that were carried out to assign a level to each driver included in the study. The iteration process continues till all the drivers are assigned a level. In the end, a hierarchy was established.

Table 6. First-level partitioning of FDI drivers using ISM

FDI drivers	Reachability set	Antecedent set	Intersection set	Level
1. Market size	1,2,3,4,6,7,8	1,2,3,4,5,6,7,8	1,2,3,4,6,7,8	Level 1
2. Infrastructure	1,2,3,4,6,7	1,2,3,4,5,7,8	1,2,3,4,7	
3.Trade openness	1,2,3,4,6,7	1,2,3,4,5,6,7,8	1,2,3,4,6,7	Level 1
4. Human capital	1,2,3,4,5,7,8	1,2,3,4,5,7,8	1,2,3,4,5,7,8	Level 1
5. Political stability	1,2,3,4,5,6,7,8	4,5,8	4,5,8	
6. Exchange rate	1,3,6,7	1,2,3,5,6,7,8	1,3,6,7	Level 1
7. Inflation	1,2,3,4,6,7	1,2,3,4,5,6,7,8	1,2,3,4,6,7	Level 1
8. Corruption	1,2,3,4,5,6,7,8	1,4,5,8	1,4,5,8	

Note: Tables 6, 7, and 8 show the iterative partitioning of the Final Reachability Matrix. In each iteration, the reachability and antecedent sets are compared to determine the levels of the drivers. The process continues across successive iterations until all drivers are assigned to a specific hierarchical level in the ISM model.

Source: own work.

Table 7. Second-level partitioning of FDI drivers using ISM

FDI drivers	Reachability set	Antecedent set	Intersection set	Level
2. Infrastructure	2	5,2,8	2	Level 2
5. Political stability	5,2,8	5,8	5,8	
8. Corruption	2,5,8	5,8	5,8	

Source: own work.

Table 8. Third-level partitioning of FDI drivers using ISM

FDI drivers	Reachability set	Antecedent set	Intersection set	Level
5. Political stability	5,8	5,8	5,8	Level 3
8. Corruption	5,8	5,8	5,8	Level 3

Source: own work.

2.4. Development of the ISM digraph

Following the completion of the level partitioning of the final Reachability Matrix, the components and their relationships can be visually represented with the formation of the ISM digraph. Level 1 comprises five drivers, namely market size (driver number 1), trade openness (driver number 3), human capital (driver number 4), exchange rate (driver number 6), and inflation (driver number 7). Due to their maximum degrees of dependence, power, and variable driving powers, these components heavily depend on the model's lower-level connections. Level II is driven solely by infrastructure (driver number 2). The bottom level, or level III, comprises Political Stability (driver number 5) and Corruption (driver number 7). Due to their significant driving force, these lower-level elements at Level II and Level III impact the model's other related components.

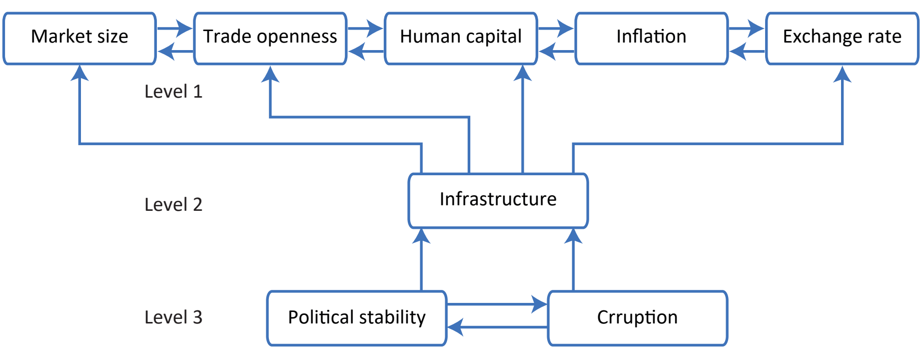

Figure 4. Interpretive Structural Modelling (ISM) digraph depicting the hierarchical relationships among FDI drivers in emerging economies

Note: Figure 4 illustrates the ISM model, which depicts the hierarchical structure of FDI drivers in emerging economies. Level 1 comprises market-related and macroeconomic factors, Level 2 highlights infrastructure as the key linkage driver, and Level 3 positions political stability and corruption as he most influential. Arrows indicate the directional relationship among drivers.

Source: own work.

2.5. MICMAC analysis

MICMAC (Matrix of Cross-Impact Multiplications Applied to Classification) analysis is used to analyse the variables' driving power and dependence power. The merger of MICMAC and ISM can help differentiate the most essential drivers/barriers from the less significant ones (Yu et al., 2020). Following the ISM analysis, all eight FDI drivers' driving and dependence powers were used to categorise them into four groups using the MICMAC technique: autonomous, linkage, dependent, and independent (also known as driving). By adding each entry to the appropriate row and column of the reachability matrix, we can determine each driver's driving and dependence power, which is depicted in Table 5. A driver's dependence power is determined by the total number of drivers who can influence it, and its driving power is determined by the total number of drivers it can affect. The MICMAC diagram in Figure 5 organises the four clusters based on the driving and dependence power of the drivers.

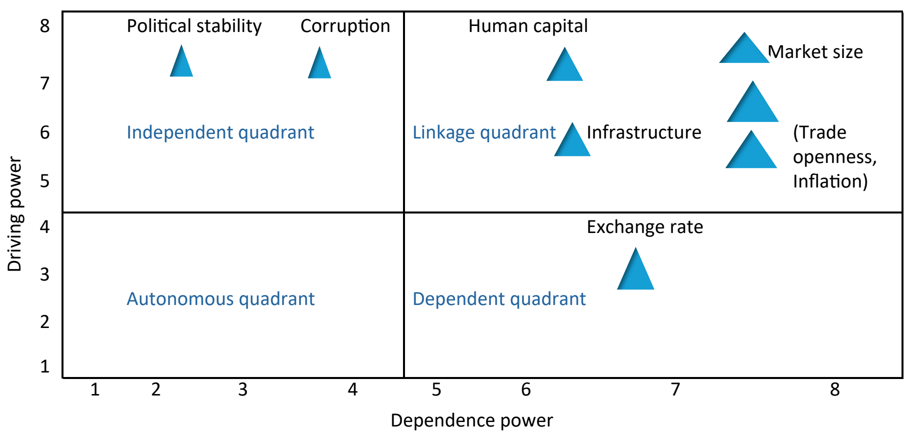

Figure 5. MICMAC analysis diagram classifying FDI drivers based on driving and dependence power

Note: Figure 5 illustrates the MICMAC analysis, classifying FDI drivers into clusters. Independent factors (X5, X8) act as strong drivers, dependent factors (X6) show high reliance, while linkage variables (X1, X2, X3, X4, X7) exhibit both high driving and dependence power.

Source: own work.

Autonomous Cluster: The factors in the Autonomous Cluster possess low driving and dependence power. They are unaffected by external factors and neither accept nor influence them, suggesting that the overall model is not significantly affected by these factors. None of the significant FDI drivers identified in this research fit into this quadrant, as shown by Figure 5, suggesting that the drivers identified for this study are all closely interrelated.

Linkage Cluster: This cluster identifies elements with high driving and dependence power. The variables that fall into this quadrant are considered unsteady; any action taken on these variables will affect other variables and have a feedback effect on themselves (Yadav & Barve, 2015). Five drivers (driver number 1: Market size, driver number 2: Infrastructure, driver number 3: Trade openness, driver number 4: Human capital, and driver number 7: Inflation) fall into this cluster. Notably, most of the FDI drivers were discovered to belong to the linkage cluster in the current study; as a result, managing and controlling these drivers would need a significant amount of work because of their feedback effect.

Dependent Cluster: This cluster includes elements of high dependence and low driving power. In the current study, only one driver (driver number 6: Exchange rate) falls into this quadrant, as Figure 5 reflects. Other drivers with significant driving power influence the variables in this quadrant. Dependent drivers will typically follow suit, provided the drivers with good driving abilities are appropriately controlled.

Independent/Driving Cluster: Independent factors, also referred to as driving factors, are those that are high on driving power and low on dependence power. These elements have the potential to impact every other factor in the system. Two drivers (driver number 5: Political stability and driver number 8: Corruption) fall into this cluster. These factors should be dealt with as a priority over all other elements of the system. The factors that are part of this cluster are significant in understanding the system's structure, as most drivers depend on them.

3. Discussion and implications of the study

FDI drivers play a critical role in attracting investment across developed and emerging economies. With emerging markets securing a rising share of global FDI, understanding the interplay of determinants becomes crucial. This study analysed eight major drivers using ISM to establish a three-level hierarchical framework and MICMAC to categorise them into clusters. The findings emphasise Political stability, Corruption, and Infrastructure as the most influential factors, shaping the dynamics of other drivers such as Market size, Trade openness, Human capital, and Exchange rate.

Political stability emerged as a core determinant, corroborating earlier studies that highlight its decisive role in shaping investors' confidence (Busse et al., 2011; Wang, 2009). Instability, on the other hand, deters investment by increasing risks and uncertainties. Corruption was identified as another major barrier, distorting resource allocation and eroding investor trust.

Transparent, accountable governance significantly enhances FDI prospects, a finding consistent with Smarzynska Javorcik and Wei (2002) and Rohwer (2009). Infrastructure, positioned as a linkage factor, exerts a cascading influence on multiple drivers, underscoring its centrality in facilitating business operations and enabling economic growth.

In addition to the findings of previous conceptual literature, an increasing number of quantitative studies confirm the major roles played by political stability, corruption, and infrastructure in the determination of inflows of FDI. Indicatively, Sabir et al. (2019) use system GMM panel estimation between developed and developing countries to show that major indicators of institutional quality, including political stability and corruption control, have a positive and statistically significant impact on the FDI inflows and institutional influences vary between income groups. Similarly, Faruq (2023) observes that panel regression models show that political stability positively affects FDI attractiveness in emerging Asian economies, which further strengthens the central role of political conditions in quantitative FDI determinants studies.

There is also quantitative evidence regarding the importance of infrastructure: the most recent panel data findings suggest that infrastructure and allied economic integration variables have a significant positive effect on FDI inflows between country pairs worldwide (Rithi et al., 2025). Furthermore, Lestari et al. (2022) demonstrate that financial development supports FDI, but corruption may undermine its efficacy, particularly in its interplay with the economic conditions, also marked by the subtle phases of governance in the attraction of foreign capital.

In other works on panel estimation, corruption and quality of public services are observed to be crucial to determining the level of FDI, and they complement the macroeconomic determinants of FDI, e.g., exports and exchange rate dynamics (Sujit et al., 2020). Lastly, developing economy evidence points at possible feedback effects, in which FDI itself can foster political stability under particular circumstances, depicting convoluted causal relations between political establishments and investment streams (Okara, 2023).

Taken together, these quantitative findings are consistent with and extend our ISM-MICMAC results by affirming the exquisite role of political stability, corruption and infrastructure in various empirical settings and justify the appropriateness of employing a methodology that considers structural dependencies among FDI determinants.

The ISM-MICMAC analysis further showed that market size, trade openness, human capital, exchange rate, and inflation function largely as linking factors at the first level, reinforcing each other in a self-perpetuating cycle. While these variables are essential, their impact is mediated by foundational enablers like infrastructure and political stability. Exchange rate volatility, placed in the dependent cluster, reflects its sensitivity to broader macroeconomic shifts influenced by other drivers.

Beyond clarifying interdependencies, this research offers theoretical contributions by applying the integrated ISM–MICMAC approach in the FDI context. It moves beyond isolated variable analysis, presenting a structured system-based model that helps scholars and policymakers visualise how changes in one determinant reverberate across others. This offers a practical framework for prioritising policy measures in emerging economies where resources are often constrained.

From a managerial perspective, the findings provide actionable insights for policymakers and investment authorities. Strengthening governance mechanisms, combating corruption, and ensuring political stability should be prioritized, as they form the foundation upon which other factors build. Similarly, investment in quality infrastructure creates multiplier effects. Enhancing trade openness, human capital development, and overall competitiveness. The ISM–MICMAC framework enables decision-makers to allocate resources strategically and design adaptive, future-oriented policies to attract and sustain FDI inflows.

In sum, the study highlights that while multiple drivers influence FDI, emerging economies must focus first on stabilising political conditions, curbing corruption, and investing in infrastructure. These act as pivotal levers, shaping the broader ecosystem of investment determinants. By providing both theoretical clarity and managerial guidance, this research contributes to strengthening the strategic roadmap for enhancing FDI inflows into emerging economies.

Conclusions

Compared to earlier studies on this subject, this study claims that by using an ISM-MICMAC methodology, the originality of the interdependencies amongst the FDI drivers facilitates a better understanding and applicability. This new understanding sheds essential light on the significance and impact of particular FDI drivers and how they may influence an economy's ability to attract more FDI. Numerous results from this study add to the body of literature and offer policymakers and scholars fresh perspectives.

Whether an economy is developed, emerging, or less developed, FDI drivers play a crucial role in drawing FDI to it. According to recent statistics, emerging economies are attracting a significant portion of global FDI. Therefore, it is even more crucial to investigate the factors that lead to this massive FDI inflow into these economies. This study closes a knowledge gap by carefully examining the contextual connections among the significant drivers. Firstly, eight primary drivers of FDI were identified with the assistance of a thorough literature review and brainstorming meetings with an expert team from academia and industry. The driving and dependent powers of each of these drivers were then determined. Next, based on the interactions between the drivers, an ISM structure with three levels was created. Using ISM, our study helps to comprehend the impact of various drivers better. We use a MICMAC analysis to characterize their impact further. The identified drivers were then separated into four clusters: two drivers in independent clusters, one driver in the dependent cluster, five drivers in the linkage cluster, and no driver in the autonomous cluster.

The findings of this study indicate that the FDI drivers are interrelated and significantly impact attracting FDI to any given nation. The findings demonstrate that political stability, corruption, and infrastructure are the three factors that have the most impact on all the other elements of the model. In the hierarchy of ISM, they are positioned at the lower level (level 2 and level 3). The success of numerous other components in the model is based on these three elements. It is advised that policymakers pay particular attention to these three FDI drivers because they significantly impact effective outcomes. This study contributes significantly to our understanding of factors that influence FDI and as a viable paradigm to assist scholars and practitioners in understanding this subject more thoroughly.

These fundamental elements can be strategically addressed to greatly boost these economies' appeal to FDI. Authorities should thus adopt a comprehensive strategy and ensure that developments in diverse elements are coordinated for the best possible outcome. The promotion of economic development and growth can be achieved by sustaining FDI inflows through the coordination of policies aimed at enhancing interdependencies. Policymakers must adopt an integrated strategy that acknowledges the interdependencies between various factors in addition to addressing each one to guarantee a stable, transparent, and growth-oriented investment climate.

While the ISM–MICMAC approach is valuable, it has some limitations. It is also dependent on the expert judgement, which can create subjectivity and biasness. In addition, the research paper focuses on an aggregate perspective of Foreign Direct Investment and fails to differentiate between the various sources and ownership patterns of the FDI. In practice, foreign investors can attach different weights to certain determinants, as an example, the institutional quality and corruption may be more important in the eyes of the private investors of developed economies than state-owned enterprises of emerging economies. Moreover, as the research focused on emerging economies, the results cannot be directly applied to the situation in developed countries. Last but not least, the methodology is useful in exposing structural interdependencies among drivers but not in quantifying the strength or statistical significance of such relationships.

The paper also provides a number of research opportunities in the future. Subsequent research might further break down FDI by the source country, type of ownership, or motivation to invest to determine the extent to which the relative significance of FDI drivers can vary among distinct investor profiles. Moreover, cross-national or regional comparisons might help to shed more light on the context. In future studies, it might also be possible to develop greater analytical strength through the combination of ISM-MICMAC and quantitative methods like Partial Least Squares Structural Equation Modelling (PLS-SEM).

References

Abbas, A., Moosa, I., & Ramiah, V. (2021). The contribution of human capital to foreign direct investment inflows in developing countries. *Journal of Intellectual Capital*, *23*(1), 9–26. https://doi.org/10.1108/JIC-12-2020-0388

Adhikary, B. K. (2017). Factors influencing foreign direct investment in South Asian economies: A comparative analysis. *South Asian Journal of Business Studies*, *6*(1), 8–37.

Agudze, K., & Ibhagui, O. (2021). Inflation and FDI in industrialized and developing economies. *International Review of Applied Economics*, *35*(5), 749–764. https://doi.org/10.1080/02692171.2020.1853683

Albahouth, A. A., & Tahir, M. (2024). The relationship between trade openness and FDI inflows: Evidence-based insights from ASEAN region. *Economies*, *12*(8), 208. https://doi.org/10.3390/economies12080208

Al-Kasasbeh, O., Alzghoul, A., & Alghraibeh, K. (2022). Global FDI inflows and outflows in emerging economies post-COVID-19 era. *Future Business Journal*, *8*(1), 53.

Ansari, M. F., Kharb, R. K., Luthra, S., Shimmi, S. L., & Chatterji, S. (2013). Analysis of drivers to implement solar power installations in India using interpretive structural modeling technique. *Renewable and Sustainable Energy Reviews*, *27*, 163–174.

Anyanwu, J. C. (2012). Why does foreign direct investment go where it goes?: New evidence from African countries. *Annals of Economics and Finance*, *13*(2), 425–462.

Ayomitunde, A. T., Ganiyu, A. B., Sokunbi, G. M., & Adebola, B. Y. (2020). The determinants of foreign direct investment inflows in Nigeria: An empirical investigation. *Acta Universitatis Danubius. Economica*, *16*(3).

Bashir, H., & Ojiako, U. (2020). An integrated ISM–MICMAC approach for modelling and analysing dependencies among engineering parameters in the early design phase. *Journal of Engineering Design*, *31*(8–9), 461–483.

Bhujabal, P., Sethi, N., & Padhan, P. C. (2024). Effect of institutional quality on FDI inflows in South Asian and Southeast Asian countries. *Heliyon*, *10*(5), e27060. https://doi.org/10.1016/j.heliyon.2024.e27060

Blomstrom, M., & Kokko, A. (2001). Foreign direct investment and spillovers of technology. *International Journal of Technology Management*, *22*(5–6), 435–454.

Busse, M., Nunnenkamp, P., & Spatareanu, M. (2011). Foreign direct investment and labour rights: A panel analysis of bilateral FDI flows. *Applied Economics Letters*, *18*(2), 149–152.

Chowdhury, N. A., Ali, S. M., Paul, S. K., Mahtab, Z., & Kabir, G. (2020). A hierarchical model for critical success factors in apparel supply chain. *Business Process Management Journal*, *26*(7), 1761–1788.

Dewi, F. D., & Septriani, S. (2023). Analysis of factors affecting foreign direct investment in ASEAN countries. *ISAR Journal of Economics and Business Management*, *1*(6), 36–40.

Dobrota, M., Zornić, N., & Marković, A. (2021). FDI time series forecasts: Evidence from emerging markets. *Management: Journal of Sustainable Business and Management Solutions in Emerging Economies*, *26*(2), 77–88.

Elfakhani, S., & Mackie, W. (2015). An analysis of net FDI drivers in BRIC countries. *Competitiveness Review*, *25*(1), 98–132.

Faruq, A. T. M. (2023). *The determinants of foreign direct investment (FDI): A panel data analysis for the emerging Asian economies*. arXiv preprint arXiv:2307.07037

Gan, X., Chang, R., Zuo, J., Wen, T., & Zillante, G. (2018). Barriers to the transition towards off-site construction in China: An interpretive structural modeling approach. *Journal of Cleaner Production*, *197*, 8–18.

Gorynia, M., Nowak, J., Trąpczyński, P., & Wolniak, R. (2015). Does outward FDI by Polish multinationals support existing theory? Findings from a quantitative study. *Economics and Business Review*, *1*(4), 84–101. https://doi.org/10.18559/ebr.2015.4.6

Hughes, D. L., Rana, N. P., & Dwivedi, Y. K. (2020). Elucidation of IS project success factors: An interpretive structural modelling approach. *Annals of Operations Research*, *285*, 35–66.

Ibarra-Olivo, J. E., Neise, T., Breul, M., & Wrana, J. (2024). FDI and human capital development: A tale of two Southeast Asian economies. *Journal of International Business Policy*, *7*(3), 314–336. https://doi.org/10.1057/s42214-024-00186-3

Imran, M., & Rashid, A. (2023). The empirical determinants of foreign direct investment episodes. *Global Journal of Emerging Market Economies*, *15*(3), 409–435. https://doi.org/10.1177/09749101221116720

Kechagia, P., & Metaxas, T. (2022). FDI and institutions in BRIC and CIVETS countries: An empirical investigation. *Economies*, *10*(4), 77. https://doi.org/10.3390/economies10040077

Kingori, Z. I. (2022). *Effect of infrastructure on FDI behaviour in Kenya: A growth nexus analysis* [doctoral dissertation]. University of South Africa.

Kiptoo, J. (2024). The influence of political stability on foreign direct investment (FDI). *International Journal of Developing Country Studies*, *6*(1), 74–86.

Kumari, R., & Sharma, A. K. (2017). Determinants of foreign direct investment in developing countries: A panel data study. *International Journal of Emerging Markets*, *12*(4), 658–682.

Lestari, D., Lesmana, D., Yudaruddin, Y. A., & Yudaruddin, R. (2022). The impact of financial development and corruption on foreign direct investment in developing countries. *Investment Management & Financial Innovations*, *19*(2), 211–220. https://doi.org/10.21511/imfi.19(2).2022.18

Maryam, J., & Mittal, A. (2020). Foreign direct investment into BRICS: An empirical analysis. *Transnational Corporations Review*, *12*(1), 1–9.

Moraghen, W., Seetanah, B., & Sookia, N. U. H. (2023). The impact of exchange rate and exchange rate volatility on Mauritius foreign direct investment: A sector-wise analysis. *International Journal of Finance & Economics*, *28*(1), 208–224. https://doi.org/10.1002/ijfe.2416

Muhammad, S. D., Azu, N. P., & Oko, N. F. (2018). Influence of real exchange rate and volatility on FDI inflow in Nigeria. *International Business Research*, *11*(6), 73–82.

Nandal, V., Kumar, R., & Singh, S. K. (2019). Barriers identification and analysis of solar power implementation in Indian thermal power plants: An interpretative structural modeling approach. *Renewable and Sustainable Energy Reviews*, *114*, 109330.

Nazzal, A., Niñerola, A. M., Sanchez-Rebull, M. V., & Barbera-Marine, M. G. (2025). A systematic mapping review of foreign direct investment by multinational corporations in emerging economies. *Humanities and Social Sciences Communications*, *12*(1), 266. https://doi.org/10.1057/s41599-025-04571-y

Okara, A. (2023). Does foreign direct investment promote political stability? Evidence from developing economies. *Economic Modelling*, *123*, 106249. https://doi.org/10.1016/j.econmod.2023.106249

Osei, M. J., & Kim, J. (2020). Foreign direct investment and economic growth: Is more financial development better? *Economic Modelling*, *93*, 154–161. https://doi.org/10.1016/j.econmod.2020.07.009

Peeters, B., Kiratli, N., & Semeijn, J. (2019). A driver analysis for distributed recycling of 3D printing waste: Taking the maker movement perspective. *Journal of Cleaner Production*, *241*, 118313.

Qureshi, F., Qureshi, S., Vo, X. V., & Junejo, I. (2021). Revisiting the nexus among foreign direct investment, corruption and growth in developing and developed markets. *Borsa Istanbul Review*, *21*(1), 80–91. https://doi.org/10.1016/j.bir.2020.08.001

Rehman, F. U., Sohag, K., & Saeed, T. (2023). Impact of sectoral infrastructure on export and foreign direct investment inflow: Evidence from selected South Asian economies by applying a new global infrastructure index. *Journal of Infrastructure Systems*, *29*(1), 04022045. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000724

Rohwer, A. (2009). Measuring corruption: A comparison between the transparency international's corruption perceptions index and the World Bank's worldwide governance indicators. *CESifo DICE Report*, *7*(3), 42–52.

Rithi, S. R., Nomani, A., Khan, M., & Manoharan, M. (2025). Determinants influencing the flow of FDI in India: An empirical insight using panel data. *Humanities and Social Sciences Communications*, *12*(1), 1497. https://doi.org/10.1057/s41599-025-05726-7

Sabir, S., Rafique, A., & Abbas, K. (2019). Institutions and FDI: Evidence from developed and developing countries. *Financial Innovation*, *5*(1), 1–20.

Saini, N., & Singhania, M. (2018). Determinants of FDI in developed and developing countries: A quantitative analysis using GMM. *Journal of Economic Studies*, *45*(2), 348–382.

Sasana, H., & Fathoni, S. (2019). Determinant of foreign direct investment inflows in ASEAN Countries. *Jurnal Ekonomi dan Kebijakan*, *12*(2), 253–266.

Saxena, J., Sushil, P., & Vrat, P. (1990). Impact of indirect relationships in classification of variables—a MICMAC analysis for energy conservation. *Systems Research*, *7*(4), 245–253.

Smarzynska Javorcik, B. S., & Wei, S. J. (2002). *Corruption and cross-border investment: Firm level evidence*. William Davidson Institute Working Papers, 494. https://econpapers.repec.org/paper/wdipapers/2002-494.htm

Soylu, Ö. B. (2019). Do foreign direct investment and savings promote economic growth in Poland? *Economics and Business Review*, *5*(4), 3–22.

Sujit, K. S., Kumar, B. R., & Oberoi, S. S. (2020). Impact of macroeconomic, governance and risk factors on FDI intensity—an empirical analysis. *Journal of Risk and Financial Management*, *13*(12), 304. https://doi.org/10.3390/jrfm13120304

Trąpczyński, P. (2013). Does foreign direct investment theory explain subsidiary performance?: A critical literature review. *Economics and Business Review*, *13*(2), 47–64. https://doi.org/10.18559/ebr.2013.2.665

Ullah, I., & Khan, M. A. (2017). Institutional quality and foreign direct investment inflows: Evidence from Asian countries. *Journal of Economic Studies*, *44*(6), 1030–1050.

Wagner, C., & Delios, A. (2023). The determinants of inward FDI in India in the 2000s. *Journal of Indian Business Research*, *15*(3), 431–465.

Wang, M. (2009). Manufacturing FDI and economic growth: Evidence from Asian economies. *Applied Economics*, *41*(8), 991–1002.

Warfield, J. N. (1974). Developing subsystem matrices in structural modeling. *IEEE Transactions on Systems, Man, and Cybernetics*, (1), 74–80.

Watson, R. H. (1978). Interpretive structural modeling—A useful tool for technology assessment? *Technological forecasting and social change*, *11*(2), 165–185.

World Investment Report. (2006). *FDI from developing and transition economies: Implications for development.* UNCTAD.

World Investment Report. (2025). *International investment in the digital economy*. UNCTAD.

Yadav, D. K., & Barve, A. (2015). Analysis of critical success factors of humanitarian supply chain: An application of interpretive structural modeling. *International Journal of Disaster Risk Reduction*, *12*, 213–225.

Yu, T., Liang, X., & Wang, Y. (2020). Factors affecting the utilization of big data in construction projects. *Journal of Construction Engineering and Management*, *146*(5), 04020032. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001807

Zhou, F., Lim, M. K., He, Y., Lin, Y., & Chen, S. (2019). End-of-life vehicle (ELV) recycling management: Improving performance using an ISM approach. *Journal of Cleaner Production*, *228*, 231–243.

PUEB PRESS

Economics and Business Review
https://www.ebr.edu.pl

Vol. 12 (1), 2026: 81-104
https://doi.org/10.18559/ebr.2026.1.2293

Political connection and corporate ESG performance: Evidence from China

Congming Ding[1]   Xuezhenzi Hu[2]

Abstract

ESG has attracted widespread attention in China's capital markets. This study investigates the impact of corporate executives' political connections on firms' ESG performance in China. Using panel data from A-share listed companies between 2009 and 2022, this study empirically tests whether politically connected executives influence ESG ratings. The results show a significant positive association between political connections and ESG scores. Mechanism analysis reveals that such connections improve ESG performance by enhancing media scrutiny, alleviating financing constraints, and increasing access to government subsidies. To address endogeneity concerns, we employ Two-Stage Least Squares (2SLS) regression, confirming the robustness of the findings. These results highlight the role of political capital in promoting sustainable corporate practices.

Keywords

- political connection
- ESG performance
- media monitoring
- financing constraints

JEL codes: G3, M21, Q56.

Article received 20 May 2025, accepted 20 January 2026.

Suggested citation: Ding, C., & Hu, X. (2026). Political connection and corporate ESG performance: Evidence from China. *Economics and Business Review*, *12*(1), 81-104. https://doi.org/10.18559/ebr.2026.1.2293

[1] School of Public Policy and Administration, Chongqing University, 174 Shazheng Street, Shapingba, Chongqing, 400044, China, hayekting@126.com, https://orcid.org/0000-0002-4377-8037.

[2] School of Public Policy and Administration, Chongqing University, 174 Shazheng Street, Shapingba, Chongqing, 400044, China, corresponding author: 202401021075T@stu.cqu.edu.cn, https://orcid.org/0009-0003-5511-8233.

Introduction

Sustainable development has become a global priority, with ESG standards gaining prominence and rising to become core principles in corporate strategy and investment (Cao et al., 2023; Gonçalves et al., 2023; Pedersen et al., 2021). As the world's largest emitter of carbon dioxide, China accounted for 32.1% of global emissions from combustible fuels in 2023 (IEA, 2024), thus bearing critical responsibility in global climate mitigation efforts. At the 2020 United Nations General Assembly, China made a clear commitment to carbon emissions peaking before 2030 and carbon neutrality being achieved before 2060. Chinese enterprises, operating within this macro context of immense pressure and commitment, must play a decisive role in this transition. Against this backdrop, understanding the drivers of corporate ESG performance within China's unique institutional environment becomes critically important.

In emerging economies like China, informal institutions such as political connections often serve as vital supplements for market-supporting mechanisms. This is particularly pertinent in the Chinese context, characterised by a significant role of the state in resource allocation and a governance model where formal market-supporting institutions are still evolving. Under such institutional arrangements, informal networks and relationships naturally emerge as crucial channels for firms to navigate the business environment, secure resources, and interpret policy directions. While numerous scholars have examined how political connections influence corporate behaviour (Chan et al., 2012; Faccio, 2006; Florackis et al., 2023; Hu et al., 2024; Tsai et al., 2019; L. Wang & You, 2022), another group has explored the determinants of corporate ESG performance (Pedersen et al., 2021; S. Wang et al., 2025; Xie & Lv, 2022). However, a gap persists at the intersection of these two research paths. Although the resource advantages conferred by political connections have been well-documented, their impact on non-profit, long-term outcomes like ESG performance remains understudied. As a channel for resource acquisition, do political connections affect corporate ESG performance? Do they elevate or diminish ESG rating scores? What are the transmission mechanisms through which political connections affect corporate ESG performance? These questions are particularly pertinent in the Chinese context, where the government plays a central role in resource allocation and policy implementation, and corporate governance discourse is deeply embedded within a unique political framework. Examining how firms navigate and leverage political connections within such a context is therefore not merely an option but a necessity for understanding corporate behaviour.

This study aims to bridge this gap by empirically investigating the impact of political connections on corporate ESG performance in China. Based on a dataset of Chinese A-share listed companies from 2009 to 2022, this study employs a panel regression model with firm and year fixed effects to identify this relationship. Our findings indicate that political connections significantly enhance corporate ESG performance. Three primary mechanisms are identified: strengthening media monitoring, alleviating financing constraints, and increasing government subsidies. Moreover, we document significant effect heterogeneity, with the impact being more pronounced in non-SOEs, northern China, SMEs, and high-tech industries.

This paper makes a twofold contribution. Firstly, it extends ESG research beyond financial impacts (Pedersen et al., 2021) to reveal institutional drivers of corporate ESG behaviour in developing countries and their performance-enhancing mechanisms. Secondly, by exploring the link between political connections and ESG outcomes, it offers new insights into how political embeddedness shapes operational strategies in contexts characterised by distinct governance models. These findings inform policy approaches for guiding firms toward strengthened social responsibility, environmental stewardship, and sustainable development.

The paper is structured as follows: Section 1 provides a literature review; Section 2 outlines the research methods and data; Section 3 and 4 reveal the empirical results; Section 5 discusses the research findings. The final section concludes the research.

1. Literature review

ESG is a key indicator system for evaluating corporate sustainability (Cao et al., 2023; Gonçalves et al., 2023) and is widely used in such areas as investment decision-making, corporate management, and social responsibility fulfilment. ESG not only concerns long-term profitability, but also highlights a company's active commitment to and performance in environmental protection, social responsibility, and corporate governance.

Existing research on the drivers of corporate ESG performance can be broadly categorised into several streams. Firstly, a substantial corpus of literature highlights the pivotal function of formal institutions, including environmental regulations and government policies, in influencing corporate ESG conduct (Qiu & Yin, 2019). Secondly, studies have emphasised the pressures and expectations from key stakeholders, including investors, consumers, and the public, which compel firms to adopt better ESG practices to maintain legitimacy and reputation (Gu, 2024). Thirdly, the economic implications of ESG are a central focus. It has been demonstrated by scholars that strong ESG performance has the capacity to reduce financing costs (Fang & Hu, 2023) and to influence investment decisions (D. Tang & Jin, 2023). Nevertheless, the direct impact of ESG on financial performance remains a complex and contentious issue (S. Wang et al., 2022). While these studies offer valuable insights, they have predominantly focused on market incentives or formal regulatory pressures, and have largely been confined to research contexts in developed economies. In contrast, emerging economies such as China, where formal institutions are still developing and the state plays a key role in allocating resources, have not been widely studied. The influence that informal institutions, particularly political relationships at the corporate level, exercise on ESG in these countries remains a significant yet under-explored domain.

Government-business relationships refer to the multi-layered interactive networks between government and the market, government and enterprises, and even officials and entrepreneurs (Yang & Su, 2021). Through this network, enterprises gain access to greater government resource allocation and support—including credit financing facilities, tax incentives, and industry access advantages (Chan et al., 2012; L. Wang & You, 2022)—while also obtaining advance information on official transitions and policy shifts, thereby effectively mitigating risks stemming from political uncertainty (Alam et al., 2023). Simultaneously, government-business relationships elevate corporate visibility at both governmental and public levels. Media and societal scrutiny intensify regarding environmental investments and social responsibility practices. Consequently, under dual pressure from public opinion and societal expectations, companies increase environmental protection and social responsibility expenditures, yielding significant outcomes in environmental investments and performance (Chu et al., 2025; Hu et al., 2024). Of course, overreliance on political-business networks can introduce agency costs and governance challenges. They may lead to the misallocation of resources (W. Zhang et al., 2013) and weaken internal corporate governance mechanisms (S. Tang & Sun, 2014). For private firms, close political ties have been associated with lower operational efficiency, though this effect may diminish with improvements in the institutional environment (Deng & Zeng, 2009). In certain contexts, such as overcapacity industries, political connections can even trigger a "resource curse", exacerbating corporate financialisation and distracting from substantive operations (Mao et al., 2022). Therefore, maintaining necessary interactions while safeguarding against power abuse is crucial for enhancing ESG performance.

Compared with developed countries, China's ESG-themed indices started late, and the number of issuances, though initially small, has grown significantly in recent years. The development of enterprises cannot only rely on their efforts but also requires the support of the external environment, especially in developing countries where formal institutional rules and external markets are imperfect. The government's policy direction and the closeness to politics can affect the firm operation. In this context, we predict that a firm's close political connection with the government will help the firm's investment and performance on ESG.

2. Research methods and data

This study utilises a sample of firms listed on China's A-share market from 2009 to 2022.[3] ESG performance data were sourced from the SNSI ESG[4] database, which is widely recognised and applied in the market and provides valuable ESG information for companies, investors, and other stakeholders. The political connection data were derived from executive background information provided in the "Executive Personal Information" section of the Chinese Securities Market and Accounting Research Database (CSMAR). The samples were processed according to the following requirements: retain the samples with a normal listing in the observation period and exclude the abnormal samples such as Special Treatment (ST) firms; retain the rest of the companies except for the financial industry; and exclude companies with missing ESG and political connection data. Finally, unbalanced panel data consisting of 36,901 sample observations from 3932 listed companies are obtained. All continuous variables were winsorised at 1% and 99% levels to reduce the impact of outliers.

Following Xie and Lv (2022), this study adopts the SNSI ESG rating score, which includes environmental, social, and governance data, as a measure of firms' ESG performance. Consistent with Z. Liu et al. (2020), this study defines the political connection dummy variable (PC) as 1 if a firm's chairman or general manager has previously served as a government official, CPPCC member, NPC deputy, or party representative at any administrative level. Otherwise, the PC is assigned a value of 0. The emphasis on the chairman and general manager arises from their crucial roles in corporate decision-making and governance.

Drawing from the prior literature (Hu et al., 2024; Tsai et al., 2019; S. Wang et al., 2025), this study includes a set of control variables to account for other determinants of ESG performance. Firm age (Lnage) controls for accumulated experience, positing that older firms may have more mature systems for long-term ESG management. Firm size (Size), measured by the number

[3] The sample period starts from 2009 because the SNSI ESG ratings are initiated by Sino-Securities Index Information Service (Shanghai) Co. Ltd., and the earliest ratings of A-share listed companies start from 2009 and end in 2022, due to data availability and comparability.

[4] SNSI ESG Rating is based on the core connotation and development characteristics of ESG, combined with the internationally recognized ESG evaluation framework, and taking into account China's characteristics and practical experience. The rating system builds a four-level indicator system from top to bottom, including three first-level pillar indicators, 16 second-level subject indicators, 44 third-level topic indicators, and nearly 80 fourth-level underlying indicators. The rating result is a nine-grade "AAA-C" rating for the subject, and the scores of the ESG overall score, the first-level indicators, the second-level indicators, and the third-level indicators are all between the standard scores of 0–100, and the higher the scores, the better the subject's performance on the indicators. The ESG composite score is the main indicator for testing the hypotheses. In addition, the scores of the individual ESG pillars were used.

of employees, accounts for the greater resources and public scrutiny larger firms face, which can influence their ESG investments. Return on assets (ROA) is included as financially sound firms likely have more slack resources for ESG initiatives. Leverage ratio (Leverage) controls for the potential constraint high debt imposes on non-essential investments. Cash holding level (Cash) captures financial flexibility, with higher liquidity expected to facilitate ESG spending. Shareholding concentration (Sh_Conc) addresses how governance structure may either enable or impede ESG policy implementation. Growth (Growth) and capital intensity (CI) control for the firm's strategic focus and investment structure, which may affect its propensity and capacity to fund long-term ESG projects. Table 1 and Table 2 provide the descriptive statistics and the results of correlation analysis and multiple collinearity tests for the main variables used in the empirical analysis, respectively.

Table 1. Descriptive statistics

Variable type	Variables	*N*	Mean	SD	Min	Max
Dependent variable	ESG rating score (ESG)	36901	73.06	5.224	56.74	84.06
Key explanatory variable	Political connection (PC)	36901	0.303	0.459	0	1
Mechanism variables	Media monitoring (Media)	36901	0.355	0.429	–1	1
	Financing constraints (KZ)	33574	1.098	2.453	–11.33	13.66
	Government subsidies (Grant)	35909	17.07	1.985	6.399	21.28
Control variables	Firm age (Lnage)	36901	2.093	0.888	0	3.332
	Firm size (Size)	36608	5.062	9.903	0.084	70.37
	Return on assets (ROA)	36900	0.037	0.064	–0.263	0.205
	Leverage ratio (Leverage)	36900	0.427	0.210	0.052	0.940
	Cash holding level (Cash)	36900	0.206	0.147	0.016	0.704
	Shareholder concentration (Sh_Conc)	36901	53.42	15.49	19.67	88.58
	Growth (Growth)	36812	0.376	1.016	–0.682	7.209
	Capital intensity (CI)	36883	2.534	2.177	0.399	14.53

Source: SNSI ESG and CSMAR databases.

Table 1 presents ESG and PC, the key variables in our study. ESG rating scores vary significantly across companies, ranging from a minimum of 56.74 to a maximum of 84.06, with a mean of 73.06. This indicates that the overall ESG performance of the sample companies is at a relatively high level. The

Table 2. Correlation matrix and collinearity analysis

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(1) ESG	1.000												
(2) PC	0.052*	1.000											
(3) Media	0.126*	–0.033*	1.000										
(4) KZ	–0.179*	0.001	–0.077*	1.000									
(5) Grant	0.151*	0.003	0.144*	–0.024*	1.000								
(6) Lnage	–0.131*	–0.014*	–0.007	0.298*	0.083*	1.000							
(7) Size	0.164*	0.031*	0.026*	0.033*	0.341*	0.159*	1.000						
(8) ROA	0.233*	0.027*	0.067*	–0.529*	0.046*	–0.215*	0.025*	1.000					
(9) Leverage	–0.112*	0.009	–0.044*	0.633*	0.129*	0.393*	0.250*	–0.371*	1.000				
(10) Cash	0.124*	–0.047*	0.019*	–0.605*	–0.055*	–0.325*	–0.105*	0.262*	–0.436*	1.000			
(11) Sh_Conc	0.154*	0.007	0.003	–0.209*	0.058*	–0.343*	0.175*	0.213*	–0.090*	0.129*	1.000		
(12) Growth	–0.003	0.003	–0.048*	0.067*	–0.080*	0.055*	–0.064*	–0.005	0.081*	0.017*	–0.012	1.000	
(13) CI	–0.066*	0.019*	–0.055*	0.086*	–0.134*	0.100*	–0.148*	–0.189*	–0.022*	–0.027*	–0.051*	0.178*	1.000
VIF	1.11	1.01	1.04	2.83	1.20	1.33	1.34	1.51	2.01	1.66	1.23	1.06	1.13

Note: * indicates significance at the 1% level.

Source: own work.

mean value of the political connections variable (PC) is 0.303, indicating that nearly one-third of executives in listed companies possess political connections. This suggests that under China's unique institutional environment during its transition period, establishing connections with the government serves as a crucial means for enterprises to pursue development.

The correlation analysis results in Table 2 indicate that the correlation coefficients between variables are generally low. The last row of Table 2 also reports the variance inflation factors (VIF) for each variable, all of which are found to be less than 3. This demonstrates that the fixed-effects model constructed in this paper does not exhibit significant multicollinearity issues.

To investigate the relationship between political connections and corporate ESG performance, this study employs a panel regression model with firm and year fixed effects as primary empirical strategy. Our baseline model is specified as follows:

$$ESG_{it} = \alpha_0 + \alpha_1 PC_{it} + \sum_{j=1}^{k} \delta_j Control_{it} + id_i + year_t + \mu_{it} \quad (1)$$

where ESG_{it} is the firm's ESG rating scores, PC_{it} is the degree of political connection of the firm, i and t are the listed firm and the year, respectively, id_i is an individual firm fixed effect, $year_t$ is a time-fixed effect, μ_{it} is a randomised disturbance term. The coefficient α_1 captures the effect of political connections on ESG performance. Based on the research hypothesis, α_1 is expected to be positive, indicating a beneficial impact of political connections on ESG performance.

The set of control variables, $Control_{it}$ is selected based on the established literature on corporate ESG and political connections (Hu et al., 2024; Tsai et al., 2019; S. Wang et al., 2025). They include: firm age (lnage), firm size (Size), return on assets (ROA), leverage ratio (Lev), cash holding level (Cash), shareholding concentration (Shrcr), growth (growth), capital intensity (CI). The same set of controls is consistently applied across all regression models in Tables 3 to 7 to ensure the comparability and coherence of our findings.

3. Results and discussion

3.1. Baseline regression

Table 3 presents the results of the effect of political connection on ESG performance. The coefficients on PC in columns (1) through (4) are all significantly positive. Column (4) indicates that after controlling for a range of variables, adding individual and year fixed effects, and clustering for industry and prefecture, the results show that firms with at least one chairman or general manager who has had political experience tend to outperform their peers in terms of ESG performance. These results provide initial support for the underlying research hypothesis that political connections are associated with higher ESG performance.

Table 3. Baseline regression results n ESG

Variable	(1)	(2)	(3)	(4)
PC	0.590***	0.160*	0.493***	0.177**
	(0.059)	(0.090)	(0.057)	(0.080)
Lnage			–0.425***	–0.873***
			(0.035)	(0.195)
Size			0.093***	0.101***
			(0.003)	(0.014)
ROA			14.279***	3.945***
			(0.462)	(0.702)
Lev			–0.861***	–2.964***
			(0.155)	(0.427)
Cash			1.903***	0.799*
			(0.202)	(0.440)
Shrcr			0.017***	0.009
			(0.002)	(0.007)
Growth			0.075***	–0.061
			(0.026)	(0.045)
CI			0.007	–0.074*
			(0.013)	(0.042)
Constant	72.879***	73.014***	71.836***	75.027***
	(0.033)	(0.034)	(0.163)	(0.690)
Observations	36901	36741	36519	36348
R^2	0.003	0.537	0.099	0.555
Firm FE	NO	YES	NO	YES
Year FE	NO	YES	NO	YES

Notes: Clustered robust standard errors of the estimated coefficients at the industry and prefecture level are in parentheses; ***, **, and * indicate 1%, 5%, and 10% significance levels, respectively.

Source: own work.

The signs and significance of the control variables are largely consistent with theoretical expectations. For instance, Size and ROA exhibit positive and significant coefficients, suggesting that larger and more profitable firms tend to have better ESG performance, likely due to greater resources and managerial slack. The results for these controls enhance our confidence in the model specification.

3.2. Robustness tests

The possibility of reverse causality poses a significant challenge to establishing a causal interpretation of our findings. While we hypothesise that political connections (PC) enhance ESG performance, it is equally plausible that firms with superior (or inferior) ESG records may self-select into forming political ties. For instance, firms anticipating regulatory scrutiny due to poor environmental practices may proactively appoint politically connected executives to mitigate potential penalties or to gain preferential access to policy information (Florackis et al., 2023). This strategic appointment creates an endogeneity problem, as unobserved firm characteristics driving both the formation of political connections and ESG outcomes would bias the OLS estimates.

To address this identification concern, we employ an instrumental variable (IV) approach. Our instrument is the mean level of political connections of other firms within the same prefecture-level city and industry (Luo & Liu, 2019). The results of the two-stage least squares (2SLS) estimation are presented in Columns (1) and (2) of Table 4.

The first-stage *F*-statistic is well above the conventional threshold of 10, indicating a strong instrument. Most importantly, as shown in Column (2), the coefficient on the instrumented PC variable remains positive and statistically significant at the 5% level. This finding provides stronger causal evidence that political connections exert a positive influence on corporate ESG performance, even after accounting for the potential bias introduced by reverse causality.

To address potential concerns that our results are driven by evolving regulatory frameworks in the early stages of ESG development in China, we conduct a robustness check by restricting the sample period. Specifically, we shorten the time window to the period from 2015 to 2022. This choice is motivated by a key institutional change: the establishment of Green Finance Committee of the China Society of Finance in 2015, which marked a critical inflection point and the beginning of a more formalized and standardized era for ESG practices in the country (W. Wang, 2025). By focusing on this later period, we effectively test whether our main findings hold in a regulatory environment where ESG concepts became more mainstream and measurable.

Table 4. Robustness tests

	(1)	(2)	(3)	(4)
	Instrumental variable approach		Shortened time window	Alternative explanatory variable
	PC	ESG	ESG	ESG
PC_Level				0.049*
				(0.025)
PC		2.998**	0.253**	
		(1.237)	(0.108)	
IV	0.599***			
	(0.059)			
Observations	36348	36348	24791	36348
R^2			0.630	0.555
F	14.220			
Controls	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Note: The control variables are Lnage, Size, ROA, Leverage, Cash, Sh_Conc, Growth, CI.

Source: own work.

The regression results, presented in Column (3) of Table 4, show that the coefficient of PC remains positive and statistically significant. Notably, the point estimate is larger than that in the benchmark regression. This finding not only confirms the robustness of the positive relationship between political connection and ESG performance but also provides suggestive evidence that the role of political connections has become more pronounced as national and public attention to ESG has intensified post-2015. The consistency of our core result across different sample periods strengthens the credibility of our primary conclusion.

Furthermore, we alleviate concerns regarding any measurement error by employing an alternative measure for our key independent variable–political connection. Specifically, we use the political connection level (PC_Level) of corporate executives to replace the previous political connection (PC) (Fan et al., 2007). As shown in Column (4) of Table 4, the coefficient on PC_Level remains positive and significant at the 10% level, reinforcing that our findings are not sensitive to the specific operationalisation of the political connection construct.

4. Mechanism analysis and extended analysis

4.1. Mechanism analysis

Close political connections can bring potential various resources to enterprises, solve the problem of resource demand in the process of enterprise development, and then promote the operation and development of enterprises, which also helps to improve the ESG performance of enterprises; at the same time, close political connections also put the enterprises under more open and transparent media supervision, which motivates them to engage in more ESG behaviours. This study explores the impact of political connections on corporate ESG performance through three channels: media monitoring, alleviating financing constraints, and providing political resources. To further verify whether there is a mechanism effect, the following model (2) is designed for these purposes:

$$Mechanism_{it} = \beta_0 + \beta_1 PC_{it} + \sum_{j=1}^{k} \delta_j Control_{it} + id_i + year_t + \mu_{it} \quad (2)$$

In this specification, $Mechanism_{it}$ represents the mechanism variables, such as Media, KZ, and Grant, and the rest of the variables are explained as in equation (1).

4.1.1. Media monitoring

Legitimacy theory suggests (Al-Twaijry et al., 2003; Bednar, 2012) that the business activities of enterprises need to conform to the order and norms of society. As ESG ratings become a focus of social and national attention, firms strive to portray an environmentally friendly and responsible image through green innovation, fulfilment of social responsibility, and improvement of corporate governance levels in order to gain social acceptance and enhance legitimacy. If a corporate executive is involved in politics, he or she will inevitably receive more attention from the public. This is particularly salient in the era of digital media. From the perspective of agenda-setting theory, the media tends to pursue socially sensitive topics. As noted by Hu et al. (2024), once executives establish political connections, their information is eagerly reported, exposing them to heightened stakeholder scrutiny and reputational risk. Based on this situation, the executives of the firm, whether for optimizing the firm or their personal image (Hamdi, 2024), and under reputational pressure from media exposure (Hu et al., 2024), will pay more attention to ESG, make more investments, and take the lead in ESG performance to cater to the concerns and expectations of the outside world about the firm.

The role of media as an external governance and monitoring mechanism has been documented in various contexts, including China (Li et al., 2022). While acknowledging the state's overarching influence on traditional media in China, the media landscape is not monolithic. Two key nuances support the relevance of media monitoring: firstly, at the operational level, local media and financial news outlets often retain a degree of autonomy in reporting on specific corporate conduct, especially regarding environmental and social issues that align with national policy priorities. Secondly, and more critically, the rise of social media, such as TikTok and RedNote, has created an alternative and vibrant arena for public oversight. Numerous studies have shown that social media in China acts as a powerful channel for exposing corporate environmental violations, labour disputes, and governance scandals, generating substantial public pressure that compels firms to respond (T. Ding & Chen, 2025; D. Liu, 2025; Sun, 2021; Y. Zhang, 2023). Therefore, for a politically connected firm, maintaining a positive image across both traditional and social media becomes a strategic imperative to manage reputational risk.

Furthermore, positive media coverage itself can be a valuable asset. It enhances corporate image and reputation, creating a consistent positive signal that builds stakeholder confidence. This improved reputation can, in turn, help alleviate corporate financing constraints (Capelle-Blancard & Petit, 2019)and managers can no longer ignore their impact on firm value. In this paper, we investigate the extent and the determinants of the stock market's reaction following ordinary news related to environmental, social and governance issuesthe so-called ESG factors. To that purpose, we use an original database provided by Covalence EthicalQuote. Our empirical analysis is based on about 33,000 ESG news (positive or negative, forming a virtuous cycle that supports ESG investments.

The Janis-Fadner coefficient (Janis & Fadner, 1943) of media monitoring is an indicator that quantifies the net tonality of media coverage. It is based on the number of positive and negative media reports about an enterprise or event, and is used to measure the bias and intensity of monitoring. This paper uses the number of positive, negative and neutral media reports from the China Research Data Service (CNRDS) financial database to construct a media monitoring indicator using the $J-F$ coefficient. The calculation method is as follows:

$$J-F=\begin{cases}\dfrac{e^2-ec}{t^2} & \text{if} \quad e>c\\ \dfrac{ec-c^2}{t^2} & \text{if} \quad e<c\\ 0 & \text{if} \quad e=c\end{cases}$$

where e is the number of positive reports, c is the number of negative reports, and t is the sum of the number of positive and negative reports. The value of the $J-F$ coefficient ranges from –1 to 1. When there are more positive reports about an enterprise, the $J-F$ coefficient approaches 1; when there are more negative reports about an enterprise, the $J-F$ coefficient approaches –1.

Table 5 examines whether media monitoring, financing constraints, and government subsidies serve as transmission channels. The models here replace the dependent variable with each mechanism variable.

Table 5. Transmission mechanism test

	(1)	(2)	(3)	(4)
	ESG	**Media**	**KZ**	**Grant**
PC	0.177**	0.023**	–0.092***	0.141***
	(0.080)	(0.011)	(0.030)	(0.043)
Observations	36348	36348	32978	35411
R^2	0.555	0.311	0.815	0.597
Controls	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Note: The control variables are lnage, Size, ROA, Lev, Cash, Shrcr, Growth, CI.

Source: own work.

Column (2) of Table 5 shows that the coefficient of the effect of political connection on corporate media monitoring ($J-F$ coefficient) is significantly positive, indicating that politically connected firms garner more favourable net media coverage. This finding is consistent with the logic that such firms are incentivized to actively manage their media image. The positive coverage likely reflects both their efforts to showcase ESG-aligned behaviours and the media's tendency to report on prominent, politically linked entities. Firms, in turn, tend to invest more in ESG not only to maintain this favourable image but also to pre-empt potential negative scrutiny, as predicted by reputational pressure theory (Hu et al., 2024), thereby further improving their ESG rating scores.

4.1.2. Financing constraints

During the critical period of China's economic transformation and development, the capital market system is still imperfect, and the conditions for enterprises to go public are harsh, making it difficult for most enterprises to obtain all the funds needed for R&D projects through equity financing.

A good relationship with the government can significantly reduce firms' reliance on equity financing (Boubakri et al., 2012), enabling them to show potential advantages in obtaining bank loans, especially long-term loans (Chan et al., 2012; Fan et al., 2008). Firms that are well-capitalised, less costly, and have relatively fewer financing constraints are more inclined to play a greater role in the ESG sector, as these firms have more resources to devote to ESG.

This paper employs the KZ index as a proxy variable for financing constraints. Proposed by Kaplan and Zingales (1997), the KZ index aims to assess the difficulty of an enterprise's access to funds through a comprehensive analysis of a series of financial indicators. The calculation method of the KZ index is relatively complex and involves financial data, such as an enterprise's total assets at the beginning of the period (Asset), operating net cash flow (OCF), cash dividends (Dividends), cash holding level (Cash), asset-liability ratio (Leverage), and Tobin's Q. The construction method is as follows:

$$KZ = b1 \times \frac{OCF}{Asset} + b2 \times Leverage + b3 \times \frac{Dividends}{Asset} + b4 \times \frac{Cash}{Asset} + b5 \times \text{Tobin's}Q$$

In the formula, OCF, Dividends, and Cash are all normalised using total assets at the beginning of the period, Leverage is the asset-liability ratio, and Tobin's Q is the Tobin's Q value, which is the ratio of the market value of the enterprise to the replacement cost of its assets. The higher the KZ index, the more severe the financing constraints faced by the enterprise and the more limited its financing ability. Referring to the study by Lamont et al. (2001), the values of $b1 \sim b5$ are −1.001909, 3.139193, −39.3678, −1.314759, 0.2826389.

Column (3) of Table 5 shows that political connection has a significantly negative effect on the financing constraints index of firms. Given that a larger financing constraint index implies a higher degree of financing constraint, this result suggests that firms with political connections face lower financing constraints compared to firms without political connections and, thus, are able to invest more in ESG.

4.1.3. Political resources

From the perspective of the enterprise's development, the formation of political connections makes it easier for the enterprise to obtain subsidies and capital (Tsai et al., 2019), which promotes its expansion and growth and directly affects its investment behaviour. According to signalling theory, government subsidies, as a signal of favourable investment, can help enterprises attract more external resources and have a positive impact on their financial and operational activities. Stakeholder theory emphasises that enterprises will comprehensively consider and respect the legitimate interests of all stakeholders in the management process (Donaldson & Preston, 1995). Therefore, firms with stronger political connections are usually more responsive to ESG policy orientation in order to maintain a good cooperative relationship with the government.

Government agencies provide various forms of financial assistance or support to enterprises. These subsidies usually take the form of tax reductions or exemptions, grants, or other forms of non-monetary benefits, and are intended to support enterprises' development, innovation, poverty alleviation, social responsibility, and other activities. The total amount of government subsidies received across all enterprise projects constitutes the enterprise's government subsidies, represented by the variable Grant.

Column (4) of Table 5 shows that the effect of political connections on government subsidies is significantly positive, which suggests that politically connected firms are able to obtain more government support and directly promote their ESG investments. These findings corroborate our theoretical framework that political connections facilitate ESG investment through these three pathways.

4.2. Extended analysis

4.2.1. Structural analysis

The effects of political connections on different factors of ESG may differ, according to the composition of the three pillars of ESG in the SNSI ESG ratings geared towards exploring the effect of political connection on the three pillars, respectively. The regression results are shown in Table 6.

Table 6. Structural analysis

	(1)	(2)	(3)	(4)
	ESG	**E**	**S**	**G**
PC	0.177**	0.108	0.328*	0.070
	(0.080)	(0.123)	(0.195)	(0.147)
Observations	36348	36348	36348	36348
R^2	0.555	0.612	0.574	0.499
Controls	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Note: The control variables are lnage, Size, ROA, Lev, Cash, Shrcr, Growth, CI.

Source: own work.

The coefficient of the impact of political connection on CSR in column (3) is significantly positive, and the coefficients of the impact of political connection on corporate environmental protection and corporate governance in columns (2) and (4) are not significant. The above regression results illustrate the impact of political connection on ESG, mainly through improving corporate social responsibility.

4.2.2. Heterogeneity analysis

Ownership structure, regional differences, firm size, and industry characteristics can be considered to play an important role in the mechanisms that influence the role of political connections on firms' ESG performance. This is because significant differences exist in resource acquisition and management capabilities between state-owned and non-state-owned firms (Liang et al., 2012), which may lead to different degrees of politically involved executives' roles in different environments. There are significant differences in political-business relationships in China's southern and northern regions,[5] with firms in the north having stronger political-business relationships and the south being more market-oriented, with the government establishing generally weak ties with firms (C. Ding et al., 2024), which may lead to differential effects of political connections in different regions. Large-scale firms, which are rarely resource-constrained, experience a limited marginal effect from additional policy resources on their ESG performance. In contrast, for small-scale firms, such resources represent a critical input, yielding a substantially more significant impact. There are significant differences between high-tech and non-high-tech industries in terms of their development models, resource dependence, and regulatory environments, which makes it possible for political connections to impact differently on the ESG performance of the two types of industries. Based on this, categorical regression analysis is carried out on each of these dimensions to explore the differential impact of political connections.

Table 7 explores the heterogeneous effects across different sub-samples. The stark contrast in the magnitude and significance of the PC coefficient across these groups reveals that the effect of political connections is not uni-

[5] According to the National Bureau of Statistics and economic geography to the Qinling and Huaihe River as the boundary of the division of the North and South standards, excluding Hong Kong, Macao and Taiwan, this paper treats Jiangsu, Zhejiang, Shanghai, Anhui, Fujian, Jiangxi, Hubei, Hunan, Guangdong, Guangxi Zhuang Autonomous Region, Hainan, Sichuan, Chongqing, Guizhou, Yunnan and Tibet Autonomous Region as the southern region; Beijing, Tianjin, Hebei Province, Shanxi, Inner Mongolia Autonomous Region, Liaoning, Jilin, Heilongjiang, Shandong, Henan, Shaanxi, Gansu, Qinghai, Ningxia Hui Autonomous Region and Xinjiang Uygur Autonomous Region are treated as the northern region.

Table 7. Heterogeneity analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	SOE	Non-SOE	North	South	Small	Large	High-tech	Non-High-tech
	ESG	ESG	ESG	ESG	ESG	ESG	ESG	ESG
PC	0.101	0.227**	0.407***	0.080	0.282**	0.097	0.259**	0.029
	(0.215)	(0.092)	(0.145)	(0.096)	(0.115)	(0.100)	(0.107)	(0.111)
Observation	4446	31108	11315	25033	17939	18026	19570	16726
R^2	0.628	0.571	0.569	0.551	0.596	0.573	0.549	0.586
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES

Note: The control variables are Lnage, Size, ROA, Leverage, Cash, Sh_Conc, Growth, CI.

Source: own work.

form. It is particularly pronounced in non-state-owned firms, northern local firms, small-scale firms, and high-tech firms, which aligns with our contention that these firms rely more heavily on informal institutions like political ties to compensate for their inherent resource or institutional disadvantages.

5. Discussion

This study provides compelling evidence that political connections play a significant role in the ESG performance of Chinese enterprises. Our findings reveal the complex role that informal institutions play in promoting corporate sustainability within emerging economies.

The literature presents the dual nature of political connections, highlighting both their potential for resource misallocation (W. Zhang et al., 2013) and their role in resource provision (Chan et al., 2012; L. Wang & You, 2022). Our findings reconcile this tension by demonstrating that, within the specific context of China's top-down sustainability drive, political connections can be channelled to drive substantive ESG investment. By strengthening media monitoring, easing financing constraints, and increasing government subsidies, political connections simultaneously bolster firms' ESG capabilities through resource allocation and strengthen their incentives via external pressure. This dual effect demonstrates that within the Chinese context, political connections transcend mere resource acquisition tools to become vehicles for value creation aligned with national strategies such as green development. The mechanisms we identify forge a direct link between the established literature and our findings. The roles of eased financing constraints and increased government subsidies provide empirical validation for the resource-based view of political connections, showing how accessed resources (Chan et al., 2012) are strategically allocated to meet long-term ESG goals. Concurrently, the significant role of media monitoring integrates this literature with stakeholder and legitimacy theories (Gu, 2024), revealing that the heightened visibility of connected firms triggers a pressure to conform to societal expectations on ESG, a nuanced channel previously underexplored in the political connections domain.

This study offers multiple theoretical contributions: firstly, by empirically demonstrating that acquired resources can be allocated towards long-term non-financial objectives such as ESG, it extends the boundaries of the resource-based theory of political connections, transcending traditional research focused on short-term financial gains. Secondly, by empirically validating the specific causal pathways between these elements, this study bridges the gap between political connectedness and ESG literature. While existing ESG research often emphasizes formal regulations or market pressures, this study reveals that political embeddedness, though an informal factor, is also a crucial driver of corporate production and operations.

Heterogeneity findings further deepen our understanding. The stronger effects observed in non-state-owned enterprises, SMEs, and high-tech firms indicate that political connections serve as a compensatory mechanism for companies lacking the inherent resource advantages of state-owned enterprises or operating in non-priority national development sectors. This aligns with institutional theory, where firms proactively employ political strategies to secure legitimacy and resources within specific institutional vacuums. This finding provides a critical nuance to the existing narrative. While prior research has often associated close political ties with inefficiencies or a "resource curse" for private firms (Deng & Zeng, 2009; Mao et al., 2022), our results suggest that under strong national policy signals, such as the promotion of ESG, these very connections can be transformed. For resource-constrained firms, they become a strategic tool not for rent-seeking but for overcoming barriers to making substantive sustainability investments, thereby enhancing their legitimacy and long-term resilience.

Our findings offer significant practical and policy implications. Regulators should leverage these findings to target supervision and design incentives. Policies should reward demonstrable ESG outcomes rather than political affiliations to ensure connections facilitate genuine sustainable development. Managers, particularly in non-SOEs, SMEs, and high-tech industries where the benefit is strongest, should proactively manage their political relationships not as an end in itself but as a strategic asset to facilitate long-term ESG investments. For example, they can leverage political connections to secure necessary funding for green technologies or to better understand evolving regulatory expectations. However, they must also be mindful of the associated reputational risks and increased public scrutiny.

Conclusions

This study establishes a significant positive association between the political connections of Chinese corporate executives and ESG performance. Through rigorous empirical testing, we identify three core mechanisms: strengthened media monitoring, alleviated financing constraints, and increased government subsidies. We also identify key boundary conditions related to ownership, region, firm size, and industry.

This study's principal theoretical contribution lies in integrating political strategy literature with corporate sustainability theory, revealing how informal institutions shape ESG outcomes in emerging markets. Practically, it provides regulators with clear guidance on directing corporate political capital towards public goods, while also charting a course for managers seeking to leverage networks for long-term value creation.

In sum, this paper emphasises that in developing countries with imperfect formal systems, corporate ESG performance is not solely attributable to the individual behaviour of micro-enterprises but is mixed with the influence of complex factors, including political factors.

Despite the contributions of this study, several limitations remain, presenting opportunities for future research. For instance, although we employed instrumental variables to address endogeneity issues, the possibility of omitted variable bias cannot be entirely ruled out. Subsequent research may incorporate more granular firm-level variables or adopt alternative identification strategies. Regarding causal identification, future studies could employ more direct event study methods to discern the impact of political connections on corporate ESG performance. In terms of data, while the political relationship metrics in this study were meticulously constructed, future research could leverage more refined datasets to broaden the research dimensions. This would not only capture the existence of relationships but also allow for a more precise measurement of their intensity, thereby reducing measurement errors and enhancing the reliability of causal inferences.

References

Alam, A. W., Farjana, A., & Houston, R. (2023). State-level economic policy uncertainty (EPU) and firm financial stability: Is there any political insurance? *Economics Letters*, *225*, 111027. https://doi.org/10.1016/j.econlet.2023.111027

Al-Twaijry, A. A., Brierley, J. A., & Gwilliam, D. R. (2003). The development of internal audit in Saudi Arabia: An institutional theory perspective. *Critical Perspectives on Accounting*, *14*(5), 507–531. https://doi.org/10.1016/S1045-2354(02)00158-2

Bednar, M. K. (2012). Watchdog or lapdog? A behavioral view of the media as a corporate governance mechanism. *Academy of Management Journal*, *55*(1), 131–150. https://doi.org/10.5465/amj.2009.0862

Boubakri, N., Guedhami, O., Mishra, D., & Saffar, W. (2012). Political connections and the cost of equity capital. *Journal of Corporate Finance*, *18*(3), 541–559. https://doi.org/10.1016/j.jcorpfin.2012.02.005

Cao, M., Duan, K., & Ibrahim, H. (2023). Green investments and their impact on ESG ratings: An evidence from China. *Economics Letters*, *232*, 111365. https://doi.org/10.1016/j.econlet.2023.111365

Capelle-Blancard, G., & Petit, A. (2019). Every little helps? ESG news and stock market reaction. *Journal of Business Ethics*, *157*(2), 543–565. https://doi.org/10.1007/s10551-017-3667-3

Chan, K. S., Dang, V. Q. T., & Yan, I. K. M. (2012). Chinese firms' political connection, ownership, and financing constraints. *Economics Letters*, *115*(2), 164–167. https://doi.org/10.1016/j.econlet.2011.12.008

Chu, P., Hou, X., & Xie, G. (2025). Political connections and bias in ESG news. *Journal of Asian Economics*, *97*, 101877. https://doi.org/10.1016/j.asieco.2025.101877

Deng, J., & Zeng, Y. (2009). Can Political Connection Improve the Performance of Private Enterprises. *China Industrial Economics*, *02*, 98–108. https://link.oversea.cnki.net/doi/10.19581/j.cnki.ciejournal.2009.02.010

Ding, C., Wang, C., & Chen, H. (2024). Optimize the doing business environment to promote balanced economic development between North and South: Evolution from limited economic access order to open economic access order. *Journal of Quantitative & Technological Economics*, *41*(1), 110–129. https://www.jqte.net/sljjjsjjyj/ch/reader/view_abstract.aspx?flag=1&file_no=20240106&journal_id=sljjjsjjyj

Ding, T., & Chen, L. (2025). Research on the influence mechanism of social media communication on ESG performance of listed companies. *China Journal of Commerce*, *34*(13), 149–153. https://link.oversea.cnki.net/doi/10.19699/j.cnki.issn2096-0298.2025.13.149

Donaldson, T., & Preston, L. E. (1995). The stakeholder theory of the corporation: Concepts, evidence, and implications. *Academy of Management Review*, *20*(1), 65–91. https://doi.org/10.5465/AMR.1995.28081377

Faccio, M. (2006). Politically connected firms. *American Economic Review*, *96*(1), 369–386. https://www.jstor.org/stable/30034371

Fan, J. P., Rui, O. M., & Zhao, M. (2008). Public governance and corporate finance: Evidence from corruption cases. *Journal of Comparative Economics*, *36*(3), 343–364. https://doi.org/10.1016/j.jce.2008.05.001

Fan, J. P., Wong, T. J., & Zhang, T. (2007). Politically connected CEOs, corporate governance, and Post-IPO performance of China's newly partially privatized firms. *Journal of Financial Economics*, *84*(2), 330–357. https://doi.org/10.1016/j.jfineco.2006.03.008

Fang, X., & Hu, D. (2023). Corporate ESG performance and innovation-evidence from A-share listed companies. *Economic Research Journal*, *58*(2), 91–106.

Florackis, C., Fu, X., & Wang, J. (2023). Political connections, environmental violations and punishment: Evidence from heavily polluting firms. *International Review of Financial Analysis*, *88*, 102698. https://doi.org/10.1016/j.irfa.2023.102698

Gonçalves, T. C., Barros, V., & Avelar, J. V. (2023). Environmental, social and governance scores in Europe: What drives financial performance for larger firms? *Economics and Business Letters*, *12*(2), 121–131. https://doi.org/10.17811/ebl.12.2.2023.121-131

Gu, J. (2024). Investor attention and ESG performance: Lessons from China's manufacturing industry. *Journal of Environmental Management*, *355*, 120483. https://doi.org/10.1016/j.jenvman.2024.120483

Hamdi, B. (2024). Political connections of independent directors and earnings quality: The case of French firms. *Economics Letters*, *244*, 111996. https://doi.org/10.1016/j.econlet.2024.111996

Hu, Y., Cheng, H., & Li, H. (2024). Political connections and corporate ESG performance. *Sustainable Economies*, *2*(2), 93. https://doi.org/10.62617/se.v2i2.93

IEA. (2024). *CO_2 emissions in 2023*. International Energy Agency. https://www.iea.org/reports/co2-emissions-in-2023

Janis, I. L., & Fadner, R. H. (1943). A coefficient of imbalance for content analysis. *Psychometrika*, *8*(2), 105–119. https://doi.org/10.1007/BF02288695

Kaplan, S. N., & Zingales, L. (1997). Do investment-cash flow sensitivities provide useful measures of financing constraints? *Quarterly Journal of Economics*, *112*(1), 169–215. https://doi.org/10.1162/003355397555163

Lamont, O., Polk, C., & Saaá-Requejo, J. (2001). Financial constraints and stock returns. *Review of Financial Studies*, *14*(2), 529–554. https://doi.org/10.1093/rfs/14.2.529

Li, Z., Shao, Y., Li, Z., & Li, S. (2022). ESG information disclosure,media supervision and corporate financing constraints. *Scientific Decision Making*, *7*, 1–26.

Liang, X., Lu, X., & Wang, L. (2012). Outward internationalization of private enterprises in China: The effect of competitive advantages and disadvantages compared to home market rivals. *Journal of World Business*, *47*(1), 134–144. https://doi.org/10.1016/j.jwb.2011.02.002

Liu, D. (2025). The impacts faced by corporate reputation management in the social media environment and corresponding countermeasures. *Proceedings of the Second Academic Symposium on Intelligent Engineering and Economic Development*, 4, 513–516. https://link.oversea.cnki.net/doi/10.26914/c.cnkihy.2025.036960

Liu, Z., Zhang, Y., & Zhang, L. (2020). Anti-corruption and corporate risk-taking—on the improvement of business-politics relationship. *Research on Economics and*

Management, *41*(5), 107–124. https://link.oversea.cnki.net/doi/10.13502/j.cnki.issn1000-7636.2020.05.008

Luo, X., & Liu, W. (2019). Political connection and firm's environmental violation punishment: Sheltering or monitoring—evidence from IPE database. *Journal of Shanxi University of Finance and Economics*, *41*(10), 85–99. https://kns.cnki.net/kcms2/article/abstract?v=CeTIW1Ci4b0KXtz1tBMECdw4Jj9YcT_dW7t3Ahv51FJ2SKMnSPZPyO2tWR_8kpOdG6q9LCLIYfe9kBw2UvTfgnNqQsZF1MUrR0vp0E_-1PvNxKy8-Pt50h4lyJUZmJVqYqvIlSCIk6E_AMpjAadDSCpa7FnHF3e22n9wTZmab_o=&uniplatform=NZKPT&language=CHS

Mao, Z., Tan, M., & Zeng, S. (2022). Political connection exacerbates corporate financialization? Evidence from China's private enterprises. *Journal of Hunan University*, *36*(2), 35–42. https://link.oversea.cnki.net/doi/10.16339/j.cnki.hdxbskb.2022.02.006

Pedersen, L. H., Fitzgibbons, S., & Pomorski, L. (2021). Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics*, *142*(2), 572–597. https://doi.org/10.1016/j.jfineco.2020.11.001

Qiu, M., & Yin, H. (2019). An analysis of enterprises' financing cost with ESG performance under the background of ecological civilization construction. *Journal of Quantitative and Technical Economics*, *36*, 108–123. https://www.jqte.net/sljjjsjjyj/ch/reader/view_abstract.aspx?flag=1&file_no=20190307&journal_id=sljjjsjjyj

Sun, R. (2021). *Research on the impact mechanism of public participation on enterprise emission control from the perspective of new media*.

Tang, D., & Jin, X. (2023). ESG Investor behavior and related research frontiers in the context of carbon neutrality: Overview and extension. *Economic Research Journal*, *58*(09), 190–208.

Tang, S., & Sun, Z. (2014). The political connections, the CEO's salary, and firm's future management performances. *Management World*, *5*, 93–105. https://link.oversea.cnki.net/doi/10.19744/j.cnki.11-1235/f.2014.05.008

Tsai, L. C., Zhang, R., & Zhao, C. (2019). Political connections, network centrality and firm innovation. *Finance Research Letters*, *28*, 180–184. https://doi.org/10.1016/j.frl.2018.04.016

Wang, L., & You, K. (2022). The impact of political connections on corporate tax burden: Evidence from the Chinese market. *Finance Research Letters*, *47*, 102944. https://doi.org/10.1016/j.frl.2022.102944

Wang, S., Hu, H., & Fan, H. (2025). Does supply chain finance improve firms' ESG performance? *Economics Letters*, *247*, 112098. https://doi.org/10.1016/j.econlet.2024.112098

Wang, S., Tian, Y., & Dang, L. (2022). ESG implementation, competition strategy and financial performance of industrial enterprises. *Accounting Research*, *3*, 77–92. https://doi.org/10.3969/j.issn.1003-2886.2022.03.006

Wang, W. (2025). Ten years of green finance development in China. *China Finance*, (13), 69–71.

Xie, H., & Lv, X. (2022). Responsible multinational investment: ESG and OFDI in China. *China Finance and Economic Review*, *14*(3), 3–22. https://doi.org/10.1515/cfer-2025-0013

Yang, L., & Su, S. (2021). Dynamic mechanism of formation of new type of government-business relationship—based on provincial panel data of A-share listed companies in China. *Journal of Beijing University of Posts and Telecommunications*, *23*(1), 41–50. https://journalsk.bupt.edu.cn/CN/Y2021/V23/I1/41

Zhang, W., Zhang, S., & Li, B. (2013). Political connections, the characteristics of enterprises' M&A and the performance of M&A. *Nankai Business Review*, *16*(02), 64–74.

Zhang, Y. (2023). *Internet public opinion, agency costs, and corporate performance: A research from the perspective of dual principal-agent*. Xidian University.

PUEB PRESS
Economics and Business Review
https://www.ebr.edu.pl
Vol. 12 (1), 2026: 105-133
https://doi.org/10.18559/ebr.2026.1.2746

Liquidity risk and liquidity timing in the cross-section of Indian equity mutual fund returns

Suresh Kumar[1]   Hyder Ali[2]

Abstract

This study examines how aggregate market liquidity influences the cross-section of Indian equity mutual fund returns through two mechanisms: (1) funds' long-run exposure to liquidity risk, and (2) managers' time-varying liquidity timing. Using a comprehensive sample from 2007–2024, we estimate rolling liquidity betas, form portfolios sorted by liquidity exposure, and compute a high-minus-low liquidity-beta return spread. The liquidity premium is positive and economically meaningful in tranquil and recovery regimes, but weakens or vanishes during systemic stress, consistent with state-dependent liquidity pricing. Adding a traded equity-liquidity factor to standard benchmarks explains a meaningful portion of the spread, while an independently constructed timing factor captures an additional 55%–64%, highlighting the importance of conditional beta management. Timing effects are concentrated among high-liquidity-beta funds, smoothing returns in normal markets but offering limited protection in crises. Findings are robust to alternative benchmarks, flow-adjusted timing specifications, and post-COVID subperiod definitions.

Keywords

- liquidity risk
- liquidity timing
- asset pricing
- mutual funds
- emerging markets

JEL codes: G11, G12, E44, E47, C53.

Article received 2 November 2025, accepted 3 March 2026.

Suggested citation: Kumar, S., & Ali, H. (2026) Liquidity risk and liquidity timing in the cross-section of Indian equity mutual fund returns. *Economics and Business Review*, *12*(1), 105-133. https://doi.org/10.18559/ebr.2026.1.2746

[1] Department of Business Administration, Sukkur IBA University, Sukkur 65200, Pakistan, corresponding author: sureshkumar@iba-suk.edu.pk, https://orcid.org/0000-0002-1655-5452.

[2] Department of Business Administration, Sukkur IBA University, Sukkur 65200, Pakistan, hyder.ali.phd@gmail.com, https://orcid.org/0000-0003-2820-8455.

Introduction

Illiquidity is expected to command a return premium, compensating investors for trading frictions, execution risk, and the difficulty of converting positions into cash when liquidity is scarce. Foundational work formalises two complementary channels: a level effect, whereby fewer liquid assets earn higher average returns, and a covariance effect, whereby innovations in aggregate liquidity are priced as systematic risk (Acharya & Pedersen, 2005; Amihud & Mendelson, 1986; Liu, 2006; Pástor & Stambaugh, 2003). There is substantial literature documenting these channels across asset classes and international markets (Chaieb et al., 2021; Dang & Nguyen, 2020; Kang et al., 2019; Zhu et al., 2023), underscoring liquidity as a core, non-diversifiable state variable for expected returns and risk premia.

Liquidity risk becomes most salient precisely when diversification is least effective. During systemic episodes, market-wide liquidity deteriorates, trading costs rise, and funding constraints bind—conditions under which exposures to liquidity shocks can amplify losses. Evidence from 2007–2009 shows that liquidity-sensitive strategies suffered disproportionately during the Global Financial Crisis (Idzorek et al., 2012; Lou & Sadka, 2011). The COVID-19 shock renewed attention to the link between open-end fund structure, redemption pressure, and market fragility: flows can force asset sales into illiquid markets, propagating price dislocations and weakening the usual relation between liquidity and compensation (Jiang et al., 2022; Y. Ma et al., 2022). Parallel evidence connects liquidity premia to funding markets, currencies, and broader anomaly returns (Bechtel et al., 2023; Söderlind & Somogyi, 2025; Virk & Butt, 2022), reinforcing the view that liquidity is priced primarily through its behaviour in bad states (X. Ma et al., 2021; Wu, 2019).

For mutual funds, the empirical picture is nuanced because funds can both bear liquidity risk and adjust it over time. Prior studies shows that funds with higher liquidity betas tend to earn higher average returns (Dong et al., 2019; Foran & O'Sullivan, 2014), consistent with compensation for bearing systematic liquidity risk. At the same time, static exposures alone rarely explain observed performance differences: investment style, flow sensitivity, and investor composition can affect both portfolio choice and the extent to which funds transmit or absorb liquidity shocks. Studies across delegated portfolio management suggest that timing ability—systematically altering exposures in response to liquidity states—may exist but is concentrated and difficult to identify cleanly from mechanical trading induced by flows (Aiken & Kang, 2023; Bodson et al., 2013; Wattanatorn et al., 2020).

These considerations motivate a sharper question: Do equity mutual funds merely load on liquidity risk, or can they also time liquidity conditions in a way that is distinct from flow-induced trading? A natural empirical framework allows market exposure to vary with liquidity deviations (Cao et al., 2013). If timing is systematic, it should be measurable in fund-level state-dependent exposures, and it should have investable implications for cross-sectional spreads.

We revisit this issue for Indian equity mutual funds. There are several reasons why India is a particularly informative setting for liquidity risk and state-dependent beta management for institutional reasons that map directly into liquidity exposure and timing. Firstly, the mutual fund sector has expanded rapidly: industry assets under management (AUM) grew to INR 53.4 lakh crore by March 2024, the number of unique mutual fund investors rose to about 4.5 crore, and mutual fund purchases became predominantly digital (about 89%–90% of transactions by FY2024). These features can support persistent inflows in normal times yet more synchronised redemptions during stress, tightening constraints precisely when funds would need to rebalance (AMFI–CRISIL, 2024). Secondly, Indian equities trade predominantly in electronic, order-driven limit order books; prior evidence documents commonality in liquidity on the National Stock Exchange, implying that market-wide liquidity shocks can be pervasive and rapidly transmitted (G. Kumar & Misra, 2018). Thirdly, the institutional environment has evolved through regulatory reforms that standardise fund categories and constrain style drift, strengthening cross-fund comparability and the interpretation of cross-sectional sorts (Securities and Exchange Board of India, 2017). Finally, crisis-era interventions underscore the perceived systemic importance of mutual fund liquidity transformation: the Reserve Bank of India opened a special liquidity facility for mutual funds in April 2020 amid redemption-related strains, and FSAP-style assessments discuss liquidity risks in investment funds (IMF, 2025; Reserve Bank of India, 2020).

The sample period spans heterogeneous stress and recovery episodes (the GFC, the COVID-19 shock, and the post-pandemic years). Despite its economic importance, systematic evidence on liquidity risk premia and liquidity timing in Indian equity mutual funds remains limited, leaving open whether the conditional liquidity–return relation documented in developed markets extends to an emerging market where liquidity dynamics, investor composition, and intermediation frictions differ materially from the canonical US benchmark.

Our empirical design separates static exposure from dynamic timing. First, we estimate each fund's liquidity beta using rolling windows and form quintile portfolios sorted on liquidity exposure. We then construct a high-minus-low liquidity-beta spread (HMLiq) as a baseline measure of the liquidity premium. Next, we estimate liquidity-timing specifications in which market beta varies with liquidity states (Cao et al., 2013), and we construct an investable Timer factor from lagged timing-coefficient sorts. This factor-based decomposition allows us to quantify how much of the HMLiq spread reflects standard risk compensation and how much is attributable to systematic, state- dependent beta management.

Three findings summarise the results. Firstly, HMLiq is positive and statistically reliable in tranquil and recovery regimes but attenuates or disappears during systemic stress, consistent with conditional liquidity pricing. Secondly, augmenting benchmark models with a traded equity liquidity factor explains a meaningful share of HMLiq, while the independently constructed Timer factor absorbs an additional 55%–64% of the spread, indicating that timing is a first-order component of liquidity-based return differentials. Thirdly, timing effects are concentrated among high-liquidity-beta funds: they smooth return paths in normal markets but provide limited protection under acute stress, which is consistent with the broader view that liquidity transformation is most fragile precisely when aggregate liquidity deteriorates.

The paper's contribution is threefold. Firstly, we provide systematic evidence on liquidity exposure and liquidity timing in an emerging-market mutual fund setting, documenting a regime-dependent liquidity premium over 2007–2024. Secondly, by con- structing an investable Timer factor from lagged timing signals, we operationalise a timing channel and quantify its role in explaining the cross-sectional liquidity spread. Thirdly, we align inference with contemporary fund-performance standards and identification concerns raised in the literature: we evaluate robustness under modern factor benchmarks (Carhart-4 and Fama–French-5), incorporate flow-based controls and standard fund-characteristic controls (size, turnover, fees, and age) to distinguish timing from mechanically-induced beta shifts, refine the post-COVID segmentation, and supplement crisis-window inference with daily evidence for early 2020. Together, these steps strengthen the interpretation of timing as a distinct mechanism and clarify the conditions under which liquidity compensation is economically meaningful.

The remainder of the paper is organised as follows: Section 1 reviews the related literature; Section 2 describes the data, sample construction, and variable definitions; Section 3 presents the main results and robustness analyses; Section 4 discusses implications and mechanisms and last section concludes.

1. Literature review

This study sits at the intersection of (1) liquidity as a priced state variable in asset markets and (2) liquidity transformation and fragility in open-ended delegated portfolios. We first review the asset-pricing foundations of liquidity risk, then summarize evidence on liquidity premia and liquidity management in mutual funds, and finally position our contribution—liquidity exposure versus liquidity timing—within the emerging debate on crisis-state behaviour and fund resilience.

1.1. Liquidity risk

Liquidity risk—the possibility that trades cannot be executed quickly and at low cost without materially affecting prices—is central to both asset pricing and financial stability. Two channels underpin its role in expected returns. The first is a level channel: less liquid securities earn higher average returns as compensation for trading frictions, inventory risk, and delayed execution (Amihud & Mendelson, 1991; Lee et al., 2022; Liu, 2006). The second is a covariance channel: innovations in aggregate market liquidity are priced as a systematic risk factor, so portfolios with higher exposure (liquidity beta) require additional compensation (Acharya & Pedersen, 2005; X. Ma et al., 2021; Pástor & Stambaugh, 2003; Shih & Su, 2016). Liquidity-augmented factor frameworks confirm that the covariance channel is particularly salient in downturns and high-volatility regimes, consistent with liquidity behaving as a bad-state risk that is difficult to diversify away from (X. Ma et al., 2021; Wu, 2019).

A central theme in the recent literature is that crises reveal the economic content of liquidity risk. During systemic episodes, market-wide liquidity deteriorates and trading constraints bind; consequently, sensitivity to liquidity shocks can amplify losses (Idzorek et al., 2012; Lou & Sadka, 2011). The COVID-19 episode further highlighted how liquidity shocks can propagate through intermediaries: open-ended funds under redemption pressure may transmit stress to asset prices and weaken the usual relation between liquidity and compensation (Jiang et al., 2022; Y. Ma et al., 2022). Related work shows that liquidity premia extend beyond equities: liquidity risk is priced in currencies (Söderlind & Somogyi, 2025), helps organise a broad set of anomaly returns and hedging demands (Virk & Butt, 2022), and exhibits important asymmetries in downside states (Palwishah et al., 2024). Finally, funding conditions and market liquidity are tightly linked: exposure to liquidity needs and rollover risk carries persistent pricing implications (Bechtel et al., 2023). Taken together, these results motivate an empirical design that distinguishes normal-state liquidity compensation from crisis-state behaviour, and that evaluates whether any apparent premia survive when liquidity becomes scarce.

Institutional structure shapes how liquidity risk materialises. Ownership concentration and commonality in holdings can amplify liquidity co-movement (Sensoy, 2017), while macro and policy environments transmit liquidity conditions across markets (Hassanein, 2022). Relatedly, policy uncertainty has been shown to contain forecasting information for broad risk premia (including the equity premium), suggesting another channel through which policy regimes may shape expected returns (Ali & Naz, 2025a).

Conversely, concentrated ownership can mitigate fire-sale pressure in certain settings (Giannetti & Jotikasthira, 2024). These institutional insights are particularly relevant for mutual funds, where portfolio choice, investor composition, and redemption design jointly determine how liquidity shocks affect performance.

1.2. Liquidity risk premium in mutual funds

Open-ended funds are structurally exposed to liquidity risk because they offer investors frequent redemption while holding assets that can be costly to liquidate. This liquidity transformation is a long-recognized vulnerability and a recurring focus of policy and stability debates (Chernenko & Sunderam, 2016). In stress states, redemptions can induce funds to sell their most liquid holdings first—a "reverse flight to liquidity"— transmitting pressure to asset prices and creating externalities for remaining investors (Jiang et al., 2022; Y. Ma et al., 2022). This mechanism makes liquidity management (cash buffers, trading schedules, and proactive rebalancing) economically meaningful for both performance and resilience.

Empirically, mutual-fund liquidity premia operate through two distinct routes: exposure and management. On the exposure side, funds with higher liquidity betas earn higher average returns in normal times, consistent with compensation for bearing systematic liquidity risk (Dong et al., 2019; Foran & O'Sullivan, 2014). However, a recurring finding is that standard traded liquidity factors explain only part of the high-minus-low liquidity-beta spread (HMLiq), leaving residual components that invite alternative interpretations (Dong et al., 2019). State conditioning sharpens this picture: excluding extreme illiquidity periods strengthens the positive liquidity beta–performance relation (Sadka, 2010), and fund characteristics such as turnover, age, and managerial attributes can mediate the strength of liquidity premia (Goyenko, 2012). These results suggest that the liquidity premium in mutual funds is inherently state-dependent and potentially entangled with managerial decisions and investor flows.

On the management side, researchers ask whether funds time liquidity by adjusting exposures dynamically with liquidity conditions. A common empirical approach allows market beta to co-move with deviations in aggregate liquidity (Cao et al., 2013). Within this framework, timing ability has been documented for certain fund segments (e.g., top performers or bank-affiliated complexes), consistent with information advantages and organizational resources (Alam & Ansari, 2020; Bodson et al., 2013; Wattanatorn et al., 2020; Wattanatorn & Tansupswatdikul, 2019). At the same time, the evidence emphasises scarcity and the concentration of timing skill: even in hedge funds, aggregate liquidity timing is limited and ability is concentrated in a small subset of managers (Aiken & Kang, 2023). This uneven distribution underscores the empirical challenge: measured “timing” can reflect genuine anticipatory rebalancing, but it can also reflect mechanically-induced beta changes arising from flows and forced trading during stress episodes.

1.3. Post-COVID evidence and identification challenges

Post-COVID-19 work clarifies both the fragility mechanism and the identification challenge relevant for timing. During March 2020, funds facing redemptions reallocated in ways consistent with reverse flight to liquidity, amplifying stress (Y. Ma et al., 2022). Bond-market evidence similarly shows that fund illiquidity is closely linked to fragility in asset prices during stress (Jiang et al., 2022). At the institutional margin, ownership concentration can mitigate flow-induced price pressure, highlighting that market structure shapes the propagation of liquidity shocks (Giannetti & Jotikasthira, 2024). These studies jointly imply that timing estimates can be contaminated by flow-driven mechanics precisely in the crisis states where liquidity risk is most economically meaningful. More broadly, asymmetric downside liquidity exposures (Palwishah et al., 2024) and liquidity-based hedging demands (Virk & Butt, 2022) underscore how liquidity premia should be evaluated with explicit attention to stress regimes and the short-sample inference problems that arise in acute episodes.

1.4. Positioning and implications for our study

The literature yields two robust messages that directly motivate our design. Firstly, liquidity compensation is conditional: it is most informative when evaluated across regimes rather than averaged across the full sample. This regime dependence echoes conditional asset-pricing and return-predictability evidence in which betas and premia vary with observable state variables (Ali, 2021). Secondly, mutual funds are not passive carriers of liquidity risk—observed spreads can reflect both static exposure and dynamic management, and crisis-period behaviour may be confounded by flows and forced trading.

Our study contributes to these debates in three ways. (1) We provide systematic evidence on liquidity premia and liquidity timing for Indian equity mutual funds over 2007–2024, an economically important emerging-market setting, where institutional frictions and liquidity dynamics may differ from developed markets. (2) Building on the state-dependent beta framework (Cao et al., 2013), we separate exposure from timing by pairing liquidity-beta sorted portfolios (HMLiq) with an investable timing factor (Timer) constructed from lagged timing signals, enabling a transparent decomposition of the cross-sectional spread. (3) We interpret the results through the lens of the recent fragility literature (Jiang et al., 2022; Y. Ma et al., 2022) and the evidence that timing skill is scarce and concentrated (Aiken & Kang, 2023; S. Kumar et al., 2023), motivating robustness and identification checks that distinguish timing from flow-driven mechanics and emphasise regime-specific inference. Prior Indian evidence remains limited and mixed (Alam & Ansari, 2020); extending the analysis through 2024 and explicitly separating exposure from timing helps clarify when liquidity risk is compensated, how much of the spread reflects dynamic beta management, and why systemic stress remains difficult to hedge.

2. Variables, data, and methods

2.1. Liquidity risk measures

India-specific RM, SMB, and HML factors are obtained from the Indian Finance Database (IFD). The risk-free rate $R_{f,t}$ is the 1-month Treasury bill yield, converted to a simple monthly rate from the quoted annualised value. Market excess return is defined as $R_{m,t} - R_{f,t}$. All series extend through December 2024 to match the fund sample window.

For stock j on day d, Amihud illiquidity is:

$$ILLIQ_{j,d} = \frac{\left|R_{j,d}\right|}{Dvol_{j,d}}$$

where $R_{j,d}$ is the daily return and $Dvol_{j,d}$ is rupee trading value. Monthly $ILLIQ_{j,t}$ is the average across trading days in month *t*. We exclude zero-return/zero-volume days, winsorise each month's cross-section at the 1st/99th percentiles, and require at least 15 valid trading days. Aggregate market liquidity is then:

$$AML_t = \frac{1}{N_t}\sum_{j=1}^{N_t} ILLIQ_{j,t}$$

where N_t is the number of eligible broad-index constituents in month t. Constituents are updated monthly to reflect membership changes.

Because liquidity is persistent, the priced component is the unexpected innovation. We estimate:

$$AML_t = a + b\,AML_{t-1} + e_t, \qquad InnAML_t \equiv \hat{e}_t$$

and define $InnAML_t$ as the innovation in aggregate market liquidity. For comparability across subperiods, $InnAML_t$ is standardised to mean zero and unit variance.

Following Lou and Sadka (2011), the liquidity beta of fund i is estimated from:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{1,i}\,(R_{m,t} - R_{f,t}) + \beta_{2,i}\,InnAML_t + \epsilon_{i,t} \tag{1}$$

We estimate $\beta_{2,i}$ using a rolling 24-month window and sort funds each January into quintiles by their most recent $\beta_{2,i}$: PoF1 (highest liquidity beta) through PoF5 (lowest). Quintile portfolios are equally weighted, held for 12 months, and rebalanced annually. The liquidity premium at the fund level is captured by PoF1–PoF5 (“HMLiq”), a zero-investment portfolio that isolates compensation for exposure to liquidity innovations.

2.2. Liquidity timing and flows

To test for liquidity timing, we allow market beta to vary with deviations of aggregate liquidity from its recent mean (Cao et al., 2013):

$$\beta_{m,i} = \beta_{0m,i} + \gamma_{m,i}(Liq_{m,t} - \overline{Liq_m}) \tag{2}$$

where $Liq_{m,t}$ is the aggregate liquidity state and $\overline{Liq_m}$ is its rolling mean (36 months). Substituting (2) into a standard factor benchmark yields:

$$\begin{aligned} R_{i,t} - R_{f,t} &= \alpha_i + \left[\beta_{0m,i} + \gamma_{m,i}(Liq_{m,t} - \overline{Liq_m})\right] \\ &(R_{m,t} - R_{f,t}) + \beta_{s,i} SMB_t + \beta_{v,i} HML_t + \epsilon_{i,t} \end{aligned} \tag{3}$$

the coefficient $\gamma_{m,i}$ measures liquidity timing: $\gamma_{m,i} > 0$ indicates that the fund raises (lowers) its market exposure when liquidity is above (below) trend.

To address the concern that measured timing may reflect mechanically-induced beta changes arising from flows, we compute net flows using the standard AUM decomposition:

$$Net\ Flow_{i,t} = \frac{AUM_{i,t} - AUM_{i,t-1}(1 + R_{i,t})}{AUM_{i,t-1}}$$

Appendix C in the Supplementary material augments the timing regressions with flow interactions and outflow-state terms to separate timing from flow-driven trading pressure. In addition, Appendix G in the Supplementary material examines whether estimated liquidity exposure and timing are systematically explained by standard fund characteristics (size, turnover, expense ratios, and age) using panel regressions with fund and month fixed effects and conditional double-sorts.

2.3. Mutual funds data

Our empirical analysis uses a comprehensive panel of actively managed Indian equity mutual funds from November 2007 through December 2024. This horizon spans multiple liquidity environments and market regimes, enabling regime-specific inference on both liquidity risk exposure and liquidity timing. For consistency across the paper, we define the following subperiods: Nov 2007–Dec 2024, Nov 2007–May 2009, Jun 2009–Dec 2019, Jan 2020–Jun 2020, and Jul 2020–Dec 2024. Because the post-pandemic era contains heterogeneous phases, we further split Jul 2020–Dec 2024 into finer subperiods in Appendix D in the Supplementary Material.

Fund-level data are obtained from Morningstar India, supplemented with disclosures from the Association of Mutual Funds in India (AMFI) and fund fact sheets for cross-validation of classifications and share-class attributes. The sample includes actively managed equity open-ended funds and excludes index funds, ETFs, sector/thematic funds, funds-of-funds, and closed-end products. For each fund, we collect total-return NAVs, assets under management (AUM), and available fund characteristics (e.g., expense ratios and portfolio attributes when disclosed). In addition, we compute fund flows using the standard AUM-based flow decomposition described below; flows play an explicit role in the robustness and identification checks in Appendix C in the Supplementary material.

Market-wide liquidity measures are constructed from daily stock-level trading data for constituents of a broad Indian equity index (BSE 500 constituents), using daily prices and trading value to compute Amihud-style illiquidity at the stock level and then aggregating to the market level. This construction ensures that the liquidity state is derived bottom-up (stock → market) and is consistent across regimes.

The initial sample contains 546 share classes. We impose the following screens: (1) the fund must be an open-ended, actively managed equity fund; (2) at least 80% of assets must be allocated to domestic Indian equities; (3) at least 36 consecutive months of return history are required to support rolling-window exposure estimation and annual portfolio formation; and (4) funds must satisfy a minimum AUM screen at formation to mitigate return distortions and imprecise beta estimates associated with very small funds.

The baseline AUM screen sets a minimum of USD 5 million (converted from INR using month-end exchange rates), following the methodology outlined by

Kacperczyk et al. (2008). Because the appropriate cutoff can be market-dependent, we do not rely on this threshold mechanically: Appendix A in the Supplementary Material replicates the full analysis under alternative minimum-AUM screens and distribution-based cutoffs (excluding the bottom tail of the Indian AUM distribution) and shows that the main conclusions are not driven by small-fund behaviour.

Where multiple share classes exist for a given portfolio, we consolidate at the portfolio level by retaining the oldest unhedged accumulation class to avoid overweighting fund families with multiple fee variants. Funds enter the panel once they satisfy history requirements and remain until termination/merger when data are available. This design mitigates mechanical survivorship effects and ensures that portfolio formation reflects the information set available at each formation date.

All NAVs and distributions are recorded in INR. Monthly returns are calculated from end-of-month total-return NAVs with reinvested distributions. Total-return series are net of ongoing expense ratios and management fees, and gross of front/back loads. Factors and state variables (market, SMB, HML, and liquidity) are aligned to the same calendar month.

We benchmark the fund universe to the BSE 500 index and report descriptive statis- tics for funds and the benchmark across standard horizons and for the COVID/post-COVID period. Table 1 summarises return and risk characteristics for the final sample of 208 funds.

We drop months with missing NAVs, reinvest distributions, remove data errors (e.g., non-positive NAVs), and winsorise monthly returns at the 0.5% and 99.5% tails. For stock-level liquidity inputs, we require at least 15 valid trading days per month for inclusion in monthly illiquidity measures.

2.4. Estimation and robustness checks

Funds are sorted each January using rolling-window estimates, and quintile portfolios are held from February to the following January. We compute monthly returns for PoF1–PoF5 and HMLiq and evaluate performance using CAPM, FF3F, and liquidity/timing-augmented models.

Time-series regressions report heteroskedasticity- and autocorrelation-consistent Newey–West statistics. For spread portfolios and overlapping formation procedures, HAC corrections are applied analogously.

The empirical design is complemented by a set of robustness checks that map directly to the main identification concerns, presented and discussed in the Supplementary material. Appendix A reports alternative minimum-AUM screens and value-weighted portfolio results. Appendix B reports modern benchmark models (Carhart-4 and FF5) with liquidity and Timer augmenta-

Table 1. Descriptive statistics of Indian equity funds and BSE 500 (Nov 2007–Dec 2024; subperiods and horizons)

	1 year		5 years		10 years		COVID & post-COVID (2020–2024)	
	Funds	**BSE500**	**Funds**	**BSE500**	**Funds**	**BSE500**	**Funds**	**BSE500**
Return analysis								
Total return (%)	7.28	8.97	101.75	119.99	244.34	214.01	68.42	72.15
Annualized mean return (%)	11.54	14.19	14.15	26.04	21.81	20.61	13.72	14.95
Annualized mean excess return (%)	–2.29	–1.16	–9.24	–1.11	–2.91	–2.31	–1.85	–1.12
Risk								
Annualised standard deviation (%)	12.76	13.47	16.88	17.39	18.62	19.75	19.25	20.11
Annualised downside risk (%)	8.19	8.72	12.14	12.49	14.35	14.21	13.81	13.95
Annualised tracking error (%)	2.04		2.23		2.49		2.62	
Risk/return								
Sharpe measure	0.12	0.24	0.54	0.59	0.91	0.83	0.58	0.61
Jensen alpha (%)	–1.60	–0.88	1.13				0.97	
Information ratio	–0.81	–0.42	0.80	0.28			0.74	0.31
Treynor measure	0.02	0.03	0.09	0.12	0.28	0.26	0.19	0.21
Correlation	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99

Notes: Descriptive statistics are based on 208 actively managed Indian equity funds benchmarked against the BSE 500. Returns are INR total returns (net of ongoing fees and gross of loads). Standard horizons (1y, 5y, 10y) are complemented by COVID and post-COVID performance (2020–2024).

Source: own work.

tions. Appendix C implements flow-adjusted timing specifications. Appendix D refines the post-COVID segmentation. Appendix E reports Timer diagnostics (correlations, VIFs, and orthogonalized Timer). Appendix F assesses sensitivity to rolling-window length and supplements crisis-window inference using daily evidence for Jan–Jun 2020. Appendix G adds fund-characteristic controls (size, turnover, expense ratios, age) and conditional double-sorts to assess whether liquidity exposure and timing estimates are mechanically explained by fund type rather than discretionary beta management.

3. Results and analysis

This section evaluates whether exposure to innovations in aggregate market liquidity is compensated in Indian equity mutual fund returns and whether such compensation is state-dependent. Quintile portfolios (PoFs) are formed each January by sorting funds on their *ex ante* liquidity betas from (1) estimated over the prior 24 months; portfolios are equally weighted, held for 12 months, and rebalanced annually. By construction, PoF1 has the highest liquidity beta (greatest exposure to liquidity innovations), while PoF5 has the lowest (often negative) exposure.

Period labels are harmonised across text and tables using Nov 2007–Dec 2024, Nov 2007–May 2009, Jun 2009–Dec 2019, Jan 2020–Jun 2020, and Jul 2020–Dec 2024. Because the post-pandemic era contains heterogeneous phases, Appendix D in the Supplementary material further refines Jul 2020–Dec 2024 into finer subperiods. Inference in the acute COVID window (Jan 2020–Jun 2020) is inherently noisy due to the short sample; Appendix F in the Supplementary material therefore supplements monthly results with daily evidence and rolling-window sensitivity.

3.1. Descriptive properties and the liquidity-beta premium

Table 2 reports average monthly returns and ex-post liquidity betas for the PoFs across subperiods. Three patterns are central. Firstly, liquidity betas decline monotonically from PoF1 to PoF5 in each subperiod, confirming that the preformation sort generates portfolios with ordered liquidity exposure. This monotonicity is a necessary diagnostic for interpreting PoF spreads as related to systematic liquidity risk rather than a sorting artifact. Secondly, in longer samples such as Jun 2009–Dec 2019 and Jul 2020–Dec 2024, average returns are higher for the high-liquidity-beta portfolios relative to the low-liquidity-beta portfolios, consistent with a positive liquidity premium outside acute stress. Thirdly, in short crisis windows, the cross-sectional return ordering is less stable and estimates are less precise, consistent with the view

Table 2. Mean returns of portfolios sorted by liquidity-innovation exposure

	PoF1	PoF2	PoF3	PoF4	PoF5
Panel A: Nov 2007–Dec 2024					
Mean	0.717	0.625	0.445	0.383	0.215
Standard deviation	4.030	3.660	3.380	3.080	2.350
Alpha	0.593	0.491	0.460	0.402	0.175
Liquidity beta	0.950	0.750	0.440	–0.150	–0.370
Panel B: Jun 2009–Dec 2019					
Mean	0.540	0.517	0.480	0.390	0.318
Standard deviation	3.050	2.930	2.690	2.350	2.030
Alpha	0.454	0.370	0.315	0.225	0.178
Liquidity beta	0.870	0.700	0.370	–0.090	–0.290
Panel C: Nov 2007–May 2009					
Mean	1.254	0.974	0.676	0.334	0.021
Standard deviation	6.580	5.840	5.050	4.880	4.470
Alpha	1.020	0.730	0.440	0.080	–0.110
Liquidity beta	1.170	0.980	0.620	–0.110	–0.430
Panel D: Jan 2020–Jun 2020					
Mean	0.310	0.280	0.240	0.190	0.150
Standard deviation	7.200	6.950	6.500	6.300	6.100
Alpha	0.200	0.170	0.150	0.120	0.080
Liquidity beta	1.300	1.050	0.780	0.200	–0.100
Panel E: Jul 2020–Dec 2024					
Mean	0.480	0.460	0.430	0.390	0.350
Standard deviation	3.600	3.400	3.200	2.900	2.700
Alpha	0.350	0.310	0.280	0.220	0.180
Liquidity beta	0.920	0.740	0.410	–0.050	–0.220

Notes: Funds are sorted each January into five equally weighted quintile portfolios (PoF1 = highest, PoF5 = lowest liquidity beta) using the prior 24 months of estimates from (1). The table reports average monthly returns (%) and ex post liquidity betas by subperiod. Market excess return is $R_{m,t} - R_{f,t}$; the liquidity factor is the innovation in aggregate market liquidity ($InnAMLt$).

Source: own work.

that liquidity compensation is conditional and may compress when common shocks dominate and market functioning deteriorates.

Taken together, the descriptive evidence is consistent with conditional liquidity pricing: bearing liquidity-innovation exposure is associated with economically meaningful return differences in normal and recovery regimes, while acute stress episodes attenuate or eliminate the premium. Subsequent sub-sections evaluate whether these differences are absorbed by traded liquidity risk (exposure) and whether a separate timing channel contributes to the spread (management).

3.2. Fund-level liquidity risk premium (HMLiq)

We summarise the fund-level liquidity premium as the high-minus-low return spread (PoF1−PoF5), denoted HMLiq. Table 3 reports the mean monthly spread, its volatility and HAC *t*-statistic, the correlation between PoF1 and PoF5, and the intercept/slope from regressing PoF1 on PoF5.

Table 3. Fund-level liquidity risk premium (HMLiq = PoF1−PoF5)

	Nov 2007–Dec 2024	**Jun 2009–Dec 2019**	**Nov 2007–May 2009**	**Jan 2020–Jun 2020**	**Jul 2020–Dec 2024**
Mean	0.480	0.470	0.520	0.120	0.310
Standard deviation	2.680	2.570	2.980	3.900	2.400
t-statistic (HAC)	2.680	2.050	0.920	0.420	2.030
Correlation	0.930	0.920	0.980	0.950	0.910
Cumulative	17.520	59.400	15.720	1.800	11.160
Alpha (PoF1 on PoF5)	0.480	0.470	0.520	0.100	0.280
t-statistic (HAC)	3.940	2.440	1.130	0.390	2.100
Beta (PoF1 on PoF5)	0.940	0.940	1.130	1.200	0.960

Notes: Funds are sorted each January by liquidity beta (prior 24 months). The table reports monthly high-minus-low (HMLiq) spreads. Cumulative is the arithmetic sum of monthly spreads.

Source: own work.

HMLiq is positive in the aggregate sample and is most reliably estimated in the longer normal and recovery subperiods. In crisis windows, estimates are weaker and less stable, consistent with conditional liquidity pricing and the dominance of common shocks during systemic stress. The high correlation between PoF1 and PoF5 (ρ typically above 0.90) indicates that HMLiq isolates a comparatively subtle component of expected returns, reinforcing the role of factor attribution and disciplined inference.

Because the Indian mutual fund cross-section includes a long tail of small funds, Appendix A in the Supplementary material replicates the full pipeline under alternative minimum-AUM screens and distribution-based cutoffs, and compares equal-weighted to value-weighted portfolios. These checks ensure that the HMLiq pattern is not driven by an arbitrary AUM threshold or by small-fund behaviour.

3.3. Risk-adjusted performance with a traded liquidity factor

We next assess whether HMLiq reflects compensation for exposure to a traded stock-level liquidity factor or residual performance beyond such exposure. Table 4 reports alphas for each quintile portfolio and for HMLiq under CAPM and FF3F, as well as liquidity-augmented variants that add the traded liquidity factor (Liq). A decline in the HMLiq alpha after adding Liq indicates that traded liquidity risk explains part of the spread.

Adding Liq attenuates HMLiq in the aggregate sample and in normal/recovery regimes, indicating that a nontrivial component of the liquidity-beta spread is attributable to traded liquidity exposure. Crisis-period estimates are less stable and statistically weaker, consistent with the broader evidence that liquidity premia are not reliably earned under systemic stress.

Because mutual-fund performance evaluation commonly controls for Momentum and (in newer frameworks) Profitability and Investment, Appendix B in the Supplementary material re-estimates HMLiq alphas under Carhart-4 and Fama–French-5, each optionally augmented with Liq and Timer. Richer benchmarks attenuate alphas but preserve the paper's central regime-dependent conclusions.

Table 4. Risk-adjusted performance of liquidity-beta sorted portfolios of funds (Nov 2007–Dec 2024, subperiods)

Model	Port1	Port2	Port3	Port4	Port5	HMLiq	Liq.
Panel A: Nov 2007–Dec 2024							
CAPM	0.396	0.272	0.153	0.064	0.015	0.381	–
	(1.440)	(1.000)	(0.620)	(0.250)	(0.050)	(2.210)	–
CAPM+Liq.	0.304	0.213	0.129	0.041	0.032	0.272	0.562
	(1.110)	(0.780)	(0.510)	(0.160)	(0.060)	(1.380)	(3.880)
FF3F	0.398	0.278	0.158	0.067	0.024	0.359	–
	(1.490)	(1.030)	(0.640)	(0.260)	(0.080)	(2.320)	–
FF3F+Liq.	0.322	0.232	0.137	0.046	0.037	0.285	0.470
	(1.200)	(0.860)	(0.550)	(0.170)	(0.120)	(1.380)	(3.380)

Model	Port1	Port2	Port3	Port4	Port5	HMLiq	Liq.
Panel B: Jun 2009–Dec 2019							
CAPM	0.484	0.314	0.196	0.107	–0.014	0.498	–
	(3.940)	(3.090)	(2.970)	(1.580)	(0.150)	(3.050)	–
CAPM+Liq.	0.373	0.248	0.190	0.119	0.014	0.359	0.612
	(3.150)	(2.700)	(2.800)	(1.700)	(0.150)	(2.260)	(3.850)
FF3F	0.454	0.301	0.188	0.104	–0.012	0.466	–
	(4.020)	(3.360)	(2.830)	(1.530)	(0.130)	(3.100)	–
FF3F+Liq.	0.376	0.268	0.183	0.108	0.003	0.373	0.418
	(3.340)	(3.160)	(2.670)	(1.530)	(0.030)	(2.470)	(2.630)
Panel C: Nov 2007–May 2009							
CAPM	–0.456	–0.486	–0.628	–0.622	–0.692	0.236	–
	(–0.230)	(–0.240)	(–0.340)	(–0.320)	(–0.350)	(0.540)	–
CAPM+Liq.	–0.409	–0.440	–0.589	–0.583	–0.655	0.246	0.189
	(–0.200)	(–0.220)	(–0.320)	(–0.290)	(–0.330)	(0.550)	(–0.750)
FF3F	0.712	0.751	0.555	0.605	0.571	0.141	–
	(0.310)	(0.320)	(0.260)	(0.270)	(0.260)	(0.283)	–
FF3F+Liq.	1.115	1.148	0.995	0.955	0.910	0.205	0.206
	(0.490)	(0.500)	(0.400)	(0.430)	(0.410)	(0.433)	(0.780)
Panel D: Jan 2020–Jun 2020							
CAPM	–0.892	–0.754	–0.643	–0.611	–0.585	–0.307	–
	(–0.420)	(–0.360)	(–0.310)	(–0.300)	(–0.280)	(–0.520)	–
CAPM+Liq.	–0.801	–0.692	–0.601	–0.567	–0.542	–0.259	0.128
	(–0.380)	(–0.330)	(–0.300)	(–0.280)	(–0.260)	(–0.490)	(0.840)
FF3F	–0.741	–0.628	–0.539	–0.502	–0.478	–0.263	–
	(–0.350)	(–0.300)	(–0.260)	(–0.240)	(–0.230)	(–0.480)	–
FF3F+Liq.	–0.653	–0.571	–0.496	–0.461	–0.438	–0.215	0.102
	(–0.310)	(–0.280)	(–0.240)	(–0.230)	(–0.210)	(–0.430)	(0.690)
Panel E: Jul 2020–Dec 2024							
CAPM	0.512	0.344	0.229	0.118	0.021	0.491	–
	(2.340)	(1.980)	(1.640)	(1.120)	(0.190)	(2.270)	–
CAPM+Liq.	0.417	0.283	0.207	0.097	0.038	0.329	0.391
	(2.010)	(1.750)	(1.480)	(0.990)	(0.270)	(2.020)	(2.610)
FF3F	0.485	0.331	0.218	0.115	0.019	0.466	–
	(2.420)	(2.030)	(1.570)	(1.140)	(0.180)	(2.210)	–
FF3F+Liq.	0.392	0.278	0.196	0.098	0.036	0.307	0.342
	(1.950)	(1.690)	(1.410)	(1.010)	(0.250)	(2.010)	(2.210)

Notes: Average monthly alphas (%) for quintile portfolios (PoF1 highest to PoF5 lowest liquidity beta) and the high-minus-low spread (PoF1−PoF5).

Source: own work.

3.4. Timing ability of liquidity-beta sorted funds

The preceding evidence indicates that traded liquidity risk explains a meaningful fraction of $HMLiq$, yet a residual component remains in normal and recovery regimes. A natural presumption is that part of this residual reflects state-dependent beta management (liquidity timing). We test this mechanism in two complementary ways: (1) estimating timing coefficients in the state-dependent beta specification (3), and (2) constructing an investable Timer factor from lagged timing-coefficient sorts and using it to decompose $HMLiq$.

3.4.1. Liquidity timing coefficients

We estimate the liquidity-timing specification in (3) for each liquidity-beta quintile (PoF1– PoF5). The timing coefficient $\gamma_{m,i}$ multiplies the interaction $(Liq_{m,t} - Liq_m)(R_{m,t} - R_{f,t})$; $\gamma_{m,i} > 0$ indicates that a portfolio scales market exposure up (down) when liquidity is above (below) trend, consistent with liquidity-state-dependent beta management. Results are reported in Table 5.

Table 5. Liquidity timing in liquidity–beta sorted portfolios

	α_i	$\beta_{0m,i}$	$\beta_{s,i}$	$\beta_{v,i}$	$\gamma_{m,i}$	Adj R^2
Panel A: Nov 2007–Dec 2024						
PoF1	0.246 (2.21)	0.875 (30.27)	0.173 (5.83)	–0.025 (–1.07)	0.020 (3.31)	0.957
PoF2	0.167 (1.96)	0.907 (48.51)	0.134 (4.59)	0.022 (1.02)	0.000 (0.00)	0.975
PoF3	0.061 (0.94)	0.842 (27.81)	0.130 (3.89)	–0.030 (–1.04)	0.020 (1.62)	0.968
PoF4	0.019 (0.26)	0.861 (32.56)	0.200 (5.20)	–0.040 (–0.94)	–0.040 (–0.94)	0.946
PoF5	–0.023 (–0.17)	0.810 (13.61)	0.019 (0.55)	0.045 (1.55)	0.017 (0.94)	0.925
Panel B: Jun 2009–Dec 2019						
PoF1	0.316 (2.93)	0.813 (25.54)	0.172 (6.25)	0.012 (0.51)	0.113 (2.43)	0.914
PoF2	0.190 (2.21)	0.967 (37.30)	0.132 (7.26)	0.032 (1.64)	–0.019 (–0.49)	0.961
PoF3	0.061 (1.05)	0.944 (45.58)	0.013 (0.93)	–0.005 (–0.35)	–0.014 (–0.45)	0.973
PoF4	0.019 (0.43)	0.949 (40.63)	–0.025 (–1.76)	–0.012 (–0.73)	–0.046 (–1.63)	0.972
PoF5	–0.029 (–0.32)	1.087 (32.49)	–0.040 (–2.01)	–0.012 (–0.67)	–0.046 (–1.16)	0.964

	α_i	$\beta_{0m,i}$	$\beta_{s,i}$	$\beta_{v,i}$	$\gamma_{m,i}$	Adj R^2
Panel C: Nov 2007–May 2009						
PoF1	−0.158 (−0.48)	0.893 (27.45)	0.182 (2.06)	−0.057 (−1.03)	0.033 (3.88)	0.986
PoF2	−0.015 (−0.05)	0.891 (43.89)	0.204 (2.56)	−0.090 (−1.98)	0.019 (3.33)	0.991
PoF3	−0.278 (−0.94)	0.794 (32.30)	0.102 (1.31)	−0.054 (−1.09)	0.007 (1.17)	0.990
PoF4	−0.153 (−0.41)	0.832 (32.46)	0.152 (1.53)	−0.070 (−1.20)	−0.003 (−0.52)	0.987
PoF5	−0.204 (−0.35)	0.669 (15.37)	0.034 (1.97)	−0.077 (−0.78)	−0.012 (−1.34)	0.964
Panel D: Jan 2020–Jun 2020						
PoF1	−0.312 (−0.95)	0.842 (11.23)	0.208 (2.01)	−0.044 (−0.55)	0.041 (1.78)	0.942
PoF2	−0.274 (−0.87)	0.864 (12.74)	0.191 (2.24)	−0.031 (−0.44)	0.027 (1.42)	0.936
PoF3	−0.198 (−0.64)	0.833 (10.48)	0.154 (1.75)	−0.026 (−0.39)	0.010 (0.55)	0.927
PoF4	−0.146 (−0.48)	0.790 (9.67)	0.120 (1.23)	−0.015 (−0.22)	−0.005 (−0.27)	0.912
PoF5	−0.121 (−0.43)	0.758 (8.92)	0.082 (0.97)	−0.009 (−0.14)	−0.014 (−0.69)	0.903
Panel E: Jul 2020–Dec 2024						
PoF1	0.284 (2.57)	0.889 (28.91)	0.165 (6.12)	−0.022 (−0.84)	0.036 (2.16)	0.963
PoF2	0.177 (1.97)	0.901 (31.42)	0.137 (5.47)	0.014 (0.65)	0.012 (1.01)	0.958
PoF3	0.062 (1.01)	0.872 (26.37)	0.128 (3.98)	−0.018 (−0.72)	0.009 (0.74)	0.952
PoF4	0.019 (0.34)	0.847 (24.26)	0.112 (3.55)	−0.020 (−0.83)	−0.007 (−0.59)	0.948
PoF5	−0.027 (−0.39)	0.803 (22.91)	0.098 (3.12)	0.021 (0.92)	−0.010 (−0.83)	0.941

Source: own work.

Two results are most salient. Firstly, timing loadings are concentrated in the high-liquidity-beta portfolios (PoF1 and, in some periods, PoF2), consistent with the idea that timing is most relevant where exposure is highest. Secondly, timing estimates are stronger and more precisely estimated in longer samples, while short crisis windows yield inherently noisy inference. To assess whether these timing estimates simply proxy for fund type, Appendix G in the Supplementary material relates γ (and β^{liq}) to lagged flows, size, turnover, fees, and age with fund and month fixed effects; flows remain an important predictor, but timing is not mechanically subsumed by standard characteristics.

Because contemporaneous interaction terms may capture both anticipatory behaviour and rapid reaction, Appendix C in the Supplementary material augments the timing regression with flow controls and outflow-state interactions to distinguish discretionary timing from mechanically-induced beta shifts driven by flows. Appendix E in the Supplementary material reports Timer diagnostics (correlations, VIFs, and orthogonalised Timer) to address multicollinearity and stress-proxy concerns.

3.5. Timing-adjusted performance: The timer decomposition

We now ask whether the residual HMLiq performance reflects state-dependent beta management rather than static exposure alone. Using the rolling-window estimates from (3), we estimate fund-level timing coefficients $\gamma_{m,i}$, sort funds into quintiles by $\gamma_{m,i}$, and form an investable Timer factor as the return on the highest-minus-lowest timing portfolio. We then re-estimate alphas for the liquidity-beta portfolios and for HMLiq after augmenting CAPM and FF3F with Timer. If timing captures an economically important component of HMLiq, the HMLiq alpha should attenuate materially once Timer is included. Results are reported in Table 6.

Table 6. Liquidity timing skills–adjusted performance of liquidity–beta sorted portfolios

Model	Port1	Port2	Port3	Port4	Port5	Port1–Port5	Timer
Panel A: Nov 2007–Dec 2024							
CAPM	0.396 (1.44)	0.272 (1.00)	0.153 (0.62)	0.064 (0.25)	0.015 (0.05)	0.381 (2.21)	–
CAPM+Timer	0.366 (1.29)	0.325 (1.15)	0.223 (0.88)	0.147 (0.55)	0.230 (0.79)	0.136 (0.79)	0.618 (7.56)
FF3F	0.398 (1.49)	0.278 (1.03)	0.158 (0.64)	0.067 (0.26)	0.024 (0.08)	0.359 (2.32)	–
FF3F+Timer	0.392 (1.40)	0.346 (1.25)	0.232 (0.91)	0.152 (0.57)	0.229 (0.78)	0.162 (1.01)	0.540 (5.31)
Panel B: Jun 2009–Dec 2019							
CAPM	0.484 (3.94)	0.314 (3.09)	0.196 (2.97)	0.107 (1.58)	–0.014 (0.15)	0.498 (3.05)	–
CAPM+Timer	0.416 (3.38)	0.299 (3.14)	0.188 (2.77)	0.134 (1.94)	0.055 (0.60)	0.362 (2.28)	0.406 (3.86)

Model	Port1	Port2	Port3	Port4	Port5	Port1–Port5	Timer
FF3F	0.454 (4.02)	0.301 (3.36)	0.188 (2.83)	0.104 (1.53)	–0.012 (0.13)	0.466 (3.10)	–
FF3F+Timer	0.411 (3.59)	0.305 (3.58)	0.182 (2.67)	0.127 (1.83)	0.048 (0.53)	0.363 (2.45)	0.324 (3.25)
Panel C: Nov 2007–May 2009							
CAPM	–0.456 (–0.23)	–0.486 (–0.24)	–0.628 (–0.34)	–0.622 (–0.32)	–0.692 (–0.35)	0.236 (0.54)	–
CAPM+Timer	–1.332 (–0.47)	–1.272 (–0.45)	–1.171 (–0.45)	–1.216 (–0.44)	–1.182 (–0.30)	–0.520 (–0.75)	0.109 (0.06)
FF3F	0.712 (0.31)	0.751 (0.32)	0.555 (0.26)	0.605 (0.27)	0.571 (0.26)	0.141 (0.28)	–
FF3F+Timer	–0.990 (–0.35)	–0.889 (–0.31)	–0.800 (–0.31)	–0.832 (–0.29)	–0.386 (–0.14)	–0.606 (–0.88)	0.858 (1.84)
Panel D: Jan 2020–Jun 2020							
CAPM	–0.285 (–0.42)	–0.231 (–0.36)	–0.194 (–0.31)	–0.151 (–0.24)	–0.143 (–0.23)	–0.142 (–0.58)	–
CAPM+Timer	–0.398 (–0.61)	–0.342 (–0.55)	–0.291 (–0.49)	–0.232 (–0.38)	–0.211 (–0.34)	–0.176 (–0.72)	0.312 (2.14)
FF3F	–0.196 (–0.28)	–0.165 (–0.26)	–0.152 (–0.24)	–0.134 (–0.21)	–0.119 (–0.20)	–0.077 (–0.36)	–
FF3F+Timer	–0.317 (–0.47)	–0.278 (–0.44)	–0.239 (–0.40)	–0.195 (–0.32)	–0.171 (–0.28)	–0.168 (–0.66)	0.254 (1.87)
Panel E: Jul 2020–Dec 2024							
CAPM	0.412 (2.89)	0.297 (2.36)	0.204 (1.79)	0.128 (1.04)	0.024 (0.22)	0.388 (2.14)	–
CAPM+Timer	0.339 (2.41)	0.276 (2.21)	0.187 (1.64)	0.114 (0.97)	0.018 (0.15)	0.207 (1.29)	0.521 (4.62)
FF3F	0.395 (2.74)	0.283 (2.29)	0.197 (1.73)	0.121 (0.99)	0.021 (0.19)	0.374 (2.05)	–
FF3F+Timer	0.328 (2.31)	0.265 (2.14)	0.182 (1.59)	0.108 (0.92)	0.015 (0.14)	0.196 (1.22)	0.487 (4.21)

Source: own elaboration.

In the aggregate sample, adding Timer substantially reduces the HMLiq alpha, implying that a large fraction of the liquidity-beta spread is associated with systematic liquidity-state-dependent beta management. This attenuation is also evident in Jul 2020–Dec 2024. In Jun 2009–Dec 2019, a residual component remains after conditioning on Timer, suggesting that additional priced dimensions or fund selection effects contribute to the spread. In crisis windows, both HMLiq and its timing-adjusted counterpart are statistically weak, indicating limited scope for systematic timing to offset acute systemic stress. More broadly, Appendix G in the Supplementary material shows that Timer continues to absorb a large share of HMLiq even when the spread is recomputed within low- versus high-turnover and low- versus high-expense subsamples.

Timer is constructed from lagged timing signals and rebalanced on the annual formation calendar, which mitigates look-ahead concerns. Appendices B-F in the Supplementary Material report robustness checks (modern benchmarks, flow-adjusted timing, Timer diagnostics, post-pandemic heterogeneity, and crisis-window inference). Across the full sample, liquidity exposure is associated with economically meaningful return differentials in normal and recovery regimes but not in acute stress states.

Traded liquidity risk explains a nontrivial share of the liquidity-beta spread, and a separate timing channel—captured by state-dependent betas and the investable Timer factor—accounts for an additional substantial component, particularly outside systemic stress. These conclusions are reinforced by robustness and identification checks in the appendices: alternative AUM screens and weighting (Appendix A in the Supplementary material), modern factor benchmarks (Appendix B), flow-adjusted timing (Appendix C), post-COVID subperiods (Appendix D), Timer diagnostics (Appendix E), and rolling-window and daily COVID inference (Appendix F).

4. Discussion

Our results point to a conditional liquidity-risk premium in delegated portfolios. In the long sample, the liquidity-beta spread is economically meaningful (HMLiq averages about 0.48% per month), and it is most reliably estimated outside acute stress episodes (0.47% per month in the 2009–2019 tranquil regime and 0.31% per month post-COVID). In contrast, the spread becomes statistically weak during the Global Financial Crisis and the 2020 COVID shock. This state dependence is consistent with liquidity-risk pricing frameworks in which expected returns are increasing in exposure to aggregate liquidity shocks (Acharya & Pedersen, 2005; Pástor & Stambaugh, 2003) and with emerging-market evidence that local market liquidity is a first-order driver of expected returns (Bekaert et al., 2007). The novelty here is that we document these patterns in a large Indian mutual fund cross section, where the intermediary channel is potentially stronger because open-end funds transform underlying asset liquidity into daily redeemability.

At the same time, the crisis evidence highlights an important nuance: a standard risk-premium interpretation would predict higher compensation in bad states, yet realised cross-sectional premia compress when liquidity risk becomes systemic. A standard "risk premium" interpretation might lead one to expect higher compensation for liquidity exposure when marginal utility is high. Instead, we find that realised cross-sectional premia compress precisely when aggregate liquidity risk is most salient. A plausible reconciliation is that systemic episodes are dominated by market-wide funding and market-liquidity spirals (Brunnermeier & Pedersen, 2009), which raise return commonality and reduce dispersion across fund portfolios. In these states, cross-sectional differences in liquidity exposure may be swamped by (1) correlated liquidity shocks, (2) binding trading constraints, and (3) "flight to quality" dynamics, all of which make it difficult for high-exposure funds to earn (as opposed to require) a premium within a short crisis window.

A second contribution is to separate static liquidity exposure from dynamic beta management. Prior work documents that mutual-fund liquidity risk can show up as performance differentials that are not fully absorbed by standard traded factors (Dong et al., 2019; Foran & O'Sullivan, 2014). Our findings echo that logic: adding a traded stock-level liquidity factor attenuates HMLiq in normal/recovery regimes but does not fully eliminate it. The additional explanatory power comes from a timing mechanism. Building on liquidity timing tests in Cao et al. (2013), we show that state- dependent liquidity betas and an investable Timer factor—constructed from lagged timing-coefficient sorts—absorb a substantial incremental share of the liquidity-beta spread, particularly outside systemic stress.

This timing result is novel in two ways. Firstly, it provides a portfolio-level decomposition that clarifies why fund-level liquidity-beta spreads can persist even when traded liquidity factors only partially explain them: part of the spread behaves like *managed* exposure rather than a fixed characteristic. Secondly, the Timer factor offers a practical benchmark for performance attribution in emerging markets, where (1) liquidity is highly time-varying and (2) the mapping between stock-level liquidity factors and delegated-portfolio returns is less direct. Consistent with this interpretation, timing effects are concentrated in the high-liquidity-beta portfolios, suggesting that liquidity-sensitive mandates provide both the incentive and the opportunity to vary exposure across liquidity states.

The fact that timing does not provide reliable protection in crisis windows is informative. It aligns with the mutual-fund fragility literature, which emphasises strategic complementarities in investor redemptions and the externalities created by forced selling in illiquid markets. In particular, payoff complementarities can make outflows highly sensitive to bad performance (Chen et al., 2010), and flow-driven trading can generate fire-sale price pressure (Coval & Stafford, 2007). When market-wide illiquidity is high, these forces are amplified in more illiquid funds (Goldstein et al., 2017) and can become especially salient during systemic events such as COVID-19 (Falato et al., 2021). In our setting, such mechanisms can help explain why (1) HMLiq collapses during systemic stress and (2) the incremental contribution of timing is statistically fragile: during a liquidity shock, the marginal ability to rebalance is curtailed exactly when rebalancing is most valuable. Appendix G also indicates that turnover and fee intensity correlate with measured timing, consistent with trading frictions shaping implementation, but these characteristics do not overturn the conclusion that timing is fragile in systemic stress states.

Importantly, this interpretation also disciplines what we can claim about "skill." Although Timer is built from predictable variation in estimated betas, similar in spirit to conditional risk-premium forecasting exercises (Ali & Naz, 2025b), part of the apparent timing could still reflect mechanical effects of flows, endogenous risk-taking, or composition changes in the investable universe. The flow-adjusted timing specifications in Appendix C in the Supplementary material are designed to separate discretionary timing from mechanically-induced beta shifts. We therefore interpret Timer as evidence of a timing channel rather than a clean measure of managerial skill.

For investors and researchers, the results suggest that evaluating Indian equity mutual funds with static-factor alphas can be misleading: liquidity-related performance is state-dependent, and a meaningful component is tied to dynamic exposure management. For regulators and risk monitors, the evidence is consistent with a world in which open-end funds can be procyclical: in normal states, liquidity exposure is rewarded, while in stress states, flows and market-wide illiquidity dominate and cross-sectional premia compress.

Some limitations point to natural extensions. Firstly, liquidity is measured through innovations in aggregate market liquidity; future work could combine this with holdings-based liquidity (portfolio "illiquidity" and concentration) to better separate exposure from endogenous trading. Secondly, crisis windows are short and statistically noisy, so inference about stress-state premia should remain cautious. Thirdly, because emerging markets can undergo structural and regulatory changes over long samples, further work could study whether the liquidity exposure–timing mapping shifts across subperiods and fund categories (e.g., large-cap versus mid/small-cap mandates).

Overall, the key takeaway is that liquidity in delegated portfolios is not only a priced exposure in normal times but also a managed exposure; however, the benefits of such management are sharply limited when liquidity risk becomes systemic.

Conclusions

This paper examines how liquidity exposure and liquidity timing jointly shape the cross section of Indian equity mutual fund returns over 2007–2024, covering the Global Financial Crisis, the COVID-19 shock, and the post-pandemic recovery. The main result is that the fund-level liquidity premium is regime contingent: funds with higher exposure to innovations in aggregate market liquidity earn higher average returns in tranquil and recovery regimes, whereas the high-minus-low liquidity-beta spread (HMLiq) contracts sharply and becomes statistically weak during systemic stress.

A second key finding is that the liquidity-beta spread reflects not only static exposure but also dynamic beta management. A liquidity-augmented factor model explains a meaningful portion of HMLiq, and an independently constructed, tradable Timer factor—built from lagged timing-coefficient sorts—absorbs an additional share of the spread in the aggregate sample and especially in the post-COVID period. Timing effects are most pronounced among high-liquidity-beta portfolios, consistent with the view that liquidity-sensitive mandates are the primary locus of state-contingent exposure adjustment. However, timing does not provide reliable protection in acute stress states: when liquidity risk becomes system-wide, both the spread and timing-adjusted performance are statistically fragile.

These findings have practical implications. For investors, liquidity exposure should be treated as a regime-dependent source of expected return rather than a crisis hedge: it is rewarded in normal and recovery environments but offers limited insulation when liquidity deteriorates abruptly. For asset managers, the evidence supports the value of monitoring liquidity states and adjusting exposure in a disciplined manner, while also highlighting the limits of such strategies under market-wide stress. For policymakers, the results reinforce the importance of tools that reduce redemption externalities and curb fire-sale dynamics, since fund-level actions alone are unlikely to offset aggregate liquidity shocks.

Several extensions could further strengthen identification and external validity. Future work can incorporate richer benchmark models (e.g., Momentum, Profitability, and Investment) and holdings-based measures of portfolio liquidity to better distinguish residual performance from omitted risks. Higher-frequency estimation and explicit modelling of flows and trading costs would help separate anticipatory timing from rapid reaction. Finally, linking timing capacity to microstructure conditions (depth, resilience, and price impact) can clarify when liquidity management is effective and when structural fragility dominates outcomes.

References

Acharya, V. V., & Pedersen, L. H. (2005). Asset pricing with liquidity risk. *Journal of Financial Economics*, *77*(2), 375–410. https://doi.org/10.1016/j.jfineco.2004.06.007

Aiken, A. L., & Kang, M. (2023). Hedge fund manager timing and selectivity skill over time. A holdings-based estimate. *Finance Research Letters*, *58*, 104439. https://doi.org/10.1016/j.frl.2023.104439

Alam, M., & Ansari, V. A. (2020). Mutual fund managers' market timing abilities: Indian evidence. *Journal of Asset Management*, *21*(4), 342–354. https://doi.org/10.1057/s41260-020-00166-1

Ali, H. (2021). *Empirical studies on the conditional CAPM and equity premium prediction*. The University of Manchester.

Ali, H., & Naz, S. (2025a). Forecasting equity premium in the face of climate policy uncertainty. *Journal of Forecasting*, *44*(2), 513–546. https://doi.org/10.1002/for.3206

Ali, H., & Naz, S. (2025b). Out-of-sample equity premium prediction: A voting approach to forecast combination. *Annals of Finance*, *21*(3), 243–281. https://doi.org/10.1007/s10436-025-00466-9

AMFI–CRISIL. (2024). *AMFI–CRISIL Factbook 2024*. Association of Mutual Funds in India & CRISIL Intelligence.

Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics*, *17*(2), 223–249. https://doi.org/10.1016/0304-405X(86)90065-6

Amihud, Y., & Mendelson, H. (1991). Liquidity, asset prices and financial policy. *Financial Analysts Journal*, *47*(6), 56–66. https://doi.org/10.2469/faj.v47.n6.56

Bechtel, A., Ranaldo, A., & Wrampelmeyer, J. (2023). Liquidity risk and funding cost. *Review of Finance*, *27*(2), 399–422. https://doi.org/10.1093/rof/rfac020

Bekaert, G., Harvey, C. R., & Lundblad, C. T. (2007). Liquidity and expected returns: Lessons from emerging markets. *Review of Financial Studies*, *20*(6), 1783–1831. https://doi.org/10.1093/rfs/hhm030

Bodson, L., Cavenaile, L., & Sougné, D. (2013). A global approach to mutual funds market timing ability. *Journal of Empirical Finance*, *20*, 96–101.

Brunnermeier, M. K., & Pedersen, L. H. (2009). Market liquidity and funding liquidity. *Review of Financial Studies*, *22*(6), 2201–2238. https://doi.org/10.1093/rfs/hhn098

Cao, C., Simin, T. T., & Wang, Y. (2013). Do mutual fund managers time market liquidity? *Journal of Financial Markets*, *16*(2), 279–307. https://doi.org/10.1016/j.finmar.2012.10.004

Chaieb, I., Errunza, V., & Langlois, H. (2021). How is liquidity priced in global markets? *Review of Financial Studies*, *34*(9), 4216–4268. https://doi.org/10.1093/rfs/hhaa125

Chen, Q., Goldstein, I., & Jiang, W. (2010). Payoff complementarities and financial fragility: Evidence from mutual fund outflows. *Journal of Financial Economics*, *97*(2), 239–262. https://doi.org/10.1016/j.jfineco.2010.03.016

Chernenko, S., & Sunderam, A. (2016). *Liquidity transformation in asset management: Evidence from the cash holdings of mutual funds*. NBER Working Paper, 22391. https://doi.org/10.3386/w22391

Coval, J., & Stafford, E. (2007). Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics*, *86*(2), 479–512. https://doi.org/10.1016/j.jfineco.2006.09.007

Dang, T. L., & Nguyen, T. M. H. (2020). Liquidity risk and stock performance during the financial crisis. *Research in International Business and Finance*, *52*, 101165. https://doi.org/10.1016/j.ribaf.2019.101165

Dong, X., Feng, S., & Sadka, R. (2019). Liquidity risk and mutual fund performance. *Management Science*, *65*(3), 1020–1041. https://doi.org/10.1287/mnsc.2017.2851

Falato, A., Goldstein, I., & Hortaçsu, A. (2021). Financial fragility in the COVID-19 crisis: The case of investment funds in corporate bond markets. *Journal of Monetary Economics*, *123*, 35–52. https://doi.org/10.1016/j.jmoneco.2021.07.001

Foran, J., & O'Sullivan, N. (2014). Liquidity risk and the performance of UK mutual funds. *International Review of Financial Analysis*, *35*, 178–189. https://doi.org/10.1016/j.irfa.2014.09.001

Giannetti, M., & Jotikasthira, C. (2024). Bond price fragility and the structure of the mutual fund industry. *Review of Financial Studies*, *37*(7), 2063–2109. https://doi.org/10.1093/rfs/hhad095

Goldstein, I., Jiang, H., & Ng, D. (2017). Investor flows and fragility in corporate bond funds. *Journal of Financial Economics*, *126*(3), 592–613. https://doi.org/10.1016/j.jfineco.2016.11.007

Goyenko, R. (2012). *Treasury liquidity, funding liquidity and asset returns.* https://doi.org/10.2139/ssrn.2023187

Hassanein, M. (2022). Systematic market risk adjusted for liquidity cost in the banking sector. *Archives of Business Research*, *10*(6), 232–246. https://doi.org/10.14738/abr.106.12513

Idzorek, T. M., Xiong, J. X., & Ibbotson, R. G. (2012). The liquidity style of mutual funds. *Financial Analysts Journal*, *68*(6), 38–53. https://doi.org/10.2469/faj.v68.n6.3

IMF (International Monetary Fund). (2025). *India: Financial sector assessment program—financial system stability assessment*. IMF Staff Country Report, 2025/055. https://doi.org/10.5089/9798229001748.002

Jiang, H., Li, Y., Sun, Z., & Wang, A. (2022). Does mutual fund illiquidity introduce fragility into asset prices? Evidence from the corporate bond market. *Journal of Financial Economics*, *143*(1), 277–302. https://doi.org/10.1016/j.jfineco.2021.05.022

Kacperczyk, M., Sialm, C., & Zheng, L. (2008). Unobserved actions of mutual funds. *The Review of Financial Studies*, *21*(6), 2379–2416. https://doi.org/10.1093/rfs/hhl041

Kang, W., Li, N., & Zhang, H. (2019). Information uncertainty and the pricing of liquidity. *Journal of Empirical Finance*, *54*, 77–96. https://doi.org/10.1016/j.jempfin.2019.08.005

Kumar, G., & Misra, A. K. (2018). Commonality in liquidity: Evidence from India's national stock exchange. *Journal of Asian Economics*, *59*, 1–15. https://doi.org/10.1016/j.asieco.2018.09.001

Kumar, S., Nor, E., Nik Azman, N. H. N., & Ali, H. (2023). Commonality in liquidity: Evidence from emerging Asian actively-managed equity funds. *Asian Academy of Management Journal of Accounting and Finance*, *19*(2), 233–257.

Lee, H. T., Lee, B. S., & Jang, B. G. (2022). Stock prices, changes in liquidity, and liquidity premia. *Finance Research Letters*, *48*, 102894. https://doi.org/10.1016/j.frl.2022.102894

Liu, W. (2006). A liquidity-augmented capital asset pricing model. *Journal of Financial Economics*, *82*(3), 631–671. https://doi.org/10.1016/j.jfineco.2005.10.001

Lou, X., & Sadka, R. (2011). Liquidity level or liquidity risk? Evidence from the financial crisis. *Financial Analysts Journal*, *67*(3), 51–62. https://doi.org/10.2469/faj.v67.n3.5

Ma, X., Zhang, X., & Liu, W. (2021). Further tests of asset pricing models: Liquidity risk matters. *Economic Modelling*, *95*, 255–273. https://doi.org/10.1016/j.econmod.2020.12.013

Ma, Y., Xiao, K., & Zeng, Y. (2022). Mutual fund liquidity transformation and reverse flight to liquidity. *Review of Financial Studies*, *35*(10), 4674–4711. https://doi.org/10.1093/rfs/hhac007

Palwishah, R., Kashif, M., Rehman, M. U., & Al-Faryan, M. A. S. (2024). Asymmetric liquidity risk and currency returns before and during COVID-19 pandemic. *International Review of Financial Analysis*, *91*, 102919. https://doi.org/10.1016/j.irfa.2023.102919

Pástor, L., & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, *111*(3), 642–685. https://doi.org/10.1086/374184

Reserve Bank of India. (2020, April). *RBI Announces INR 50,000 crore Special Liquidity Facility for Mutual Funds (SLF-MF).* Press Release 2019-2020/2276. https://www.rbi.org.in/commonman/english/Scripts/PressReleases.aspx?Id=3210

Sadka, R. (2010). Liquidity risk and the cross-section of hedge-fund returns. *Journal of Financial Economics*, *98*(1), 54–71. https://doi.org/10.1016/j.jfineco.2010.05.001

Securities and Exchange Board of India. (2017, October 6). *Categorization and rationalization of mutual fund schemes.* Circular. https://www.sebi.gov.in/legal/circulars/oct-2017/categorization-and-rationalization-of-mutual-fund-schemes_36199.html

Sensoy, A. (2017). Firm size, ownership structure, and systematic liquidity risk: The case of an emerging market. *Journal of Financial Stability*, *31*, 62–80. https://doi.org/10.1016/j.jfs.2017.06.007

Shih, Y., & Su, X. (2016). Liquidity in up and down markets for asset pricing: Evidence from the Taiwan Stock Market. *Asia-Pacific Journal of Financial Studies*, *45*(5), 729–754. https://doi.org/10.1111/ajfs.12150

Söderlind, P., & Somogyi, F. (2025). Liquidity risk and currency premia. *Management Science*, *71*(1), 518–537. https://doi.org/10.1287/mnsc.2023.01031

Virk, N. S., & Butt, H. A. (2022). Asset pricing anomalies: Liquidity risk hedgers or liquidity risk spreaders? *International Review of Financial Analysis*, *81*, 102104. https://doi.org/10.1016/j.irfa.2022.102104

Wattanatorn, W., Padungsaksawasdi, C., Chunhachinda, P., & Nathaphan, S. (2020). Mutual fund liquidity timing ability in the higher moment framework. *Research in International Business and Finance*, *51*, 101105. https://doi.org/10.1016/j.ribaf.2019.101105

Wattanatorn, W., & Tansupswatdikul, P. (2019). An ability to forecast market liquidity—evidence from South East Asia mutual fund industry. *Journal of Finance and Data Science*, *5*(1), 22–32. https://doi.org/10.1016/j.jfds.2018.10.002

Wu, Y. (2019). Asset pricing with extreme liquidity risk. *Journal of Empirical Finance*, *54*, 143–165. https://doi.org/10.1016/j.jempfin.2019.09.002

Zhu, S., Liu, H., & Boryniec, T. (2023). Long-term reversal and value effects: The role of liquidity risk. *Applied Economics*, *55*(14), 1546–1566. https://doi.org/10.1080/00036846.2022.2097634

PUEB PRESS

Economics and Business Review
https://www.ebr.edu.pl

Vol. 12 (1), 2026: 135-158
https://doi.org/10.18559/ebr.2026.1.2652

Forecasting cryptocurrencies in turbulent times: Evidence on parsimony versus model complexity

Anna Tatarczak[1] Oleksandra Humeniuk[2]

Abstract

This study examines short-term return forecasting for Bitcoin, Ethereum, and Litecoin over 2020–2024, comparing autoregressive benchmarks with Kitchen Sink and VARX-type models using point and density accuracy measures supported by Diebold–Mariano and Model Confidence Set inference. The results demonstrate that the AR(1) benchmark and parsimonious specifications incorporating cryptocurrency-specific variables consistently outperform the more elaborate linear frameworks considered, while the inclusion of macro-financial predictors offers limited benefits. Findings highlight the robustness of autoregressive dynamics for short-term cryptocurrency forecasting and underscore the importance of parsimony over model complexity. These results are consistent with a market environment characterised by high structural uncertainty, sentiment-driven trading and rapidly shifting regimes, in which additional macro-financial information contributes little to forecastability beyond short-run return momentum and crypto-specific volatility.

Keywords

- cryptocurrencies
- financial forecasting
- time series models
- emerging markets
- financial markets

JEL codes: C32, C53, G15, G17, E44.

Article received 27 September 2025, accepted 17 February 2026.

Suggested citation: Tatarczak, A., & Humeniuk, O. (2026). Forecasting cryptocurrencies in turbulent times: Evidence on parsimony versus model complexity. *Economics and Business Review*, *12*(1), 135-158. https://doi.org/10.18559/ebr.2026.1.2652

[1] Maria Curie-Sklodowska University, 5 M. Curie-Skłodowskiej Square, 20-031 Lublin, Poland, corresponding author: anna.tatarczak@umcs.pl, https://orcid.org/0000-0001-8573-5791.

[2] Maria Curie-Sklodowska University, 5 M. Curie-Skłodowskiej Square, 20-031 Lublin, Poland, oleksandrahumeniuk5@gmail.com, https://orcid.org/0009-0001-8476-1987.

Introduction

The increasing popularity of cryptocurrencies has fundamentally reshaped financial markets by introducing decentralised, transparent, yet highly volatile digital assets (Arnone, 2024; Lipton, 2021; Wątorek et al., 2023). These characteristics challenge traditional frameworks of value, exchange, and financial forecasting, making cryptocurrencies a focal point for both academic inquiry and investment practice. In this context, time series forecasting has become particularly important, as it provides tools for managing risk, designing investment strategies, and understanding broader market dynamics. Cryptocurrency markets share many statistical properties with traditional financial assets—including non-normal return distributions, volatility clustering, and non-linear dependence structures—while typically exhibiting these features with greater intensity and instability, particularly in terms of volatility levels and regime persistence (Bouchaud, 2020; Sewell, 2011). This instability is consistent with a market environment in which sentiment and disagreement dynamics in online discussions may contribute to future volatility, although the effect appears weaker for Bitcoin than for equities (Akarsu & Yilmaz, 2024). In addition, continuous 24/7 trading, evolving regulation, and rapid technological shifts further complicate modelling and forecasting efforts (Corbet et al., 2018; Zetzsche et al., 2017).

Since Bitcoin's introduction (Nakamoto, 2008), the cryptocurrency ecosystem has expanded to include assets such as Ethereum and Litecoin, with features like smart contracts and faster transactions (Buterin, 2013), thereby increasing the demand for accurate short-term forecasting. Existing studies employ both univariate and multivariate approaches (Antar, 2025; Garay et al., 2024), with mixed evidence on the benefits of incorporating macro-financial predictors (Campbell et al., 1997; Catania et al., 2018). Recent findings suggest that such variables often fail to improve short-horizon forecasts (Agarwal et al., 2024; Casarin et al., 2015; Conlon et al., 2021; Wronka, 2022), underscoring an ongoing trade-off between model complexity and forecasting reliability.

Recent syntheses document rapid growth and thematic diversification in cryptocurrency research; however, they also highlight fragmentation and the lack of integrative short-horizon forecasting frameworks (Jalal et al., 2021; Yue et al., 2021). While sustainability-focused reviews extend the agenda toward ESG issues, they do not assess forecasting gains at daily horizons (Alqudah et al., 2023). At the same time, methodological surveys note the widespread adoption of deep learning models without systematic benchmarking against parsimonious time-series baselines using formal inference (J. Zhang et al., 2024), leaving the incremental value of model complexity and external predictors unresolved. Motivated by these insights, this study examines short-term forecastability as a reflection of underlying market mechanisms. Existing empirical evidence shows that cryptocurrency prices are heavily influenced by sentiment, speculative trading, and regime-dependent market conditions rather than by slow-moving macroeconomic fundamentals (Jalal et al., 2025; Yue et al., 2021). At the same time, increasingly complex forecasting architectures have been adopted without clear evidence that they yield economically meaningful gains (C. Zhang et al., 2024). Against this backdrop, comparing parsimonious autoregressive benchmarks with multivariate and predictor-rich models can provide insight into whether short-term predictability is structurally constrained.

This study investigates the performance of several forecasting models—both univariate and multivariate—for predicting short-term returns of Bitcoin (BTC), Ethereum (ETH), and Litecoin (LTC) over the 2020–2024 period. The 2020–2024 sample deliberately covers an exceptionally turbulent phase associated with the COVID-19 pandemic, geopolitical tensions and the global inflation surge, which provides a natural stress test for forecasting models under extreme uncertainty. Rather than assuming that cryptocurrencies behave differently from other financial assets in such conditions, the paper verifies whether added model complexity offers predictive gains when even institutional macroeconomic forecasts perform poorly. Specifically, it compares benchmark autoregressive models with more complex approaches, including Kitchen Sink (KS) regression and Vector Autoregressive models with exogenous variables (VARX). The empirical study investigates whether model complexity improves short-horizon return and density forecasts relative to simple autoregressive benchmarks, and whether external predictors generate statistically significant gains in predictive accuracy. Model evaluation uses both point metrics—mean squared error (MSE), mean absolute error (MAE), and mean absolute deviation (MAD)—and probabilistic metrics, such as the Log Score (LS). Statistical significance of forecast improvements is assessed using the Diebold–Mariano test (Diebold & Mariano, 1995) and Model Confidence Set (Hansen et al., 2011) procedures. The contribution of this paper lies in its comparative analysis of forecasting models applied to cryptocurrency returns, using daily data spanning the period 2020–2024. The study assesses predictive accuracy under highly volatile and structurally unstable market conditions, comparing univariate and multivariate models from autoregressive benchmarks to Kitchen Sink and VARX specifications. By contrasting crypto-specific dynamics with models incorporating macro-financial and volatility predictors, it evaluates the incremental value of exogenous information for short-term forecasting. Model performance is assessed using point and density measures (MSE, MAE, MAD, Log Score) with formal inference via Diebold–Mariano and Model Confidence Set procedures, providing a robust multidimensional evaluation of forecasting accuracy.

The rest of the paper is organised as follows: Section 1 is devoted to the literature review on cryptocurrency time series. Section 2 provides information about data and methodology: data sources and processing, variable construction, model specifications, and the forecasting and evaluation design. In Section 3, we present our empirical findings. Section 4 is devoted to the discussion of results and their implications. Section 5 discusses study limitations and avenues for future research. The last section comprises a critical summary and conclusions.

1. Literature review

Financial time series consist of sequential observations of asset prices and related indicators and are central to forecasting and risk management (Campbell et al., 1997; Fan & Yao, 2003; Taylor, 2008; Tsay, 2005). Their modelling is complicated by stylised facts such as non-normal, heavy-tailed return distributions, excess kurtosis and skewness (Cont, 2001; Mandelbrot, 1963; Rachev et al., 2005; Sewell, 2011), weak linear autocorrelation, and persistent volatility clustering in squared or absolute returns (Bollerslev, 1986). Additional stylised facts include the leverage effect (Black, 1976; Christie, 1982), long memory in volatility (Baillie et al., 1996; Mills & Markellos, 2008), and non-linear dependence structures (Granger & Teräsvirta, 1993; Hamilton, 1994), implying intrinsic uncertainty in financial time series and motivating the use of advanced econometric and machine learning approaches (Sezer et al., 2020; C. Zhang et al., 2024). In cryptocurrency markets, this uncertainty is further reinforced by evidence of time-varying efficiency and episodic return predictability, consistent with the adaptive market hypothesis (Karasiński, 2023). Traditional models such as ARIMA and GARCH remain central for modelling short-term dynamics and volatility (Sezer et al., 2020; C. Zhang et al., 2024), but their linear structure and restrictive assumptions limit their ability to capture non-linearities, heavy tails, and structural breaks (Sezer et al., 2020). These limitations are especially pronounced in cryptocurrency markets, which are decentralised (Nakamoto, 2008), highly volatile and sentiment-driven (Bouoiyour et al., 2015), fragmented across exchanges (Feng et al., 2018; Gandal & Halaburda, 2014), continuously traded (Corbet et al., 2018), and subject to fragmented regulatory oversight (Zetzsche et al., 2017). Given these complexities, traditional econometric models often underperform in cryptocurrency settings, motivating the use of machine learning methods such as ANNs and SVR to capture non-linear and high-dimensional dependencies (Jiang, 2021; L. Zhang et al., 2017). More recently, deep learning models have gained prominence due to their ability to jointly model linear and non-linear structures directly from data (Bengio, 2012; Bouteska et al., 2024; C. Zhang et al., 2024). Architectures including LSTM, convolutional–recurrent models, and Transformers have demonstrated strong predictive performance in financial and cryptocurrency time series (Hu et al., 2021; C. Zhang et al., 2024), with LSTM remaining a standard benchmark and newer architectures offering greater flexibility at higher computational cost (Sezer et al., 2020; Zhang et al., 2024).

Recent bibliometric studies support this picture of a fragmented and methodologically heterogeneous literature on cryptocurrencies. Yue et al. (2021) show that research on the economic effects of cryptocurrencies has evolved from early work on technological foundations and miner behaviour toward analyses of price formation, risk management, and the macroeconomic implications of digital assets, but argue that the underlying transmission mechanisms and theoretical frameworks remain underdeveloped. Complementary evidence from Jalal et al. (2025) indicates that business and finance research is organised around four main streams—determinants of returns, market efficiency, (de)diversification and herding, and regulation and governance—yet many contributions rely on overlapping datasets and parallel empirical designs, which limits cumulative progress and integrative assessments of forecasting performance. More recently, Alqudah et al. (2023) document a rapid expansion of work at the intersection of cryptocurrencies and ESG, highlighting concerns about environmental externalities, speculative trading, and the long-term sustainability of digital assets as an investment class. At the same time, reviews of deep learning applications in financial forecasting emphasise that increasingly complex architectures—LSTMs, convolutional–recurrent hybrids, Transformers, and related models—have become the default choice for price prediction, even though their incremental benefits over simpler time-series benchmarks are not always systematically evaluated, especially with respect to economic interpretability and model risk (C. Zhang et al., 2024). These syntheses collectively suggest that a key unresolved issue is not only how to forecast cryptocurrency prices, but whether forecast gains delivered by complex models are economically meaningful and theoretically consistent with the underlying market structure. In this context, linear time-series models retain an important role as benchmarks with well-understood statistical properties, against which the incremental value of more complex non-linear and deep learning architectures can be assessed.

2. Methodology and data

2.1. Dataset and model specification

The empirical analysis is based on daily data covering the period from 1 January 2020 to 1 December 2024. This period is characterised by pronounced regime shifts and market-wide stress episodes, so the resulting forecasts should be interpreted as conditional on a high-volatility, crisis-like environment rather than as representative of more tranquil phases of cryptocurrency trading. The empirical strategy deliberately focuses on linear specifications (AR, VAR, VARX, KS), providing a consistent framework for formal forecast evaluation and serving as a reference point for future comparisons with non-linear and deep-learning models. Non-linear econometric specifications (such as TAR, STAR or Markov-switching models), volatility models (such as GARCH or stochastic volatility frameworks), and machine-learning or deep learning architectures (such as LSTM networks, transformers or ensemble models) are therefore intentionally excluded from the empirical design in order to keep the benchmark coherent and econometrically tractable over the turbulent 2020–2024 period. The analysis focuses on Bitcoin (BTC), Ethereum (ETH), and Litecoin (LTC) because they represent the longest-standing and most liquid segments of the cryptocurrency market, ensuring reliable high-frequency data and reducing concerns regarding idiosyncratic exchange-specific noise. BTC and ETH jointly account for the majority of market capitalisation and trading volume, making them the dominant drivers of systemic crypto-asset dynamics. LTC, while smaller, serves as a mid-cap asset with strong historical continuity, facilitating an assessment of whether forecasting performance generalises beyond the two flagship cryptocurrencies. Together, these three assets provide a balanced and representative sample of the cryptocurrency market. The dataset comprises time series for three leading cryptocurrencies—BTC, ETH, and LTC—alongside a range of macro-financial indicators that are used as potential predictors in the multivariate model configurations. All financial and macroeconomic data were retrieved from Investing.com, while additional background information on cryptocurrencies was sourced from CoinMarketCap.

While the analysis relies on spot market data from cryptocurrency exchanges, it is important to note that the investment landscape evolved substantially during 2021–2024 with the introduction and rapid expansion of exchange-traded products (ETFs/ETNs) tracking BTC and ETH on regulated exchanges such as XETR, XLON, XNAS, XNYS, and XWAR. These instruments attract a broader and more heterogeneous investor base and may follow distinct patterns in price discovery and volatility transmission. As they fall outside the scope of this study, the empirical results should be interpreted as reflecting the dynamics of spot markets rather than the behaviour of regulated exchange-traded cryptocurrency products.

Cryptocurrency-specific variables include daily closing prices, high–low price spreads, and trading volumes. These were transformed to logarithmic returns to stabilise variance and approximate stationarity. Volatility proxies were constructed as the natural logarithm of the daily high–low spread. The macro-financial variables incorporated into the analysis include the US 5-year credit default swap index (CDS_5y), the STOXX Europe 600 Index (ES_600), the Dow Jones US Gold Mining Index (GLD), the Nikkei 225 Index (NK225), the S&P 500

Index (SP500), the Dow Jones Commodity Index – Silver Subcomponent (SV), and the CBOE Volatility Index (VIX). These indicators were selected to reflect global equity market conditions, credit risk perceptions, commodity market dynamics, and investor sentiment. To address non-stationarity, most of the financial variables were transformed using the first difference of their logarithmic values. The selection of macro-financial predictors was guided by established categories that capture global risk sentiment (VIX), equity market conditions across major regions (S&P 500, STOXX Europe 600, Nikkei 225), commodity-related hedging channels (Gold Mining Index, Silver Subindex), and credit risk perceptions (US 5-year CDS). These variables are standard benchmarks in the short-horizon forecasting literature and serve to represent broad asset-class linkages without overwhelming the model with highly collinear predictors. Indicators such as NASDAQ, the US Dollar Index, interest rates or inflation measures were not included because they convey information that is strongly correlated with the equity and credit indices already present, increasing the risk of overfitting in daily-frequency models. Similarly, on-chain data (e.g., hash rate, transaction fees, miner revenue) were excluded to maintain a consistent daily sampling frequency and to ensure comparability across assets, but their incorporation presents a valuable extension for future research.

The forecasting framework comprises univariate and multivariate models. Univariate specifications include an AR(1) benchmark, a full Kitchen Sink (KS) regression with macro-financial and crypto-specific predictors, a reduced KS-noregr model excluding macro-financial variables, and an Avg forecast combining KS and KS-noregr. The multivariate set consists of a benchmark AR(1) estimated separately (M1), a VAR(3) capturing cross-asset dynamics (M2), a VAR with endogenous volatility proxies based on high–low spreads (M3), and a VARX(3) incorporating macro-financial variables as exogenous inputs (M4). All models are estimated using an expanding-window scheme to approximate real-time forecasting and limit overfitting.

2.2. Forecasting strategy and evaluation metrics

The analysis evaluates two types of forecasts: point forecasts, which estimate the conditional expectation of returns, and density forecasts, which provide the full predictive distribution. Forecasts are generated for one- to seven-day horizons ($h = 1$ to $h = 7$), and the performance of each model is assessed using both point and density forecast accuracy measures.

Point forecasts are evaluated using the following metrics: mean squared error (MSE) measures the average squared difference between forecasted and actual returns. It penalises larger forecast errors more heavily due to squaring, which makes it particularly sensitive to outliers. Mean absolute deviation (MAD) calculates the average absolute difference between forecasted and actual values. MAD also serves as an alternative to MSE due to its lower sensitivity to outliers. This metric provides a more robust evaluation by treating all forecast errors equally. The mean absolute error (MAE) calculates the average absolute difference between the predicted and the actual values, treating all errors equally regardless of their direction. A lower MAE value indicates a more accurate forecasting model. In contrast to metrics such as the mean squared error (MSE), MAE is less sensitive to large individual errors and is thus more robust to the presence of outliers.

Density forecasts are assessed using the Log Score (LS). It measures how well the forecasted distribution assigns probability to the observed outcome. Given the heavy-tailed nature of cryptocurrency returns and the presence of extreme downside events, the evaluation framework is extended to include tail-sensitive risk measures. In addition to symmetric loss functions and the Log Score, we compute one-sided left-tail Value at Risk (VaR) and Expected Shortfall (ES) at the 5% and 1% significance levels. Predictive distributions are approximated using the same parametric assumptions adopted for density forecasts, ensuring internal consistency across evaluation criteria. VaR forecasts are assessed using unconditional and conditional coverage tests, while ES accuracy is evaluated on exceedance days. This extension allows the forecast comparison to explicitly account for downside risk characteristics that are central to cryptocurrency markets. To compare forecasting models statistically, the study employs two formal evaluation procedures. The Diebold–Mariano (DM) test (Diebold & Mariano, 1995) examines the null hypothesis of equal forecast accuracy between two competing models. This test is applied separately for MSE (point forecasts) and LS (density forecasts). The Model Confidence Set (MCS) procedure (Hansen et al., 2011) identifies the subset of models that are statistically indistinguishable in performance at a given confidence level. The procedure iteratively removes the least accurate model until the null hypothesis of equal predictive accuracy can no longer be rejected. Together, these metrics and statistical tests provide a rigorous framework for evaluating model performance across both forecast horizons and evaluation dimensions.

3. Empirical results

3.1. Univariate forecast evaluation

This subsection assesses out-of-sample point forecast accuracy of four univariate models—AR(1), KS, KS-noregr, and Avg—applied to BTC, ETH, and LTC using an expanding window over horizons from h = 1 to h = 7.

Performance is evaluated using MSE, with AR(1) as the benchmark, KS including all predictors, KS-noregr restricting attention to autoregressive and crypto-specific variables, and Avg combining KS and KS-noregr forecasts, allowing the incremental value of macro-financial information and model complexity to be assessed.

As shown in Table 1, the AR(1) benchmark performs robustly for Bitcoin and Ethereum across all horizons, while the full KS specification systematically underperforms, consistent with overfitting from noisy macro-financial predictors. In contrast, the parsimonious KS-noregr and Avg models dominate KS and, particularly for Ethereum and Litecoin, often match or outperform AR(1). The Model Confidence Set applied to MSE excludes KS for all assets and retains AR(1), KS-noregr, and Avg, while Diebold–Mariano tests yield few significant differences but remain directionally consistent with these results.

Table 1. Mean squared error for univariate models (in %)

h	1	2	3	4	5	6	7
Bitcoin							
AR1	6.11	6.02	6.01	6.01	6.03	6.03	6.04
KS	6.20	6.13	6.13	6.16	6.16	6.16	6.15
KS-noregr	6.12	6.07	6.09	6.08	6.10	6.07	6.12
Avg	6.12	6.07	6.09	6.08	6.10	6.07	6.12
Ethereum							
AR1	7.78	7.62	7.60	7.62	7.64	7.66	7.68
KS	7.99	7.81	7.81	7.86	7.87	7.92	7.93
KS-noregr	7.77	7.62	7.58	7.64	7.63	7.68	7.67
Avg	7.77	7.62	7.58	7.64	7.63	7.68	7.67
Litecoin							
AR1	10.91	10.53	10.48	10.51	10.53	10.55	10.57
KS	11.12	10.96	10.96	10.99	10.99	11.05	11.03
KS-noregr	10.79	10.65	10.50	10.55	10.47	10.54	10.47
Avg	10.78	10.65	10.50	10.55	10.47	10.54	10.47

Source: own work.

In addition to MSE, Table 2 presents the mean absolute deviation (MAD), which provides a complementary view of forecast accuracy. The results broadly confirm the MSE findings: KS-noregr and Avg yield lower errors than KS, with particularly strong performance for Litecoin.

Table 2. Mean absolute deviation over the forecast horizon (%)

h	1	2	3	4	5	6	7
Bitcoin							
AR1	1.72	1.71	1.69	1.69	1.71	1.71	1.71
KS	1.75	1.74	1.73	1.74	1.74	1.74	1.74
KS-noregr	1.71	1.70	1.70	1.70	1.70	1.70	1.70
Avg	1.71	1.69	1.70	1.70	1.69	1.69	1.69
Ethereum							
AR1	1.89	1.86	1.86	1.86	1.87	1.87	1.87
KS	1.93	1.91	1.91	1.92	1.92	1.93	1.93
KS-noregr	1.89	1.87	1.87	1.87	1.87	1.88	1.87
Avg	1.89	1.87	1.87	1.87	1.87	1.88	1.87
Litecoin							
AR1	2.22	2.18	2.17	2.17	2.17	2.18	2.18
KS	2.24	2.22	2.21	2.22	2.22	2.23	2.22
KS-noregr	2.22	2.20	2.22	2.20	2.20	2.22	2.19
Avg	2.22	2.20	2.22	2.20	2.20	2.22	2.19

Source: own work.

Diebold–Mariano tests indicate that differences in point forecast accuracy are generally not statistically significant; the only significant improvement over AR(1) occurs for Litecoin at horizon $h = 2$ using the full KS model ($p = 0.0439$). Predictor relevance analysis shows consistent patterns across assets: ES600 is the most influential macro-financial predictor for Bitcoin and Ethereum, alongside strong autoregressive spillovers—particularly from Bitcoin to Ethereum—while Litecoin is primarily driven by lagged returns of Bitcoin and Ethereum, with a secondary role for the S&P 500. Rankings based on coefficient magnitude corroborate these findings, reinforcing the hierarchical structure of influence within the cryptocurrency market.

Turning to density forecasts, Table 3 reports the Log Score (LS) values for each univariate model. The KS model consistently achieves the lowest LS values for all three cryptocurrencies, suggesting superior probabilistic accuracy. These improvements, however, are not always confirmed as statistically significant by the DM test.

The Model Confidence Set (MCS) procedure applied to Log Score losses yields a differentiated picture across assets. For Bitcoin, only the AR(1) benchmark remains in the superior set, confirming its robustness in probabilistic forecasting. In contrast, for Ethereum and Litecoin, the KS-noregr and Avg

Table 3. Log Score (LS) of univariate models over the forecast horizon

h	1	2	3	4	5	6	7
Bitcoin							
AR1	837.31	837.88	835.95	833.34	830.52	828.25	825.53
KS	834.24	834.30	832.00	828.81	826.75	824.29	822.49
KS-noregr	836.92	836.14	833.28	831.36	828.43	826.93	823.13
Avg	836.92	836.14	833.28	831.36	828.43	826.93	823.13
Ethereum							
AR1	792.84	794.53	792.92	790.13	787.48	784.95	782.20
KS	787.88	789.95	787.90	784.46	782.25	778.91	776.45
KS-noregr	793.12	794.53	793.40	789.81	787.85	784.48	782.51
Avg	793.12	794.53	793.40	789.81	787.85	784.48	782.51
Litecoin							
AR1	730.83	735.28	734.14	731.79	729.34	726.92	724.68
KS	726.24	728.08	726.06	723.49	721.39	718.53	716.91
KS-noregr	732.94	733.33	733.83	730.95	730.28	727.10	726.38
Avg	732.94	733.33	733.82	730.95	730.28	727.10	726.38

Source: own work.

specifications are retained, while the full KS model is consistently excluded due to its statistical instability. Taken together, these findings indicate that univariate models centred on crypto-specific information—lagged returns and high–low volatility proxies—offer more reliable density forecasts than specifications that indiscriminately incorporate macro-financial variables, whose short-term predictive contribution appears limited.

To account for the heavy-tailed nature of cryptocurrency returns and the presence of extreme downside events, the forecast evaluation is extended to include tail-sensitive risk measures. Table 4 reports one-day-ahead Value at Risk (VaR) and Expected Shortfall (ES) backtesting results at the 5% and 1% significance levels for all univariate specifications. In line with the point and density forecast results, the AR(1) benchmark and the parsimonious KS-noregr model exhibit the most stable tail-risk performance across assets. Their empirical exceedance rates are close to nominal levels, and both unconditional and conditional coverage tests generally fail to reject correct calibration. By contrast, the full Kitchen Sink (KS) specification displays weaker tail behaviour, particularly for Ethereum, where violations cluster over short horizons, leading to rejections of conditional coverage despite acceptable average accuracy. This indicates that the inclusion of broad macro-financial predictors may deteriorate tail-risk properties even when symmetric error metrics or Log Scores suggest comparable performance. Overall, tail-focused diagnostics reinforce the central finding of this study: parsimonious autoregressive structures and cryptocurrency-specific information deliver more robust short-term forecasts, not only in terms of average accuracy but also with respect to downside risk.

Table 4. VaR and ES backtesting results for one-day-ahead forecasts

Asset	Model	VaR 5% hit ()	CC *p*-value	VaR 1% hit (%)	CC *p*-value	ES tail loss
BTC	AR(1)	4.6	0.44	1.9	0.26	low
BTC	KS	5.7	0.26	1.9	0.26	medium
BTC	KS-noregr	4.6	0.44	1.9	0.26	low
ETH	AR(1)	4.4	0.92	0.8	0.92	low
ETH	KS	6.0	0.02	0.5	0.63	high
ETH	KS-noregr	4.4	0.92	0.8	0.92	low
LTC	AR(1)	4.1	0.75	1.4	0.75	low
LTC	KS	4.4	0.43	1.4	0.75	medium
LTC	KS-noregr	4.6	0.44	1.4	0.75	low

Source: own work.

3.2. Multivariate forecast evaluation

As shown in Tables 5 and 6, the benchmark AR(1) specification (M1) attains the lowest MSE and MAE at most horizons, while the VAR(3) model (M2) performs similarly without systematic gains. Adding endogenous volatility proxies (M3) yields only marginal and statistically insignificant improvements at longer horizons, whereas the VARX(3) model with macro-financial predictors (M4) consistently underperforms, particularly beyond $h = 4$, indicating limited short-term forecasting gains from cross-asset spillovers or macroeconomic information.

Table 5. Mean squared error for multivariate models (in %)

h	1	2	3	4	5	6	7
M1	5.58	6.45	7.25	8.16	9.44	11.00	13.71
M2	5.84	6.57	7.34	8.13	9.43	11.01	13.74
M3	6.11	8.84	7.90	8.35	9.55	11.02	13.72
M4	6.93	8.13	9.38	10.61	12.37	14.34	17.86

Source: own work.

The Diebold–Mariano test for MSE does not yield statistically significant results. For the model comparisons, the null hypothesis of equal predictive accuracy between AR(1) and alternative specifications could not be rejected at the 5% significance level. The Model Confidence Set procedure indicates that M1 belongs to the Superior Set of Models at a 10% confidence level, using MSE as the loss metric. M2 and M3 occasionally join the set, particularly at short-term horizons. M4 is regularly eliminated due to its limited predictive contribution in this setup.

Table 6. Mean absolute error for multivariate models (in %)

h	**1**	**2**	**3**	**4**	**5**	**6**	**7**
M1	1.43	1.59	1.74	1.87	2.06	2.27	2.73
M2	1.51	1.63	1.79	1.85	2.06	2.25	2.73
M3	1.63	1.98	1.99	1.96	2.13	2.27	2.73
M4	1.75	1.94	2.24	2.43	2.78	3.04	3.65

Source: own work.

Diebold–Mariano tests based on MAE confirm that M1 consistently outperforms M4 across horizons, while horizon $h = 7$ yields no significant results due to limited observations. Simpler models (M1 and M2) provide more reliable short-term forecasts, and MCS results for MAE mirror those for MSE, with M1 and M2 retained in the superior set at the 10% level. Density forecast results (Table 7) further show strong AR(1) performance in terms of Log Score; although M3 and M4 occasionally improve scores at longer horizons, these gains are modest and statistically insignificant.

Table 7. Log Score (LS) of multivariate models over the forecast horizon

h	**1**	**2**	**3**	**4**	**5**	**6**	**7**
M1	2.30	2.56	2.21	2.17	2.10	2.02	1.87
M2	1.83	1.82	1.81	1.80	1.78	1.76	1.72
M3	1.86	1.75	1.78	1.77	1.75	1.74	1.71
M4	1.80	1.81	1.80	1.77	1.76	1.75	1.72

Source: own work.

Table 8 provides a breakdown of Log Scores for each cryptocurrency. Notably, Bitcoin consistently achieves higher predictive accuracy compared to Ethereum and Litecoin across all model specifications. The decline in Log Scores across forecast horizons reflects the increased difficulty of making longer-term predictions in volatile markets.

Table 8. Log Score (LS) for multivariate models over the forecast horizon for each cryptocurrency

h	1	2	3	4	5	6	7
Bitcoin							
M1	2.53	2.50	2.47	2.44	2.39	2.33	2.22
M2	1.98	1.98	1.96	1.95	1.94	1.92	1.89
M3	2.05	1.93	1.92	1.91	1.91	1.91	1.88
M4	1.95	1.96	1.95	1.92	1.93	1.92	1.89
Ethereum							
M1	2.26	2.21	2.17	2.12	2.05	1.94	1.79
M2	1.79	1.79	1.78	1.77	1.75	1.72	1.68
M3	1.86	1.69	1.75	1.75	1.73	1.71	1.68
M4	1.77	1.77	1.77	1.73	1.73	1.71	1.69
Litecoin							
M1	2.17	2.05	2.00	1.94	1.86	1.76	1.61
M2	1.72	1.70	1.68	1.67	1.65	1.62	1.58
M3	1.66	1.62	1.64	1.64	1.62	1.61	1.57
M4	1.69	1.69	1.68	1.65	1.64	1.63	1.58

Source: own work.

The Diebold–Mariano test results at the 5% significance level point out that the M1 consistently exhibits statistically superior predictive performance compared to the other models across short to medium forecast horizons (h1–h5). At longer forecast horizons (h6–h7), however, the model's predictive accuracy differences become statistically insignificant, suggesting an increase in forecast uncertainty. Detailed Diebold–Mariano test *p*-values for pairwise model comparisons across horizons and loss functions are reported in Appendix A.

The Model Confidence Set analysis also confirmed that only AR(1) consistently provided superior predictive accuracy across Bitcoin, Ethereum, and Litecoin returns. More complex models, such as VAR and VARX, were eliminated from the superior set at a 10% significance level, indicating that additional variables or model complexity did not yield better forecast performance in this application.

Figure 1 illustrates the cumulative Log Score differences for each model relative to the AR(1) benchmark. Positive values indicate better performance. M2 and M3 show minor gains at longer horizons, but M1 remains dominant across most horizons.

The analysis of cumulative adjusted Log Scores indicates that the baseline model M1 consistently outperforms its competitors across all forecast

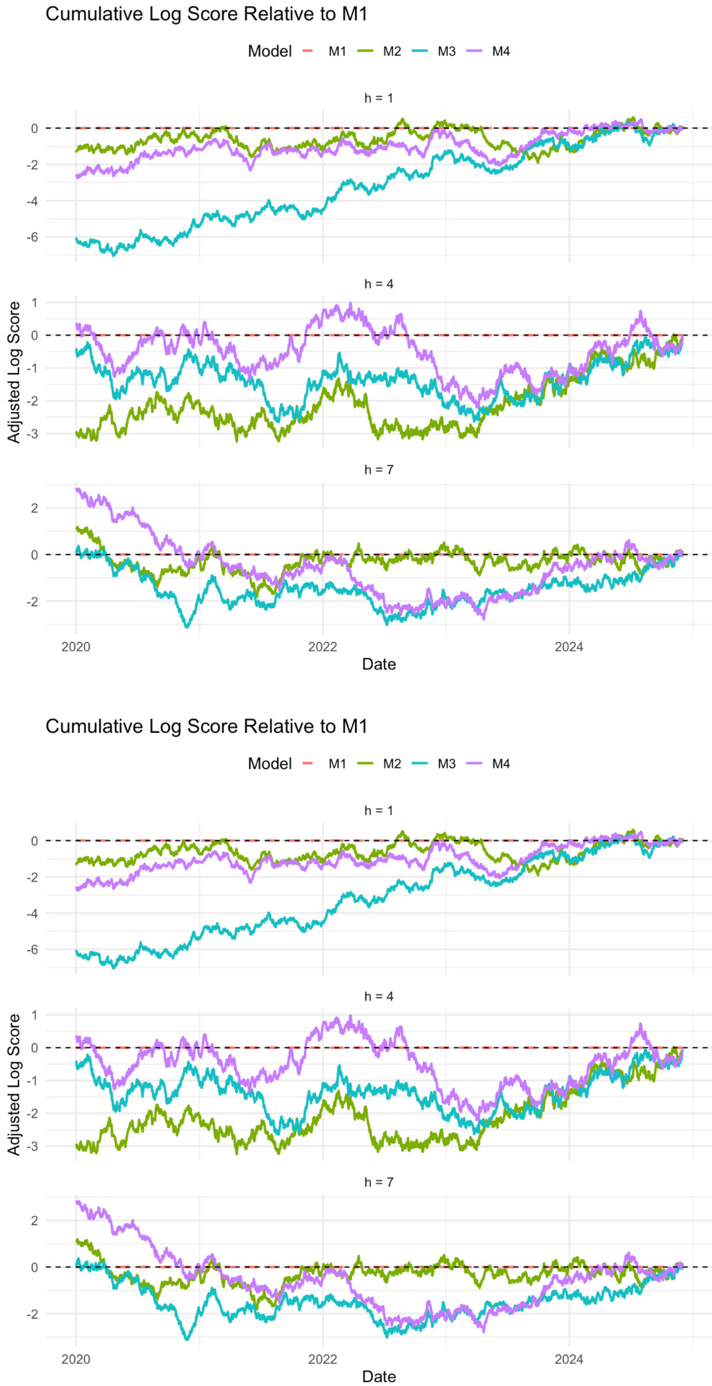

Figure 1. Cumulative Log Score visualisation

Source: own work.

horizons, showing the most significant lead at shorter horizons. In contrast, models M2 and M3 exhibit lower predictive accuracy, especially for short-term forecasts. Model M4 shows some competitiveness at medium horizons but does not consistently surpass M1. As the forecast horizons extend, the errors increase, and the performance gaps between models narrow, though M1 still retains a slight edge.

4. Discussion

The dominance of the AR(1) benchmark and closely related crypto-specific linear specifications aligns with economic evidence suggesting that short-horizon cryptocurrency returns are driven primarily by sentiment, momentum, and market microstructure rather than by macro-financial fundamentals. Bibliometric studies indicate that speculative behaviour, attention cycles and regime shifts play a central role in shaping price dynamics (Jalal et al., 2025; Yue et al., 2021). In such environments, additional predictors often introduce noise rather than signal, limiting the gains from complex multivariate structures. The weak and unstable contribution of macro variables is consistent with findings that cryptocurrencies exhibit low and highly time-varying integration with traditional financial assets, particularly during stress periods (Alqudah et al., 2023), including pandemic-related episodes, in which the safe-haven and diversification properties of cryptocurrencies appear regime-dependent and unstable (Barbu et al., 2022). Deep learning reviews also note that model sophistication cannot fully overcome structural instability and noisy patterns at daily frequencies (J. Zhang et al., 2024). Together, these mechanisms provide an economic interpretation for the strong performance of parsimonious linear specifications observed in this study. The limited contribution of macro-financial variables in our setting should not be interpreted as evidence of their irrelevance, but rather as a reflection of the low signal-to-noise ratio at daily horizons and the strong collinearity among global risk and equity indicators. More granular predictors—such as on-chain activity measures—may capture additional structure, but incorporating them requires a dedicated modelling framework beyond the scope of the present benchmark study.

From the perspective of the current forecasting frontier, non-linear and deep learning approaches are natural candidates to capture residual non-linearities and complex interactions that lie beyond linear dynamics. In this study, however, we treat linear models as economically interpretable benchmarks under stress conditions, so that subsequent work can evaluate whether additional complexity in non-linear or deep learning frameworks delivers stable and economically meaningful gains over these baselines. Accordingly, the contribution of this paper lies in establishing a rigorous post-2020 benchmark for linear time-series models rather than proposing a new state-of-the-art non-linear forecasting architecture.

The results should not be interpreted as evidence that cryptocurrency markets were more forecastable than other assets during this turbulent period. The small and often insignificant performance differentials between models mirror the widespread forecasting failures observed in institutional macroeconomic and financial projections, and indicate that, under such conditions, added econometric complexity yields no meaningful gains over parsimonious autoregressive benchmarks.

The empirical results indicate that parsimonious models based on autoregressive dynamics and crypto-specific indicators outperform more complex specifications at short horizons. The AR(1) model exhibits stable performance across all assets, while the full Kitchen Sink (KS) specification suffers from overfitting, particularly for Bitcoin and Ethereum. By contrast, the reduced KS-noregr model performs comparably to or better than AR(1) in both point and density forecasts. These findings are consistent with Campbell et al. (1997) and Catania et al. (2018), who document the limited short-term forecasting gains from macro-financial predictors.

The multivariate results confirm the dominance of the AR(1) benchmark, which consistently matches or outperforms VAR and VARX models across horizons. Gains from adding volatility proxies or macro-financial variables are generally marginal and statistically insignificant; while the VAR with high–low spreads (M3) shows occasional improvements at medium horizons, the VARX specification (M4) underperforms. These findings are consistent with Casarin et al. (2015) and Corbet et al. (2018), who document the horizon-dependent and often weak short-run predictive contributions of external financial information.

Probabilistic evaluation using the Log Score shows that while the KS model occasionally improves density forecasts—particularly for Ethereum—these gains are rarely statistically significant and are not confirmed by the Model Confidence Set, whereas AR(1) is consistently retained in the superior set. This robustness of simple autoregressive models is consistent with the findings of Adhikari and Agrawal (2014), especially in noisy financial environments.

These results are consistent with prior evidence emphasising the importance of parsimony and domain-specific model design in financial forecasting (Mills & Markellos, 2008; Sezer et al., 2020). In highly volatile and structurally unstable cryptocurrency markets (Bouoiyour et al., 2015; Sewell, 2011), complex models with noisy external predictors may reduce short-term accuracy, while the relevance of macro-financial variables is likely horizon- and target-dependent. This interpretation aligns with Baillie et al. (1996) and Jiang (2021), who show that non-linear and long-memory effects are more effectively captured over longer samples or for alternative targets such as volatility. Moreover, the limited gains from macro-financial variables in this study should be interpreted in light of the increasingly time-varying and regime-dependent integration of cryptocurrency markets with traditional financial systems. Recent evidence indicates that, particularly since 2020, cryptocurrencies have become more strongly correlated with major equity indices—especially technology stocks—and US macroeconomic conditions, although this integration remains unstable and horizon-dependent (Wątorek et al., 2023).

5. Limitations and avenues for further research

This study has several limitations that should be considered when interpreting the results. Firstly, the analysis is restricted to three major cryptocurrencies—Bitcoin, Ethereum, and Litecoin—over the period 2020–2024. The findings may therefore not generalisable to smaller tokens, stablecoins, or other digital assets with different market structures and liquidity conditions. Secondly, the dataset coincides with major regime shifts, including the COVID-19 aftermath, geopolitical tensions and the global inflation surge, which may have introduced structural breaks that the expanding-window approach does not fully capture. Since the study does not include a comparison with the more tranquil pre-2020 period (e.g., 2015–2019), we cannot formally assess whether the dominance of parsimonious models persists across different volatility regimes. Thirdly, the use of aggregated data sources may involve measurement errors. In particular, mismatches between continuous cryptocurrency trading and macro-financial indicators with business-day frequency could lead to timing inconsistencies. The analysis does not incorporate regulated exchange-traded products (ETFs/ETNs) that have emerged as major investment vehicles for BTC and ETH since 2021. Because these instruments may differ from spot markets in terms of volatility transmission, liquidity and investor composition, the generalisability of the results to the ETF/ETN segment is inherently limited. Fourthly, the study evaluates only linear and relatively simple models (AR(1), VAR, VARX, KS). Within this set, the autoregressive benchmark is restricted to an AR(1) specification; more flexible but still parsimonious time-series models such as AR(p), ARIMA, HAR-RV or GARCH-type frameworks are not explored. This restriction is deliberate: the objective is to establish a well-understood linear benchmark for the turbulent 2020–2024 period rather than to propose a new state-of-the-art forecasting architecture.

Consequently, the conclusion that the AR(1) benchmark outperforms more complex specifications should be interpreted as conditional on the restricted class of models considered here, and future research should assess whether alternative low-dimensional autoregressive or volatility models can match or surpass this baseline. Fifthly, forecasts are based on daily data and do not account for intraday patterns, transaction costs, or exchange heterogeneity, which limits their practical applicability to trading strategies. Sixthly, the forecasting horizons examined in this study are deliberately restricted to short-term windows (1–7 days). While this design is appropriate for evaluating high-frequency return predictability, it may limit the ability of multivariate and VARX-type models to demonstrate their strengths. Macro-financial variables typically operate through slower transmission channels and may exert influence over medium- or long-term horizons rather than at daily frequencies. As a result, the finding that macroeconomic predictors provide limited incremental value should be interpreted as conditional on these short horizons and not more broadly as evidence of their irrelevance. Extending the analysis to multi-week or multi-month horizons would allow future research to assess whether the predictive contribution of macro-financial information increases when the forecast window aligns more closely with the underlying economic adjustment dynamics. Finally, the reliance on standard evaluation metrics may inadequately reflect tail risk and economic significance; accordingly, the findings should be regarded as indicative rather than conclusive.

Conclusions

This study assesses the short-term forecasting performance of univariate and multivariate time series models for daily returns of Bitcoin, Ethereum, and Litecoin over 2020–2024, comparing autoregressive benchmarks with more complex specifications using standard point and density evaluation metrics supported by formal statistical inference. The findings lead to some conclusions. Firstly, the AR(1) specification provides a well-performing baseline for cryptocurrency return forecasting within the linear model class considered. Their consistent inclusion in the superior model sets across assets and horizons underscores their robustness, particularly at short horizons. Secondly, incorporating cryptocurrency-specific variables—especially lagged returns and high–low volatility spreads—yields modest but tangible improvements in forecast accuracy. The reduced Kitchen Sink model (KS-noregr), which focuses on these domain-relevant predictors, frequently outperformed more elaborate specifications, especially for Ethereum and Litecoin. Thirdly, complex multivariate models that integrate macro-financial variables (e.g., VARX) did not systematically enhance forecast accuracy. In several cases, the inclusion of exogenous predictors reduced predictive performance, reflecting issues of overfitting and the limited short-term relevance of macroeconomic signals for cryptocurrency prices. These results are consistent with prior research indicating that digital asset returns are driven more by endogenous momentum and speculative trading than by traditional financial indicators in the short run (Bouoiyour et al., 2015; Conlon et al., 2021; Corbet et al., 2018). Fourthly, density forecasts highlighted some advantages of the full KS model, particularly for Ethereum, but these gains were inconsistent and often lacked statistical significance. By contrast, the AR(1) model demonstrated stable performance across both point and density evaluations, reinforcing the value of parsimonious structures. Overall, within the turbulent post-2020 environment analysed here and within the set of linear models considered, the evidence suggests that parsimony combined with careful variable selection is more effective for short-term cryptocurrency forecasting than additional model complexity. While non-linear and machine learning models may offer gains over longer horizons or alternative targets (e.g., volatility or tail risk), the results confirm the practical relevance of autoregressive dynamics and crypto-specific features. The study provides a unified benchmarking of parsimonious autoregressive baselines against kitchen-sink and VAR/VARX models using formal inference (Diebold–Mariano; Model Confidence Set), clarifying the incremental value of crypto-specific versus macro-financial predictors. Consistent with recent reviews (Jalal et al., 2025; Yue et al., 2021; C. Zhang et al., 2024), the evidence shows that parsimony dominates at daily horizons, establishing transparent and economically interpretable benchmark models and motivating future extensions using high-frequency, on-chain, and non-linear or ensemble frameworks.

From a practical perspective, the findings suggest that parsimonious autoregressive and crypto-specific models account for well-performing tools for short-term forecasting and risk management, while the limited short-run predictive role of macro-financial variables implies that regulatory risk assessments should focus more on market microstructure, speculative behaviour, and crypto-specific volatility dynamics than on broad macroeconomic indicators.

Appendix

Table A1. Mean Squared Error (MSE)

Asset	Horizon	AR(1) vs KS	AR(1) vs KS-noregr	AR(1) vs Avg
BTC	$h = 1$	0.4321	0.6814	0.5178
BTC	$h = 3$	0.0847	0.2941	0.1736
BTC	$h = 7$	0.0632	0.2119	0.0915
ETH	$h = 1$	0.3884	0.7442	0.6017
ETH	$h = 3$	0.0173	0.5628	0.0419
ETH	$h = 7$	0.0191	0.4896	0.0384
LTC	$h = 1$	0.5176	0.8092	0.6931
LTC	$h = 3$	0.1412	0.6034	0.2289
LTC	$h = 7$	0.0924	0.4187	0.1561

Source: own work.

Table A2. Log Score (LS)

Asset	Horizon	AR(1) vs KS	AR(1) vs KS-noregr	AR(1) vs Avg
BTC	$h = 1$	0.4762	0.7028	0.6115
BTC	$h = 3$	0.0819	0.3314	0.0976
BTC	$h = 7$	0.0687	0.2885	0.0894
ETH	$h = 1$	0.4027	0.7619	0.5842
ETH	$h = 3$	0.0236	0.5941	0.0498
ETH	$h = 7$	0.0214	0.5237	0.0441
LTC	$h = 1$	0.5381	0.8124	0.7016
LTC	$h = 3$	0.1597	0.6472	0.2443
LTC	$h = 7$	0.0886	0.4593	0.1698

Source: own work.

References

Adhikari, R., & Agrawal, R. K. (2014). A combination of artificial neural network and random walk models for financial time series forecasting. *Neural Computing and Applications*, *24*(6), 1441–1449. https://doi.org/10.1007/s00521-013-1386-y

Agarwal, M., Gill, K. S., Upadhyay, D., Dangi, S., & Chythanya, K. R. (2024, April). *The evolution of cryptocurrencies: Analysis of Bitcoin, Ethereum, Bit connect and Dogecoin in comparison*. 2024 IEEE 9th International Conference for Convergence in Technology (pp. 1–6). https://doi.org/10.1109/i2ct61223.2024.10543872

Akarsu, S., & Yilmaz, N. (2024). Social media disagreement and financial markets: A comparison of stocks and Bitcoin. *Economics and Business Review*, *10*(4), 189–213. https://doi.org/10.18559/ebr.2024.4.1683

Alqudah, M., Ferruz, L., Martín, E., Qudah, H., & Hamdan, F. (2023). The sustainability of investing in cryptocurrencies: A bibliometric analysis of research trends. *International Journal of Financial Studies*, *11*(3), 93. https://doi.org/10.3390/ijfs11030093

Antar, M. (2025). Quantile analysis of Bitcoin returns: Uncovering market dynamics. *Journal of Risk Finance*, *26*(1), 122–146. https://doi.org/10.1108/jrf-05-2024-0154

Arnone, G. (2024). Introduction to cryptocurrencies and digital currencies. In G. Arnone, *Navigating the world of cryptocurrencies: Technology, economics, regulations, and future trends* (pp. 1–11). Springer Nature. https://doi.org/10.1007/978-3-031-69176-8_1

Baillie, R. T., Bollerslev, T., & Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *74*(1), 3–30.

Barbu, T. C., Boitan, I. A., & Cepoi, C. O. (2022). Are cryptocurrencies safe havens during the COVID-19 pandemic? A threshold regression perspective with pandemic-related benchmarks. *Economics and Business Review*, *8*(2), 29–49. https://doi.org/10.18559/ebr.2022.2.3

Bengio, Y. (2012). Deep learning of representations for unsupervised and transfer learning. *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*, *27*, 17–36.

Black, F. (1976). Studies of stock price volatility changes. *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economic Statistics Section*, 177–181.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31*(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1

Bouchaud, J. P. (2013). Crises and collective socio-economic phenomena: Simple models and challenges. *Journal of Statistical Physics*, *151*(3), 567–606. https://doi.org/10.1007/s10955-012-0687-3

Bouoiyour, J., Selmi, R., & Tiwari, A. K. (2015). Is Bitcoin business income or speculative foolery? New ideas through an improved frequency domain analysis. *Annals of Financial Economics*, *10*(1), 1550002. https://doi.org/10.1142/s2010495215500025

Bouteska, A., Abedin, M. Z., Hajek, P., & Yuan, K. (2024). Cryptocurrency price forecasting—a comparative analysis of ensemble learning and deep learning methods.

International Review of Financial Analysis, *92*, 103055. https://doi.org/10.1016/j.irfa.2023.103055

Buterin, V. (2013). *Ethereum whitepaper: A Next-generation smart contract and decentralized application platform*. Ethereum Foundation.

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The econometrics of financial markets*. Princeton University Press.

Casarin, R., Grassi, S., Ravazzolo, F., & van Dijk, H. K. (2015). *Dynamic predictive density combinations for large data sets in economics and finance*. Tinbergen Institute Discussion Paper, 15-084/III. https://hdl.handle.net/10419/125081

Catania, L., Grassi, S., & Ravazzolo, F. (2018). *Forecasting cryptocurrencies financial time series*. BI Norwegian Business School.

Christie, A. A. (1982). The stochastic behavior of common stock variances: Value, leverage and interest rate effects. *Journal of Financial Economics*, *10*(4), 407–432. https://doi.org/10.1016/0304-405x(82)90018-6

Conlon, T., Corbet, S., & McGee, R. J. (2021). Inflation and cryptocurrencies revisited: A time-scale analysis. *Economics Letters*, *206*, 109996. https://doi.org/10.1016/j.econlet.2021.109996

Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, *1*(2), 223–236. https://doi.org/10.1080/713665670

Corbet, S., Meegan, A., Larkin, C., Lucey, B., & Yarovaya, L. (2018). Exploring the dynamic relationships between cryptocurrencies and other financial assets. *Economics Letters*, *165*, 28–34. https://doi.org/10.1016/j.econlet.2018.01.004

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13, 253–263. https://doi.org/10.1080/07350015.1995.10524599

Fan, J., & Yao, Q. (2003). *Nonlinear time series: nonparametric and parametric methods*. Springer.

Feng, W., Wang, Y., & Zhang, Z. (2018). Informed trading in the Bitcoin market. *Finance Research Letters*, *26*, 63–70. https://doi.org/10.1016/j.frl.2017.11.009

Gandal, N., & Halaburda, H. (2014). *Competition in the cryptocurrency market*. Bank of Canada.

Garay, J., Kiayias, A., & Leonardos, N. (2024). The bitcoin backbone protocol: Analysis and applications. *Journal of the ACM*, *71*(4), 1–49. https://doi.org/10.1145/3653445

Granger, C. W. J., & Teräsvirta, T. (1993). *Modelling nonlinear economic relationships*. Oxford University Press.

Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press.

Hansen P. R., Lunde A., & Nason J. M. (2011). The model confidence set. *Econometrica*, *79*(2), 453–497. https://doi.org/10.3982/ecta5771

Hu, Z., Zhao, Y., & Khushi, M. (2021). A survey of forex and stock price prediction using deep learning. *Applied System Innovation*, *4*(1), 9. https://doi.org/10.3390/asi4010009

Jalal, R. N. U. D., Alon, I., & Paltrinieri, A. (2025). A bibliometric review of cryptocurrencies as a financial asset. *Technology Analysis & Strategic Management*, *37*(4), 432–447. https://doi.org/10.1080/09537325.2021.1939001

Jiang, W. (2021). Applications of deep learning in stock market prediction: Recent progress. *Expert Systems with Applications*, 184, 115537. https://doi.org/10.1016/j.eswa.2021.115537

Karasiński, J. (2023). The adaptive market hypothesis and the return predictability in the cryptocurrency markets. *Economics and Business Review*, *9*(1), 94–118. https://doi.org/10.18559/ebr.2023.1.4

Lipton, A. (2021). Cryptocurrencies change everything. *Quantitative Finance*, *21*(8), 1257–1262. https://doi.org/10.1080/14697688.2021.1944490

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, 36(4), 394–419. https://doi.org/10.1086/294632

Mills, T. C., & Markellos, R. N. (2008). *The econometric modelling of financial time series* (2nd ed.). Cambridge University Press.

Nakamoto, S. (2008). *Bitcoin: A peer-to-peer electronic cash System*. https://bitcoin.org/bitcoin.pdf

Rachev, S. T., Mittnik, S., Fabozzi, F. J., Focardi, S. M., & Jasic, T. (2005). *Financial econometrics: From basics to advanced modeling techniques.* John Wiley & Sons.

Sewell, M. (2011). *Characterization of financial time series*. UCL Department of Computer Science.

Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied Soft Computing*, *90*, 106181. https://doi.org/10.1016/j.asoc.2020.106181

Taylor, S. J. (2008). *Modelling financial time series*. World Scientific.

Tsay, R. S. (2005). *Analysis of financial time series* (2nd ed.). John Wiley & Sons.

Wątorek, M., Kwapień, J., & Drożdż, S. (2023). Cryptocurrencies are becoming part of the world global financial market. *Entropy*, *25*(2), 377. https://doi.org/10.3390/e25020377

Wronka, C. (2022). Money laundering through cryptocurrencies-analysis of the phenomenon and appropriate prevention measures. *Journal of Money Laundering Control*, *25*(1), 79–94. https://doi.org/10.1108/jmlc-02-2021-0017

Yue, Y., Li, X., Zhang, D., & Wang, S. (2021). How cryptocurrency affects economy? A network analysis using bibliometric methods. *International Review of Financial Analysis*, *77*, 101869. https://doi.org/10.1016/j.irfa.2021.101869

Zetzsche, D. A., Buckley, R. P., Arner, D. W., & Föhr, L. (2017). *The ICO gold rush: It's a scam, it's a bubble, it's a super challenge for regulators*. University of Luxembourg Law Working Paper, 11/2017. https://doi.org/10.2139/ssrn.3072298

Zhang, C., Sjarif, N. N. A., & Ibrahim, R. (2024). Deep learning models for price forecasting of financial time series: A review of recent advancements: 2020–2022. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *14*(1), e1519. https://doi.org/10.1002/widm.1519

Zhang, J., Cai, K., & Wen, J. (2024). A survey of deep learning applications in cryptocurrency. *Iscience, 27*(1), 108509. https://doi.org/10.1016/j.isci.2023.108509

Zhang, L., Aggarwal, C., & Qi, G. J. (2017, August). Stock price prediction via discovering multi-frequency trading patterns. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 2141–2149). https://doi.org/10.1145/3097983.3098117

PUEB PRESS

Economics and Business Review
https://www.ebr.edu.pl

Vol. 12 (1), 2026: 159-182
https://doi.org/10.18559/ebr.2026.1.2793

From digital mining to market prices: An empirical analysis of the relationship between energy consumption and price dynamics of Bitcoin and Ether

Levent Sezal[1]

Abstract

This study compares the relationships between Bitcoin and Ethereum's energy consumption and price dynamics. Using daily frequency data, Augmented Dickey–Fuller (ADF), Phillips–Perron (PP), ARDL cointegration tests, and Toda–Yamamoto causality analysis were applied to evaluate the effects of cryptocurrency markets on energy demand from both short-term and long-term perspectives. The results indicate that there is a long-term cointegration relationship between energy consumption and prices for Bitcoin, and also unidirectional causality from prices to energy consumption. In contrast, ARDL boundary test results for Ethereum revealed no long-term relationship, and causality analysis also failed to detect any directional causality between price and energy consumption. This indicates that with Ethereum's transition to a Proof-of-Stake mechanism, energy consumption has become independent of price movements. The findings reveal that the effects of cryptocurrency markets on the energy economy vary according to technology-specific structural characteristics.

Keywords

- Bitcoin
- Ethereum
- energy consumption
- cryptocurrency markets
- digital mining
- energy economy

JEL codes: C32, C58, Q55.

Article received 22 November 2025, accepted 15 January 2026.

Suggested citation: Sezal, L. (2026). From digital mining to market prices: An empirical analysis of the relationship between energy consumption and price dynamics of Bitcoin and Ether. *Economics and Business Review*, *12*(1), 159-182. https://doi.org/10.18559/ebr.2026.1.2793

[1] Vocational School of Social Sciences, Kahramanmaraş Sütçü Imam University, Kahramanmaraş, Türkiye, leventsezal@ksu.edu.tr, https://orcid.org/0000-0002-8873-7335.

Introduction

Cryptocurrencies, particularly since the emergence of Bitcoin (BTC) in 2009, have become a central topic in financial markets, the technology sector, and public policy debates. Although the distributed ledger technologies provided by cryptocurrencies are attractive due to their potential to reduce intermediary costs and enable decentralised payments, their energy and environmental costs have become a focus for both academic circles and policymakers in recent years. Due to the computationally intensive nature of Proof-of-Work (PoW) mining mechanisms, numerous studies have estimated the energy consumption of the Bitcoin network and the associated carbon emissions to be high. For example, De Vries (2018) demonstrated that the Bitcoin network has a significant electricity demand, and subsequent detailed calculations revealed that Bitcoin mining's annual electricity consumption could run to tens of thousands of GWh (De Vries, 2018; Stoll et al., 2019). Methodological studies on this subject emphasise that energy consumption estimates can vary significantly depending on the calculation method used and therefore highlight the need for both upper and lower band estimates and transparent data sources (CCAF, 2025; Gallersdörfer et al., 2020).

The price behaviour of cryptocurrency markets has also been studied extensively in academic literature. Early studies showed that cryptocurrencies are highly volatile assets exhibiting bubble tendencies and possessing price dynamics distinct from traditional assets (Cheah & Fry, 2015). Subsequent empirical analyses point to both short-term speculative movements and long-term structural changes in cryptocurrency prices; furthermore, it has been reported that volatility dependencies between cryptocurrency markets are increasing (Bouri et al., 2020; Koutmos, 2018).

The direct empirical link between energy consumption and crypto prices is more limited in the literature, but some studies have found co-movement and causality between the hashrate (the network's processing power) and prices at the time-frequency level. For example, studies examining the time-frequency relationships between hashrate, energy inputs, and prices have shown that energy costs and mining profitability can have both short- and long-term effects on prices (Das & Dutta, 2020; Fantazzini & Kolodin, 2020; Rehman & Kang, 2021). Although these findings are largely consistent, comprehensive comparisons of energy consumption indices with direct cryptocurrency price time series—particularly those using cointegration and causality frameworks—remain relatively scarce in the literature (Das & Dutta, 2020; Qin, 2023; Rehman & Kang, 2021).

Therefore, comprehensively testing the relationship between energy consumption indicators and the price dynamics of major crypto assets such as Bitcoin (BTC) and Ether (ETH) offers unique contributions to both financial economics and environmental policy discussions. Importantly, Ether refers to the crypto asset, while Ethereum denotes the underlying blockchain technology (EU Blockchain Observatory and Forum, 2023). Ether is included not to replicate Bitcoin's energy–price dynamics, but to serve as a benchmark case illustrating how transitioning the Ethereum network from Proof-of-Work to Proof-of-Stake (PoS) can fundamentally alter—and potentially eliminate—the link between market prices and energy consumption.

This study contributes to the growing literature on cryptocurrency energy economics in several important ways. Firstly, unlike most existing studies that focus primarily on Bitcoin, this paper provides a comparative analysis of Bitcoin and Ether within a unified econometric framework, allowing for a direct assessment of how different consensus mechanisms shape the price-energy relationship. Secondly, by jointly applying ARDL cointegration and Toda–Yamamoto causality tests, the study disentangles long-run equilibrium relationships from short-run causal dynamics. Thirdly, the findings offer post-Merge empirical evidence showing that Ethereum's transition to Proof-of-Stake has fundamentally altered the interaction between energy consumption and market prices. These results extend the literature by demonstrating that the energy–price nexus in cryptocurrencies is technology-specific rather than universal**,** with important implications for sustainability-oriented blockchain design.

Accordingly, this study empirically examines the long- and short-term relationships between cryptocurrency prices and energy consumption within a formal hypothesis-testing framework. In pursuit of this objective, the paper first defines the data set and variables, then applies unit root and cointegration tests (ADF, PP, and ARDL-bounds) as well as Granger (1969) Toda–Yamamoto causality analyses to identify potential long-term and directional linkages. The remainder of the paper is organised as follows: Section 1 reviews the relevant literature, Section 2 outlines the data and methodology, Section 3 presents the empirical results, and the final Section discusses the findings and draws conclusions, including the implications for policy and future research.

1. Literature review

Two main lines of research on cryptocurrencies stand out in the literature: (1) the energy consumption and environmental impacts of mining; (2) cryptocurrency price dynamics, volatility, and market linkages. This strand is discussed here in the context of crypto-energy interactions, as price dynamics and volatility directly affect mining incentives and energy demand.

In his study, de Vries (2018) focused on the energy consumption of Bitcoin mining and quantitatively assessed the contribution of the Proof-of-Work mechanism to the growing demand for electricity. He presented the concerns frequently voiced in public regarding the Bitcoin network's energy consumption with concrete data and computational approaches. He particularly emphasised the energy intensity of mining activities and its increase over time. This study is one of the first significant contributions that regularly brings the energy dimension of crypto mining into the academic research agenda.

Gallersdörfer et al. (2020) took a broader perspective, examining the energy consumption not only of Bitcoin but also of various other cryptocurrencies. This study demonstrated that other crypto assets besides Bitcoin are also significant in terms of energy consumption and that similar energy problems exist for many Proof-of-Work-based tokens. It also highlighted the limitations of the estimation methods used.

Kumar and Anandarao (2019) presented evidence on the time-frequency level regarding volatility spreads in cryptocurrency markets by combining GARCH and wavelet analyses. Their findings, particularly that volatility linkages change across different time scales (short, medium, long term), demonstrated the usefulness of wave-based methods in cryptocurrency research. On the other hand, Bouri et al. (2020), in their study on the volatility structure of cryptocurrency assets, separated the transient and persistent components of volatility and showed how the relationships between volatility components change during periods of market stress. Such studies are important for understanding the sources and persistence of volatility in cryptocurrency markets. In another study, Fantazzini and Kolodin (2020) focused on the relationship between Bitcoin's hashrate and its price, suggesting that changes in mining power could be linked to long-term price dynamics. Such analyses demonstrate that mining activities can indirectly influence market mechanisms.

Rehman and Kang (2021), meanwhile, identified time-frequency co-movement between Bitcoin's hashrate and energy commodity markets. This suggests that price changes in energy markets may be indirectly related to mining costs and, consequently, mining activities. The study is valuable in that it shows that energy-crypto interactions are influenced not only by local but also by global energy market conditions. In another study, Das and Dutta (2020) analysed the relationship between mining revenues and Bitcoin energy consumption, showing that energy consumption increases as mining profitability rises during periods of high prices. This finding directly supports the price-energy relationship, strengthening the hypothesis that price increases energy demand through mining incentives.

Marchewka-Bartkowiak and Wiśniewski (2022) examined energy-focused tokens as financial investment instruments and highlighted the intersection between energy markets and digital assets. This study emphasises that crypto assets should be examined not only for their financial returns but also for their energy market and sustainability dimensions, aligning with this study that addresses the energy consumption and price dynamics of Bitcoin and Ether together. In another study, Barbu et al. (2022) examined the safe-haven function of cryptocurrencies during the COVID-19 period using a threshold regression approach and found that cryptocurrencies exhibited different behavioural patterns depending on market conditions. These findings support the idea that crypto assets may be sensitive to regimes in terms of both financial stability and energy consumption, reinforcing the importance of the structures examined in our study using ARDL and causality analysis.

In Qin et al.'s (2023) study, the relationship between Bitcoin's energy consumption and carbon emissions in the US context was empirically examined. The study found that Bitcoin mining affected US energy consumption and carbon emissions during certain periods based on time-series analyses; however, the magnitude and direction of this impact were shown to vary depending on electricity generation sources, regional energy market conditions, regulatory interventions, and seasonal fluctuations in mining activity. This study contributes to the empirical literature by highlighting that the environmental effects of crypto mining are not uniform over time and depend on country-specific energy structures, thereby underlining the importance of contextual factors in price-energy interaction studies.

The study by Sapra et al. (2024) addressed the non-linear relationships and causality patterns between Bitcoin energy consumption, price, and the Crypto Volatility Index. Using time-frequency and causality methods, Sapra identified the asymmetric effects of price fluctuations on energy consumption, pointing to the necessity of using non-linear models in price-energy research.

Adewuyi et al. (2024) questioned whether Bitcoin's electricity consumption and carbon footprint time series conform to the random walk hypothesis. The research showed that energy consumption/CO_2 series exhibited surprising movements in some periods and that simple random walk assumptions are not always valid. In another study by Wang et al. (2024), dynamic volatility contagion between multiple cryptocurrencies and energy markets was examined. This study revealed that crypto-energy interactions may be related not only to direct energy consumption data but also to energy market prices and volatility indicators. The findings demonstrate the potential for macro-level linkages between crypto assets and energy markets and the need for flexible policies.

The study by Bilirer and Zeren (2024) is noteworthy as an empirical study centred on Turkey. In this study, Fourier Granger causality and Fourier-ADL cointegration tests were applied using weekly price, energy consumption, and CO_2 emission series for Bitcoin and Ethereum. The results indicated reciprocal effects between price and energy/CO_2 and highlighted how Fourier-based methods offer approaches beyond the median.

Although the existing literature on cryptocurrencies has expanded rapidly, it has largely evolved along two parallel strands. The first strand focuses on energy consumption and environmental impacts, examining electricity usage, carbon emissions, and sustainability concerns associated primarily with Proof-of-Work-based systems (e.g., De Vries, 2018; Krause & Tolaymat, 2018; Stoll et al., 2019). The second strand concentrates on price dynamics and market behaviour, analysing returns, volatility, and speculative characteristics of cryptocurrency prices (e.g., Cheah & Fry, 2015; Corbet et al., 2018; Koutmos, 2018). Despite the rich insights provided by both strands, studies that directly integrate energy consumption indicators with cryptocurrency price series using cointegration and causality frameworks remain relatively scarce (Das & Dutta, 2020; Qin et al., 2023; Rehman & Kang, 2021). By jointly examining energy consumption indices and price dynamics for Bitcoin and Ether within a unified empirical framework, the present study bridges this gap and contributes to a more integrated understanding of the economic and environmental dimensions of cryptocurrency markets.

Both the energy consumption and price dynamics of cryptocurrencies have attracted considerable attention in the literature; energy-focused studies quantify the environmental costs of mining, while financial studies provide in-depth empirical evidence on price volatility, volatility contagion, and market linkages. However, at the intersection of these two lines of research, comprehensive studies that systematically address the long-term cointegration and directional causality relationships between energy consumption indices and cryptocurrency prices are limited. The existing literature has mostly focused on the hashrate and price relationship or crypto-crypto volatility spillovers, and the assessment of the relationships between energy consumption indices and prices from a cointegration and causality perspective has rarely been conducted. This study contributes to the literature by testing long-term equilibrium relationships, short-term adjustment mechanisms, and directional causality between energy consumption indices and cryptocurrency prices, focusing specifically on Bitcoin (BTC) and the crypto asset Ether (ETH).

2. Dataset and variables

This study uses daily time series data covering the period from 20 May 2017 to 23 October 2025 to examine the long-term relationships and causality between energy consumption and price dynamics in the Bitcoin (BTC) and Ether (ETH) markets. The variables of the study are cryptocurrency prices and energy consumption indicators. Price data for Bitcoin and Ether are obtained from CoinMetrics, while energy consumption indicators are sourced from the Cambridge Bitcoin Electricity Consumption Index (CBECI) provided by the Cambridge Centre for Alternative Finance.

As cryptocurrency markets operate 24/7, the term "closing price" can be ambiguous. We therefore define the daily price as the data provider's daily close (end-of-day) quotation recorded at 00:00 UTC for each calendar day, following the standard convention used by CoinMetrics. Accordingly, we use the term "daily close (end-of-day) price" to reflect the continuous trading structure of cryptocurrencies. The series have been transformed into natural logarithms, as is common practice in the financial literature. This approach, which aims to measure the impact of changes in energy consumption on price formation, has also been adopted in previous studies (Krause & Tolaymat, 2018; Sapra et al., 2024).

The key variables of the study are energy consumption indicators for Bitcoin and Ethereum. Bitcoin energy consumption data was obtained from the Cambridge Bitcoin Electricity Consumption Index (CBECI) platform, recognised as the most comprehensive data source on a global scale; Ethereum energy consumption data was obtained from the Digiconomist database, covering the Ethereum Proof-of-Work (PoW) period.

Energy consumption data shows estimated electricity consumption in terawatt-hours (TWh) or kilowatt-hours (kWh) per day. It is known that energy consumption indices are not officially reported and are largely modelled and calculated based on variables such as mining hashrate, hardware energy efficiency, and geographical distribution (CCAF, 2023). Therefore, although energy consumption series are based on estimates, the CBECI and Digiconomist platforms are widely referenced by academic circles (Rehman & Kang, 2021; Sapra et al., 2024). The variables used in the study are as shown in Table 1.

Table 1. Variable definitions

Variables	Definition	Period	Frequency	Source
BTC_P	Bitcoin daily close (end--of-day) price (USD)	2017–2025	daily	CoinMetrics
ETH_P	Ether daily close (end--of-day) price (USD)	2017–2025	daily	CoinMetrics
BTC_E	Bitcoin energy consumption index (TWh)	2017–2025	daily	CBECI
ETH_E	Ethereum energy consumption index (TWh)	2017–2025	daily	CBECI

Source: own work.

The "Estimated TWh per Year" series reported under the Cambridge Bitcoin Electricity Consumption Index (CBECI) was used as the energy consumption index. This series is widely used in the literature as the most probable energy consumption estimate, taking into account hardware efficiency and hashrate distribution (Sapra et al., 2024).

3. Method

This study examines the long-term relationship and causality between energy consumption and price dynamics in the Bitcoin and Ethereum markets. To this end, the following tests were applied: (1) Augmented Dickey–Fuller (ADF) and Phillips–Perron (PP) unit root tests to determine the stationarity properties of the series; (2) the Autoregressive Distributed Lag (ARDL) cointegration test to investigate the long-term relationship between the series; (3) the Toda–Yamamoto (TY) Granger causality test to determine the direction of causality. The methodological framework is parallel to the econometric approaches frequently used in the literature for energy consumption-oriented analyses of crypto assets (Bilirer & Zeren, 2024; Mensi et al., 2022; Sapra et al., 2024).

3.1. Unit root tests

To avoid spurious regression in time series analyses, testing the stationarity of the series is a prerequisite (Enders, 2015). In this context, the stationarity levels of all series were first examined using the Augmented Dickey–Fuller (ADF) and Phillips–Perron (PP) tests. While the ADF test aims to eliminate autocorrelation by using additional lags in the model (Dickey & Fuller, 1981), the PP test has a more flexible structure against heteroscedasticity and serial correlation (Phillips & Perron, 1988). Applying both tests ensures the robustness of the stationarity results.

The stationarity of the series has been tested at the level and in the first difference. Determining the degree of stationarity is important for establishing the applicability of the ARDL model and for determining the necessary degree of integration for the Toda–Yamamoto test.

3.2. ARDL cointegration test

The Autoregressive Distributed Lag (ARDL) bounds test approach developed by Pesaran et al. (2001) was applied to examine the existence of a long-term relationship between the series. The ARDL model was preferred because it can work even when variables have different integration orders (I(0) or I(1)), it can provide reliable results in small samples, and it offers the possibility of estimating both short- and long-term coefficients simultaneously (Narayan, 2005).

In the ARDL bounds test approach, the appropriate lag lengths were first determined based on information criteria, and then the presence of cointegration was tested using the F-statistic. If the calculated F-statistic exceeds the upper critical values provided by Pesaran et al. (2001), it is concluded that there is a long-term relationship between the variables.

After estimating the long-term coefficients, short-term dynamics were analysed using an error correction model (ECM). The negative and statistically significant ECM parameter was interpreted as an indicator of the speed of return to long-term equilibrium (Banerjee et al., 1996). The ARDL Bound Test equation involving two variables to be performed in order to reveal the cointegration relationship is as follows:

$$\Delta Y_t = \beta_0 + \sum_{i=1}^{m} \beta_{1i} Y_{t-i} + \sum_{i=1}^{m} \beta_{2i} X_{t-i} + \beta_{3i} Y_{t-1} + \beta_4 X_{t-1} + \varepsilon_t \tag{1}$$

where: ΔY_t represents the dependent variable, X_t represents the independent variable, ε_t represents the error term, and m represents the optimal lag length, which is the value where the information criteria are smallest. The hypotheses regarding the existence of cointegration in the ARDL bounds test model are as follows:

H0: $\beta_3 = \beta_4 = 0$ (There is no cointegration.)
H1: $\exists \beta_i < 0,\ i = 3{,}4$ (There is cointegration.)

"In the ARDL bounds test approach, after revealing the cointegration relationship for the variables, the long-term relationship coefficients of the variables are examined. Furthermore, the existence of short-term deviations from the long-term relationship can also be examined using an error correction model. Although the primary focus of the study is on long-term relationships, the error correction model is employed to capture short-run dynamics and the speed of adjustment toward long-run equilibrium. The equation for the long-term relationship is as follows:"

$$Y_t = \beta_0 + \sum_{i=1}^{m} \beta_{1i} Y_{t-i} + \sum_{i=0}^{n} \beta_{2i} X_{t-i} + \varepsilon_t \tag{2}$$

"In the equation, Y_t is the dependent variable, X_t is the independent variable, β_0 is the constant term, ε_t is the error term, is the error term, and m and n represent the optimal lag length."

3.3. Toda–Yamamoto causality test

The augmented Granger causality approach proposed by Toda and Yamamoto (1995) was used to determine the direction of causality between energy consumption and cryptocurrency prices. This test offers significant advantages over the traditional Granger causality test because it can be applied independently of the integration degrees of the variables and possible cointegration relationships (Dolado & Lütkepohl, 1996).

The study examined the bidirectional causal relationship between both the BTC price-BTC energy consumption and the ETH price-ETH energy consumption. The results obtained enable an assessment of whether energy consumption is a factor driving price formation in the market. The VAR ($m + d\,\max$) model estimated in the Toda–Yamamoto causality approach consists of (Toda & Yamamoto, 1995):

$$Y_t = \omega + \sum_{t-1}^{m} a_{1i} x_{t-i} + \sum_{i=1}^{m} \beta_{1i} Y_{t-i} + \sum_{j=m+1}^{d\max} \delta_{1i} X_{t-i} + \sum_{j=m+1}^{d\max} \theta_{1i} Y_{t-i} + \varepsilon_{1t} \tag{3}$$

$$X_t = \varphi + \sum_{i=1}^{m} a_{2i} X_{t-i} + \sum_{i=1}^{m} \beta_{2i} Y_{t-i} + \sum_{j=m+1}^{d\max} \delta_{2i} X_{t-i} + \sum_{j=m+1}^{d\max} \theta_{2i} Y_{t-i} + \varepsilon_{2t} \tag{4}$$

"The appropriate lag length (m) can be determined using information criteria, while the maximum integration order ($d\,\max$) can be determined using unit root tests. To determine the existence of a reciprocal causality relationship between the variables, the hypotheses $\mathrm{H}_0{:}\alpha_{1i} = 0$ and $\mathrm{H}_0{:}\alpha_{2i} = 0$ are tested using the adjusted Wald test statistic. If the calculated test statistic value is greater than the X^2 table value with k degrees of freedom, the above hypotheses are rejected (Toda & Yamamoto, 1995)."

3.4. Research hypotheses

This study aims to examine the relationships between price dynamics in the Bitcoin and Ether markets and the energy consumption indices of these networks from both long-term and short-term perspectives. While price–hashrate and price–volatility relationships have been extensively studied in the literature, only a limited number of studies have systematically examined cointegration and directional causality between energy consumption indices and cryptocurrency prices (Das & Dutta, 2020; Qin et al., 2023; Rehman & Kang, 2021). By extending this line of research, the present study provides a comparative analysis of Bitcoin and Ether within a unified econometric framework, thereby highlighting how consensus mechanisms shape the price-energy relationship. Four main hypotheses are presented below.

H1: There is a long-term cointegration relationship between the Bitcoin price and the Bitcoin energy consumption index.

Theoretically, persistent price increases may lead to increased mining profitability, thereby expanding mining capacity and, consequently, energy consumption. Conversely, persistent or long-lasting changes in energy costs may affect mining margins and thus long-term pricing expectations. Therefore, a long-term relationship (cointegration) may emerge between the BTC price and the BTC energy index. Empirically, this hypothesis is tested using the ARDL bounds test.

H2: There is a long-term cointegration relationship between Ether's price during its PoW period and the Ethereum energy consumption index; however, the transition to PoS (Merge) weakens or eliminates this relationship.

Under the PoW mechanism, sustained increases in the price of ETH are expected to affect energy consumption by increasing mining incentives; therefore, a long-term relationship similar to H1 is possible for the PoW period. However, due to the dramatic drop in energy consumption after 'The Merge' in September 2022, it is likely that the cointegration structure will change or disappear.

H3: Short-term shocks in cryptocurrency prices (especially positive price shocks) increase mining profitability, leading to short-term increases in energy consumption.

As a market mechanism, sudden increases in prices can raise mining revenues and, consequently, the intensity of mining activities; this results in an increase in energy demand in the short term. This directional effect will be tested using the Toda–Yamamoto method; finding positive causality provides empirical evidence that price fluctuations create physical effects through the

Table 2. Research hypotheses summary

Hypothesis Code	Hypothesis	Test method	Expected relationship / direction
H1	There is a long-term cointegration relationship between the Bitcoin price and the Bitcoin energy consumption index	ARDL Cointegration	Long-term positive relationship
H2	There is a long-term cointegration relationship between the price of Ethereum during its PoW period and the Ethereum energy consumption index; the transition to PoS weakens or eliminates this relationship	ARDL Cointegration + Sub-Period Analysis	PoW: positive long-term relationship; PoS: weak/no relationship
H3	Short-term changes in cryptocurrency prices Granger-cause energy consumption (Price → Energy)	Toda–Yamamoto Causality	Positive short-term effect
H4	Changes in energy consumption Granger-causally affect cryptocurrency prices (Energy → Price)	Toda–Yamamoto Causality	Weak/uncertain short-term effect

Source: own work.

energy demand channel and points to important policy implications for energy supply-demand management.

H4: Changes in energy consumption (e.g., rapid increases or cost shocks) may create short-term effects on cryptocurrency prices.

Sudden increases in energy consumption or rises in energy costs can affect mining profitability and perceptions of network security, thereby altering investor expectations and consequently prices. Identifying energy → price Granger causality will provide empirical evidence regarding the channel through which shocks in energy markets are transmitted to crypto asset prices.

According to Table 2, which summarises the research hypotheses, hypotheses H1 and H2 aim to test whether there is a long-term structural equilibrium relationship between market prices and energy consumption. Hypotheses H3 and H4, on the other hand, focus on the direction of the relationship and evaluate short-term dynamics.

4. Findings

First, the stationarity properties of the Bitcoin and Ether price series and the energy consumption indices related to these assets were examined using the Augmented Dickey–Fuller (ADF) and Phillips–Perron (PP) unit root tests. After determining the stationarity structure, the ARDL bounds test approach developed by Pesaran et al. (2001) was applied to test whether there was a long-term relationship between the variables. At this stage, optimal lag lengths were determined by considering information criteria, and appropriate ARDL models were constructed to estimate the long-term coefficients. In the subsequent stage, an extended VAR model based on the Toda–Yamamoto (1995) approach was used to reveal the short-term causality relationships between the variables.

4.1. Unit root test results

This subsection examines the stationarity properties of the Bitcoin and Ether price series used in the study and the energy consumption indices related to these assets. In this context, Augmented Dickey–Fuller (ADF) and Phillips–Perron (PP) unit root tests were applied to assess whether the series are stationary at their levels or in their differences. While the ADF test addresses the autocorrelation issue with lagged difference terms, the PP test offers a more flexible structure by considering possible serial correlations and heteroskedasticity in the error terms using non-parametric methods. The combined use of both tests allows for a more robust determination of the degree of integration of the series. Table 3 presents the ADF and PP test results and provides a detailed interpretation of the findings regarding the degree of integration of the series.

Table 3. ADF and PP unit root test results

Variables	Test	Level *t*-statistic	Level *p*	1st difference *t*-statistic	1st difference *p*	Critical value (1%)	Result (I(d))
lnBTC_E	ADF	–4.0227	0.0013	–25.9362	0.0000	–3.435	I(0)
	PP	–4.1123	0.0009	–41.1083	0.0000	–3.435	I(0)
lnBTC_P	ADF	–1.6374	0.4632	–58.3172	0.0000	–3.435	I(1)
	PP	–1.6316	0.4662	–58.2538	0.0001	–3.435	I(1)
lnETH_E	ADF	–0.9510	0.7722	–18.8100	0.0000	–3.435	I(1)
	PP	–1.8902	0.3372	–107.9101	0.0001	–3.435	I(1)
lnETH_P	ADF	–2.1215	0.2363	–38.1731	0.0000	–3.435	I(1)
	PP	–2.1096	0.2410	–58.3352	0.0001	–3.435	I(1)

Note: The critical value threshold is 1%, in line with MacKinnon (1996).

Source: own work.

The ADF and PP unit root tests were applied to assess whether the series were stationary at their levels. The findings indicate that the Bitcoin energy consumption series (LNBTC_E) is stationary at its level, both according to the ADF and PP tests. In contrast, the Bitcoin price series (LNBTC_P), Ethereum energy consumption (LNETH_E), and Ether price series (LNETH_P) show the presence of a unit root at the level in both the ADF and PP tests.

ADF and PP unit root tests were applied to assess whether the series were stationary at their levels. The findings indicate that the Bitcoin energy consumption series (LNBTC_E) is stationary at its level, in both ADF and PP tests. In contrast, the Bitcoin price series (LNBTC_P), Ethereum energy consumption (LNETH_E) and Ether price series (LNETH_P) indicate the presence of a unit root at the level in both the ADF and PP tests, revealing that they are not stationary. These results validate the applicability of ARDL cointegration analysis and the methodological justification of setting the maximum degree of integration (d max) to 1 for the VAR model in the Toda–Yamamoto causality test.

4.2. ARDL bound test results

The unit root test results showed that the variables were integrated at the I(0) and I(1) levels, making the ARDL bounds test approach developed by Pesaran et al. (2001) applicable. The ARDL method allows series with different integration orders to be analysed in the same model and is widely preferred in the time-series literature due to its high estimation power in small samples. Descriptive statistics are evaluated in Table 4.

Table 4. Descriptive statistics

	LNBTC_E	LNBTC_P	LNETH_E	LNETH_P
Mean	1.740297	4.316554	–0.781668	2.931164
Median	1.731.895	4.363605	0.453821	3.130999
Maximum	2.195326	5.095823	1.287155	3.681995
Minimum	0.619043	3.281965	–5.723042	1.617629
Standard deviation	0.307732	0.454565	2.100261	0.522715
Skewness	–1.346475	–0.123947	–0.545067	–0.415975
Kurtosis	5.340976	1.915679	1.419403	1.798104
Jarque-Bera	1.632898	1.586713	4.726632	2.739421
Probability	0.000000	0.000000	0.000000	0.000000
Sum	5.356633	13286.35	–2.405191	9.019193
Sum of squared deviations	2.913880	6.357986	13568.53	8.404595
Observations	3078	3078	3078	3078

Source: own work.

The basic characteristics of the series differ in various aspects. The standard deviation values show that the volatility in Ethereum's energy consumption is relatively high compared to other series, indicating that seasonal changes in energy usage are more pronounced. The skewness and kurtosis values deviating significantly from normal distribution indicate that the series are not symmetrical and that the tail behaviour is heavier than in normal distribution.

Table 5. ARDL bound test for Bitcoin energy consumption

“Model	K	M	F-Statistic	Significance level	Lower bound	Upper bound
				1%	4.94	5.58
ARDL (10.9)	10	12	7.5573531	5%	3.62	4.16
				10%	3.02	3.51

Source: own work.

According to Table 5, the F-statistic value obtained from the ARDL(10,9) model, which examines the relationship between Bitcoin energy consumption and Bitcoin price, exceeds the upper limits of the critical values proposed by Pesaran et al. (2001) at 1% significance levels. Therefore, the null hypothesis of no cointegration is strongly rejected, and it is concluded that there is a long-term relationship between Bitcoin prices and energy consumption. This finding is consistent with the literature suggesting that Bitcoin mining is directly linked to economic incentives; as prices rise, mining becomes more profitable, leading to more miners joining the network and an increase in total energy consumption.

Table 6. ARDL Bitcoin energy consumption error correction model

Variable	Coefficient	Standard error	t-statistic	Probability
D(LNBTC_E(-1))	0.228412	0.018000	12.68948	0.0000
D(LNBTC_E(-2))	0.084282	0.018463	4.564906	0.0000
D(LNBTC_E(-3))	–0.038588	0.018327	–2.105546	0.0353
D(LNBTC_E(-4))	0.056194	0.018295	3.071488	0.0021
D(LNBTC_E(-5))	–0.046578	0.018253	–2.551816	0.0108
D(LNBTC_E(-6))	–0.068430	0.018230	–3.753721	0.0002
D(LNBTC_E(-7))	–0.149019	0.018275	–8.154469	0.0000
D(LNBTC_E(-8))	–0.014774	0.018422	–0.801944	0.4226
D(LNBTC_E(-9))	0.053736	0.017959	2.992183	0.0028
D(LNBTC_P)	–0.009857	0.010433	–0.944797	0.3448
D(LNBTC_P(-1))	–0.004102	0.010452	–0.392497	0.6947
D(LNBTC_P(-2))	–0.009433	0.010444	–0.903230	0.3665
D(LNBTC_P(-3))	–0.006628	0.010434	–0.635286	0.5253
D(LNBTC_P(-4))	0.029702	0.010426	2.848865	0.0044
D(LNBTC_P(-5))	–0.001947	0.010434	–0.186585	0.8520
D(LNBTC_P(-6))	0.000511	0.010427	0.048991	0.9609
D(LNBTC_P(-7))	–0.001867	0.010420	–0.179209	0.8578
D(LNBTC_P(-8))	0.034981	0.010407	3.361401	0.0008
CointEq(-1)*	–0.003092	0.000649	–4.763081	0.0000
R-squared	0.128387	Mean dependent variable		0.000465
Adjusted R-squared	0.123241	S.D. dependent variable		0.009810
S.E. of regression	0.009186	Akaike info criterion		–6.536199
Sum squared residuals	0.257257	Schwarz criterion		–6.498863
Log likelihood	10045.53	Hannan–Quinn criterion		–6.522785
Durbin–Watson statistic	1.999098			

Source: own work.

According to Table 6, the error correction term coefficient (–0.003092) in the model is statistically significant ($t = –4.763$; $p = 0.0000$). The negative and significant nature of the error correction term coefficient indicates that there is a long-term equilibrium relationship between the series and that short-term shocks gradually disappear in the long term. However, the coefficient's very small value indicates that the system's return to equilibrium is slow. The use of daily data, the lengthy adaptation process of mining equipment, and the difficulty adjustment occurring through block times are among the natural causes of these slow effects.

Looking at the short-term coefficients, it can be seen that a large proportion of the lags in the Bitcoin energy consumption series are significant. This indicates that energy consumption has a strong autoregressive structure and that short-term changes are largely influenced by its own internal dynamics. In contrast, the vast majority of Bitcoin price short-term coefficients are not significant; this indicates that short-term movements in the Bitcoin price do not have a rapid effect on energy consumption, and that energy consumption is shaped more by long-term price incentives.

Table 7. ARDL boundary test results for the Ethereum energy consumption model

“Model	K	M	F-Statistic	Significance level	Lower bound	Upper bound
				1%	6.84	7.84
ETH-ARDL (5.0)	10	12	1.2182841	5%	4.94	5.73
				10%	4.04	4.78

Source: own work.

According to Table 7, the F-statistic value in the ARDL (5.0) model established for Ethereum is well below even the lowest threshold of Pesaran critical values. Therefore, the null hypothesis, according to which there is no long-term relationship, cannot be rejected. In other words, there is no long-term cointegration relationship between the Ether price and Ethereum energy consumption. This finding is consistent with Ethereum's transition from the Proof-of-Work (PoW) mechanism to the Proof-of-Stake (PoS) mechanism in 2022. The PoS mechanism has dramatically reduced energy consumption and eliminated the structural relationship between price and energy consumption. The results support the notion that the energy consumption of the post-PoS Ethereum network has become independent of price changes.

According to Table 8, the long-term error correction term LNETH_E(-1) in the ARDL error correction model was not statistically significant, despite carrying the expected negative sign. This result indicates that there is no mechanism representing a return to long-term equilibrium in Ethereum's energy

Table 8. ARDL Ethereum energy consumption error correction model

Variable	Coefficient	Standard error	*t*-statistic	Probability
C	0.019212	0.018463	1.040528	0.2982
D(LNETH_E(-1))	–0.398638	0.018122	–21.99742	0.0000
D(LNETH_E(-2))	–0.377021	0.019447	–19.38731	0.0000
D(LNETH_E(-3))	–0.244722	0.020547	–11.91047	0.0000
D(LNETH_E(-4))	–0.168908	0.020970	–8.054624	0.0000
D(LNETH_E(-5))	–0.058415	0.021160	–2.760615	0.0058
D(LNETH_E(-6))	–0.082163	0.020954	–3.921122	0.0001
D(LNETH_E(-7))	–0.085698	0.020531	–4.174098	0.0000
D(LNETH_E(-8))	–0.079831	0.019380	–4.119135	0.0000
D(LNETH_E(-9))	–0.091660	0.018005	–5.090897	0.0000
CointEq(-1)*	–0.002530	0.002043	–1.238437	0.2156
R-squared	0.183425	Mean dependent variable		–0.001355
Adjusted R-squared	0.180744	S.D. dependent variable		0.233805
S.E. of regression	0.211623	Akaike info criterion		–0.264428
Sum squared residuals	136.4130	Schwarz criterion		–0.242748
Log likelihood	415.1787	Hannan–Quinn criterion		–0.256637
F-statistic	68.42150	Durbin–Watson statistic		1.997831
Prob(F-statistic)	0.000000			–6.522785

Source: own work.

consumption. In other words, it was observed that short-term shocks occurring in the series did not correct themselves based on a long-term relationship, and the system did not exhibit a tendency to return to equilibrium. This is consistent with the absence of cointegration and indicates that long-term dynamics are not affected by price changes. The fact that Ethereum's energy consumption began to follow a stable process after the transition to PoS provides an economically reasonable framework for the error correction coefficient being insignificant.

The fact that all lagged values of D(LNETH_E) in the model are highly significant indicates that Ethereum's energy consumption has a strong autoregressive structure in the short term. The sensitivity of energy consumption to past values reveals that the system is shaped by its own internal dynamics, with sudden changes being determined not by price but by previous levels of energy usage.

In contrast, the Ether price variable (LNETH_P) was found to be statistically insignificant in both level and difference equations. This indicates that Ether price fluctuations do not have any effect on energy consumption in the short term. Given the low and stable energy usage in the PoS system, it is a finding consistent with theory that price changes have no effect on energy consumption, even in the short term. The absence of a price-energy relationship for Ether should therefore be interpreted as evidence of a structural break induced by the Proof-of-Stake transition, rather than as a lack of economic relevance.

4.3. Toda–Yamamoto causality analysis results

The Toda–Yamamoto causality analysis was applied to reveal the directional relationship between energy consumption and price dynamics for Bitcoin and Ether. Previously conducted unit root and cointegration tests showed that there was no long-term relationship, particularly for Ethereum, while Bitcoin exhibited long-term dependence. However, the direction of these relationships had not yet been determined. Therefore, using the Toda–Yamamoto approach, which is unaffected by differences in the degrees of integration of the series, the analysis tested whether there was a causality flow from Bitcoin prices to energy consumption or from energy consumption to prices, and similarly, whether there were possible directional interactions between energy consumption and the price series for Ether. Below, the Wald test results obtained through the extended VAR model are presented, and the causality relationships for both cryptocurrencies are evaluated separately.

Table 9. Toda–Yamamoto causality test results-1

"Dependent variable	Independent variable	d max	k	Chi-Square test statistic	Chi-Square p-value	Relationship
LNBTC_E	LNBTC_P	8	8	2.303595	0.0033	there is a relationship
LNBTC_P	LNBTC_E	8	8	1.333014	0.1010	there is no relationship

Source: own work.

According to Table 9, the results obtained for the dependent variable LNBTC_E indicate that removing the LNBTC_P variable from the model leads to a statistically significant loss of information. This finding indicates a significant causal relationship from Bitcoin prices to energy consumption. Therefore, price changes are an important factor explaining the Bitcoin network's energy consumption. This result is fully consistent with the theoretical expectation that mining activities are sensitive to prices and that price rises increase the number of miners and energy consumption.

In contrast, no causality from energy consumption to prices was detected for the dependent variable LNBTC_P. A p-value above 10% indicates that energy consumption does not play a statistically significant role in explaining Bitcoin prices. This finding supports the view that price formation is largely determined by factors such as market expectations, liquidity, macroeconomic factors, and investor behaviour, and that energy consumption is not a variable that drives the price.

The Toda–Yamamoto causality test has revealed a unidirectional causal relationship for Bitcoin. According to the analysis results, significant causality was found from Bitcoin prices to Bitcoin energy consumption, while no causality was detected from energy consumption to prices. Due to Bitcoin's mining--based structure, price movements directly affect the profitability of mining activities, which in turn has a decisive impact on total energy consumption. The literature frequently emphasises that increases in Bitcoin prices raise energy demand by boosting investment in mining hardware and hash power (Hayes, 2016; Krause & Tolaymat, 2018). This theoretical expectation is consistent with the empirical findings of this study. The Toda–Yamamoto test identified a significant causal relationship from Bitcoin prices to energy consumption. This result indicates that price increases make mining more attractive, thus demonstrating that energy consumption is sensitive to price dynamics. Thus, hypothesis 3 has been empirically confirmed, revealing that Bitcoin prices are the fundamental factor driving the network's energy consumption.

Table 10. Toda–Yamamoto causality test results-2

"Dependent variable	Independent variable	d max	k	Chi-Square test statistic	Chi-Square p-value	Relationship
LNETH_E	LNETH_P	5	5	0.847907	0.9739	there is a relationship
LNETH_P	LNETH_E	5	5	2.387410	0.7933	there is no relationship

Source: own work.

According to Table 10, when LNETH_E (Ethereum energy consumption) is the dependent variable, the Wald test value measuring the contribution of the LNETH_P variable to the model indicates that Ether prices have no causal effect on energy consumption. This result is particularly important because Ethereum's transition from Proof-of-Work (PoW) to Proof-of-Stake (PoS) in 2022 ('The Merge') reduced energy usage by approximately 99.9% and eliminated the economically motivated relationship between price and energy consumption. Therefore, price increases in the post-PoS period do not increase mining activity or energy demand; validators' energy requirements are largely independent of price dynamics.

When LNETH_P (Ether price) is the dependent variable, the Wald statistic testing the contribution of the LNETH_E (Ethereum energy consumption) indicates that energy consumption does not have a statistically significant effect on explaining the Ether price. This result suggests that short-term Ether price movements are not directly linked to energy consumption associated with mining and validation processes, and are likely driven by factors other than energy-related dynamics. The fact that both test results have relatively high p-values clearly demonstrates that there is no bidirectional causality relationship in Ethereum. This result is consistent with Ethereum's current consensus mechanism, indicating that energy consumption is not driven by price and that changes in energy usage are not decisive for price. According to the Toda–Yamamoto causality test results, since no directional causality was detected between Ether price and energy consumption, hypothesis 4, according to which there is a causal relationship between Ether prices and Ethereum energy consumption, was rejected. This finding confirms that with Ethereum's transition to the PoS mechanism, energy consumption is largely determined by technological protocols and is no longer influenced by price-based economic incentives.

Conclusions

This study comprehensively examined the relationships between Bitcoin and Ethereum's energy consumption and price dynamics using cointegration and causality analyses. The findings obtained provide significant contributions when evaluated comparatively with the theoretical and empirical discussions that have emerged in the cryptocurrency energy economics literature, particularly in recent years, due to increased academic interest.

The results show that there is a long-term cointegration relationship between energy consumption and prices for Bitcoin. According to the ARDL bounds test results, the Bitcoin model produced an F-statistic above the critical values, and the long-term relationship was statistically confirmed. The negative and significant error correction term indicates that the short-term deviations experienced in both cryptocurrencies have returned to equilibrium. This finding is consistent with the literature on how prices affect mining profitability in Bitcoin's mining-based Proof-of-Work structure and how this determines energy consumption (Hayes, 2016; Krause & Tolaymat, 2018). Empirical results reinforce the view that prices drive energy demand in the long term.

The results obtained for Ethereum indicate a structure that is completely different from Bitcoin. The boundary test results for the ARDL model show that there is no long-term cointegration relationship between Ethereum's energy consumption and prices. The insignificance of the error correction term and the fact that the F-statistic remains below all threshold values confirm that Ethereum's energy consumption has evolved into a structure independent of price dynamics with the transition from PoW to PoS. Indeed, the dramatic drop in Ethereum's energy usage (approximately 99.9%) after 'The Merge' supports findings in the literature that the price-energy relationship has structurally weakened (de Vries, 2018).

Causality results reveal a clear divergence between the two cryptocurrencies. The Toda–Yamamoto test for Bitcoin shows a unidirectional causality relationship from Bitcoin prices to energy consumption. This result aligns with the economic incentive model, where mining activities increase during periods of rising Bitcoin prices, thereby driving up energy demand. Conversely, no causality was found in the opposite direction; that is, no effect of changes in energy consumption on Bitcoin prices was detected. In the causality analysis conducted for Ethereum, no unidirectional causality relationship was found. The absence of any causality detected from prices to energy consumption or from energy consumption to prices is consistent with the PoS mechanism making energy usage independent of network security.

When compared to previous studies in the literature, these findings offer significant parallels and new contributions. For example, the results are consistent with studies showing that Bitcoin energy consumption is affected by price movements (Corbet et al., 2022; Hayes, 2016). Similarly, the findings of this study are in line with recent research indicating that Ethereum lost its energy consumption–price relationship with its transition to PoS (de Vries, 2018). However, this study contributes to the literature by presenting a comparative analysis of both Bitcoin and Ethereum using the same dataset, the same methodological framework, and the same time frame.

The findings of this study carry various implications for policymakers and regulators. Bitcoin's price-sensitive, energy-intensive mining structure fuels sustainability debates and supports the idea that it should be directed towards more carbon-efficient technologies. Encouraging the use of renewable energy sources in Bitcoin mining, implementing carbon tax-like regulations, or establishing energy efficiency criteria can be considered among the policy objectives. For Ethereum, the reduction of energy consumption to very low levels demonstrates that the PoS mechanism is more advantageous in terms of environmental sustainability and strengthens the debate on transitioning to PoS-like mechanisms for other blockchain projects.

However, the study has some limitations. Firstly, the fact that energy consumption data is based on estimates and that different sources use different algorithms may limit the absolute accuracy of the results. Furthermore, the study only considers Bitcoin and Ethereum, and the exclusion of other PoW-based cryptocurrencies from the analysis creates a comparative diversity constraint.

There are many areas of research for future studies. Comparing different PoW and PoS-based coins with a broader sample could explain the energy consumption–price relationship more systematically. To assess potential structural changes in the long-term relationship between energy consumption and Ether prices more comprehensively, sub-period analyses surrounding the transition from PoW to PoS can be examined. Furthermore, directly incorporating carbon emission data into the model could reveal the environmental impacts of energy consumption in a more specific manner.

References

Adewuyi, O. A., Wahab, B. A., Tiwari, A. K., & Do, H. X. (2024). Do bitcoin electricity consumption and carbon footprint follow a random walk? *Journal of Environmental Management*, *367*, 121958. https://doi.org/10.1016/j.jenvman.2024.121958

Banerjee, A., Dolado, J., & Mestre, R. (2001). Error-correction mechanism tests for cointegration in a single-equation framework. *Journal of Time Series Analysis*, *19*(3), 267–283. https://doi.org/10.1111/1467-9892.00091

Barbu, T. C., Boitan, I.A. & Cepoi, C. O. (2022). Are cryptocurrencies safe havens during the COVID-19 pandemic? A threshold regression perspective with pandemic-related benchmarks. *Economics and Business Review*, *8*(2), 29–49. https://doi.org/10.18559/ebr.2022.2.3

Bilirer, M., & Zeren, F. (2024). The impact of energy consumption in crypto assets on crypto asset prices and carbon emissions: Case of Bitcoin and Ethereum. *Studii Financiare*, *28*(1), 6–25.

Bouri, E., Jawad, S., Shahzad, H., Roubaud, D., Kristoufek, L., & Lucey, B. (2020). Bitcoin, gold, and commodities as safe havens for stocks: New insight through wavelet analysis. *The Quarterly Review of Economics and Finance*, *77*, 156–164. https://doi.org/10.1016/j.qref.2020.03.004

CCAF (Cambridge Centre for Alternative Finance). (2025). *Cambridge Bitcoin Electricity Consumption Index (CBECI).* https://ccaf.io/cbeci

Cheah, E. T., & Fry, J. (2015). Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics Letters*, *130*, 32–36. https://doi.org/10.1016/j.econlet.2015.02.029

Corbet, S., Lucey, B., Urquhart, A., & Yarovaya, L. (2018). Cryptocurrencies as a financial asset: A systematic analysis. *International Review of Financial Analysis*, *62*, 182–189. https://doi.org/10.1016/j.irfa.2018.09.003

Corbet, S., Hou, Y., Hu, Y., Larkin, C., Lucey, B., & Oxley, L. (2022). Cryptocurrency liquidity and volatility interrelationships during the COVID-19 pandemic. *Finance Research Letters*, *45*, 102137. https://doi.org/10.1016/j.frl.2021.102137

Das, D., & Dutta, A. (2020). Bitcoin's energy consumption: Is it the Achilles heel to miners' revenue? *Economics Letters*, *186*, 108530. https://doi.org/10.1016/j.econlet.2019.108530

De Vries, A. (2018). Bitcoin's growing energy problem. *Joule*, *2*(5), 801–805. https://doi.org/10.1016/j.joule.2018.04.016

Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, *49*(4), 1057–1072. https://doi.org/10.2307/1912517

Dolado, J. J., & Lütkepohl, H. (1996). Making Wald tests work for cointegrated VAR systems. *Econometric Reviews*, *15*(4), 369–386. https://doi.org/10.1080/07474939608800362

Enders, W. (2015). *Applied econometric time series* (4th ed.). Wiley.

EU Blockchain Observatory and Forum. (2023). *Ethereum merge trend report.* https://blockchain-observatory.ec.europa.eu/publications/trend-report-ethereum-merge_en

Fantazzini, D., & Kolodin, N. (2020). Does the hashrate affect the bitcoin price? *Journal of Risk and Financial Management*, *13*(12), 1–32. https://doi.org/10.3390/jrfm13110263

Gallersdörfer, U., Klaaßen, L., & Stoll, C. (2020). Energy consumption of cryptocurrencies beyond Bitcoin. *Joule*, *4*(9), 1843–1846. https://doi.org/10.1016/j.joule.2020.07.013

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, *37*(3), 424–438. https://doi.org/10.2307/1912791

Hayes, A. (2016). *Cryptocurrency value formation: An empirical analysis leading to a cost of production model for valuing Bitcoin*. SSRN Working Paper. https://doi.org/10.2139/ssrn.2648366

Koutmos, D. (2018). Return and volatility spillovers among cryptocurrencies. *Economics Letters*, *173*, 122–127. https://doi.org/10.1016/j.econlet.2018.10.004

Krause, M. J., & Tolaymat, T. (2018). Quantification of energy and carbon costs for mining cryptocurrencies. *Nature Sustainability*, *1*(11), 711–718. https://doi.org/10.1038/s41893-018-0152-7

Kumar, A. S., & Anandarao, S. (2019). Volatility spillover in cryptocurrency markets: Evidence from GARCH and wavelet analysis. *Physica A: Statistical Mechanics and its Applications*, *524*, 448–458. https://doi.org/10.1016/j.physa.2019.04.154

MacKinnon, J. G. (1996). Numerical distribution functions for unit root and cointegration tests. *Journal of Applied Econometrics*, *11*(6), 601–618.

Marchewka-Bartkowiak, K., & Wiśniewski, M. (2022). Energy tokens as digital instruments of financial investment. *Economics and Business Review*, *8*(3), 109–125. https://doi.org/10.18559/ebr.2022.3.6

Mensi, W., Rehman, M. U., Al-Yahyaee, K. H., & Vo, X. V. (2023). Frequency dependence between oil futures and international stock markets and the role of gold, bonds, and uncertainty indices: Evidence from wavelet approaches. *Resources Policy*, *80*, 103161. https://doi.org/10.1016/j.resourpol.2022.103161

Narayan, P. K. (2005). The saving and investment nexus for China: Evidence from cointegration tests. *Applied Economics*, *37*(17), 1979–1990.

Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, *16*(3), 289–326. https://doi.org/10.1002/jae.616

Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, *75*(2), 335–346. https://doi.org/10.1093/biomet/75.2.335

Qin, M. , Wu, T., Ma. X., Albu, L. L., & Umar, M. (2023). Are energy consumption and carbon emission caused by Bitcoin? *Economic Analysis and Policy*, *80*, 109–120. https://doi.org/10.1016/j.eap.2023.08.004

Rehman, M. U., & Kang, S. H. (2021). A time–frequency comovement and causality relationship between Bitcoin hashrate and energy commodity markets. *Global Finance Journal*, *49*, 100576. https://doi.org/10.1016/j.gfj.2020.100576

Sapra, N., Shaikh, I., Roubaud, D., Asadi, M., & Grebinevych, O. (2024). Uncovering Bitcoin's electricity consumption relationships with volatility and price: Environmental repercussions. *Journal of Environmental Management*, *356*, 120528. https://doi.org/10.1016/j.jenvman.2024.120528

Stoll, C., Klaaßen, L., & Gallersdörfer, U. (2019). The carbon footprint of Bitcoin. *Joule*, *3*(7), 1647–1661. https://doi.org/10.1016/j.joule.2019.05.012

Toda, H. Y., & Yamamoto, T. (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics*, *66*(1–2), 225–250. https://doi.org/10.1016/0304-4076(94)01616-8

Wang, X., Fang, F., Ma, S., Xiang, L., & Xiao, Z. (2024). Dynamic volatility spillover among cryptocurrencies and energy markets: An empirical analysis based on a multilevel complex network. *North American Journal of Economics and Finance*, *69*, 102035. https://doi.org/10.1016/j.najef.2023.102035

Aims and Scope

The aim of **Economics and Business Review** is to provide a platform for academicians from all over the world to share, discuss and integrate their research in the fields of economics and finance, including both behavioural economics and finance, with a key interest in topics that are relevant for emerging market economies. The journal welcomes submissions of articles dealing with micro, mezzo and macro issues that are well founded in modern theories or based on empirical studies and which are valuable for an international readership.

Your paper your way policy

The authors are initially expected to adjust their manuscripts to meet the basic requirements presented in the submission checklist below. Once the text has been accepted for publication, authors must adhere to all guidelines available on our website: https://journals.ue.poznan.pl/ebr

Basic requirements

- The submission has not been previously published nor is it under consideration for publication elsewhere (or an explanation has been provided in Comments to the Editor).
- The submitted manuscript must be anonymous. A separate title page must also be submitted, specifying each author's affiliation, email address, and ORCID iD. Acknowledgements and references to research grants should be included on the title page.
- The manuscript should be prepared in OpenOffice, Microsoft Word, or RTF document file format.
- The length of the manuscript should not exceed 8,000 words (including figures and tables, references, and footnotes).
- The manuscript includes an abstract of 100 to 150 words and is divided into clearly distinctive sections, including Introduction and Conclusions. The Introduction should state the aim of the study, research methods, main results, and particularly the study's contribution to international literature. The final paragraph should outline the article's content.
- All tables and figures should be numbered and presented consecutively according to their order in the text. Tables and figures should be as self-explanatory as possible, so readers do not need to refer to the main text to understand the information presented. The sources of all data used in tables and figures must be specified.
- The authors should use a consistent referencing style throughout the text.
- Please add an AI declaration (as a separate section in the manuscript) specifying whether and how you used generative AI during the preparation of the manuscript.

The submission must be made via the **submission system**: https://journals.ue.poznan.pl/ebr/submission

More information and advice on the suitability and formats of manuscripts can be obtained from:

Economics and Business Review
Al. Niepodległości 10
61-875 Poznań
Poland
e-mail: secretary@ebr.edu.pl
https://journals.ue.poznan.pl/ebr

Subscription

Economics and Business Review (EBR) is published quarterly and is the successor to the Poznań University of Economics Review. The EBR is published by the Poznań University of Economics and Business Press.

Economics and Business Review is indexed and distributed in Scopus, Claritave Analytics, DOAJ, ERIH plus, ProQuest, EBSCO, CEJSH, BazEcon and Paradigm.

Subscription rates for the print version of the EBR: institutions: 1 year – €50.00; individuals: 1 year – €25.00. Single copies: institutions – €15.00; individuals – €10.00. The EBR on-line edition is free of charge.