JOURNAL OF APPLIED AND THEORETICAL SOCIAL SCIENCES

TEORİK VE UYGULAMALI SOSYAL BİLİMLER DERGİSİ

(e-ISSN: 2687-5861)

(DOI Prefix:10.37241/jatss)

International Peer-Reviewed and Open Access Electronic Journal
Uluslararası Hakemli ve Açık Erişimli Elektronik Dergi

VOLUME 8. ISSUE 1. 2026

CİLT 8. SAYI 1. 2026

PUBLISHER AND CHIEF EDITOR / SAHİBİ VE BAŞ EDİTÖR

PROF. DR. NİLGÜN SERİM

RESEARCHER-WRITER/ARAŞTIRMACI-YAZAR

Journal of Applied and Theoretical Social Sciences/Istanbul/ Turkey

editor.jatss@iccsor.com , nilgun.serim@iccsor.com

ASSISTANT EDITOR/ YARDIMCI EDİTÖR

ASST.PROF. DR / DR. ÖĞRETİM ÜYESİ NAZMİ Y.YAĞANOĞLU

ÇANAKKALE ONSEKİZ MART UNIVERSITY

Biga Faculty of Economics and Administrative Sciences/Department of Econometrics/Çanakkale/Turkey/ **nazmiyagan@comu.edu.tr**

COMMUNICATION AND COORDINATION/İLETİŞİM VE KOORDİNASYON

https://www.iccsor.com/index.php/jatss

E-mail/ E-posta:

editor.jatss@iccsor.com

Ferah M. Nihat Sami Banarlı S. 32/ 8 Usküdar, Istanbul/Türkiye

Ferah M. Nihat Sami Banarlı S. 32/ 8 Üsküdar, İstanbul, Türkiye

We are Indexed by the Following Databases / Endekslendiğimiz Veri Tabanları

TR DİZİN

Since 2024

EBSCOhost

Since 2021

EuroPub

Since 2021

ROAD DIRECTORY OF OPEN ACCESS SCHOLARLY RESOURCES

Since 2020

WorldCat®

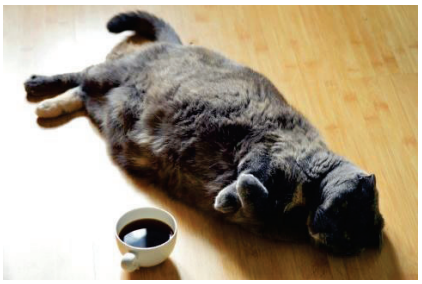
Since 2020

CORE

Since 2021

Since 2021

Since 2020



FATCAT

Since 2023

Since 2023

Member of:

Since 2019

Archived by:

Since 2022

FIELD EDITORS/ ALAN EDİTÖRLERİ

EDITORIAL and ADVISORY BOARD/ YAYIN ve DANIŞMA KURULU

By Title In Alphabetical Order/ Ünvana Göre Alfabetik Sıra ile

		of Agricultural Extension & Rural Development/ Ile-Ife, Osun State/ Nigeria	
Assoc.Prof.Dr.	Jakub BERCIK	Slovak University of Agriculture/ Faculty of Economics and Management/ the Department of Marketing and Trade/ Nitra/ Slovakia	bercik.jakubxx@gmail.com
Assoc.Prof.Dr.	Hasan DURAN	Istanbul University/Faculty of Political Sciences/Department of Political Science and International Relations/İstanbul/Turkey	hasan.duran@istanbul.edu.tr
Assoc.Prof.Dr.	Aleksander JANEŠ	University of Primorska/ Faculty of Management & Faculty of Polymer Technology/Koper/Slovenia	aleksander.janes.fm@gmail.com
Assoc.Prof.Dr.	Dushko JOSHESKI	University "Goce Delcev" at Stip, The Department of Business Administration/ Stip/North Macedonia	dushkojosheski@gmail.com
Assoc.Prof.Dr.	Adiqa KAUSAR KIANI	Federal Urdu University of Arts, Science & Technolog/Department of Economics/Islamabad/ Pakistan	adiqakian@gmail.com
Assoc.Prof.Dr.	Aroon P. MANOHARAN	Suffolk University/The Institute for Public Service (IPS)/ Sawyer Business School/Boston,/Massachusetts/ U.S.A	amanoharan@suffolk.edu
Assoc.Prof.Dr.	Jasmina MANGAFIC	University of Sarajevo/ The School of Economics and Business/Sarajevo/ Bosnia and Herzegovina	jasminamangafic111@gmail.com
Assoc.Prof.Dr.	Nelda MOUTON	North-West University, Business School, Operations Management / Supply Chain Management, North West / South Africa	nelda.mouton@nwu.ac.za
Assoc.Prof.Dr.	Edwin QUINN	Tennessee Wesleyan University/ Business Administration Program/ Athens, Tennessee/U.S.A.	ehq00@hotmail.com
Assoc.Prof.Dr.	Alexander SHEMETEV	The Russian Presidential Academy of National Economy and Public Administration/ Moscow, Saint-Petersburg/Russia	anticrisis2010@mail.ru
Assoc.Prof.Dr.	Arunkumar SIVAKUMAR	VIT-AP University/ VIT School of Business/ Tamil Nadu/India	arunkumar.sivakumar@vitap.ac.in
Asst.Prof.Dr.	Verda Davaslıgil ATMACA	Çanakkale Onsekiz Mart University/Biga Faculty of Economics and Administrative Sciences/Department of Econometrics/Çanakkale/Turkey	verdaatmaca@comu.edu.tr

CONTENTS / İÇİNDEKİLER

Research Articles/Araştırma Makaleleri

Journal of Applied and Theoretical Social Sciences

JATSS, 2026; 8(1), 1-22

First Submission:01.01.2026
Revised Submission After Review:03.03.2026
Accepted For Publication:24.03.2026
Available Online Since:25.03.2026

Research Article

Dynamic Volatility Connectedness between the Sustainability Index and Stock Market Main Sector Equity Indices

Letife Özdemir[a]

Abstract

Introduction: Assessing volatility spillovers provides guidance the decision-making processes of market participants and policymakers. In this context, identifying volatility transmitters and receivers between sustainability and traditional indices contributes to risk management.

Method: The dynamic volatility spillover among the Sustainability Index (XUSRD), the BIST-100 (XU100) and main sector indices in Turkey using the time-varying parameter vector autoregressive method. Daily data for the period 02 February 2020 to 31 July 2025 for the XUSRD, XU100, Financial (XUMAL), Industrial (XUSIN), Services (XUHIZ), and Technology (XUTEK) indices are used.

Results or Findings: The XUSRD, XU100, and XUSIN indices are volatility transmitters; while the XUMAL, XUTEK, and XUHIZ indices are volatility receivers. The high value of TCI indicates that volatility spillovers among the indices are quite strong. Volatility spillovers increases during periods of global events (the coronavirus outbreak and the Russia-Ukraine war), as well as domestic events (the February 6 earthquakes and political developments in 2025). Furthermore, volatility transfer occurs from the XUSRD to the Industrial, Services, and Technology indices at rates of 19.50%, 17.59%, 17.47%, and 14.97%, respectively.

Discussion or Conclusion: The findings indicate a high degree of dynamic volatility connectedness among the indices included in the analysis. Therefore, including these indices in the same portfolio may increase portfolio risk. Furthermore, changes in sustainability-focused expectations are seen to be rapidly transmitted to the production and service sectors, particularly the financial sector. Investors can make less risky investment decisions by paying attention to the volatility spillovers among these indices.

Keywords: sustainability index, stock market sector indices, dynamic connectedness, volatility spillower, TVP-VAR model

JEL Codes: C32, G11, G17

[a] Doç. Dr., Afyon Kocatepe Üniversity, Faculty of Economics and Administrative Sciences, Department of International Trade and Finance, Afyonkarahisar/Türkiye, letifeozdemir@aku.edu.tr, ORCID ID: https://orcid.org/0000-0002-8636-2277 (Corresponding Author)

Journal of Applied and Theoretical Social Sciences

JATSS, 2026; 8(1), 1-22

İlk Başvuru: 01.01.2026
Düzeltilmiş Makalenin Alınışı: 03.03.2026
Yayın İçin Kabul Tarihi: 24.03.2026
Online Yayın Tarihi: 25.03.2026

Araştırma Makalesi

Sürdürülebilirlik Endeksi İle Borsa ve Ana Sektör Pay Endeksleri Arasındaki Dinamik Volatilite Bağlantısı

Letife Özdemir[a]

Öz

Giriş: Volatilite yayılmasının değerlendirilmesi, piyasa katılımcılarının ve politika yapıcıların karar alma süreçlerine rehberlik eder. Bu bağlamda, sürdürülebilirlik ve geleneksel endeksler arasındaki volatilite aktarıcılarını ve alıcılarını belirlemek, risk yönetimine katkı sağlar.
Yöntem: Türkiye'de Sürdürülebilirlik Endeksi (XUSRD) ile BİST-100 (XU100) ve ana sektör endeksleri arasındaki dinamik volatilite yayılımı, zamana göre değişen parametreli vektör otoregresif yöntemiyle incelenmektedir.XUSRD, XU100, Mali (XUMAL), Sınai (XUSIN), Hizmetler (XUHİZ) ve Teknoloji (XUTEK) endekslerine ait 02.02.2020–31.07.2025 dönemi günlük veriler kullanılmaktadır.
Sonuçlar ya da Bulgular: XUSRD, XU100 ve XUSIN endeksleri volatilite yayıcısı; XUMAL, XUTEK ve XUHİZ endeksleri ise volatilite alıcısı konumundadır. TCI katsayısının yüksek olması, endeksler arasında volatilite yayılımı oldukça güçlü olduğunu işaret etmektedir. Volatilite yayılımı koronavirüs salgını ve Rusya–Ukrayna savaşı gibi küresel olayların ve 6 Şubat depremleri ve 2025 yılı siyasi gelişmeleri gibi ülke içi olayların olduğu dönemlerde artmaktadır. Ayrıca, XUSRD endeksinden Mali, Sınai, Hizmetler ve Teknoloji endekslerine sırasıyla %19,50, %17,59, %17,47 ve %14,97 oranlarında volatilite aktarımı gerçekleşmektedir.
Tartışma ya da Yapılan Çıkarımlar: Bulgular, analize dâhil edilen endeksler arasında yüksek düzeyde dinamik volatilite bağlantılılığı olduğunu ifade etmektedir. Bundan dolayı, bu endeksler aynı portföy içerisinde yer alması portföy riskini artırabilir. Ayrıca, sürdürülebilirlik odaklı beklentilerdeki değişimlerin başta finans sektörü olmak üzere üretim ve hizmet sektörlerine hızlı biçimde yansıdığı görülmektedir. Yatırımcıların, bu endeksler arasındaki volatilite yayılımına dikkat ederek daha az riskli yatırım kararları alabilirler.

Anahtar Kelimeler: sürdürülebilirlik endeksi, borsa sektör endeksleri, dinamik bağlantılılık, volatilite yayılımı, TVP-VAR modeli

JEL Kodlar: C32, G11, G17

[a] Doç. Dr., Afyon Kocatepe Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Uluslararası Ticaret ve Finansman Bölümü, Afyonkarahisar/Türkiye, letifeozdemir@aku.edu.tr, ORCID ID: https://orcid.org/0000-0002-8636-2277 (Sorumlu Yazar)

Introduction

Sustainability is described in the Brundtland Report (1987) as "meeting the needs of the present without compromising the ability of future generations to meet their own needs". Corporate sustainability represents the firm-level application of the sustainability concept and refers to the adaptation of environmental, social, and governance (ESG) principles into corporate activities in order to enable firms to maintain their existence and support their growth. ESG principles were conceptually, for the first time, provided with a systematic framework through the report "Who Cares Wins" published by the United Nations in 2004 (Mian et al., 2024). The global acceptance of ESG criteria and their incorporation into investment decision-making processes gained prominence following the establishment of the United Nations Principles for Responsible Investment in 2006 (Peillex & Comyns, 2020).

The integration of ESG principles into investment decision-making processes has facilitated the emergence of sustainability indices as a new category of equity market indices. Sustainability indices include firms that carry out activities in accordance with ESG principles (Bhue et al., 2025). In recent years, sustainability indices have gained increasing importance in financial markets and have emerged as an alternative investment option for investors in portfolio construction (Zhang et al., 2022; Aboluwodi et al., 2024). At the same time, sustainability indices are important financial indicators that guide investors within the framework of a responsible investment approach (Kurnoga et al., 2022). Unlike traditional stock market indices, these indices offer an evaluation opportunity based on sustainability principles, and their importance has been increasing day by day in both the academic literature and investment strategies. Existing studies indicate that sustainability indices exhibit strong bidirectional volatility spillovers with traditional stock market indices during periods of global crises and uncertainty (Arouri et al., 2025; Naeem et al., 2023; Nogueira & Madaleno, 2022; Malhotra, 2025; Zeng et al., 2025; Doğan et al., 2023; Erben Yavuz, 2023; Kavas, 2025; Toy, 2025).

Globalization, liberalization, and technological advancements have led financial markets to become more integrated with one another and have caused adverse developments arising in one market to be transmitted rapidly to other markets. In this context, volatility spillovers and dynamic connectedness among financial markets have become a extensively assessed subject in prior studies (Khan, 2025). Generally, during periods when economic uncertainty and volatilities specifically the coronavirus outbreak, market participants and policymakers tend to closely examine risk and return spillovers across markets in order to protect investments and ensure financial stability (Naeem et al., 2023). Existing empirical studies has been undertaken to enhance the understanding of volatility spillovers in financial markets, to assess their implications for diversification benefits, and to develop portfolio construction approaches aimed at mitigating the associated risks (Choi vd., 2021).

Volatility spillover refers to the process by which uncertainty or fluctuations occurring in a particular market are transmitted to other markets. In this context, the "volatility transmitter" refers to the market that transmits net volatility to other markets, while the "volatility receiver" refers to the market that absorbs net volatility in the system, meaning the market that is more affected by shocks in other markets. Volatility connectedness quantitatively reveals the extent and direction in which shocks arising in different markets or assets spread throughout the system (Baruník, Kočenda, & Vácha, 2017). Assessing volatility spillovers contributes significantly to the decision-making processes of market participants and policymakers. Identifying net volatility transmitters and net volatility receivers between sustainability and conventional indices enables the enhancement of effective portfolio

diversification and risk management strategies. Moreover, it provides important insights into understanding the role of sustainable finance markets within the financial system.

Although prior studies have investigated volatility spillovers between sustainability indices and financial market assets, notable gaps remain in the existing literature. The literature indicates that there is strong dynamic volatility connectedness among various sustainability indices (Bhue et al., 2025; Marin-Rodriguez et al., 2024; Umar et al., 2020; Wan et al., 2024; Shaik & Rehman, 2023); that volatility spillovers run from sustainability indices to capital market indices (Zhang et al., 2022; Maraqa & Bein, 2020; Arouri et al., 2025; Naeem et al., 2023); and that analyses of dynamic volatility spillovers between sustainability indices and sectoral indices reveal sustainability indices to act as net volatility transmitters (Malhotra, 2025; Zeng et al., 2025). In the context of Türkiye, empirical evidence on volatility spillovers involving sustainability indices remains rather limited. Empirical evidence indicates that the Sustainability Index functions as a net volatility receiver from the Dow Jones Sustainability Index (Başkaya & Özdemir, 2025; Erben Yavuz, 2023). This finding is also supported by the study of Doğan et al. (2023). Kaya (2023) and Kavas (2025) investigated volatility spillovers between the sustainability index and various commodity prices, showing that, in most cases, the sustainability index acts as a net volatility transmitter. However, only two studies focusing on the Indian and Chinese capital markets have been identified that incorporate sectoral indices into their analyses (Malhotra, 2025; Zeng et al., 2025). These studies have demonstrated that the sustainability index transmits volatility to sectoral indices. To our knowledge, no prior study has been identified for the case of Türkiye. The need for further research to identify volatility spillovers between the sustainability index and sectoral indices, and to understand how these dynamics are reflected in investor behavior, constitutes the main motivation of this study.

This paper aims to evaluate the dynamic volatility spillovers among Türkiye's the Sustainability Index, the benchmark stock market index, and the main sectoral indices by employing a time-varying parameter vector autoregressive (TVP-VAR) methodology. Volatility spillovers among the XUSRD index, the XU100 index, representing the largest equity market benchmark, and the main sectoral indices, namely Financial, Industrial, Services, and Technology, are examined in detail using daily data covering the period from 02.02.2020 to 31.07.2025. From an investor perspective, evaluating sectoral indices alongside the XU100 Index, rather than focusing solely on the XU100 Index, is of considerable importance. While the XU100 Index reflects the overall market trend, sectoral indices allow for a more detailed assessment of sector-specific behaviors and structural characteristics. Within this framework, the study focuses on the following research questions: (i) Does the Sustainability Index act as a volatility transmitter or a volatility receiver in the equity market? (ii) Which of the main sectoral indices acts as the largest net volatility receiver? (iii) To which sector does the Sustainability Index transmit volatility most strongly?

While aiming to answer these research questions, this paper is considered to add to the literature in three in notable respects. First, the inclusion of firms with high ESG performance in sustainability indices within a given sector influences the return and volatility dynamics of those sectors. Therefore, price movements and volatility shocks occurring in sustainability indices can spillover into sectoral indices. Therefore, price movements and volatility shocks occurring in sustainability indices can spillover into sectoral indices. Similarly, sector-specific developments may also influence the performance of sustainability indices. For these reasons, this study provides empirical evidence to the existing literature by examining how volatility spillovers and shocks are transmitted between sustainability and sectoral indices. Second, instead of methods based on the assumption of constant parameters, this study employs the TVP-VAR approach, which allows parameters to vary over time and enables an analysis of the

temporal evolution of connectedness. In this manner, the temporal sensitivity of Türkiye's capital markets to periods of uncertainty is identified, and market reactions are assessed in a comparative framework. Thus, identifying volatility-transmitting and volatility-receiving markets between sustainability and sectoral indices will contribute to investors' portfolio diversification and risk management decisions. Third, this paper addresses an existing gap in the literature by analyzing volatility spillovers between the XUSRD Index, the XU100 Index, and the equity indices of main sectors in the Turkish economy, namely the industrial, financial, technology, and services sectors.

The remainder of the study is structured as follows. In the second section, studies investigating dynamic volatility spillovers between the sustainability index and financial indicators are reviewed. In the third section, the data set and methodology are presented, while the findings are reported in the fourth section. In the final section, the findings are discussed and evaluated.

Literature Review

Previous studies on volatility spillovers and dynamic interconnectedness across financial markets plays a critical role in shaping investment decision-making processes. In recent years, the incorporation of ESG considerations into investment strategies has enhanced the role of sustainability indices by offering various advantages in portfolio diversification. Consequently, the number of studies examining volatility spillovers between sustainability indices and different financial markets has increased substantially.

Maraqa and Bein (2020) examine the dynamic interrelationships and volatility spillovers among sustainability equity indices and the major stock market returns of oil-importing and oil-exporting European countries, using a DCC-MGARCH framework. Their findings reveal that the dynamic relationships among sustainability equity indices, stock returns of European oil-importing/exporting countries, and oil markets differ across market conditions. Arouri et al. (2025) investigate the mutual connectedness among the S&P Green Bond Index, the S&P500 Index, and the S&P500ESG Index by employing TVP-VAR and DCC-GARCH models. The results indicate the presence of significant volatility spillovers regardless of market conditions, with green bonds acting as the primary volatility receivers, while the S&P500 and S&P500ESG indices serve as net volatility transmitters. Using the Diebold–Yilmaz and Baruník–Křehlík connectedness frameworks, Naeem et al. (2023) analyze the country-level connectedness among sustainable and Islamic equity investments across nineteen countries representing developed and emerging financial markets. Their results document moderate to strong short-term connectedness, and time-varying connectedness analyses reveal that major market-wide shocks, such as the European sovereign debt crisis, the Chinese financial crisis, and the coronavirus outbreak, significantly intensify interconnectedness among the indices. Nogueira and Madaleno (2022), employing an MGARCH model, find that the EURO STOXX Sustainability Index is highly dependent on movements in conventional stock market indices, indicating that sustainability indices function predominantly as net volatility receivers.

Studies examining the dynamic volatility connectedness between sustainability indices belonging to different markets are also included in the literature. Bhue et al. (2025) investigate information asymmetry, dynamic connectedness, and volatility spillovers between ESG equity indices in emerging markets using GARCH, FIGARCH, and TVP-VAR models. Their findings reveal the presence of leverage effects in ESG equity indices and indicate a moderate level of connectedness among the indices. In particular, ESG equity indices from the Philippines, Indonesia, Korea, Singapore, and India are found to be more sensitive to sustainability-related

innovation shocks originating from the ESG indices of Brazil, South Africa, and Mexico. Marin-Rodriguez et al. (2024) analyze the connectedness of sustainable investment indices using a quantile VAR (QVAR) framework. The results show persistent dynamic connectedness among the Sustainable Impact Index, the Paris-aligned equity index, and the green bond index. Moreover, the green bond index generally acts as a net volatility receiver, whereas the other sustainability indices function as volatility transmitters. Zhang et al. (2022) employ a DCC-GARCH-based dynamic connectedness approach to examine volatility spillovers among the MSCI World ESG equity index, the renewable energy equity index, the S&P Green Bond Index, the Dow Jones Sustainability equity index, and carbon emission futures. Their findings identify carbon emission futures as the primary volatility transmitter, followed by the ESG and the sustainability equity index.

Using a DCC-GARCH framework, Shaik and Rehman (2023) assess dynamic volatility connectedness among ESG equity indices and show that ESG indices in the Middle East and Africa as well as Latin America act as net shock transmitters, while those in the United States and the Asia-Pacific region are net volatility receivers. Umar et al. (2020) document significant and persistent volatility spillovers among ten ESG equity indices from developed and emerging markets. The connectedness among these indices exhibits a pronounced dynamic structure, particularly during periods associated with the European sovereign debt crisis, Greece-originated systemic disturbances, and the coronavirus pandemic. Wan et al. (2024), employing a TVP-VAR framework, reveal strong connectedness among ESG equity indices in both returns and volatility. While ESG indices in Europe and North America predominantly display outward spillovers, those in the Asia-Pacific region and India are characterized by inward spillovers. In addition, the coronavirus outbreak is found to have a positive effect on volatility connectedness among ESG equity indices.

Focusing on volatility spillovers between sustainability indices and sectoral indices, Malhotra (2025) finds that India's ESG and clean energy indices consistently act as net volatility transmitters, whereas sectoral indices serve as diversification assets with relatively low connectedness, based on a TVP-VAR approach. Similarly, Zeng et al. (2025) examine risk connectedness between China's sustainability index, the Chinese stock market, and seven sectoral indices using a TVP-VAR-CAViaR connectedness framework. The evolution of total risk connectedness exhibits dynamic fluctuations in both the Chinese stock market and the sustainability index, particularly during extreme events specifically the coronavirus outbreak. Their results indicate that sustainability indices play a central role in risk transmission.

A limited number of studies focusing on the Turkish context examine volatility spillovers among the XUSRD Index and stock market indices or other financial indicators, generally concluding that the XUSRD Index functions as a net volatility receiver. Doğan et al. (2023) investigate the dynamic connectedness between the XUSRD Index, the XU100 Index, the S&pGCEI, and S&P GSCI using a TVP-VAR framework. Their findings indicate that volatility originating from S&P GSCI is transmitted to the S&PGCEI, the XU100 Index, and the XUSRD Index. Erben Yavuz (2023) analyze dynamic connectedness among the XUSRD Index, the S&PGSCI, and the Dow Jones Sustainability World Index using a TVP-VAR approach, documenting that the XUSRD Index acts as a net volatility receiver. In a related study, Başkaya and Özdemir (2025) employ a DCC-GARCH model and show that volatility shocks in the S&P Dow Jones Sustainability Index reduce the volatility of the XUSRD Index.

Kaya (2023) analyze dynamic connectedness among the XUSRD Index and fossil fuel prices, including crude oil, Brent oil, heating oil, natural gas, and coal, using a TVP-VAR methodology. The evidence confirm that Brent oil, crude oil, and heating oil serve as net

volatility transmitters, whereas the XUSRD Index, natural gas, and coal predominantly act as volatility receivers. Kavas (2025) apply a TVP-VAR model to examine dynamic connectedness and volatility spillovers among the XUSRD Index, the US dollar exchange rate, Bitcoin, oil prices, and gold prices. Their findings identify Bitcoin as the largest volatility transmitter, followed by the XUSRD Index, while oil and gold emerge as the largest volatility receivers.

The existing literature also includes studies analyzing volatility spillovers between the XU100 Index and various market indicators. Specifically, Aydoğdu (2025) examine volatility spillovers among the XU100Index, Brent crude oil, the US dollar, gold, Bitcoin, and the Volatility Index; Çilek (2025) focus on the XU100 Index, interest rates, gold, and the US dollar; Erdoğan (2023) analyze volatility spillovers among interest rates, the XU100 Index, bond yields, the US dollar, gold, and CDS premiums; and Gülcemal (2025) investigate volatility spillovers among the XU100 Index, the Housing Price Index, gold, and the US dollar exchange rate. These studies generally suggest that the XU100 Index acts as a net volatility transmitter in the domestic market. In contrast, Polat (2021), examining volatility spillovers between BRICS markets and Türkiye, and Erdoğan (2024), analyzing volatility spillovers between the XU100 Index and stock market indices of eight different countries, find that the XU100 Index functions as a net volatility receiver in the global market context. Collectively, these findings suggest that while the XU100 Index acts as a volatility transmitter within the domestic market, it tends to assume the role of a volatility receiver in international market settings.

Only a limited number of studies examine volatility spillovers among sectoral indices in Türkiye. Kocaarslan (2020) analyze volatility interactions among the Technology, Industrial, Services, and Financial indices using a causality-in-variance framework, revealing substantial volatility transmission from the Technology Index to other major sectoral indices. Additionally, significant volatility spillovers are identified from the Industrial Index to the Services and Financial indices. Toy (2025), employing a TVP-VAR approach, document the presence of dynamic connectedness among the Financial, Services, Industrial, and Technology indices. Their results indicate that the Industrial and Services indices act as net volatility transmitters, whereas the Financial and Technology indices serve as net volatility receivers.

Overall, when evaluated in general terms, the number of studies examining volatility spillovers between sustainability indices and financial indicators in the existing literature is still limited. To our knowledge, no prior research has addressed volatility spillovers between the XUSRD Index, the XU100 Index, and equity indices representing key sectors of the Turkish economy, namely industrial, financial, technology, and services, within a unified framework. Moreover, few studies have jointly analyzed sustainability indices and sectoral indices, focusing exclusively on the Indian and Chinese markets. To address these gaps, the present study evaluates the dynamic connectedness structure and volatility spillovers between the XUSRD Index and major sectoral indices using a time-varying parameter vector autoregressive methodology.

Method

This paper examines volatility spillovers among the XUSRD Index, the XU100 Index, and equity indices representing key sectors of the Turkish economy, namely industrial, financial, technology, and services, using the TVP-VAR methodology developed by Antonakakis and Gabauer (2017). Daily closing price data for the selected indices were obtained from the Investing database and cover the period from 2 January 2020 to 31 July 2025. Prior to the analysis, logarithmic daily returns were derived using $100\times\ln(P_t/P_{(t-1)})$. Detailed information regarding the variables and their definitions is presented in Table 1.

Table 1

Variable Description

Variable	Description	Data Source	Period
XUSRD	BIST Sustainability Index	Investing.com	02/01/2020-31/07/2025
XU100	BİST-100 Index	Investing.com	02/01/2020-31/07/2025
XUSIN	BİST Industrial Index	Investing.com	02/01/2020-31/07/2025
XUMAL	BİST Financial Index	Investing.com	02/01/2020-31/07/2025
XUTEK	BİST Technology Index	Investing.com	02/01/2020-31/07/2025
XUHIZ	BİST Services Index	Investing.com	02/01/2020-31/07/2025

Notes. Prepared by the author.

In this paper, volatility spillovers among the XUSRD Index, the XU100 Index, and sectoral equity indices are examined using the TVP-VAR methodology. The TVP-VAR framework, which allows for the analysis of volatility spillovers and time-varying volatility dynamics, was originally developed within the connectedness framework pioneered by Diebold and Yılmaz (2009, 2012, 2014) and subsequently extended by Antonakakis and Gabauer (2017). The TVP-VAR model in light of the BIC is presented in Equations (1), (2), and (3) (Antonakakis, Gabauer, Gupta ve Plakandaras, 2018):

$$Y_t = \beta_t Y_{t-1} + \varepsilon_t \qquad \varepsilon_t \sim N(0, S_t) \qquad (1)$$

$$\beta_t = \beta_{t-1} + u_t \qquad u_t \sim N(0, R_t) \qquad (2)$$

$$Y_t = A_t \varepsilon_{t-1} + \varepsilon_t \qquad (3)$$

In Equations (1), (2), and (3), Y_t, ε_t, and $u_t$are $N \times 1$vectors, while A_t, S_t, β_t, and R_t are $N \times N$matrices. Equation (3) represents the Wold decomposition of the system. The time-varying coefficients in the vector moving average representation constitute the theoretical foundation of the connectedness index developed by Diebold and Yılmaz (2012). This index is computed using generalized impulse response functions and generalized forecast error variance decompositions (GFEVD), drawing on the frameworks proposed by Koop et al. (1996) and Pesaran and Shin (1998).

In this study, the primary focus is on the proportion of the h-step-ahead forecast error variance of variable $i$that is attributable to shocks originating from variable j. This can be expressed mathematically as follows:

$$\tilde{\varphi_{ij,t}^{g}}(h) = \frac{\sum_{t=1}^{h-1} \delta_{ij,t}^{2,g}}{\sum_{i=1}^{N} \sum_{t=1}^{h-1} \delta_{ij,t}^{2,g}} \qquad (4)$$

In Model (4), $\tilde{\phi_{ij,t}^{g}}(h)$denotes the h-step-ahead GFEVD. Building on the GFEVD, the Total Connectedness Index (TCI), which measures the overall degree of interdependence within the network, is constructed and formulated as follows:

$$C_t^g(h) = \frac{\sum_{ij=1,i\neq j}^{N} \varphi_{ij,t}^{\sim g}(h)}{\sum_{j=1}^{N} \varphi_{ij,t}^{\sim g}(h)} x100 \quad (5)$$

The TCI consists of three main components: total directional connectedness (TDC) to others, TDC from others, and net TDC. These components capture both the direction and the magnitude of interactions among variables.

First, attention is focused on spillovers transmitted from variable ito all other variables j. This measure is referred to as TDC to others and is presented in Equation (6):

$$C_{i\rightarrow j,t}^g(h) = \frac{\sum_{j=1,i\neq j}^{N} \varphi_{ji,t}^{\sim g}(h)}{\sum_{j=1}^{N} \varphi_{ji,t}^{\sim g}(h)} x100 \quad (6)$$

Second, spillovers transmitted from all variables j to variable i are computed. This concept is referred to as TDC from others and is presented in Equation (7).

$$C_{i\leftarrow j,t}^g(h) = \frac{\sum_{j=1,i\neq j}^{N} \varphi_{ij,t}^{\sim g}(h)}{\sum_{i=1}^{N} \varphi_{ij,t}^{\sim g}(h)} x100 \quad (7)$$

Third, net TDC**,** denoted as $C_{i,t}^{\tilde{g}}$, is obtained by computing the difference between TDC to others and TDC from others**.**

$$C_{i,t}^g(h) = C_{i\rightarrow j,t}^g(h) - C_{i\leftarrow j,t}^g(h) \quad (8)$$

The sign of net TDC provides information on the role of each variable in the volatility transmission mechanism. If net TDC is positive $\left(C_{i,t}^{\tilde{g}}(h) > 0\right)$, variable $i$acts as a net volatility transmitter**;** conversely, if it is negative $\left(C_{i,t}^{\tilde{g}}(h) < 0\right)$, variable $i$functions as a net volatility receiver**.** Finally, to examine bilateral interactions, net TDC is decomposed, and net pairwise directional connectedness (NPDC) is computed and expressed as follows:

$$NPDC_{ij}(h) = \frac{\varphi_{ji,t}^{\sim g}(h) - \varphi_{ij,t}^{\sim g}(h)}{N} x100 \quad (9)$$

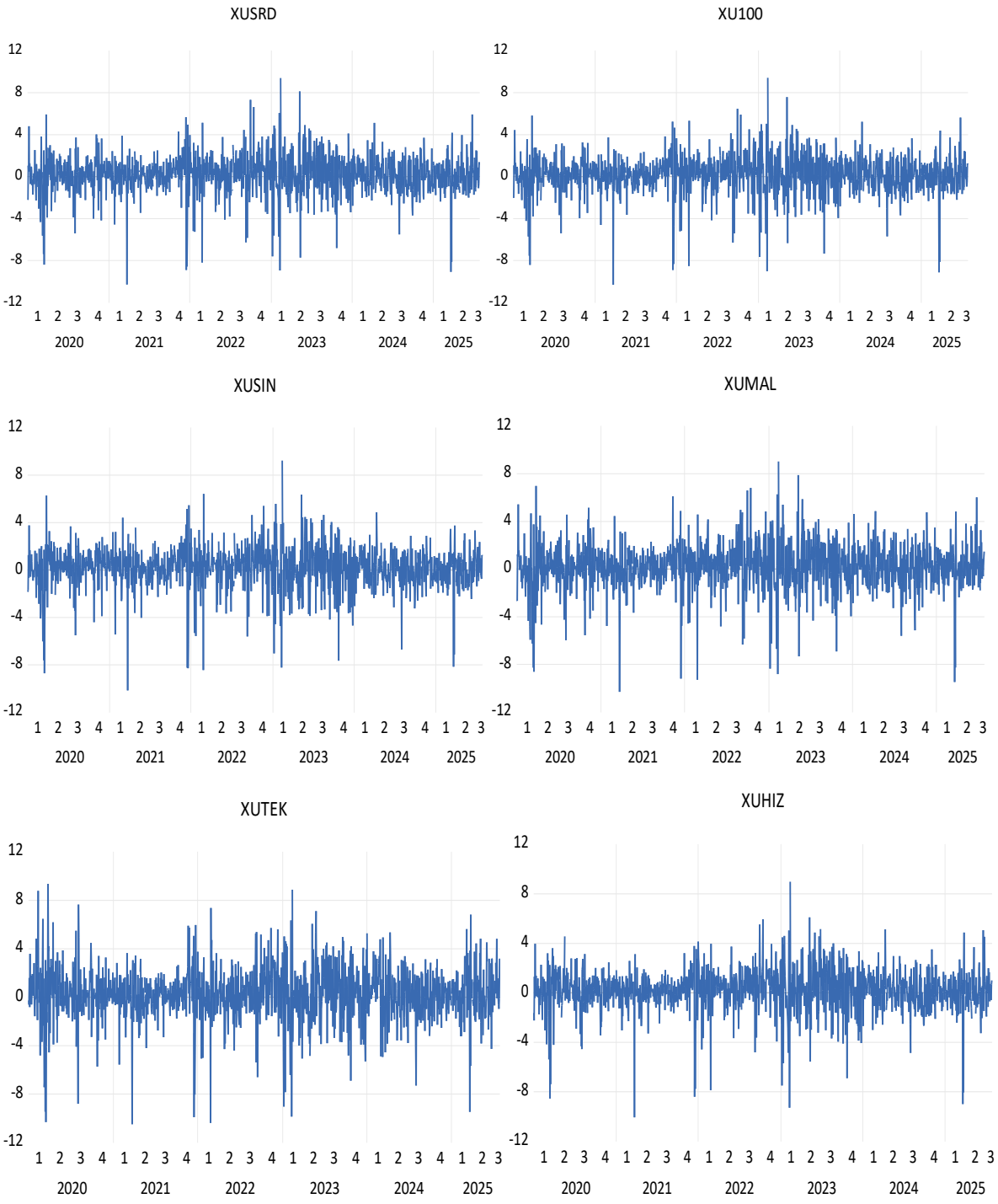
If $\text{NPDC}_{ij}(h) > 0$, variable $i$exerts an influence on variable j; conversely, if $\text{NPDC}_{ij}(h) < 0$, variable $j$exerts an influence on variable i.

Empirical Findings

The study first presents the return series of the XUSRD Index, the XU100 Index, and the sectoral equity indices in Figure 1. An examination of the graphs shows that the indices exhibit similar volatility patterns. The effects of the coronavirus outbreak in 2020, the Russia–Ukraine war in 2022, the February 6 earthquakes in 2023, and the political events in 2025, which caused volatility in the markets, can be observed in the graphs.

Figure 1.

Plots of the Return Series



Note. Created by the author in EViews 12; Arial font customization is not supported.

Summary statistics for the index returns included in the paper are reported in Table 2. The Technology equity index exhibits the highest average return among the indices considered. Variance, as a fundamental statistical measure, captures the dispersion of observations around the mean and reflects the degree of variability in the series. Among the indices, the Technology equity index displays the greatest deviation from its mean. The negative skewness values indicate that the distributions are left-skewed, while kurtosis values exceeding three suggest that the return distributions exhibit leptokurtic characteristics. The Jarque–Bera test results, based on skewness and kurtosis measures, indicate that none of the variables follow a normal distribution (Jarque ve Bera, 1987). The test statistics confirm the rejection of normality for all indices at the 1% significance level. The stationarity properties of the variables are examined using the ERS unit root test developed by Elliott, Rothenberg, and Stock (1996), and the results reveal that all variables are stationary. Finally, the Ljung–Box Q and Q^2 test confirm the presence of autocorrelation in the return series.

Table 2

Descriptive Statistics

	XUSRD	XU100	XUSIN	XUMAL	XUTEK	XUHIZ
Mean	0.166	0.16	0.163	0.167	0.228	0.178
Variance	3.415	3.27	3.151	3.826	4.848	2.926
Skewness	-0.646	-0.768	-0.854	-0.591	-0.516	-0.850
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Kurtosis	4.616	4.856	4.583	3.892	3.406	5.173
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Jarque-Bera	1334.824	1506.743	1389.345	960.882	735.627	1721.831
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ERS	-5.093	-5.547	-5.739	-4.604	-13.407	-6.066
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Q(10)	7.159	6.928	8.052	11.206	8.789	3.284
	(0.238)	(0.261)	(0.166)	(0.039)	(0.121)	(0.778)
$Q^2(10)$	139.645	130.264	150.412	113.443	171.256	167.766
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note. Created by the author using a web-based R-Shiny interface.

To examine the dynamic volatility connectedness between the variables, the TVP-VAR model is employed. The TDC relationships between the XUSRD Index, the XU100 Index, and the sectoral equity indices obtained from the model are summarized in Table 3. Dynamic connectedness reflects the transmission of volatility spillovers across the variables.

Table 3

Dynamic Correlation Relationships for Variables

	XUSRD	XU100	XUSIN	XUMAL	XUTEK	XUHIZ	From
XUSRD	19.68	19.29	16.32	17.60	11.34	17.60	80.32
XU100	18.86	19.23	16.96	16.98	11.68	16.29	80.77
XUSIN	17.59	18.69	21.14	14.16	12.78	15.63	78.86
XUMAL	19.50	19.24	14.52	21.83	10.74	14.18	78.17
XUTEK	14.97	15.79	15.81	12.76	26.40	14.26	73.60
XUHIZ	17.47	18.48	16.09	14.17	11.94	21.86	78.14
To	88.39	91.49	79.70	75.67	58.48	76.12	469.86
Inc.Own	108.07	110.71	100.85	97.50	84.88	97.98	cTCI/TCI
NET	8.07	10.71	0.85	-2.50	-15.12	-2.02	93.97/78.31

Note. Created by the author using a web-based R-Shiny interface.

Table 3 reports the extent to which volatility in each index is driven by its own shocks and by shocks originating from other indices, as well as the direction and magnitude of spillover effects across indices. The "NET" measure presented in the table captures the difference between the volatility transmitted to other variables and the volatility received from them. A positive NET value indicates that the variable acts as a net volatility transmitter, whereas a negative value implies that it functions as a net volatility receiver. According to Table 3, the XUSRD Index, the XU100 Index, and the Industrial Index are identified as net volatility transmitters, while the Financial, Technology, and Services sector indices emerge as net volatility receivers. The XU100 Index is the largest volatility transmitter (10.71%), followed by the XUSRD Index (8.07%), whereas the Technology sector index is the largest volatility receiver (15.12%).

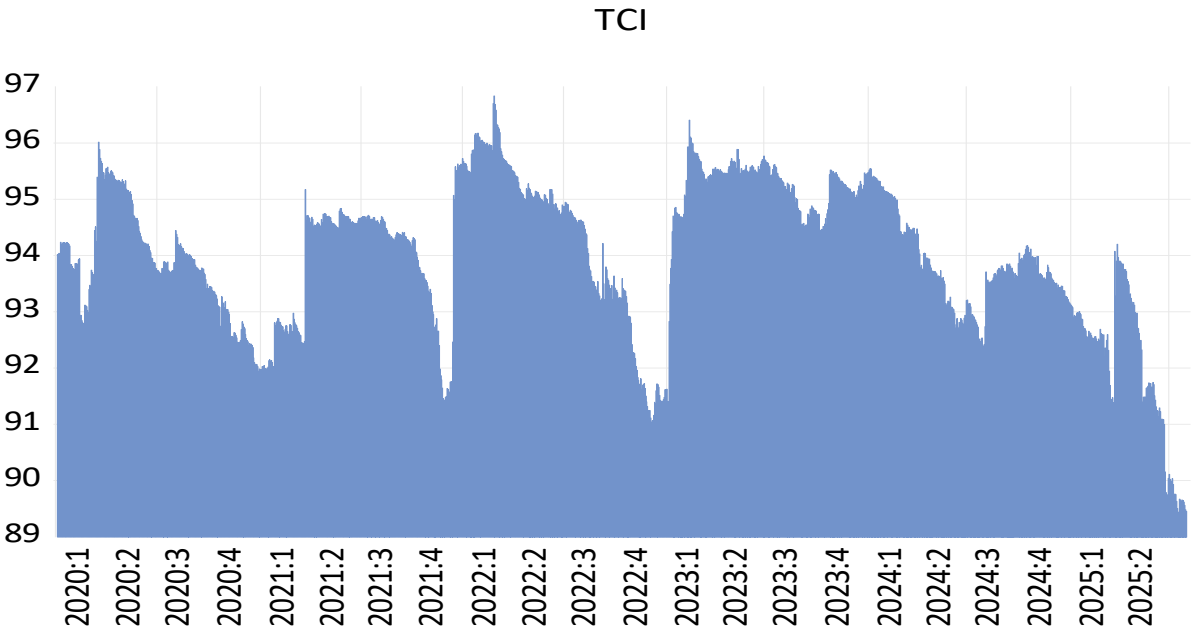
The findings reported in Table 3 also allow for an assessment of the sources of volatility fluctuations across the variables. In this context, volatility fluctuations observed in the XUSRD Index are largely driven by its own innovations (19.68%), followed by shocks originating from the XU100 Index (19.29%). The Technology sector index exerts the weakest influence on the XUSRD Index, accounting for only 11.34% of its volatility. For the XU100Index, 19.23% of volatility fluctuations stem from its own shocks, while 18.86% originate from the XUSRD Index, 16.98% from the Financial Index, 16.96% from the Industrial Index, 16.29% from the Services Index, and 11.68% from the Technology Index. Volatility fluctuations in the Industrial Index are primarily driven by its own shocks (21.14%), followed by shocks from the XU100 Index (18.69%) and the XUSRD Index (17.59%). Similarly, volatility fluctuations in the Financial Index are mainly explained by its own shocks (21.83%), while shocks from the XU100 Index (19.24%) and the XUSRD Index (19.50%) also play a substantial role. For the Technology Index, 26.40% of volatility fluctuations are attributed to its own innovations, whereas 15.79% and 14.97% are driven by shocks from the XU100 and XUSRD indices, respectively. Finally, volatility fluctuations in the Services Index are largely explained by its own shocks (21.86%), followed by shocks from the XU100Index (18.48%) and the XUSRD Index (17.47%).

Over the entire sample period, the XUSRD Index transmits volatility most strongly to the Financial, XU100, Industrial, Services, and Technology indices, respectively. Conversely, it receives volatility primarily from the XU100 Index, followed by the Financial, Services, Industrial, and Technology indices.

Finally, the TCI between the variables is estimated at 78.31% over the sample period, indicating a high degree of interconnectedness across the indices. Figure 2 illustrates the evolution of the TDC among the XUSRD Index, the XU100 Index, and the sectoral equity indices.

Figure 2

Total Dynamic Connectedness

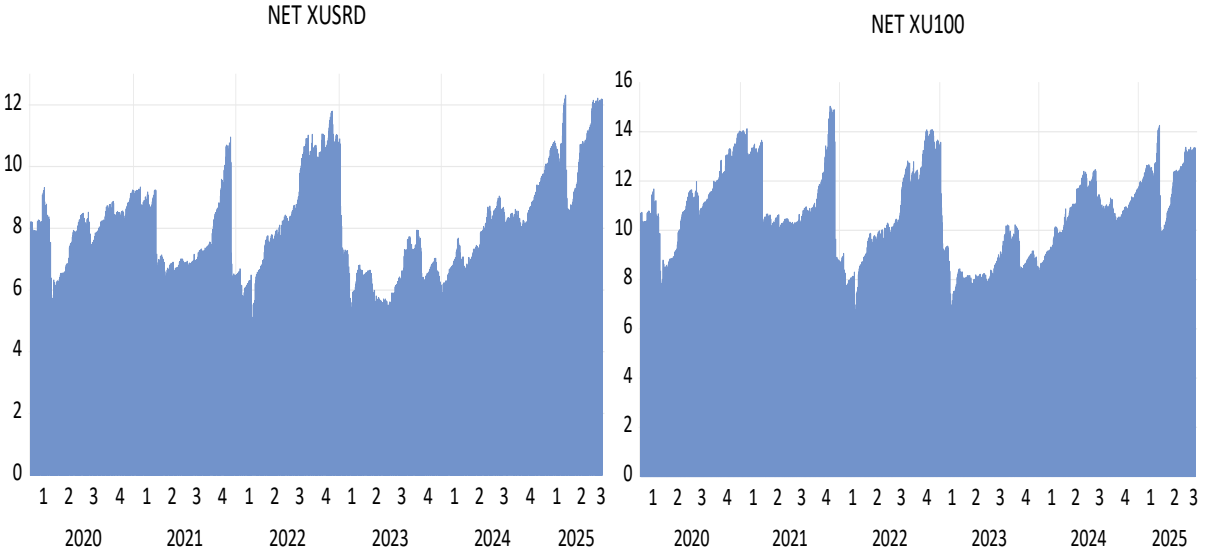


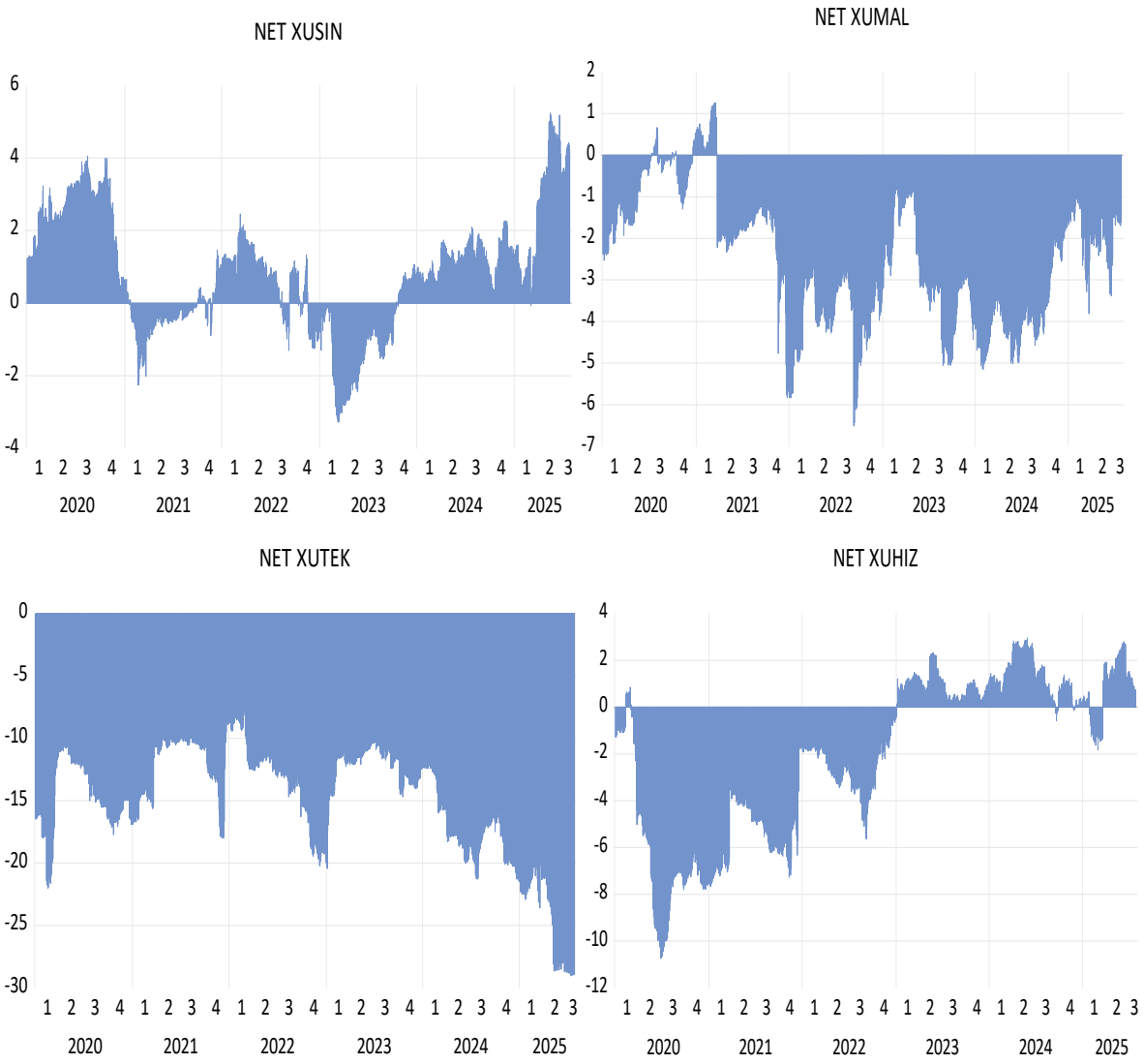
Note. Created by the author in EViews 12; Arial font customization is not supported.

An examination of the figure reveals that dynamic connectedness among the variables exhibits periods of increase and decrease over time. Periods characterized by heightened total connectedness and intensified volatility spillovers across variables coincide with the onset of the coronavirus outbreak in early 2020, the outbreak of the Russia–Ukraine war in early 2022, the earthquakes of 6 February 2023, and the period of political events in the second quarter of 2025. Notably, volatility spillovers across the indices reached their highest level following the earthquakes that occurred on 6 February 2023. To provide a clearer depiction of volatility transmission among the variables, the net total directional connectedness plots are presented in Figure 3.

Figure 3

Dynamic Net Directional Connectedness



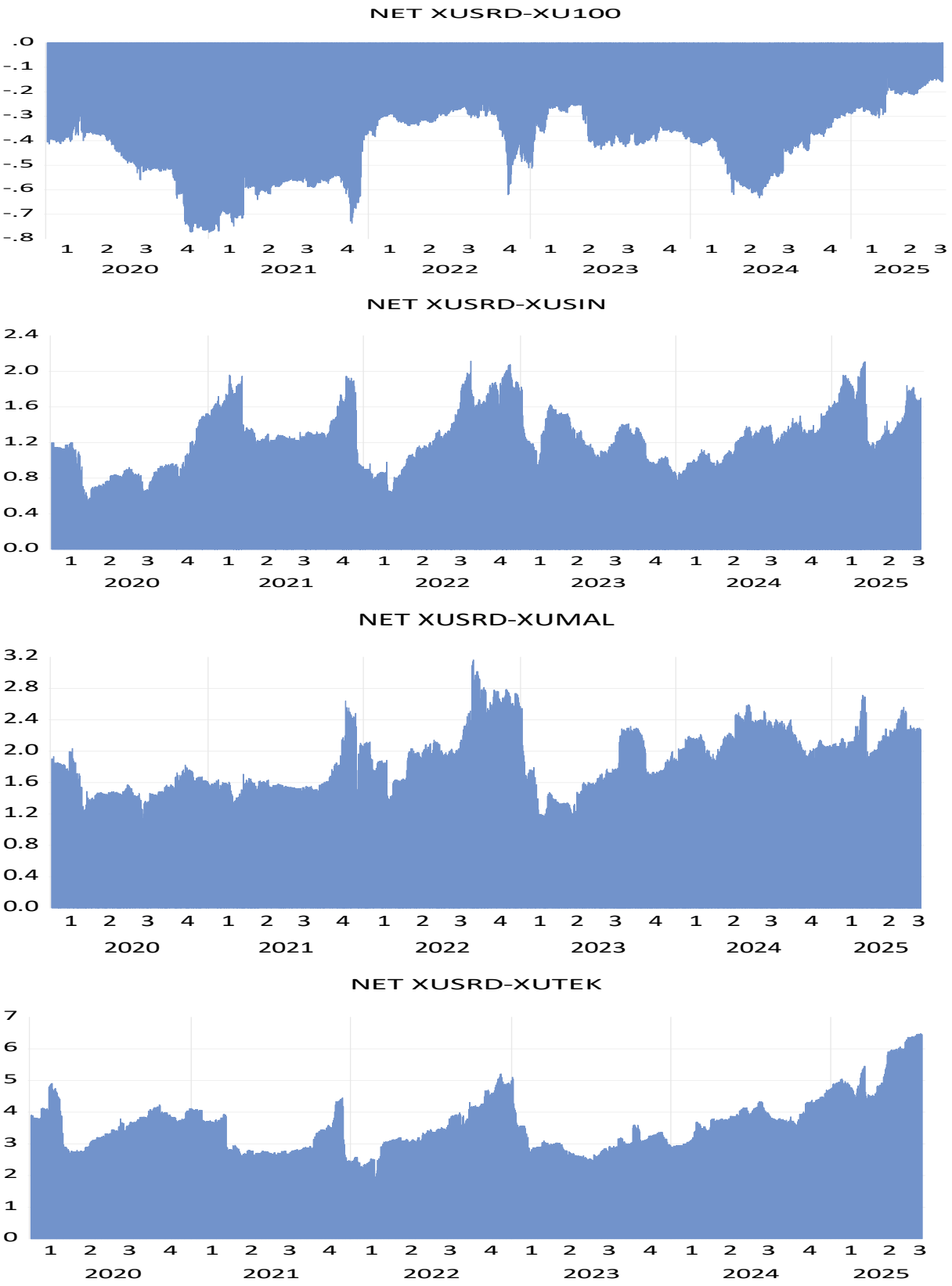


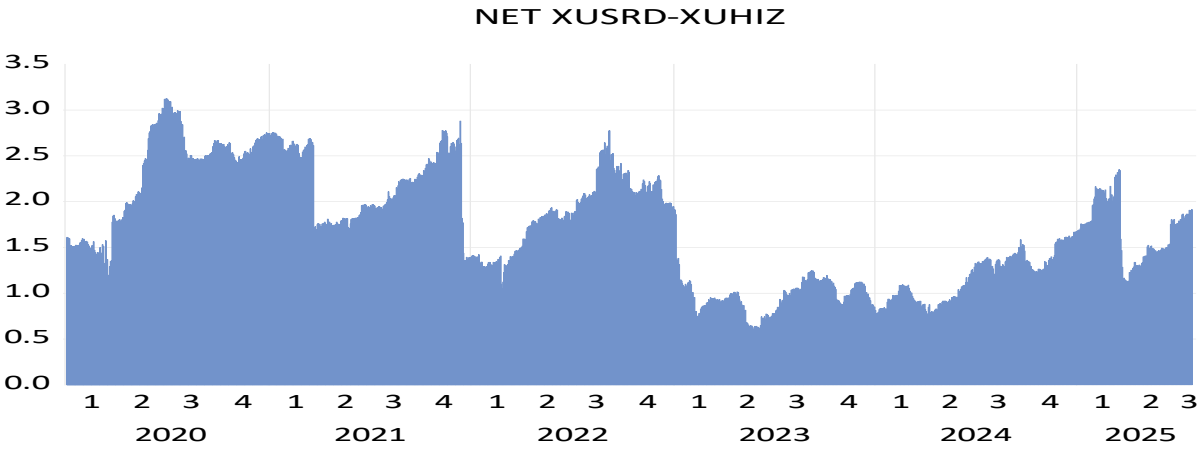
Note. Note. Created by the author in EViews 12; Arial font customization is not supported.

According to the directional connectedness analysis, the dark-shaded area in Figure 3 remaining above the horizontal axis indicates that the corresponding variable acts as a net volatility transmitter during the sample period, whereas values below the horizontal axis imply a net volatility receiver. The graphical evidence suggests that both the XUSRD Index and the XU100 Index function as net volatility transmitters. The Industrial equity index acts as a volatility receiver in 2021 and 2023, while serving as a volatility transmitter during the remaining periods. An examination of the Financial equity index indicates that it behaves as a volatility transmitter only at the beginning of 2021, whereas it assumes the role of a volatility receiver throughout the rest of the sample period. The Services equity index functions as a volatility receiver between 2020 and 2022, but shifts to a volatility-transmitting role after 2023. In contrast, the Technology equity index consistently remains a net volatility receiver throughout the entire period. The net pairwise dynamic connectedness relationships illustrating bilateral volatility spillovers between the XUSRD Index and the other indices are presented in Figure 4.

Figure 4.

Net Pairwise Dynamic Connectedness



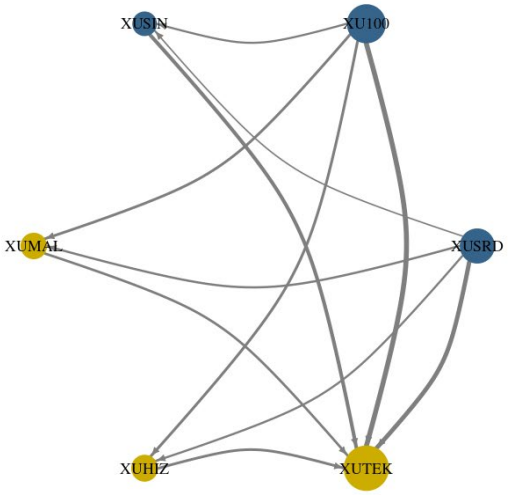


Note. Created by the author in EViews 12; Arial font customization is not supported.

In the net pairwise dynamic connectedness plots, the dark-shaded area lying above the horizontal axis indicates that volatility is transmitted from the first index to the second, whereas an area below the horizontal axis implies that volatility is transmitted from the second index to the first. An examination of the plots reveals a volatility spillover from the XU100 index to the XUSRD Index, while volatility is transmitted from the XUSRD Index to the Industrial, Financial, Services, and Technology stock indices. To further illustrate the direction and strength of volatility spillovers among the XUSRD Index, the XU100 index, and the sectoral indices, the net connectedness network across variables is presented in Figure 5.

Figure 5

Connectedness Network among Variables



Note. Created by the author using a web-based R-Shiny interface. Arial font customization is not supported.

In Figure 5, the indices shown in blue act as net volatility transmitters, whereas those shown in yellow are net volatility receivers over the sample period. The varying sizes of the circles representing the indices are proportional to the magnitude of volatility transmitted and received by each variable. The arrows depicting volatility spillovers indicate the direction of transmission, while their thickness reflects the strength of the spillover effects. An examination of the network reveals that the XU100, XUSRD, and Industrial sector indices function as net

volatility transmitters, whereas the Financial, Services, and Technology sector indices are positioned as net volatility receivers. The XU100index emerges as the largest volatility transmitter, followed by the XUSRD Index as the second most influential transmitter. These indices transmit the highest level of volatility to the Technology sector index, which is also identified as the sector index receiving the greatest amount of volatility. The values of the net connectedness indicator presented in Table 3 for the dynamic connectedness relationships among the variables are consistent with the network outputs shown in Figure 5.

Conclusion

Dynamic connectedness and volatility spillovers across markets constitute one of the core research areas that has attracted considerable attention in the finance literature. This study examines in detail the volatility spillovers between Türkiye's sustainability index, the stock market benchmark index, and the main sector indices. In this study, the volatility spillovers among Türkiye's sustainability index (XUSRD Index), the largest equity market index (XU100 Index), and the main sectoral equity indices, Financial, Industrial, Services, and Technology are analyzed using daily data for the period from 02 February 2020 to 31 July 2025. In the empirical analyses, a TVP-VAR model with a dynamic parameter structure is adopted, and volatility spillovers are evaluated within this framework. The TVP-VAR model reveals the dynamic structure of volatility spillovers among variables and provides the opportunity to examine their evolution over time.

The net connectedness indicator, a key outputs of the TVP-VAR model, identifies whether variables act as net volatility transmitters or net volatility receivers over time. According to the net connectedness outcomes, the XUSRD, XU100, and Industrial indices function as net volatility transmitters, whereas the Financial, Technology, and Services indices are net volatility receivers. The XU100 Index is identified as the strongest volatility transmitter, with a value of 10.71%, while the XUSRD Index ranks second with 8.07%. The index receiving the highest level of volatility is the Technology equity index, with a value of 15.12%. Findings show that the sustainability index acts as a volatility transmitter is consistent with the results reported by Malhotra (2025), Zeng et al. (2025), Kavas (2025), and Arouri et al. (2025). In contrast to these studies, Doğan et al. (2023), Erben Yavuz (2023), and Başkaya and Özdemir (2025) found that the XUSRD Index acts as a volatility receiver. These studies analyzed the interactions between the XUSRD Index and global indices, including the S&P and Dow Jones Sustainability indices, indicating that the XUSRD Index, as an emerging market index, is affected by international financial markets. Another implication of the net connectedness index, namely the role of the Technology equity index as a volatility receiver, corroborates the results of Toy (2025).

The estimated TCI of 78.31% indicates a very strong level of interaction among the indices. This indicates a high degree of dynamic volatility connectedness among the indices included in the analysis and suggests that holding these indices together in the same portfolio may increase portfolio risk.

According to the total dynamic connectedness plot, volatility spillovers among the indices exhibit periods of increase and decrease over time. The dynamic net directional connectedness evidence show that the XUSRD and XU100 indices generally act as volatility transmitters, while the Technology Index functions as a volatility receiver. The Industrial, Financial, and Services indices assume alternating roles as volatility transmitters and receivers across different subperiods. Periods of elevated total connectedness and stronger volatility

interactions among the indices are observed to coincide with episodes of major economic, geopolitical, and political shocks. Within this context, particularly notable are the onset of the coronavirus outbreak in early 2020, the outbreak of the Russia-Ukraine war in early 2022, the February 6, 2023 earthquakes, and political developments in the second half of 2025. The highest level of volatility spillovers is observed during the February 2023 earthquake period. Findings reveal that major global and domestic shocks amplify volatility spillovers across markets is in line with the empirical evidence reported by Arouri et al. (2025), Naeem et al. (2023), Nogueira and Madaleno (2022), Bhue et al. (2025), Marin-Rodriguez et al. (2024), Zhang et al. (2022), Shaik and Rehman (2023), Wan et al. (2024), Malhotra (2025), Zeng et al. (2025), Doğan et al. (2023), Erben Yavuz (2023), Kaya (2023), Kavas (2025), and Toy (2025). The results suggest that, during adverse periods, the connectedness among indices does not remain constant but rather exhibits a dynamic structure characterized by fluctuations over time. This implies that investors should account for the changing connectedness structure among indices during crisis periods.

Finally, the TVP-VAR outcomes show that volatility originating from the XUSRD Index is transmitted by 19.50% to the Financial Index, 17.59% to the Industrial Index, 17.47% to the Services Index, and 14.97% to the Technology Index. These findings indicate that changes in sustainability policies and ESG-focused expectations have a significant impact on the real sector, and that fluctuations in the sustainability index are rapidly transmitted to the production and services sectors, particularly the financial sector. In particular, the Financial index exhibits a higher sensitivity to such sustainability-driven uncertainties.

By shedding light on the volatility spillover structure between Türkiye's sustainability index and the main sectoral indices, this study provides valuable insights for portfolio diversification and risk management strategies. Future research could examine the volatility structure between sustainability indices and equity market indices in both developed and emerging markets within a broader framework. Moreover, including alternative investment instruments beyond equity indices in the analysis would provide valuable information for risk management strategies.

References

Aboluwodi, D., Muzindutsi, P. F., & Nomlala, B. (2024). Investor sentiments and performance of selected ESG indices in BRICS markets during bull and bear conditions. *Investment Analysts Journal, 53*(3), 231–248. https://doi.org/10.1080/10293523.2024.2397889

Antonakakis, N., & Gabauer, D. (2017). Refined measures of dynamic connectedness based on TVP-VAR (MPRA Paper No. 78282). Munich Personal RePEc Archive.

Antonakakis, N., Gabauer, D., Gupta, R., & Plakandaras, V. (2018). Dynamic connectedness of uncertainty across developed economies: A time-varying approach. *Economics Letters, 166*, 63–75. https://doi.org/10.1016/j.econlet.2018.02.011

Arouri, M., Mhadhbi, M., & Shahrour, M. H. (2025). Dynamic connectedness and hedging effectiveness between green bonds, ESG indices, and traditional assets. *European Financial Management, 31*, 1704–1719. https://doi.org/10.1111/eufm.12561

Aydoğdu, A. (2025). Risk şokları ve Türkiye'deki finansal varlıklar arasındaki yayılım etkisinin TVP-VAR dayalı wavelet uyum analizi ile incelenmesi. *Ekonomi Politika ve Finans Araştırmaları Dergisi, 10*(1), 303–331. https://doi.org/10.30784/epfad.1595233

Baruník, J., Kočenda, E., & Vácha, L. (2017). Asymmetric volatility connectedness on the forex market. *Journal of International Money and Finance*, 77, 39-56. https://doi.org/10.1016/j.jimonfin.2017.06.003

Başkaya, H., & Özdemir, A. (2025). Bazı sürdürülebilirlik endekslerinin volatilite modelleriyle incelenmesi. *Selçuk Üniversitesi Sosyal Bilimler Meslek Yüksekokulu Dergisi, 28*(1), 60–77. https://doi.org/10.29249/selcuksbmyd.1619942

Bhue, R., Gartia, U., Panda, A. K., & Nanda, S. (2025). Sustainable equity index dynamic: Connectedness and information asymmetry in emerging markets. *Business Strategy & Development, 8*(1), e70072. https://doi.org/10.1002/bsd2.70072

Choi, K. H., McIver, R. P., & Ferraro, S. (2021). Dynamic volatility spillover and network connectedness across ASX sector markets. *Journal of Economics and Finance, 45*, 677–691. https://doi.org/10.1007/s12197-021-09544-w

Çilek, A. (2025). Interaction and volatility spillover among selected financial assets in Türkiye. *Akademik İzdüşüm Dergisi, 10*(1), 101–131.

Diebold, F. X., & Yılmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *The Economic Journal, 119*(534), 158–171.

Diebold, F. X., & Yılmaz, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting, 28*(1), 57–66.

Diebold, F. X., & Yılmaz, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics, 182*(1), 119–134.

Doğan, M., Raikhan, S., Zhanar, N., & Gulbagda, B. (2023). Analysis of dynamic connectedness relationships among clean energy, carbon emission allowance, and BIST indexes. *Sustainability, 15*(7), Article 6025. https://doi.org/10.3390/su15076025

Elliott, G., Rothenberg, T. J., & Stock, J. H. (1996). Efficient tests for an autoregressive unit root. *Econometrica, 64*, 813–836. https://doi.org/10.2307/2171846

Erben Yavuz, A. (2023). Temiz enerji, sürdürülebilir ve BIST endeksleri arasındaki ilişkilerin analizi: TVP-VAR yaklaşımı. *İşletme Akademisi Dergisi, 4*(3), 339–354.

Erdoğan, B. (2023). The volatility relationship among financial assets: TVP-VAR model. International *Journal of Business and Economic Studies, 5*(4), 225–237. https://doi.org/10.54821/uiecd.1392184

Erdoğan, B. (2024). Bitcoin, petrol ile borsalar arasındaki volatilite analizi. *İstatistik Araştırma Dergisi*, *14*(1), 19–35.

Gülcemal, T. (2025). Hisse senedi, konut piyasası, altın ve döviz kuru arasındaki dinamik bağlantılılık ilişkisinin analizi: Türkiye örneği. *Üçüncü Sektör Sosyal Ekonomi Dergisi, 60*(3), 2729–2742. https://doi.org/10.63556/tisej.2025.1575

Jarque, C. M., & Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review, 55*, 163–172. https://doi.org/10.2307/1403192

Kavas, Ü. Y. B. (2025). Finansal enstrümanların BIST sürdürülebilirlik endeksi üzerindeki dinamik etkilerinin TVP-VAR modeliyle araştırılması. *Mali Çözüm Dergisi, 35*, 1201–1225.

Kaya, M. (2023). BIST sürdürülebilirlik endeksi ile fosil yakıt fiyatları arasındaki ilişkinin analizi. *Abant Sosyal Bilimler Dergisi, 23*(3), 1475–1495. https://doi.org/10.11616/asbi.1327883

Khan, I. (2025). An analysis of stock markets integration and dynamics of volatility spillover in emerging nations. *Journal of Economic and Administrative Sciences, 41*(3), 861–878. https://doi.org/10.1108/JEAS-10-2022-0236

Koop, G., Pesaran, M. H., & Potter, S. M. (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics, 74*(1), 119–147.

Kurnoga N., Šimurina, N., & Fučkan, F. (2022). Performance differences between ESG indices and conventional market indices: A multivariate analysis of indices. *Zagreb International Review of Economics and Business, 25*(s1), 85–103. https://doi.org/10.2478/zireb-2022-0026

Malhotra, P. (2025). Environmental, social and governance (ESG) investing and Indian capital market indices: Dynamic connectedness and risk management strategies. *Journal of Capital Markets Studies, 9*(2), 191–213. https://doi.org/10.1108/JCMS-12-2024-0096

Maraqa, B., & Bein, M. (2020). Dynamic interrelationship and volatility spillover among sustainability stock markets, major European conventional indices, and international crude oil. *Sustainability, 12*(9), 3908. https://doi.org/10.3390/su12093908

Marin-Rodriguez, N. J., Gonzalez-Ruiz, J. D., & Botero, S. (2024). Connectedness between sustainable investment indexes: The QVAR approach. *Economies, 12*(7), 170. https://doi.org/10.3390/economies12070170

Mian, H. R., Hewage, K., & Sadiq, R. (2024). Responsible financing and investment: Identification, development, and assessment of environmental, social, and governance

(ESG) metrics. *Sustainable Futures, 8*, Article 100246. https://doi.org/10.1016/j.sftr.2024.100246

Naeem, M. A., Farid, S., Arif, M., Paltrinieri, A., & Alharthi, M. (2023). COVID-19 and connectedness between sustainable and Islamic equity markets. *Borsa Istanbul Review, 23*(1), 1–21. https://doi.org/10.1016/j.bir.2022.06.001

Nogueira, M. C., & Madaleno, M. (2022). Are sustainability indices infected by the volatility of stock indices? Analysis before and after the COVID-19 pandemic. *Sustainability, 14*, 15434. https://doi.org/10.3390/su142215434

Pesaran, H. H., & Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics Letters, 58*(1), 17–29.

Peillex, J., & Comyns, B. (2020). Why do financial companies adopt the United Nations principles for responsible investment? *Comptabilité Contrôle Audit, 26*(1), 79–117. https://doi.org/10.3917/cca.261.0079

Polat, O. (2021). Dynamic network connectedness of BRICS equity markets during the COVID-19 era. *Ömer Halisdemir Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 14*(4), 1486–1498. https://doi.org/10.25287/ohuiibf.936124

Shaik, M., & Rehman, M. Z. (2023). The dynamic volatility connectedness of major environmental, social, and governance (ESG) stock indices: Evidence based on DCC-GARCH model. *Asia-Pacific Financial Markets, 30,* 231–246. https://doi.org/10.1007/s10690-022-09393-5

Toy, A. (2025). BIST ana sektör endeksleri arasındaki bağlantılılık ilişkileri: TVP-VAR yaklaşımı. *Uluslararası Akademik Birikim Dergisi, 8*(2), 160–170.

Umar, Z., Kenourgios, D., & Papathanasiou, S. (2020). The static and dynamic connectedness of environmental, social, and governance investments: International evidence. *Economic Modelling, 93*, 112–124. https://doi.org/10.1016/j.econmod.2020.08.007

Wan, J., Yin, L., & Wu, Y. (2024). Return and volatility connectedness across global ESG stock indexes: Evidence from the time-frequency domain analysis. *International Review of Economics & Finance, 89*, 397–428. https://doi.org/10.1016/j.iref.2023.10.038

Brundtland, G. (1987). Report of the world commission on environment and development: our common future. *United Nations General Assembly document A/42/427.* https://cdn.iuc.edu.tr/FileHandler2.ashx?f=rapor_638035862645304306.pdf

Zeng, H., Benkraiem, R., Abedin, M. Z., & Hajek, P. (2025). Transitioning to sustainability: Dynamic spillovers between sustainability indices and Chinese stock market. *European Financial Management, 31*(5), 1742–1770. https://doi.org/10.1111/eufm.12560

Zhang, W., He, X., & Hamori, S. (2022). Volatility spillover and investment strategies among sustainability-related financial indexes. *International Review of Financial Analysis*, *83*, Article 102223. https://doi.org/10.1016/j.irfa.2022.102223

Information About the Article/Makale Hakkında Bilgiler

The Ethical Rules for Research and Publication / Araştırma ve Yayın Etiği

The author declared that the ethical rules for research and publication followed while preparing the article.

Yazar makale hazırlanırken araştırma ve yayın etiğine uyulduğunu beyan etmiştir.

Conflict of Interests/ Çıkar Çatışması

The author have no conflict of interest to declare.

Yazar çıkar çatışması bildirmemiştir.

Grant Support/ Finansal Destek

The author declared that this study has received no financial support.

Yazar bu çalışma için finansal destek almadığını beyan etmiştir.

Author Contributions/ Yazar Katkıları

The draft process of the manuscript/ Taslağın Hazırlanma Süreci L.Ö, Data Collection/Verilerin Toplanması L.Ö, Writing the Manuscript/ Makalenin Yazılması L.Ö, Submit, Revision and Resubmit Process/ Başvuru, Düzeltme ve Yeniden Başvuru Süreci L.Ö.

Journal of Applied and Theoretical Social Sciences

JATSS, 2026; 8(1), 23-44

First Submission:01.01.2026
Revised Submission After Review:09.03.2026
Accepted For Publication:24.03.2026
Available Online Since:25.03.2026

Research Article

The Functional Composition of Public Expenditure and Inflation Dynamics in Türkiye: A Fourier Toda–Yamamoto Causality Analysis under Structural Breaks

Lütfü Sizer [a]

Abstract

Introduction: This study examines the relationship between inflation and the functional composition of public expenditures in Türkiye, emphasizing that fiscal policy affects price dynamics not only through its size but also through its allocation across expenditure categories. Given Türkiye's experience with persistent inflation and policy regime shifts, understanding how different public spending components interact with inflation is of particular importance.

Method: The analysis employs quarterly data covering the period 2006Q1–2025Q2 and includes the shares of general public services, economic affairs, health, education, and social protection expenditures in GDP. To account for mixed integration orders and smooth structural changes without prior knowledge of break dates, the study applies the Fourier-augmented Toda–Yamamoto causality approach.

Results or Findings: The results reveal that the inflationary effects of public expenditures differ across functional categories. Unidirectional causality is identified from General Public Services and Health expenditures to inflation. In contrast, bidirectional causal relationships are detected between inflation and Economic Affairs, Education, and Social Protection expenditures, indicating the presence of feedback mechanisms.

Discussion or Conclusion: The findings suggest that inflation dynamics in Türkiye depend critically on the functional composition of public spending rather than aggregate expenditure alone. From a policy perspective, expenditure reallocation strategies that consider category-specific inflationary effects may enhance the effectiveness of fiscal policy in supporting price stability and macroeconomic sustainability.

Keywords: public expenditures, inflation dynamics, fourier toda–yamamoto

JEL Codes: B22, C01, C22

[a] Asst. Prof. Dr., Dicle University, Faculty of Economics and Administrative Sciences, Department of Economics, Diyarbakır/Türkiye, lutfu.sizer@dicle.edu.tr, ORCID ID: https://orcid.org/0000-0002-9605-4286 (Corresponding Author)

Journal of Applied and Theoretical Social Sciences

JATSS, 2026; 8(1), 23-44

İlk Başvuru: 01.01.2026
Düzeltilmiş Makalenin Alınışı: 09.03.2026
Yayın İçin Kabul Tarihi: 24.03.2026
Online Yayın Tarihi: 25.03.2026

Araştırma Makalesi

Türkiye'de Kamu Harcamalarının Fonksiyonel Bileşimi ve Enflasyon Dinamikleri: Yapısal Kırılmalar Altında Fourier Toda–Yamamoto Nedensellik Analizi

Lütfü Sizer[a]

Öz

Giriş: Bu çalışma, Türkiye'de enflasyon ile kamu harcamalarının fonksiyonel bileşimi arasındaki ilişkiyi incelemekte ve maliye politikasının fiyat dinamiklerini yalnızca harcama büyüklüğü üzerinden değil, aynı zamanda harcamaların fonksiyonel dağılımı yoluyla da etkilediğini vurgulamaktadır. Türkiye'nin kalıcı enflasyon sorunları ve politika rejimi değişimleri dikkate alındığında, farklı kamu harcama kalemlerinin enflasyonla nasıl etkileşime girdiğinin anlaşılması büyük önem taşımaktadır.

Yöntem: Analizde, 2006Q1–2025Q2 dönemini kapsayan üç aylık veriler kullanılmış ve Genel Kamu Hizmetleri, Ekonomik İşler, Sağlık, Eğitim ve Sosyal Koruma harcamalarının GSYİH içindeki payları dikkate alınmıştır. Değişkenlerin farklı bütünleşme derecelerine sahip olması ve önceden belirlenemeyen yumuşak yapısal kırılmaların varlığı nedeniyle Fourier genişletilmiş Toda–Yamamoto nedensellik yaklaşımı uygulanmıştır.

Sonuçlar ya da Bulgular: Elde edilen bulgular, kamu harcamalarının enflasyon üzerindeki etkilerinin harcama kalemlerine göre farklılaştığını göstermektedir. Genel Kamu Hizmetleri ve Sağlık harcamalarından enflasyona doğru tek yönlü nedensellik ilişkisi tespit edilirken; Ekonomik İşler, Eğitim ve Sosyal Koruma harcamaları ile enflasyon arasında çift yönlü nedensellik ilişkileri belirlenmiştir. Bu durum, bazı harcama kalemlerinde geri besleme mekanizmalarının varlığına işaret etmektedir.

Tartışma ya da Yapılan Çıkarımlar: Sonuçlar, Türkiye'de enflasyon dinamiklerinin yalnızca toplam kamu harcaması düzeyine değil, harcamaların fonksiyonel bileşimine de güçlü biçimde bağlı olduğunu ortaya koymaktadır. Politika açısından, harcama kalemlerine özgü enflasyonist etkileri dikkate alan bir kamu harcaması bileşimi, fiyat istikrarını desteklemede maliye politikasının etkinliğini artırabilir.

Anahtar Kelimeler: kamu harcamaları, enflasyon dinamikleri, fourier toda–yamamoto

JEL Kodlar: B22, C01, C22

[a] Dr. Öğr. Üyesi, Dicle Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İktisat Bölümü, Diyarbakır/Türkiye, lutfu.sizer@dicle.edu.tr, ORCID ID: https://orcid.org/0000-0002-9605-4286

Introduction

Inflation remains one of the most persistent challenges to macroeconomic stability and sustainable growth. Particularly in developing economies, inflation dynamics cannot be explained solely by monetary factors such as money supply growth or exchange rate pass-through. Instead, fiscal policy—and more specifically the structure and composition of public expenditures—plays a crucial role in shaping price dynamics. This perspective aligns with theoretical frameworks that emphasize the interaction between fiscal policy regimes and price level determination, commonly referred to as the Fiscal Theory of the Price Level (FTPL) (Leeper, 1991; Sims, 2013; Woodford, 1995). In countries such as Türkiye, which has experienced recurrent episodes of high and volatile inflation, the coherence between fiscal policy and price stability has long been a central issue of economic analysis (Anand & van Wijnbergen, 1989). Recent empirical research focusing on Türkiye also highlights the importance of fiscal policy dynamics in shaping inflation outcomes. For instance, Ekincı, Saygılı, and Yılmaz (2025) find that specific fiscal instruments, particularly indirect taxes and operational government expenditures, are significantly associated with long-run inflation dynamics in Türkiye.

While the relationship between public expenditures and inflation has been widely examined in the literature, much of the empirical evidence relies on aggregate measures of government spending or fiscal deficits. Such an approach implicitly assumes that different categories of public expenditures exert homogeneous effects on inflation, thereby overlooking the fact that government spending influences the economy through multiple and heterogeneous channels. In contrast, the public finance literature has long emphasized the importance of functional classifications of government expenditures. Therefore, analyzing the inflationary effects of public spending solely through aggregate expenditure measures may be misleading. A focus on the functional composition of public expenditures provides a more informative framework for understanding the underlying fiscal–inflation nexus. Building on Musgrave's (1959) functional approach to public finance, the Classification of the Functions of Government (COFOG) provides a standardized framework that distinguishes government spending according to its primary objectives, including general public services, economic affairs, health, education, and social protection (Eurostat, 2019; United Nations Statistics Division [UNSD], n.d.).

The main research questions addressed in this study are as follows:

1. How does the functional composition of public expenditures affect inflation dynamics in Türkiye?
2. Which categories of public expenditures exhibit unidirectional or bidirectional causality with inflation?
3. How do structural shifts in fiscal policy regimes influence the relationship between public expenditures and inflation?

The functional composition of public expenditures is theoretically relevant for inflation dynamics because each spending category operates through distinct demand- and supply-side mechanisms. General Public Services expenditures mainly encompass administrative services, public governance, and debt-related items. These expenditures are typically associated with short-run demand pressures and fiscal sustainability concerns, which may influence inflation expectations and price formation processes. In the Fiscal Theory of the Price Level (FTPL)

framework, fiscal imbalances and the financing of government obligations can directly affect the price level, even in the absence of monetary accommodation (Leeper, 1991; Sims, 2013; Woodford, 1995). Empirical evidence for Türkiye also suggests that the interaction between government spending, fiscal financing, and inflation has been a persistent feature of macroeconomic outcomes (Anand & van Wijnbergen, 1989).

By contrast, economic affairs expenditures include public investments and interventions aimed at supporting productive sectors such as transportation, energy, agriculture, and industry. These expenditures are not limited to stimulating aggregate demand; they also enhance productive capacity and infrastructure, thereby influencing inflation through supply-side channels. Endogenous growth models highlight that productive public spending can raise private sector productivity and long-term output, potentially mitigating inflationary pressures over time (Barro, 1990). Recent empirical studies further show that the inflationary impact of fiscal expansions critically depends on the composition of public spending, with investment-oriented expenditures exerting weaker—or even disinflationary—effects compared to consumption-oriented spending (Klein & Linnemann, 2023). Evidence from developing economies also indicates that inflation outcomes under different fiscal regimes are closely linked to the allocation of public expenditures across functional categories (Apeti et al., 2023).

Health services and education services expenditures occupy a distinct position within the functional classification of government spending, as they are directly related to human capital formation. From a theoretical perspective, health is considered a form of human capital that enhances labor productivity and effective labor supply, thereby expanding an economy's productive capacity (Grossman, 1972). Similarly, education expenditures improve workforce skills and technological adaptability, reinforcing long-term growth and supply-side resilience (Barro, 1990; Aghion & Howitt, 1998). Consequently, expenditures on health and education services may exert a dual effect on inflation: while they may generate short-run demand pressures, their long-run impact may be disinflationary through productivity and capacity-enhancing channels. This perspective suggests that analyzing health and education spending solely as expenditure aggregates may obscure their broader macroeconomic role.

Social protection expenditures, including social transfers, pensions, unemployment benefits, and income support programs, primarily function as automatic stabilizers. By smoothing household income and consumption over the business cycle, social protection mechanisms can mitigate output volatility and economic downturns. However, during periods of high inflation, expansionary social transfers may amplify demand-side pressures and contribute to price increases, particularly when not accompanied by corresponding productivity gains (Dolls et al., 2009). As a result, the relationship between social protection expenditures and inflation is inherently context-dependent and may involve feedback effects between prices, fiscal policy, and household income dynamics.

Despite the growing recognition of composition effects in fiscal policy, the existing literature remains relatively limited in its application of a comprehensive functional classification framework that simultaneously examines general public services, economic affairs, health services, education services, and social protection in relation to inflation. Most empirical studies focus on either aggregate public spending or isolated expenditure categories, thereby failing to capture the broader interaction between expenditure composition and inflation dynamics (Eurostat, 2019; UNSD, n.d.). Yet recent evidence underscores that fiscal shocks can have markedly different inflationary consequences depending on how public resources are allocated across functions (Klein & Linnemann, 2023; Apeti et al., 2023).

Against this backdrop, the present study investigates the causal relationships between inflation and the functional components of public expenditures—general public services, economic affairs, health services, education services, and social protection—in Türkiye. Given the country's experience with structural transformations, policy regime shifts, and persistent inflationary pressures, these relationships are likely to be nonlinear and subject to gradual structural changes over time. To address these features, the study employs the Fourier Toda–Yamamoto causality approach, which allows for smooth structural shifts without requiring prior knowledge of their timing or form. By combining a functional expenditure perspective with an advanced causality framework, this study aims to contribute to the literature by providing a more nuanced understanding of how the composition of public spending shapes inflation dynamics in Türkiye. In this respect, the study contributes to the literature in two main ways. First, it provides a comprehensive analysis of the functional composition of public expenditures based on the COFOG framework within a unified empirical setting. Second, by incorporating smooth structural shifts through the Fourier Toda–Yamamoto approach, the study offers a more flexible framework for examining fiscal–inflation dynamics in an economy characterized by recurring inflationary regimes such as Türkiye.

Literature Review

The empirical literature examining the relationship between public expenditures and inflation is diverse in terms of country coverage, methodological approaches, and the classification of government spending. While a substantial body of research focuses on aggregate public expenditure or specific spending categories, fewer studies employ a functional classification framework that enables a systematic comparison across different public expenditure components. To provide a structured overview of the existing evidence, Table 1 summarizes the key studies that investigate the links between public spending, its composition, and inflation dynamics.

Table 1

Summary of the Literature on Public Expenditure Components and Inflation

Authors (Year)	Country	Expenditure Component(s)	Methodology	Main Findings
Anand & van Wijnbergen (1989)	Türkiye	Aggregate public expenditure	Analytical & empirical	Fiscal financing of public spending plays a significant role in inflation dynamics.
De Haan & Zelhorst (1990)	Developing countries	Government expenditure	Cross-country regression	Expansionary fiscal policy contributes to inflation persistence.
Barro (1990)	Cross-country	Productive public spending	Endogenous growth model	Productive spending enhances supply capacity and may reduce inflationary pressures.
Bleaney (1996)	Developing countries	Government consumption	Panel analysis	Government consumption growth is positively related to inflation.
Ruge-Murcia (1999)	Brazil	Government expenditure (including interest payments), budget deficit, seigniorage-financed spending	Markov-switching regime model	Government expenditure is the main driver of high inflation; shifts to high-spending regimes significantly increase inflation and money growth, while stabilization programs without fiscal reform fail to curb inflation.
Alavirad & Athawale (2005)	Iran	Government expenditure	Time series analysis	Long-run relationship exists between government spending and inflation.

Catão & Terrones (2005)	Cross-country	Fiscal deficits & public spending	Panel econometrics	Fiscal imbalances are a key determinant of inflation in developing countries.
Olubusoye & Oyaromade (2008)	Nigeria	Government expenditure	VAR & Granger causality	Government expenditure Granger-causes inflation in the long run.
Afonso & Jalles (2012)	OECD & EU countries	Government spending	Panel regressions	Public spending may generate inflationary pressures depending on regime.
Tiwari, A.K., Tiwari, A.P., & Pandey (2012)	India	Total government expenditure, fiscal deficit, money supply (M3)	Dolado–Lütkepohl (1996) augmented VAR Granger causality test; standard Granger causality analysis	Findings vary by method; government expenditure (and partly money supply) drives fiscal deficit, while inflation shows no causal role in India.
Lin & Chu (2013)	91 countries	Public debt & expenditure	Panel threshold model	Inflationary effects of fiscal policy depend on spending and debt levels.
Nguyen (2015)	Asian economies	Fiscal deficit, government expenditure, money supply (M2)	Panel GMM	Fiscal deficit, government expenditure, and interest rates are robust determinants of inflation, while the impact of money M2 is significant only under PMG estimation.
Akobi, Umeora & Atueyi (2021)	Nigeria	Agriculture, education, health, and telecommunications expenditures	Johansen cointegration and ECM	Health and telecommunications expenditures have significant positive effects on inflation, while education and agriculture expenditures show positive but insignificant effects, indicating heterogeneous inflationary impacts of public expenditure components.
Sequeira (2021)	Cross-country	Education expenditure	Panel econometrics	Education spending influences macroeconomic performance, including inflation indirectly.
Barro & Bianchi (2023)	OECD countries	Fiscal expenditure composition	Panel macroeconomic analysis	Fiscal policy and expenditure composition play a significant role in post-pandemic inflation dynamics across OECD countries.
Klein & Linnemann (2023)	OECD countries	Spending composition	VAR analysis	Inflationary effects of fiscal policy depend on expenditure composition.
Apeti et al. (2023)	Developing countries	Public spending composition	Panel regressions	Composition of public expenditure matters for inflation targeting performance.

Raza, Laurentjoye, Byrialsen & Valdecantos (2023)	Denmark	Fiscal and macroeconomic policy channels (aggregate public spending)	Stock-flow consistent macroeconomic model	Coordinated fiscal and monetary policies play a crucial role in mitigating inflationary pressures, indicating that fiscal policy contributes to inflation dynamics through aggregate demand and income channels.
Serin & Demir (2024)	Türkiye	Fiscal policy & inflation	Fourier-ADL & Fourier TY	Fiscal regime changes affect inflation dynamics.
Tuncer (2024)	Türkiye	Total public expenditure	ARDL & TY causality	Evidence of causality between public spending and inflation.
Diğer & Uyar (2025)	Türkiye	Health expenditures (COFOG – Health)	Wavelet coherence analysis	The relationship between health expenditures and inflation exhibits heterogeneous interdependence across provinces and time horizons, indicating that sector-specific public expenditures generate differentiated inflationary effects.
Özmen Bol & Cantürk (2025)	Türkiye	Health expenditure	Descriptive & policy analysis	High inflation erodes real value of health spending.
Cevik & Miryugin (2024)	140 countries (panel)	Fiscal policy shocks	Panel econometric analysis	Expansionary fiscal policy shocks generate inflationary pressures in the short and medium term, particularly in developing economies.

Note. Created by the author.

The studies summarized in Table 1 highlight that the relationship between fiscal policy and inflation is far from uniform. Existing research suggests that the inflationary consequences of government spending depend not only on the magnitude of public expenditures but also on their functional composition. Some categories of spending may generate short-run demand pressures, while others may contribute to productive capacity and long-run macroeconomic stability. These findings indicate that different components of public expenditures may exhibit heterogeneous interactions with inflation dynamics.

Building on these theoretical insights and empirical findings, the present study formulates the following hypotheses regarding the relationship between inflation and the functional composition of public expenditures:

H1: The functional composition of public expenditures has a significant relationship with inflation dynamics.

H2: Different categories of public expenditures exhibit heterogeneous causal relationships with inflation.

H3: Structural shifts in fiscal policy regimes influence the causal interaction between public expenditures and inflation.

Econometric Methodology

This section outlines the econometric procedures employed to investigate the dynamic relationships between public expenditure components and inflation in Türkiye. The analysis proceeds in two stages. First, the integration properties of the variables are examined using both

conventional and Fourier-based unit root tests. Second, the direction of causality among the variables is analyzed using the Fourier Toda–Yamamoto causality framework, which accounts for potential structural changes in the data-generating process.

Unit Root Tests: ADF and Fourier ADF

Standard unit root tests such as the Augmented Dickey–Fuller (1979) (ADF) and Phillips–Perron (1988) (PP) tests are widely used to assess the stochastic properties of macroeconomic time series. However, these conventional tests do not explicitly consider structural changes in the deterministic components of the series. When such changes are present, the power of traditional unit root tests may be substantially reduced, leading to biased inferences regarding stationarity. Perron (1989) highlighted that ignoring structural breaks—particularly in the intercept or trend—can distort unit root test outcomes.

Subsequent studies developed unit root testing procedures that allow structural breaks to be determined endogenously. Prominent examples include Zivot and Andrews (1992), Perron (1994), Lumsdaine and Papell (1997), and Lee and Strazicich (2003, 2013), which permit one or multiple sharp breaks in the deterministic structure of the series. While these approaches represent important methodological advances, they require assumptions regarding the number, timing, or form of structural breaks, which may limit their applicability in practice.

To address this limitation, Becker, Enders, and Lee (2004, 2006) proposed a flexible framework based on Fourier approximations that captures structural changes without requiring prior knowledge of their number or timing. Building on this approach, Enders and Lee (2012) extended the conventional ADF test by incorporating low-frequency sine and cosine terms into the deterministic component, resulting in the Fourier Augmented Dickey–Fuller (FADF) unit root test. This method is particularly well-suited for modeling smooth and gradual structural changes commonly observed in macroeconomic series.

In the Fourier ADF framework, the deterministic component is specified as a time-varying function:

$$\alpha(t) = \alpha_0 + \gamma_1 \sin\left(\frac{2\pi kt}{T}\right) + \gamma_2 \cos\left(\frac{2\pi kt}{T}\right), \tag{1}$$

where k denotes the Fourier frequency, t is the time index, and T represents the sample size. The sine and cosine terms flexibly approximate unknown structural changes in the intercept and trend of the series.

By incorporating this deterministic structure into the ADF regression, the Fourier ADF test equation can be written as (Enders & Lee, 2012):

$$\Delta y_t = \alpha_1 + \delta t + \beta y_{t-1} + \gamma_1 \sin\left(\frac{2\pi kt}{T}\right) + \gamma_2 \cos\left(\frac{2\pi kt}{T}\right) + \sum_{i=1}^{p} \phi_i \, \Delta y_{t-i} + u_t. \tag{2}$$

In practice, the Fourier ADF model is estimated over a range of frequencies, typically $1 \leq k \leq 5$, and the optimal frequency is selected based on the minimum residual sum of squares. If the Fourier terms are jointly significant, unit root inference relies on the FADF test statistics. Otherwise, the conventional ADF results are considered more appropriate. In this study, both ADF and FADF tests are employed to ensure robust identification of the integration properties of the variables by accounting for linear dynamics as well as potential smooth structural changes.

Causality Analysis: Fourier Toda–Yamamoto Approach

After establishing the integration properties of the variables, the causal relationships among public expenditure components and inflation are examined using the Fourier Toda–Yamamoto (TY) causality test. Conventional Granger causality tests may produce unreliable results when variables are non-stationary or integrated of different orders, as the associated Wald statistics may not follow their standard asymptotic distributions. These issues are further exacerbated in the presence of structural changes. To overcome such limitations, this study adopts the causality testing framework proposed by Toda and Yamamoto (1995).

The TY approach conducts causality testing within a vector autoregressive (VAR) model estimated in levels, regardless of whether the variables are stationary or cointegrated. The method involves estimating an augmented VAR model of order $p + d_{\max}$, where p denotes the optimal lag length selected by information criteria and $d_{\max}$ represents the maximum integration order among the variables. By augmenting the VAR system with additional lags equal to $d_{\max}$, the TY procedure ensures the validity of Wald test statistics used for causal inference.

Despite its advantages, the standard TY test does not explicitly account for structural changes in the deterministic components of the model. To address this shortcoming, Nazlıoğlu et al. (2016) proposed a Fourier version of the TY causality test, which incorporates smooth structural shifts using low-frequency trigonometric terms. In this context, the use of the Fourier approximation provides a flexible way of capturing smooth structural changes without requiring prior identification of break dates. This feature is particularly useful for economies such as Türkiye where macroeconomic policies and inflation regimes have experienced gradual shifts over time. The deterministic component of the model is specified as:

$$Z_t = \alpha_0 + \alpha_1 t + \alpha_2 \sin\left(\frac{2\pi k t}{T}\right) + \alpha_3 \cos\left(\frac{2\pi k t}{T}\right), \tag{3}$$

where k denotes the Fourier frequency and T is the sample size. These terms allow the model to approximate unknown forms of gradual structural change without requiring explicit identification of break dates.

Incorporating the Fourier terms, the resulting Fourier VAR model of order $p + d_{\max}$ is given by:

$$y_t = \alpha_0 + \alpha_1 t + \alpha_2 \sin\left(\frac{2\pi k t}{T}\right) + \alpha_3 \cos\left(\frac{2\pi k t}{T}\right) + \sum_{i=1}^{p+d_{\max}} J_i \, y_{t-i} + \varepsilon_t, \tag{4}$$

where y_t denotes the vector of endogenous variables, J_i represents coefficient matrices, and ε_t is a vector of independently and identically distributed error terms.

A key step in implementing the Fourier TY causality test is the joint selection of the optimal lag length p and Fourier frequency k. Following Nazlıoğlu et al. (2016), this study employs the Schwarz Information Criterion (SIC) to determine the parameter combination that minimizes information loss while maintaining model parsimony.

The dataset consists of quarterly observations covering the period 2006Q1–2025Q2, yielding 78 observations. Although this sample size is sufficient for VAR-based inference, the Turkish economy experienced multiple policy shifts and inflation regimes during the sample period. In this context, the Fourier TY approach is particularly suitable, as it combines robustness to mixed integration orders with the ability to capture smooth, time-varying structural changes. Accordingly, this framework provides a reliable empirical strategy for

analyzing the directional predictability between the functional composition of public expenditures and inflation dynamics in Türkiye.

Data Set and Descriptive Statistics

This study employs quarterly data covering the period 2006Q1–2025Q2 to examine the relationship between the functional composition of public expenditures and inflation dynamics in Türkiye. The sample period is determined by data availability and ensures consistency across all variables included in the analysis. All variables are obtained from official public institutions, ensuring data reliability and transparency.

Inflation is derived from the Consumer Price Index (CPI) obtained from the Electronic Data Delivery System (EVDS) of the Central Bank of the Republic of Türkiye. Although the CPI is available at a monthly frequency, inflation is computed as the year-on-year logarithmic change in the CPI and subsequently converted to quarterly frequency by selecting end-of-quarter observations. Public expenditure variables are obtained from the Ministry of Treasury and Finance of the Republic of Türkiye, published by the General Directorate of Public Accounts, and are expressed as ratios to Gross Domestic Product (GDP).

To ensure clarity and consistency throughout the empirical analysis, all variables are denoted using abbreviated notation. Variable definitions, transformations, and data sources are summarized in Table 2.

Table 2

Variable Definitions, Notation, and Data Sources

Notation	Variable	Definition / Measurement	Source
INF	Inflation	Year-on-year CPI inflation (quarterly)	Electronic Data Delivery System (EVDS)
GEN	General Public Services	General public services expenditures / GDP	Ministry of Treasury and Finance
ECO	Economic Affairs	Economic affairs and services expenditures / GDP	Ministry of Treasury and Finance
HEA	Health Services	Health expenditures / GDP	Ministry of Treasury and Finance
EDU	Education Services	Education expenditures / GDP	Ministry of Treasury and Finance
SOC	Social Protection	Social security and social assistance expenditures / GDP	Ministry of Treasury and Finance

Note. Created by the author.

All public expenditure variables are included in the analysis in level form, as they represent proportional fiscal measures. Inflation (INF) is modeled in growth-rate form. This specification is consistent with the econometric framework adopted in the study and preserves the economic interpretation of each variable.

The relationship between inflation and the functional composition of public expenditures is summarized by the following general functional form:

$$INF_t = f(GEN_t, ECO_t, HEA_t, EDU_t, SOC_t)$$

Based on this functional relationship, the corresponding econometric model can be expressed as:

$$INF_t = \alpha_0 + \alpha_1 GEN_t + \alpha_2 ECO_t + \alpha_3 HEA_t + \alpha_4 EDU_t + \alpha_5 SOC_t + \varepsilon_t$$

where, α_0 is the constant term, α_1–α_5 represent the coefficients associated with each expenditure category, and ε_t denotes the error term.

Descriptive Statistics

This subsection presents a brief overview of the statistical properties of the variables used in the analysis. Table 3 summarizes the descriptive statistics for inflation and functional public expenditure components (ECO, EDU, GEN, HEA, and SOC) over the period 2006Q1–2025Q2, based on 78 quarterly observations. The statistics indicate substantial volatility in inflation relative to public expenditure variables, reflecting Türkiye's pronounced inflationary dynamics during the sample period. While expenditure shares generally exhibit smoother behavior, noticeable heterogeneity is observed across functional categories, with social protection and general public services displaying comparatively higher variability. The distributional characteristics further suggest deviations from normality in all variables except the education variable, underscoring the relevance of econometric methods capable of accommodating non-linear dynamics and potential structural changes in the subsequent analysis.

Table 3

Descriptive Statistics

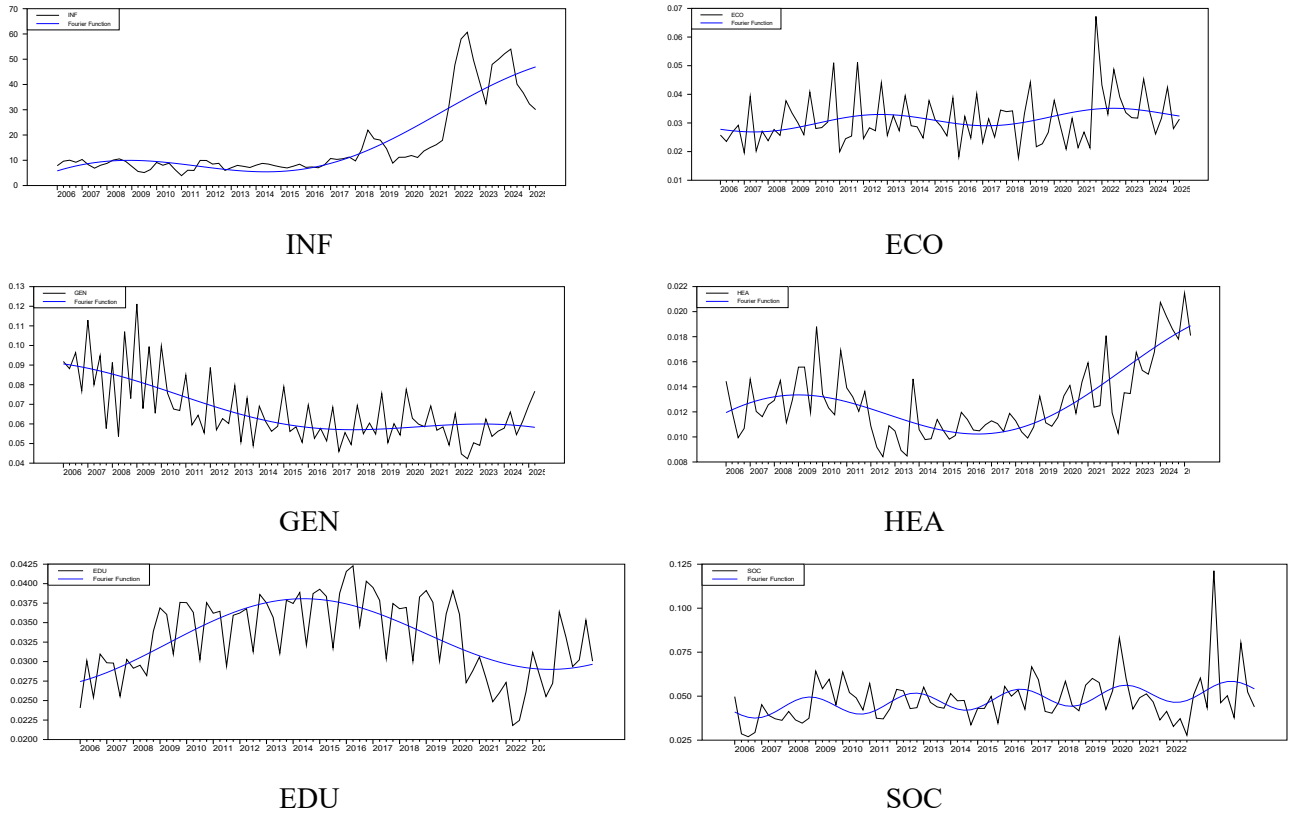
Statistic	INF	ECO	EDU	GEN	HEA	SOC
Mean	16.261	0.031	0.033	0.067	0.013	0.048
Maximum	60.679	0.067	0.042	0.121	0.021	0.121
Minimum	3.909	0.018	0.022	0.042	0.008	0.027
Std. Dev.	14.736	0.0087	0.0050	0.0165	0.0029	0.0135
Skewness	1.716	1.317	−0.269	1.212	1.012	2.336
Kurtosis	4.599	5.712	2.013	4.127	3.494	12.980
Jarque–Bera	46.607	46.454	4.109	23.231	14.097	394.612
Probability	0.000	0.000	0.128	0.000	0.001	0.000

INF denotes quarterly year-on-year CPI inflation. ECO, EDU, GEN, HEA, and SOC represent functional public expenditure components expressed as ratios to GDP.

Note. Created by the author using Eviews 12 econometric software.

Figure 1

Time path graph of the Variables



INF ECO

GEN HEA

EDU SOC

Note. Created by the author in WinRATS; Arial font customization is not supported.

Unit Root Test Results

The stationarity properties of the variables included in the analysis, namely inflation and the functional components of public expenditures (GEN, ECO, HEA, EDU, and SOC), are examined using two complementary unit root testing procedures. First, the conventional Augmented Dickey–Fuller (1979) (ADF) test is employed to assess stationarity under standard deterministic specifications. Second, the Fourier Augmented Dickey–Fuller (FADF) unit root test, proposed by Enders and Lee (2012), is applied to account for potential smooth structural changes in the data-generating process. The results obtained from the ADF and FADF tests are reported in Table 4 and Table 5, respectively.

Table 4

ADF Unit Root Test Results

	Constant		Constant and Trend	
Variables	**Test Statistic**	**5% Critical Value**	**Test Statistic**	**5% Critical Value**
INF	-1.055 (5)	-2.902	-2.128 (5)	-3.473
GEN	-1.909 (3)	-2.901	-1.305 (3)	-3.471
ECO	-2.864 (3)	-2.901	-3.093 (3)	-3.471

HEA	0.201 (3)	-2.901	-0.314 (3)	-3.471
EDU	-1.943 (4)	-2.901	-2.140 (4)	-3.472
SOC	-7.525 (0)	-2.899	-8.112** (0)	-3.469
ΔINF	-3.612** (1)	-2.902	-3.583** (4)	-3.471
ΔGEN	-10.866**(2)	-2.901	-11.029**(2)	-3.471
ΔECO	-13.213**(2)	-2.901	-13.125**(2)	-3.471
ΔHEA	-10.488**(2)	-2.901	-10.674**(2)	-3.471
ΔEDU	-3.301**(3)	-2.901	-3.321**(3)	-3.471
ΔSOC	-10.127** (2)	-2.901	-10.075**(2)	-3.471

** indicates significance at 5% significance level. Values in parentheses indicate the appropriate lag length determined according to Schwarz information criteria.

Note. Created by the author using Eviews 12 econometric software.

Table 5

Fourier ADF Unit Root Test Results

Models	**Constant**			**Constant and Trend**		
Variables	**k**	**FADF stat.**	**F (Fourier)**	**k**	**FADF stat.**	**F (Fourier)**
INF	1	-3.016	3.031	4	-3.152	1.643
GEN	1	-3.559*	2.971	1	-4.421**	4.859
ECO	2	-10.227***	3.349	2	-10.373***	2.041
HEA	2	-5.452***	6.576*	1	-7.709***	16.176***
EDU	1	-5.542***	9.060**	1	-5.564***	8.707**
SOC	5	-8.136***	3.607	5	-8.666***	3.376

Note. Created by the author using WinRATS econometric software.

FADF denotes the Fourier Dickey–Fuller test statistic. k represents the selected Fourier frequency that minimizes the residual sum of squares. The F (Fourier) statistic tests the joint significance of the sine and cosine terms. Critical values for the FADF test statistics are obtained from Enders and Lee (2012) and depend on the deterministic specification (constant or constant and trend) and the chosen Fourier frequency. For models with a constant, critical values are taken from Table 1b in Enders and Lee (2012), while for models with a constant and trend, critical values are taken from Table 1a. Values in parentheses indicate the optimal lag length selected according to the (SIC). ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 4 and Table 5 report the results of the conventional Augmented Dickey–Fuller (1979) (ADF) and the Fourier Augmented Dickey–Fuller (FADF) unit root tests, respectively. According to the ADF test results, the inflation variable (INF) and most public expenditure components (GEN, ECO, HEA, and EDU) fail to reject the null hypothesis of a unit root at levels under both constant and constant-with-trend specifications, whereas social protection expenditures (SOC) appear to be stationary when a deterministic trend is included. However, all variables become stationary after first differencing, indicating that they are predominantly integrated of order one, I(1).

The FADF test results largely corroborate the findings of the conventional ADF test while additionally accounting for smooth structural changes in the data-generating process. The joint significance of the Fourier sine and cosine terms is confirmed for several variables, particularly ECO, HEA, EDU, and SOC, implying the presence of smooth structural shifts. Based on the FADF test statistics and the corresponding critical values reported by Enders and Lee (2012), GEN, ECO, HEA, EDU, and SOC are found to be stationary at levels under at least one deterministic specification, while INF remains non-stationary in level form. Overall, the combined evidence from the ADF and FADF tests suggests a mixed integration structure, with variables being either I(0) or I(1), thereby justifying the use of the Fourier Toda–Yamamoto causality framework in the subsequent analysis.

The unit root test results indicate that the variables are integrated of different orders, with most series being I(1) while some exhibit stationarity at levels when structural changes are taken into account. This mixed integration structure renders conventional cointegration-based causality tests inappropriate. However, the Toda–Yamamoto causality framework and its Fourier-augmented extension allow for valid inference regardless of the integration properties of the variables. Accordingly, the causal relationships between inflation and the functional components of public expenditures are examined using the Fourier Toda–Yamamoto causality test.

Fourier Toda–Yamamoto Causality Results

As discussed in the previous subsection, the divergence between the conventional ADF and Fourier ADF unit root test results stems from their distinct methodological foundations. While the standard ADF test assumes a stable deterministic structure over time, the Fourier ADF framework explicitly allows for smooth and gradual structural changes through low-frequency trigonometric terms. This feature is particularly important in the context of the Turkish economy, where policy regimes and inflation dynamics have undergone notable transformations over the sample period.

To avoid potential pre-testing bias and ensure valid inference in the presence of variables with mixed integration orders, the maximum order of integration among the variables is set to $d_{max} = 1$. Given that the Fourier terms are found to be statistically significant for several series, indicating the presence of smooth structural shifts, the (Fourier TY) causality approach is adopted in the subsequent analysis.

Prior to implementing the Fourier TY causality test, the optimal lag length of the underlying VAR model is determined using standard information criteria. Following the principle of parsimony and placing particular emphasis on the Schwarz Information Criterion—commonly preferred in relatively moderate sample sizes—the optimal lag length is selected as $p = 2$. Accordingly, an augmented VAR($p + d_{max}$) = VAR(3) model is estimated in levels. This specification ensures that the Wald-type test statistics used for causality inference retain their standard asymptotic properties, regardless of the integration or cointegration characteristics of the variables.

Within this framework, the Fourier TY causality test is conducted by incorporating sine and cosine terms into the deterministic component of the VAR model to capture smooth structural changes without requiring prior knowledge of their timing, number, or functional form. By combining the robustness of the Toda–Yamamoto procedure with the flexibility of the Fourier approximation, this approach provides a reliable empirical strategy for examining the directional relationships between inflation and the functional components of public expenditures in Türkiye.

Table 6

Toda–Yamamoto and Fourier Toda–Yamamoto Causality Test Results

	TY		Fourier TY		
Variables	**Wald Stat.**	**prob.**	**k**	**Wald Stat.**	**prob.**
ECO → INF	7.834387	0.021	1	9.039234	0.002
INF → ECO	4.316515	0.116	1	7.32622	0.006
GEN → INF	25.13544	0.000	1	4.037685	0.044
INF → GEN	2.043379	0.365	1	1.114889	0.291
HEA → INF	9.254539	0.009	1	5.704799	0.016
INF → HEA	32.11856	0.000	1	0.985036	0.321
EDU → INF	22.09303	0.000	1	5.839843	0.016
INF → EDU	5.770043	0.056	1	4.099142	0.043
SOC → INF	19.87962	0.000	1	7.446898	0.006
INF → SOC	12.98707	0.001	1	6.486484	0.011

Note. Created by the author using Eviews 12 econometric software.

Table 6 reports the causality results obtained from both the standard Toda–Yamamoto (TY) and the Fourier Toda–Yamamoto (Fourier TY) approaches. The comparison of the two methodologies reveals that accounting for smooth structural changes substantially enriches the causal structure between inflation and the functional components of public expenditures in Türkiye. These results indicate that the effects of public spending on inflation differ across expenditure categories and that, for certain components, the inflation–expenditure relationship involves feedback mechanisms. These findings highlight that the inflation–expenditure relationship is heterogeneous across expenditure categories, with some components exhibiting feedback mechanisms while others reflect a more direct fiscal influence on inflation dynamics.

The conventional TY results indicate unidirectional causality running from economic affairs, general public services, health, education, and social protection expenditures to inflation, while reverse causality from inflation is detected only for selected spending categories. However, once smooth structural shifts are incorporated through the Fourier TY framework, the causal relationships become more pronounced and, in several cases, bidirectional.

In particular, the Fourier TY results reveal bidirectional causality between inflation and economic affairs, education, and social protection expenditures, suggesting the presence of feedback mechanisms in which inflation both influences and is influenced by these spending components. By contrast, causality from general public services and health expenditures to inflation remains predominantly unidirectional, implying that these categories exert inflationary effects mainly through structural and demand-driven channels rather than short-run feedback adjustments.

Overall, the findings demonstrate that ignoring structural changes may conceal important causal linkages in fiscal–inflation dynamics. The Fourier TY approach uncovers more robust causal relationships compared to the standard TY test, highlighting the critical role of public expenditure composition in shaping inflation dynamics in Türkiye. It should be noted that the Toda–Yamamoto framework identifies directional predictability rather than the magnitude or persistence of effects. Therefore, the results should be interpreted as evidence of causal linkages rather than precise quantitative estimates of fiscal impacts on inflation.

Conclusion

This study examined the causal links between inflation and the functional composition of public expenditures in Türkiye over the period 2006Q1–2025Q2 by using COFOG-based expenditure shares in GDP. Motivated by the Fiscal Theory of the Price Level perspective and the possibility that Türkiye's inflation process is shaped not only by monetary factors but also by fiscal regimes and expenditure allocation, the analysis employed a two-stage econometric strategy. First, the integration properties of the series were assessed through conventional ADF and Fourier ADF unit root tests. Second, to obtain valid causal inference under mixed integration orders and potential smooth structural shifts, the study implemented the Fourier-augmented Toda–Yamamoto causality framework.

The unit root evidence confirmed that the variables exhibit a mixed integration structure. While the conventional ADF results suggested that most variables are I(1), the Fourier ADF test indicated that several expenditure components display stationarity in levels once smooth structural changes are accounted for, whereas inflation remains non-stationary at levels. This divergence is consistent with the methodological insight that ignoring structural change may distort stationarity inference in macroeconomic time series. Given this mixed integration environment and the presence of regime changes in Türkiye's macroeconomic history, the Fourier Toda–Yamamoto approach provided an appropriate and robust setting for causality analysis.

The causality findings underscore that the inflation–fiscal nexus in Türkiye is not homogeneous across expenditure functions and becomes clearer once smooth structural shifts are incorporated. The Fourier TY results indicate that inflation is significantly affected by economic affairs, general public services, health, education, and social protection expenditures. Importantly, the causal structure differs across categories. For general public services and health, causality runs predominantly from expenditures to inflation, implying that these items shape inflation dynamics mainly through expenditure-side or structural channels rather than short-run feedback adjustments. In contrast, economic affairs, education, and social protection display bidirectional causality with inflation, suggesting feedback mechanisms whereby these spending components both influence inflation and adjust endogenously to inflationary conditions.

These results contribute to the broader literature in several ways. First, the finding that multiple public spending components Granger-cause inflation aligns with the long-standing view that fiscal policy is a key driver of inflation in developing and high-inflation settings (e.g., the fiscal-financing emphasis in Türkiye in Anand & van Wijnbergen, 1989; and the broader evidence on fiscal imbalances and inflation in Catão & Terrones, 2005). Second, the bidirectional relationships uncovered for social protection and education are consistent with the idea that certain expenditure categories operate as both policy instruments and automatic or quasi-automatic responses to macroeconomic stress—particularly in economies where inflation redistributes income and triggers compensatory fiscal measures. This feedback structure complements the panel evidence that inflationary consequences depend on fiscal composition and regime features (Afonso & Jalles, 2012; Lin & Chu, 2013) and is in line with more recent arguments that the inflationary impact of fiscal policy depends critically on how spending is allocated across functions (Klein & Linnemann, 2023; Apeti et al., 2023). Third, the stronger causal patterns obtained under Fourier TY relative to the standard TY test reinforce the methodological point that structural changes can conceal key fiscal–inflation linkages—an issue that is particularly relevant for Türkiye given multiple policy shifts and inflation regimes within the sample period. In this sense, the evidence also resonates with recent Türkiye-focused

studies emphasizing the importance of fiscal regime changes and nonlinear dynamics (Serin & Demir, 2024; Tuncer, 2024).

From a policy perspective, the results imply that the effectiveness of fiscal policy in supporting price stability depends not only on the size of public spending but also on its functional composition. The unidirectional effects from general public services to inflation suggest that expenditure items closely related to administration and debt-related services may exert inflationary pressure through demand and fiscal-sustainability channels, highlighting the importance of expenditure discipline and credibility. At the same time, the bidirectional relationships for economic affairs, education, and social protection point to a more complex interaction: inflation can trigger fiscal adjustments in these categories, while changes in these expenditures can feed back into inflation dynamics. Accordingly, expenditure reallocation strategies aimed at strengthening productive capacity and limiting inflationary pressures should take into account this feedback nature and the timing of fiscal responses under high inflation.

Beyond its empirical findings, the study contributes to policy discussions by emphasizing that inflation control strategies cannot be designed solely around aggregate fiscal contraction. Instead, policymakers should consider the functional allocation of public expenditures and their heterogeneous transmission mechanisms. In this respect, effective inflation management in Türkiye requires coordinated fiscal–monetary frameworks in which expenditure composition is aligned with price stability objectives.

Finally, the study has limitations that open avenues for further research. While the Fourier TY framework provides robust evidence on directional predictability under smooth structural shifts, causality tests do not quantify the magnitude of effects or distinguish short- and long-run transmission channels. Future work may complement these results with structural VAR/FAVAR approaches, local projections, or regime-dependent models that explicitly identify fiscal shocks and quantify dynamic impulse responses by expenditure function. Extending the analysis to alternative inflation measures and incorporating financing-side variables (e.g., debt dynamics, tax revenues, or monetary accommodation) may further refine the interpretation within the FTPL context.

Overall, the evidence presented here supports the view that in Türkiye inflation dynamics are closely linked to the functional composition of public expenditures and that accounting for structural change is essential to uncover the underlying causal structure. The findings highlight that compositional fiscal policy—rather than aggregate spending alone—should be treated as a core element of policy design in pursuing durable price stability. The conclusion now includes more detailed policy recommendations, focusing on the importance of fiscal–monetary coordination and differentiating between efficient and inefficient public expenditures for effective inflation control.

References

Afonso, A., & Jalles, J. T. (2012). Measuring the success of fiscal consolidations. *Applied Economics, 44*(9), 1053–1061. https://doi.org/10.1080/09603107.2011.637894

Akobi, K. C., Umeora, C. E., & Atueyi, C. L. (2021). Government expenditure and inflation rate in Nigeria. *International Journal of Business Systems and Economics*, *13*(4), 276-293. https://www.arcnjournals.org/images/ARCN-IJBSE-13-4-18.pdf

Alavirad, A., & Athawale, S. (2005). The impact of the budget deficit on inflation in the Islamic Republic of Iran. *OPEC Review, 29*(1), 37–49. https://doi.org/10.1111/j.0277-0180.2005.00142.x

Aghion, P., & Howitt, P. (1998). *Endogenous growth theory*. MIT Press.

Anand, R., & van Wijnbergen, S. (1989). Inflation and the financing of government expenditure: An introductory analysis with an application to Turkey. *The World Bank Economic Review, 3*(1), 17–38. https://doi.org/10.1093/wber/3.1.17

Apeti, A. K., Combes, J.-L., & Minea, A. (2023). Inflation targeting and the composition of public expenditure: Evidence from developing countries. *Journal of Macroeconomics, 76*, Article 103523. https://doi.org/10.1016/j.jmacro.2023.103523

Barro, R. J. (1990). Government spending in a simple model of endogenous growth. *Journal of Political Economy, 98*(5), 103–125. https://doi.org/10.1086/261726

Barro, R. J., & Bianchi, F. (2023, November). *Fiscal influences on inflation in OECD countries, 2020–2023* (NBER Working Paper No. 31838). National Bureau of Economic Research. https://doi.org/10.3386/w31838

Becker, R., Enders, W., & Lee, J. (2004). A general test for time dependence in parameters. *Journal of Applied Econometrics, 19*(7), 899–906. https://doi.org/10.1002/jae.787

Becker, R., Enders, W., & Lee, J. (2006). A stationarity test in the presence of an unknown number of smooth breaks. *Journal of Time Series Analysis, 27*(3), 381–409. https://doi.org/10.1111/j.1467-9892.2006.00478.x

Bleaney, M. (1996). Central bank independence, wage-bargaining structure, and macroeconomic performance in OECD countries. *Oxford Economic Papers, 48*(1), 20–38. https://doi.org/10.1093/oxfordjournals.oep.a028559

Catão, L. A. V., & Terrones, M. E. (2005). Fiscal deficits and inflation. *Journal of Monetary Economics, 52*(3), 529–554. https://doi.org/10.1016/j.jmoneco.2004.06.003

Cevik, S., Miryugin, F. (2025). It's never different: Fiscal policy shocks and inflation. *Comparative Economic Studies, 67*, 186–220 https://doi.org/10.1057/s41294-024-00238-w

De Haan, J., & Zelhorst, D. (1990). The impact of government deficits on money growth in developing countries. *Journal of International Money and Finance, 9*(4), 455–469. https://doi.org/10.1016/0261-5606(90)90022-R

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association, 74*(366), 427–431. https://doi.org/10.1080/01621459.1979.10482531

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

Diğer, H., & Uyar, U. (2025). The relationship between health expenditures and ınflation by provinces in Türkiye: Wavelet coherence analysis. *Journal of Mehmet Akif Ersoy University Economics and Administrative Sciences Faculty*, *12*(1), 271-292. https://doi.org/10.30798/makuiibf.1505841

Dolls, M., Fuest, C., & Peichl, A. (2009). Automatic stabilizers and economic crisis: US vs. Europe. *Journal of Public Economics, 96*(3–4), 279–294. https://doi.org/10.1016/j.jpubeco.2011.11.001

Ekinci, M. F., Saygılı, H., & Yılmaz, H. H. (2025). Fiscal policy and long-run inflation dynamics: Evidence from a major emerging market economy. *Applied Economics*, 1-15. https://doi.org/10.1080/00036846.2025.2532881

Enders, W., & Lee, J. (2012). The flexible Fourier form and Dickey–Fuller type unit root tests. *Economics Letters, 117*(1), 196–199. https://doi.org/10.1016/j.econlet.2012.04.081

Eurostat. (2019). *Government expenditure by function (COFOG).* Retrieved December 25, 2025, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Government_expenditure_by_function_(COFOG)

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy, 80*(2), 223–255. https://doi.org/10.1086/259880

Klein, M., & Linnemann, L. (2023). The composition of public spending and inflationary effects. *Journal of Public Economics,* 219, Article 104836. https://doi.org/10.1016/j.jpubeco.2023.104836

Leeper, E. M. (1991). Equilibria under "active" and "passive" monetary and fiscal policies. *Journal of Monetary Economics, 27*(1), 129–147. https://doi.org/10.1016/0304-3932(91)90007-B

Lee, J., & Strazicich, M. C. (2003). Minimum Lagrange multiplier unit root test with two structural breaks. *Review of Economics and Statistics, 85*(4), 1082–1089. https://doi.org/10.1162/003465303772815961

Lee, J., & Strazicich, M. C. (2013). Minimum LM unit root test with one structural break. *Economics Bulletin, 33*(4), 2483–2492.

Lin, H.-Y., & Chu, H.-P. (2013). Are fiscal deficits inflationary? *Journal of International Money and Finance,* 32, 214–233. https://doi.org/10.1016/j.jimonfin.2012.04.006

Lumsdaine, R. L., & Papell, D. H. (1997). Multiple trend breaks and the unit-root hypothesis. *Review of Economics and Statistics, 79*(2), 212–218. https://doi.org/10.1162/003465397556791

Musgrave, R. A. (1959). *The theory of public finance: A study in public economy*. McGraw-Hill.

Nazlıoğlu, Ş., Gormus, N. A., & Soytaş, U. (2016). Oil prices and real estate investment trusts (REITs): Gradual-shift causality and volatility transmission analysis. *Energy Economics, 60*, 168–175. https://doi.org/10.1016/j.eneco.2016.09.009

Nguyen, V. B. (2015). Effects of fiscal deficit and money M2 supply on inflation: Evidence from selected economies of Asia. *Journal of Economics, Finance and Administrative Science*, *20*(38), 49-53. https://hdl.handle.net/10419/179760

Olubusoye, O. E., & Oyaromade, R. (2008, August). *Modelling the inflation process in Nigeria*. African Economic Research Consortium. Working Paper No. 182, https://aercafrica.org

Özmen Bol, Z., & Cantürk, S. (2025). Changes in health expenditure relative to inflation: Prospective implications for Türkiye. *Uluslararası İktisadi ve İdari İncelemeler Dergisi*, (49), 37-50. https://doi.org/10.18092/ulikidince.1608149

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica, 57*(6), 1361–1401. https://doi.org/10.2307/1913712

Perron, P. (1994). Trend, unit root and structural change in macroeconomic time series. In B. B. Rao (Ed.), *Cointegration for the Applied Economist* (pp. 113–146). St. Martin's Press.

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

Raza, H., Laurentjoye, T., Byrialsen, M. R., & Valdecantos, S. (2023). Inflation and the role of macroeconomic policies: A model for the case of Denmark. *Structural Change and Economic Dynamics*, *67*, 32-43. https://doi.org/10.1016/j.strueco.2023.06.006

Ruge-Murcia, F. J. (1999). Government expenditure and the dynamics of high inflation. *Journal of Development Economics*, *58*(2), 333-358. https://doi.org/10.1016/S0304-3878(98)00117-5

Sequeira, T. N. (2021). Inflation, economic growth and education expenditure. *Economic Modelling*, 96, 123–136. https://doi.org/10.1016/j.econmod.2020.12.004

Serin, Ş. C., & Demir, M. (2024). Testing the fiscal theory of price level for Türkiye with Fourier-based empirical approaches. *Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Dergisi*, *19*(1), 49-72. https://doi.org/10.17153/oguiibf.1329943

Sims, C. A. (2013). Paper money. *American Economic Review, 103*(2), 563–584. https://doi.org/10.1257/aer.103.2.563

Tiwari, A. K., Tiwari, A. P., & Pandey, B. (2012). Fiscal deficit and inflation: What causes what? The case of India. *Journal of International Business and Economy*, *13*(1), 57-81.

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

Tuncer, İ. (2024). Kamu harcamaları ve enflasyon ilişkisi: Türkiye için ampirik bir analiz. *Maliye Dergisi*, 186, 45–67. https://dergipark.org.tr/tr/pub/maliye

United Nations Statistics Division. (n.d.). *Classification of the functions of government (COFOG)*. Retrieved December 25, 2025, from https://unstats.un.org/unsd/classifications/Econ/COFOG

Woodford, M. (1995). Price-level determinacy without control of a monetary aggregate. *Carnegie-Rochester Conference Series on Public Policy, 43*, 1–46. https://doi.org/10.1016/0167-2231(95)90035-0

Zivot, E., & Andrews, D. W. K. (1992). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business & Economic Statistics, 10*(3), 251–270. https://doi.org/10.1080/07350015.1992.10509904

Information About the Article/Makale Hakkında Bilgiler

The Ethical Rules for Research and Publication / Araştırma ve Yayın Etiği

The author declared that the ethical rules for research and publication followed while preparing the article.

Yazar makale hazırlanırken araştırma ve yayın etiğine uyulduğunu beyan etmiştir.

Conflict of Interests/ Çıkar Çatışması

The author have no conflict of interest to declare.

Yazar çıkar çatışması bildirmemiştir.

Grant Support/ Finansal Destek

The author declared that this study has received no financial support.

Yazar bu çalışma için finansal destek almadığını beyan etmiştir.

Author Contributions/ Yazar Katkıları

The draft process of the manuscript/ Taslağın Hazırlanma Süreci L.S, Data Collection/Verilerin Toplanması L.S, Writing the Manuscript/ Makalenin Yazılması L.S, Submit, Revision and Resubmit Process/ Başvuru, Düzeltme ve Yeniden Başvuru Süreci L.S.

Journal of Applied and Theoretical Social Sciences

JATSS, 2026; 8(1), 45-77

First Submission:18.01.2026
Revised Submission After Review:04.03.2026
Accepted For Publication:24.03.2026
Available Online Since:25.03.2026

Research Article

Crises, Protests, and Political Turbulence: An Event Study of Risk Appetite in Türkiye

Ali Özer[a]

Abstract

Introduction: This study investigates major whether economic, political, geopolitical, and social events in Türkiye shape investor risk appetite and whether the market response differs y across event types. Understanding these dynamics is important for characterizing risk-on/risk-off regimes and for designing event-contingent risk management and policy responses.

Method: We conduct an event study covering selected events over 2008–2025. Risk appetite is measured by an equally weighted composite index (TRRAI) combining sovereign credit risk (5-year CDS), the exchange rate (USD/TRY), equity returns (BIST-100), and BIST volatility. We estimate abnormal outcomes using (i) an MSCI World–based market model for BIST-100 and (ii) a VIX-based abnormal-change model for TRRAI to control for global risk sentiment. We compare reactions across multiple event windows and decompose TRRAI to quantify component-level contributions.

Results or Findings: Global systemic shocks (e.g., the global financial crisis and COVID-19) generate larger and more persistent risk-off responses in both BIST-100 and TRRAI, in wider windows. Domestic uncertainty and institution- or security-related shocks (e.g., protests, institutional disruptions, and the coup attempt) significantly depress risk appetite in shorter windows, while policy-regime shifts and macroprudential interventions exhibit more asymmetric effects in magnitude and persistence. Events involving uncertainty resolution—such as election processes—tend to trigger milder and more heterogeneous short-run reactions.

Discussion or Conclusion: The evidence indicates that risk-off dynamics transmit through distinct channels depending on event type: CDS and FX dominate in some episodes, equities in others, and volatility is particularly salient during global shocks.

Keywords: risk appetite, event study, cds, Bist-100, Türkiye

JEL Codes: G12, G14, E44, F30

[a] Assoc. Prof. Dr, Düzce University, Faculty of Business Administration, Department of Business Administration, Düzce/Turkey, aliozer@duzce.edu.tr , ORCID ID: https://orcid.org/0009-0004-8857-3360 (Corresponding Author)

Journal of Applied and Theoretical Social Sciences

JATSS, 2026; 8(1), 45-77

İlk Başvuru: 18.01.2026
Düzeltilmiş Makalenin Alınışı: 04.03.2026
Yayın İçin Kabul Tarihi: 24.03.2026
Online Yayın Tarihi: 25.03.2026

Araştırma Makalesi

Krizler, Protestolar ve Siyasi Çalkantılar: Türkiye'de Risk İştahına İlişkin Bir Olay Çalışması

Ali Özer[a]

Öz

Giriş: Bu çalışma, Türkiye'deki başlıca ekonomik, politik, jeopolitik ve sosyal olayların yatırımcıların risk iştahını biçimlendirip biçimlendirmediğini ve piyasa tepkisinin olay türlerine göre e farklılaşıp farklılaşmadığını araştırmaktadır. Bu dinamiklerin anlaşılması; risk-on/risk-off rejimlerinin karakterize edilmesi ve olaylara bağlı risk yönetimi ile politika tepkilerinin tasarlanması açısından önem taşımaktadır.

Yöntem: 2008–2025 dönemini kapsayan ve seçilmiş olaylardan oluşan bir olay çalışması yürütülmüştür. Risk iştahı, ülke riskini (5 yıllık CDS), döviz kurunu (USD/TRY), hisse senedi getirilerini (BIST-100) ve BIST volatilitesini birleştiren eşit ağırlıklı bileşik bir endeks (TRRAI) ile ölçülmüştür. Anormal sonuçlar (i) BIST-100 için MSCI World'e dayalı piyasa modeli ve (ii) küresel risk duyarlılığını kontrol etmek amacıyla TRRAI için VIX'e dayalı anormal değişim modeli kullanılarak tahmin edilmiştir. Tepkiler farklı olay pencereleri boyunca karşılaştırılmış; ayrıca TRRAI ayrıştırılarak bileşen düzeyinde katkılar ortaya konmuştur.

Sonuçlar ya da Bulgular: Küresel sistemik şoklar (ör. küresel finansal kriz ve COVID-19), daha geniş olay pencerelerinde, hem BIST-100 hem de TRRAI'de daha büyük ve daha kalıcı risk-off tepkileri üretmektedir. Yurt içi belirsizlik ile kurum- veya güvenlik- temelli şoklar (ör. protestolar, kurumsal kesintiler ve darbe girişimi) kısa pencerelerde risk iştahını anlamlı ölçüde azaltmaktadır. Buna karşılık politika rejimi değişimleri ve makroihtiyati müdahaleler büyüklük ve kalıcılık bakımından daha asimetrik etkiler sergilemektedir. Belirsizliğin çözülmesine işaret eden süreçler—seçimler gibi—genellikle daha ılımlı ve kısa dönemde daha heterojen tepkilerle ilişkilidir.

Tartışma ya da Yapılan Çıkarımlar: Bulgular, risk-off dinamiklerinin olay türüne bağlı olarak farklı kanallar üzerinden aktarıldığını göstermektedir: Bazı dönemlerde CDS ve döviz kuru, diğerlerinde hisse senetleri belirleyici olurken; volatilite özellikle küresel şoklar sırasında daha baskın bir rol oynamaktadır.

Anahtar Kelimeler: risk iştahı, olay çalışması, cds, Bist-100, Türkiye

JEL Kodlar: G12, G14, E44, F30

[a] Doç. Dr, Düzce Universitesi, İşletme Fakültesi, İşletme Bölümü, Düzce/Türkiye, aliozer@duzce.edu.tr, ORCID ID: https://orcid.org/0009-0004-8857-3360 (Sorumlu Yazar)

Introduction

Financial markets cannot be explained solely by macroeconomic fundamentals; investor perceptions, expectations, psychology, and socio-political developments often shape pricing in material ways. Risk appetite is a central concept that reflects investors' willingness to bear risk, and shifts in perceived risk are frequently priced simultaneously through multiple channels such as equities, exchange rates, and sovereign credit risk premia. Over the post-2008 period, Türkiye has experienced global shocks, institutional ruptures, mass protests, security/terror shocks, geopolitical tensions, and monetary-policy regime shifts, all of which have increased market volatility and repeatedly reshaped risk perceptions (Baker & Wurgler, 2007; Shiller, 2015). It is also emphasized that, when global risk rises synchronously, "global common volatility" shocks can strengthen cross-country synchronized risk-off behavior—so the impact of local events may amplify through interactions with the global risk regime (Engle & Campos-Martins, 2023). In this sense, Türkiye provides a suitable setting in which changes in risk appetite can be observed concurrently across multiple market segments.

Traditional finance theories, including Modern Portfolio Theory (Markowitz, 1952), the Capital Asset Pricing Model (Sharpe, 1964; Lintner, 1965), and the Arbitrage Pricing Theory (Ross, 1976), emphasize the risk–return relationship under the assumptions of rational investors and informationally efficient markets. However, crises and price anomalies observed since the 1980s have highlighted the explanatory limits of purely rational paradigms. Behavioral finance, building on prospect theory (Kahneman & Tversky, 1979), introduced a complementary perspective by explicitly analyzing how psychological mechanisms influence investment decisions. This line of research shows that non-rational behavior and psychological factors can materially affect market prices, making investor perceptions and risk attitudes central to understanding financial decision-making. Consequently, investor sentiment, risk perceptions, and risk appetite have become widely studied constructs in the literature (Baker & Wurgler, 2007; Tetlock, 2007). For Türkiye in particular, recent evidence that links the sentiment/perception channel to foreign portfolio flows supports the view that measuring risk appetite should not be confined to "fundamentals" alone (Güneş, et al., 2024).

Behavioral finance further suggests that investors do not always act fully rationally, and that changes in sentiment, attention, and perceived uncertainty can generate measurable effects on asset prices and risk premia. This perspective implies that risk appetite may fluctuate not only with "fundamentals," but also through information and perception channels such as news flow and uncertainty shocks (Baker & Wurgler, 2007; Tetlock, 2007).

Relatedly, the policy-uncertainty and geopolitical-risk literatures indicate that financial pricing responds to event shocks not only via macroeconomic fundamentals but also through uncertainty itself. Baker, Bloom, and Davis (2016) document strong links between news-based economic policy uncertainty shocks and both financial and real variables, while Caldara and Iacoviello (2022) show that increases in news-based geopolitical risk can generate downside risks to economic activity and financial markets. This framework strengthens the theoretical motivation for the event categories emphasized in this study, particularly institutional/constitutional disruptions and geopolitical/foreign-policy episodes (Baker et al., 2016; Caldara & Iacoviello, 2022). In addition, there is direct empirical evidence that mass protests and social unrest can affect financial markets. The IMF's "Pricing Protest" study reports that social unrest events can systematically reduce cumulative abnormal returns in equity markets and that the effect may be more pronounced in emerging economies, which is consistent with an expectation of short-run risk-off effects from large-scale protest shocks in Türkiye (Barrett et al., 2021). More recent and expanded evidence (using multi-country, multi-

event settings) also confirms that the negative market response to social unrest can be statistically systematic (Barrett et al., 2024).

A critical step in event-study designs is the accurate dating of when an event is first incorporated into prices. With digitalization and real-time news dissemination, political and social developments can reach broad investor audiences rapidly, prompting fast updates in expectations. In this context, the news-flow and investor-attention channel can amplify the impact of event shocks on risk perceptions and risk appetite (Bollen et al., 2011; Ranco et al., 2015). This dynamic is particularly relevant in emerging markets such as Türkiye. Accordingly, this study defines the event day ($t = 0$) as the date when the event is first reflected in a strong and salient wave of news coverage, and it additionally conducts sensitivity checks using alternative event-date definitions. When the event occurs outside a trading session or on a non-trading day, we operationally set $t = 0$ to the first subsequent trading day to capture the earliest point at which market prices can incorporate the information. Appendix A reports, for each event, the calendar event date, the event day used in the analysis ($t = 0$), the implied shift in trading days, and an operational timing classification (Trading-hours-or-same-day versus After-close-or-non-trading-day). Moreover, recent methodological discussions on event-study implementation details (event-window choice, overlapping-event risk, interpretive biases) explicitly recommend building designs that are transparent and replicable (Eden et al., 2022).

In emerging markets, risk-off episodes are often associated with sharp reversals in global portfolio flows, exerting simultaneous pressure on sovereign risk premia (CDS), exchange rates, and equity markets. In Türkiye, changes in risk appetite are frequently observed concurrently in (i) the sovereign risk premium (CDS), (ii) exchange-rate dynamics (USD/TRY), (iii) equity pricing (BIST), and (iv) market volatility. Therefore, measuring risk appetite and analyzing event-driven reactions in this setting requires a framework that jointly tracks multiple market channels rather than relying on a single asset class. This multi-market approach is also consistent with the financial stress index literature, which argues that financial stress cannot be adequately captured by one indicator and that composite measures constructed from several market segments better represent stress regimes (Illing & Liu, 2006; Holló et al., 2012). Illing and Liu (2006), for example, emphasize that composite indicators derived from different segments can better characterize stress states, while Holló, Kremer, and Lo Duca (2012) propose a composite methodology to measure systemic stress by combining multiple market blocks. This literature provides a theoretical basis for constructing TRRAI as a multi-component index spanning CDS, FX, equities, and volatility (Illing & Liu, 2006; Holló et al., 2012). Evidence that CDS–exchange rate interactions and spillovers can intensify during "market fear" periods supports the need to track multiple channels jointly (Feng et al., 2021). In addition, evidence that connectedness between Turkish CDS volatility and FX/equity derivatives markets can vary across volatility regimes suggests that event shocks may be transmitted with different intensities across multiple markets (Gök et al., 2023).

This study treats risk appetite as a latent construct and aims to measure Türkiye-specific risk-on/risk-off regimes through multiple market components, and to test the short-run pricing effects of selected political, social, and macro-financial events using an event-study methodology. Specifically, we report and compare (i) event-window reactions in equity pricing using BIST-100 and (ii) reactions in a holistic risk-appetite gauge, TRRAI, which combines CDS, exchange rates, equities, and volatility. The primary contribution is to make risk-on/risk-off dynamics comparable for the same event set by jointly reporting a single-market response (BIST-100) and a multi-market response (TRRAI) within a unified event-study design. This contribution is also consistent with evidence from Türkiye that examines—using daily data—

the relationship between equity markets and portfolio flows/monetary-policy responses (with controls such as CDS and FX), and it suggests that an event-based design can be powerful in capturing policy-regime shifts (Kartal et al., 2022).

The event-study methodology is a standard approach that quantifies the impact of a discrete event by measuring abnormal returns or abnormal changes around an event window. MacKinlay (1997) provides a systematic methodological framework, while Brown and Warner (1985) show that, with daily data, conventional tests typically perform well in short windows. Accordingly, this study applies an abnormal return–cumulative abnormal return (AR–CAR) framework for BIST-100 and measures abnormal changes in TRRAI across multiple windows, consistent with established practice (Brown & Warner, 1985; MacKinlay, 1997). To control for global risk sentiment, we use the VIX index, a widely employed proxy for expected volatility derived from option prices; its construction is documented in Cboe methodology materials (Cboe Global Markets, 2022). Beyond being an options-implied volatility gauge, the VIX also embeds a variance-risk-premium component and is therefore informative about time-varying risk compensation and market uncertainty. (Bekaert & Hoerova, 2014). In this study, ex-ante classification of events by their "dominant transmission channel" (i.e., not labeling ex post by looking at AR–CAR results) is a design preference aimed at reducing category ambiguity and hindsight bias (Eden et al., 2022).

The analysis is organized around the following research questions:
(RQ1) Do risk appetite (TRRAI) and equity pricing (BIST-100) respond in the same direction in the short run around selected events in Türkiye?
(RQ2) Do event effects differ systematically across event types (e.g., crisis/global shocks, protests, security shocks, geopolitical events, monetary-policy/macroprudential interventions)?
(RQ3) As the event window expands, do estimated reactions attenuate (transitory shocks) or cumulate (persistent shocks)?

These questions are evaluated using two complementary outputs for the same event set: (i) AR–CAR measures around event windows for BIST-100 and (ii) abnormal-change measures around event windows for TRRAI, which integrates CDS, FX, equities, and volatility. This dual-output design enables a direct comparison between pricing responses in a single asset class (equities) and a broader multi-market transmission (TRRAI). Although BIST returns are included as one of the components of the TRRAI index, the study also reports a separate AR–CAR event-study analysis for the BIST-100 index. These two measures serve distinct analytical purposes. The BIST AR–CAR analysis captures the immediate price reaction of the equity market and therefore represents a single-market pricing response to the event. In contrast, TRRAI is constructed as a multi-market indicator that aggregates information from sovereign risk (CDS spreads), exchange-rate dynamics (USD/TRY), equity performance (BIST returns), and market volatility. As a result, TRRAI reflects broader shifts in investor risk perception across several financial market segments rather than a single asset class. Reporting both measures allows the analysis to distinguish between equity-specific repricing and broader systemic risk-perception adjustments across markets. Importantly, the interpretation does not rely on the independence of these indicators; rather, their joint evaluation helps identify whether risk transmission operates primarily through equities or through credit, currency, and volatility channels. For RQ3, window-length sensitivity is assessed by examining how the magnitude and persistence of AR–CAR and TRRAI abnormal changes vary across short and wider windows; cumulative strengthening with wider windows is interpreted as evidence of persistence, whereas attenuation indicates more transitory effects.

The hypotheses tested in this framework are based on the expectations that (i) stress- and shock-type events induce contemporaneous risk-off responses in both TRRAI and BIST-100, (ii) the magnitude and persistence of responses differ by event type, and (iii) responses display attenuation or accumulation as a function of window length.

The remainder of the paper is structured as follows. The next section describes the dataset, event classification, and the construction of TRRAI. The methodology section details the event-window design and estimation approach. The results section presents event-level and event-type comparisons. The final section concludes with implications for policy and event-based risk management and outlines directions for future research.

Data and Methodology

Event set and classification approach

All series used in this study are at daily frequency. Daily close data are obtained for the BIST-100 index and the USD/TRY exchange rate, as well as Türkiye's 5-year sovereign CDS spread. We additionally use the VIX as a proxy for global risk sentiment and the MSCI World index as a benchmark for global equity market movements; these are also collected at daily close from the respective data providers. All series are synchronized using the intersection of common trading days. Observations corresponding to non-overlapping holidays across markets are removed. The same aligned calendar is used consistently for the estimation window and for each event window in the event-study design.

We define major post-2008 episodes that are likely to affect risk appetite in Türkiye as events. The post-2008 start date is driven by data availability constraints for the variables employed. To enable systematic analysis, each event is assigned to a single category based on its dominant transmission channel (i.e., each event is mapped to exactly one type). The key rationale is that, although a given event may plausibly activate multiple channels simultaneously (for example, geopolitical tensions may affect both the exchange rate and inflation expectations), multi-label classification can introduce category ambiguity and complicate inference in an empirical event-study setting.

Dominant transmission channel refers to an ex-ante assessment of which risk component the event is most directly expected to trigger at the initial stage, given the nature of the event and how it is reflected in public information flows. In practice, this corresponds to determining whether the event primarily activates institutional/political uncertainty, domestic security risk, geopolitical/foreign-policy risk, a monetary-policy or macroprudential regime change, a macro-data threshold surprise, a natural-disaster/supply shock, or a globally synchronized risk-off shock. Importantly, the classification is not determined ex post by inspecting AR–CAR results. Instead, the event-date definition is fixed and the coding is based on the information set available about the event's characteristics and contemporaneous public narrative.

The coding scheme is designed to be consistent with: (i) the news-based policy-uncertainty framework documenting the pricing effects of institutional and policy disruptions (Baker et al., 2016); (ii) the news-based geopolitical risk approach for wars, interstate tensions, and foreign-policy shocks (Caldara & Iacoviello, 2022); (iii) empirical evidence on social unrest and protests affecting financial markets in an event-study setting (Barrett et al., 2021); and (iv) standard event-study methodology (Brown & Warner, 1985; MacKinlay, 1997).

Classification decisions follow a practical set of rules. An event is coded as: (a) domestic financial crisis/market stress if it directly signals stress in the financial system; (b) monetary policy/macroprudential if it involves a policy-rate regime change or a macroprudential instrument shift; (c) geopolitics/foreign policy if it features cross-border conflict, military operations, or major international tensions; (d) security/terror if it involves domestic security shocks such as terror attacks or a coup attempt; (e) protests/social unrest if it is primarily a mass protest episode; (f) institutional/constitutional if it materially alters constitutional or institutional architecture; (g) elections/political power dynamics if it is election-related; (h) macro data shock if it corresponds to a threshold-type macroeconomic announcement; (i) natural disaster for large-scale disasters such as earthquakes; and (j) global shock for synchronized global risk-off episodes. The emphasis in separating monetary-policy events is motivated by the well-established point that the unexpected component of policy actions can transmit strongly into financial prices (Kuttner, 2001).

We classify events into ten mutually exclusive event-type categories (E1–E10) based on the dominant ex-ante transmission channel. To make this coding logic transparent, Table 1 defines the event-type scheme (E1–E10). The subsequent section presents the event list classified under this scheme in Table 2.

Table 1

Event-Type Scheme (E1–E10)

Code	Event type	Scope (summary)
E1	Domestic Financial Crisis & Market Stress	Stress regimes such as banking, FX, or sovereign-debt crises
E2	Elections & Political Power Dynamics	National/local elections; shifts in political balance and related uncertainty
E3	Institutional/Constitutional–Legal Change	Referendums, system changes, major institutional ruptures
E4	Protests & Social Unrest	Mass protests and broad-based social unrest episodes
E5	Security & Terror	Domestic security shocks such as terror attacks and coup attempts
E6	Geopolitics/Foreign Policy & Military Operations	Wars, cross-border operations, and major international tensions
E7	Monetary Policy, Macroprudential Measures & Financial Regulation	Policy-rate regime shifts, macroprudential tools (e.g., FX-protected deposits), CBRT decisions, and major market/regulatory interventions
E8	Macro Data Shock	Threshold-type announcements/breaks in key macro indicators (e.g., inflation)
E9	Natural Disaster	Large-scale disasters such as earthquakes
E10	Global Shock	Synchronized global risk-off episodes (e.g., global financial or health shocks)

Note. Created by the author using microsoft excel statistical software.

In accordance with the diagram in Table 1, the event set used in the study and the type of each event assigned to the dominant channel are reported in Table 2.

Table 2

Event set and type coding

Date	Event title	Event type	Rationale for classification
2008-09-15	Global financial crisis (post-Lehman shock)	E10	Collapse in global risk appetite → EM outflows / BIST volatility channel
2010-09-12	Constitutional Referendum	E3	Change in institutional/judicial–executive architecture → institutional risk-perception channel
2011-03-15	Syrian civil war / onset of uprising	E6	Persistent border-related geopolitical risk, migration, and security pressure → risk-premium channel
2013-05-28	Gezi Park protests (onset)	E4	Widespread protest wave → domestic political uncertainty and investor-perception channel
2013-12-17	17–25 December episode (first wave)	E3	Institutional–political crisis (perceived executive–judiciary conflict) → risk-premium channel
2015-07-20	Suruç attack	E5	Domestic security shock → flight-to-safety / volatility channel
2015-10-10	Ankara Train Station attack	E5	Large-scale terror shock → security premium and expectations channel
2015-11-24	Downing of Russian jet	E6	Sudden foreign-relations crisis → sanctions/tourism channel
2016-06-28	Atatürk Airport attack	E5	Security shock with tourism-revenue channel
2016-07-15	Coup attempt	E5	Extreme political-regime shock (domestic security) → sovereign risk (CDS) / FX channel
2017-04-16	Presidential system referendum	E3	Change in governance system → institutional framework / predictability channel
2018-08-10	2018 FX and debt crisis (critical wave)	E1	Domestic FX/debt stress → financial fragility and flight-to-safety channel
2019-10-09	Operation Peace Spring	E6	Military operation → foreign-policy and sanctions-risk channel
2020-03-11	COVID-19 (first case announcement)	E10	Global health shock → growth expectations, liquidity, and risk-off channel
2020-08-10	Eastern Mediterranean tensions (NAVTEX / Oruç Reis phase)	E6	Foreign-policy/energy tensions → sanctions and risk-premium channel
2021-04-22	Thodex crisis (withdrawals halted / collapse)	E7	Financial-market/regulatory shock (crypto platform failure) → trust and regulation-expectations channel
2021-12-20	Introduction of FX-protected deposits (announcement)	E7	Macroprudential / FX-regime intervention → FX volatility and dollarization channel
2022-09-05	Inflation exceeds 80% (Aug CPI release)	E8	Data-threshold shock → real returns, policy expectations, risk premium
2023-02-06	Kahramanmaraş earthquakes	E9	Large-scale supply shock and fiscal/reconstruction implications → growth–inflation–risk-premium channel
2023-05-28	Presidential election (runoff)	E2	Closure of political uncertainty / repricing → risk premium and policy-expectations channel
2023-06-22	Rate-hike cycle (first hike)	E7	Signal of monetary-policy regime shift → FX / CDS / capital-flows channel
2023-07-20	Policy rate hike (continuation)	E7	Continued tightening → expectations and risk-premium channel
2024-03-31	2024 local elections	E2	Perceived domestic political-balance shift → risk premium and expectations channel

Date	Event title	Event type	Rationale for classification
2025-08-23	Termination of FX-protected deposits	E7	Closure of a policy instrument (normalization) → TRY demand / FX-dynamics channel

Note. Created by the author using microsoft excel statistical software.

Table 2 outlines the distribution of events by type and the definitions of "event day" to be used in the empirical analysis. The following subsection introduces the AR–CAR event study methodology used to measure the impact of these events on financial variables.

AR–CAR (Event Study) methodology

To quantify market reactions to events, this study employs the standard event-study framework (MacKinlay, 1997; Brown & Warner, 1985). The core idea is to define the deviation of the realized return from its "normal" (expected) level around the event date as the abnormal return (AR) and to cumulate these deviations over the event window to obtain the cumulative abnormal return (CAR).

Prior work shows that, when implemented with daily data, conventional tests generally exhibit good statistical properties and can deliver reliable inference, particularly in short event windows (Brown & Warner, 1985). In line with this practice, we specify the estimation window and event windows as follows:

Estimation window: [−250, −30] trading days
Event day: t = 0
Event windows: baseline [−10, +10]; robustness windows [−1, +1] and [−3, +3]

This design allows the effects of the events listed in Table 2 to be assessed over alternative time scales. After defining event windows, we specify the normal-return model and compute AR and CAR as follows.

Return definition

For BIST-100, the daily log return is defined as:

$$R_t^{BIST} = \ln(P_t) - \ln(P_{t-1}) \quad (1)$$

where P_t denotes the closing value of the index on day t.

Normal return model (market model)

We use the standard market model for daily returns:

$$R_t^{BIST} = \alpha + \beta R_t^m + \varepsilon_t \quad (2)$$

where R_t^m is the reference series representing the global market return (MSCI World in this study). The parameters α and β are estimated by OLS over the estimation window [−250, −30] (MacKinlay, 1997).

Abnormal return (AR) and cumulative abnormal return (CAR)

The abnormal return during the event window is computed as:

$$AR_t^{BIST} = R_t^{BIST} - (\hat{\alpha} + \hat{\beta} R_t^m) \quad (3)$$

For an event window $[\tau_1, \tau_2]$, the cumulative abnormal return is:

$$\mathrm{CAR}^{\mathrm{BIST}}{}_{[\tau 1,\tau 2]} = \sum_{t=\tau 1}^{\tau 2} AR_t^{BIST} \quad (4)$$

The baseline event window is [−10, +10], and robustness results are additionally reported for [−1, +1] and [−3, +3]. Up to this point, the methodology captures event-driven reactions in the equity market. However, because risk appetite is inherently multidimensional and cannot be reduced to a single market, the next subsection motivates how risk appetite is measured (single proxy versus composite index) and proposes an index construction that is suitable for this study.

Measuring risk appetite: single proxy or composite index?

In emerging-market economies such as Türkiye, which are highly sensitive to external financing conditions, risk appetite is a latent phenomenon that cannot be directly observed in a single market. Instead, it materializes simultaneously across multiple segments—most notably sovereign credit risk, the exchange rate, and the equity market. Accordingly, the objective of a composite-index design is to extract the common component of indicators that are expected to co-move during risk-off episodes. The financial stress and risk-index literature emphasizes that composite indices combining measures such as credit risk premia (CDS/spreads), exchange-rate dynamics, and equity returns/volatility can reduce measurement error relative to single-indicator proxies and can distinguish stress regimes more consistently. In the context of Türkiye, jointly tracking CDS, the exchange rate, and equity-market volatility helps capture shifts in country risk perceptions through multiple market channels.

In this study, the component set is chosen based on a clear economic mapping. The CDS series captures the sovereign risk premium; USD/TRY reflects perceived risk toward the Turkish lira; BIST returns proxy the equity-market manifestation of risk-on/risk-off pricing; and BIST volatility represents elevated uncertainty and stress that typically intensify during risk-off periods (Illing & Liu, 2006; Holló et al., 2012; Ekinci, 2013).

The main analysis adopts the composite-index approach, thereby integrating event effects transmitted through credit, FX, and equity channels into a single dependent variable. To preserve transparency and interpretability, the underlying components (CDS, USD/TRY, BIST returns, and volatility) are also reported as part of robustness and interpretation. The index components and sign harmonization are presented in Table 3.

Table 3

Risk appetite index (TRRAI)

Component	Symbol	Measure	Risk-off direction	Note
5-year Türkiye CDS (USD)	(CDS_t)	Daily change	↑	Sovereign risk premium
USD/TRY	(FX_t)	Daily log change	↑	TRY depreciation is risk-off
BIST-100	(EQ_t)	Daily log return	↓	Equity up-moves are risk-on; sign is reversed
BIST volatility	(VOL_t)	20-trading-day rolling standard deviation	↑	Higher volatility is risk-off

Note. Created by the author

The sign harmonization in Table 3 follows a simple rationale. Increases in CDS, the exchange rate, and volatility reinforce risk-off behavior, whereas equity returns proxy risk-on conditions and are therefore included with the opposite sign. In the next step, the four

components are standardized to a common scale and combined using equal weights. As a result, TRRAI is designed to summarize the common risk-appetite component reflected jointly across credit, FX, equity, and volatility channels into a single series.

Composite risk appetite index: TRRAI (Turkey Risk-Return Appetite Index)

In this study, risk appetite is measured by TRRAI, an equally weighted composite index constructed from four market-based components. The index is built in three steps: (i) calendar alignment and sample initialization, (ii) standardization (z-scores), and (iii) sign-harmonized aggregation under equal weights.

(i) Data alignment and sample start

TRRAI is computed at daily frequency. Because the volatility component, VOL_t, is defined as a 20-trading-day rolling standard deviation, the earliest day on which the index can be computed is the first observation after at least 20 trading days. The sample therefore starts on the first date for which all four components are simultaneously available, and all series are aligned on a common trading-day calendar.

(ii) Standardization (z-scores)

Because the components are measured on different scales, each component is transformed into a comparable metric using z-score standardization. For each component:

$$z_{k,t}=\frac{x_{k,t}-\mu_k}{\sigma_k} \tag{5}$$

where $x_{k,t}\in\{ CDS_t, \Delta\ln(FX_t), R_t^{BIST}, VOL_t \}$; μ_k and σ_k denote the sample mean and sample standard deviation of component k, respectively. Since R_t^{BIST} is already defined as a daily log return, no additional differencing is applied.

(iii) Equal-weighted composite stress and risk appetite index

We first compute a composite "stress" measure that is increasing in the risk-off direction. The components are combined using equal weights, $w_k=1/4$. Equal weighting is adopted as the baseline specification because it (a) improves interpretability and (b) mitigates data-driven overfitting risks associated with estimating weights. In addition, equal weights provide a transparent benchmark in which each market block (CDS, FX, equities, and volatility) contributes symmetrically to the composite risk-perception indicator. Data-driven weighting schemes—such as principal component analysis (PCA) or latent-factor approaches—may produce weights that are highly sensitive to sample-specific volatility structures and may mechanically assign excessive influence to the most volatile component. Such behavior can reduce interpretability and comparability across samples. For these reasons, the present study retains equal weighting as the baseline specification and treats alternative weighting approaches as conceptually relevant but not necessary for the core identification strategy in the event-study setting. This approach is consistent with the composite financial stress literature emphasizing transparent aggregation across market segments (Illing & Liu, 2006; Holló et al., 2012). Accordingly, the Turkey composite stress index is defined as:

$$FSI_t^{TR}=\frac{1}{4}\left(z_{CDS,t}+z_{FX,t}-z_{EQ,t}+z_{VOL,t}\right) \tag{6}$$

where the equity component enters with a negative sign to ensure sign harmonization in the risk-off direction. Finally, to make the main index interpretable as "higher values = higher risk appetite (risk-on)," we invert the sign and define TRRAI as:

$$TRRAI_t = -FSI_t^{TR} \quad (7)$$

Under this definition, increases in TRRAI indicate a decline in overall risk perceptions for Turkish assets (risk-on), whereas decreases in TRRAI correspond to a strengthening of flight-to-safety behavior (risk-off).

Measuring abnormal effects: VIX-referenced TRRAI and MSCI World–referenced AR–CAR for BIST-100

In this study, the "normal" component is defined using different benchmark series for the two dependent variables:

- Risk appetite index (TRRAI): adjusted using the VIX, which proxies global risk sentiment.
- BIST-100 returns: adjusted using the MSCI World index, which represents global equity market returns.

The motivation for this distinction is as follows. TRRAI is a Türkiye-specific risk-on/risk-off measure constructed from multiple domestic market segments; therefore, a benchmark capturing exogenous shifts in global risk appetite (VIX) provides a more direct external reference for its normal variation. By contrast, BIST-100 is an equity-market return series, for which specifying a market model against a global equity benchmark (MSCI World) is standard practice. Specifically, over the estimation window [−250,−30], we estimate the benchmark model $TRRAI_t = \gamma+\delta VIX_t+\varepsilon_t$. We define abnormal TRRAI movements as the residual component $A\Delta TRRAI_t=\varepsilon_t$ and cumulate it within each event window to obtain cumulative abnormal changes (CAΔTRRAI).

Results

Descriptive statistics for the variables used in this study are reported in Table 4. All variables are observed at daily frequency, and after applying the relevant transformations and aligning the calendar, the final sample consists of 4,187 observations. The sample covers a sufficiently long post-event period for the final event in the dataset, ensuring that the widest event window [−10,+10] is fully observable for all events and that no forward-dated (out-of-sample) observations are required for AR–CAR or CAΔTRRAI calculations. The fact that all variables share the same number of observations indicates that the series are aligned on a common trading-day calendar and that the dataset is balanced.

Table 4

Descriptive statistics

Variables	N	Mean	Std. dev.	Min	Max
TRRAI (4-component index, level)	4187	0.001182	0.556362	-8.048016	5.432600
CDS (daily change)	4187	0,00057	0,03364	-0,22092	0,49108
USD/TRY (Δln)	4187	0.000842	0.010120	-0.293974	0.147066
BIST-100 (log return)	4187	0.000807	0.017517	-0.362394	0.094219
BIST-100 volatility (20-day rolling s.d.)	4187	0.015566	0.008098	0.005252	0.093804
VIX (Δln)	4187	-0.000123	0.079898	-0.442449	1.031882
MSCI World (log return)	4187	0.000395	0.013647	-0.343223	0.109644
TRRAI (Δln / daily change)	4187	0.000046	0.067968	-2.198158	1.908242

All series are aligned on a common trading-day calendar. CDS is measured as a daily change. BIST-100 and MSCI World are defined as daily log returns (Δln). USD/TRY, VIX, and TRRAI are defined as daily log changes (Δln). BIST-100 volatility is computed as the 20-trading-day rolling standard deviation of returns.

Note. Created by the author using microsoft excel and stata statistical softwares.

Table 4 summarizes the descriptive properties of the aligned series. The near-zero mean of TRRAI in levels is consistent with its standardized construction, while its wide range indicates that the sample period features pronounced risk-on and risk-off regimes and that risk appetite can fluctuate sharply over time. The daily-change series for TRRAI further suggests that risk appetite generates meaningful shocks at the daily frequency and can therefore be informative for capturing short-run dynamics.

Changes in the sovereign CDS measure exhibit a limited average drift, but the wide dispersion and high volatility indicate that country risk perceptions are susceptible to abrupt day-to-day jumps and are highly shock-sensitive. Similarly, the broad distributions observed for USD/TRY and VIX changes suggest that FX-market conditions and global uncertainty can deteriorate and recover rapidly in specific episodes. MSCI World returns, used to control for global market conditions, display a more stable distribution overall, while still reflecting occasional sharp price movements at the global level.

Daily log returns on the BIST-100 show a limited mean drift but substantial variation, indicating notable day-to-day volatility in the equity market. The fact that the minimum return is more extreme than the maximum implies that negative shocks can produce sharper losses, consistent with relatively pronounced downside tail risk. BIST-100 volatility, computed as a rolling standard deviation of returns, exhibits marked episodic increases, suggesting a regime-like character in which volatility can rise quickly during stress periods and provides a complementary gauge of market risk. Overall, the descriptive statistics indicate a sample period characterized by high volatility, regime shifts, and asymmetric responses, and they support the use of TRRAI as a discriminating composite measure for analyzing risk-appetite dynamics in Türkiye.

Finally, we analyze the BIST-100 response around the political, economic, and geopolitical events defined in Section 2 using the event-study methodology. For each event, cumulative abnormal returns (CAR) computed over alternative windows are reported in Table 5. Using short [−1, +1], medium [−3, +3], and wide [−10, +10] windows allows us to examine

both immediate and more persistent reactions and to assess whether shocks to the BIST-100 appear transitory or cumulative over time.

Table 5

BIST-100 event effects (CAR)

Type	Event	CAR [-1,+1]	t-stat. [-1,+1]	CAR [-3,+3]	t-stat. [-3,+3]	CAR [-10,+10]	t-stat. [-10,+10]
E10	Global financial crisis	-9.008	-2.203	-22.205	-3.555	-4.084	-0.377
E3	Constitutional Referendum	0.952	0.352	1.469	0.356	4.598	0.643
E6	Syrian civil war	4.123	1.855	3.069	0.904	2.954	0.502
E4	Gezi Park protests	-4.688	-2.450	-6.915	-2.366	-17.854	-3.527
E3	17–25 December episode	-6.414	-2.439	-7.923	-1.972	-13.203	-1.898
E5	Suruç attack	-2.316	-1.064	-3.935	-1.184	-5.219	-0.906
E5	Ankara Train Station attack	-0.963	-0.434	1.766	0.545	5.876	1.008
E6	Downing of Russian jet	-6.755	-2.982	-7.008	-1.926	-6.482	-1.013
E5	Atatürk Airport attack	-3.259	-1.188	-1.903	-0.461	-0.388	-0.053
E5	Coup attempt (15 July)	-8.532	-2.921	-8.290	-1.869	-7.278	-0.937
E3	Presidential system referendum	-0.946	-0.331	-0.329	-0.078	3.133	0.399
E1	2018 FX and debt crisis	-4.018	-1.934	-3.505	-1.105	-1.403	-0.255
E6	Operation Peace Spring	-2.714	-0.919	-4.281	-0.958	-8.876	-1.135
E10	COVID-19	-5.379	-2.361	-13.250	-3.808	-14.723	-2.443
E6	Eastern Mediterranean tensions	7.099	3.002	0.058	0.016	-9.192	-1.469
E7	Thodex crisis	-0.273	-0.123	-2.729	-0.803	-3.033	-0.515
E7	Introduction of FPD	-5.663	-1.908	-8.620	-1.921	-2.642	-0.333
E8	Inflation exceeds 80%	7.094	2.216	8.639	1.766	5.269	0.622
E9	Kahramanmaraş earthquakes	-6.671	-2.172	-2.851	-0.608	-14.864	-1.830
E2	Presidential election (runoff)	8.726	2.240	10.425	1.752	3.184	0.309
E7	Rate-hike cycle (first hike)	5.618	1.434	3.272	0.547	10.269	0.991
E7	Policy rate hike (continuation)	4.132	0.983	0.208	0.032	10.258	0.923
E2	2024 local elections	2.079	0.575	-0.314	-0.057	2.376	0.248
E7	Termination of FPD	1.667	0.582	3.438	0.786	-4.802	-0.634

FPD is FX-protected deposits

Note. Created by the author using microsoft excel and stata statistical softwares.

The results in Table 5 indicate that the BIST-100 exhibits heterogeneous and asymmetric responses across event types. Systemic shocks such as the global financial crisis and the COVID-19 pandemic generate pronounced negative CARs, particularly in the short and medium windows. Notably, the strong negative CAR following the first COVID-19 case announcement persists in the wide window, suggesting that global uncertainty shocks induce broad-based risk-off behavior in Türkiye's equity market and that the repricing is not confined to the immediate aftermath. Similarly, the clearly negative medium-window CAR for the global financial crisis indicates rapid transmission from deteriorating global risk appetite into BIST-100 pricing.

By contrast, events related to political processes—such as the constitutional referendum and presidential elections—tend to deliver positive or modest CARs in short and medium

windows. The consistently positive CARs around the constitutional referendum across windows suggest that expectations of uncertainty resolution may be priced favorably by market participants, although statistical power may be limited in some windows. Likewise, the positive short- and medium-window CARs around the presidential runoff election are consistent with a "closure of uncertainty" channel that can compress risk premia in the short run. The weakening of the effect in the wide window, however, suggests that the impact of political developments may often be concentrated over short-to-medium horizons.

For social unrest, terror attacks, and security-related shocks, the BIST-100 generally produces negative CARs in short windows. The Gezi Park protests and the 17–25 December episode display negative CARs in both short and wider windows, consistent with the view that investors interpret such episodes as high-uncertainty shocks. The coup attempt is associated with strongly negative CARs in short and medium windows, highlighting how domestic security shocks can rapidly worsen country-risk perceptions and trigger flight-to-safety dynamics. At the same time, for some events the effect attenuates or changes sign as the window widens, suggesting that markets can partially adjust over time and that recovery dynamics may emerge.

Results around macro-financial crises and policy interventions appear sensitive to the time horizon. The 2018 FX and debt crisis yields negative CARs in the short window, but the attenuation in wider windows may indicate partial pre-event pricing or a relatively rapid adjustment mechanism in the equity market. For monetary tightening steps (rate hikes), positive short-window CARs and the persistence of positive CARs in the wide window suggest that, in some episodes, tightening signals can be interpreted as uncertainty-reducing and credibility-enhancing. In contrast, while the termination of FX-protected deposits is associated with positive short and medium-window CARs, the wide-window effect turns negative, pointing to a more complex, time-varying pricing process that is likely sensitive to concurrent developments.

Overall, the AR–CAR evidence in Table 5 suggests that the BIST-100 reacts strongly and predominantly negatively to global systemic shocks and domestic events that elevate uncertainty, whereas developments perceived to reduce uncertainty or improve policy clarity can be priced more positively. These findings indicate that risk perceptions in Türkiye's equity market are event-specific, time-varying, and asymmetric, and they provide a useful benchmark for interpreting the corresponding results based on the composite risk appetite index (TRRAI).

Table 6 reports the response of the Turkey Risk-Return Appetite Index (TRRAI) around selected political, economic, and geopolitical events using cumulative abnormal changes (CAΔTRRAI) computed over alternative event windows. Within the event-study framework, short [−1, +1], medium [−3, +3], and wide [−10, +10] windows are used to examine both immediate and more persistent dynamics in risk appetite. Positive CAΔTRRAI values indicate increasing risk appetite (risk-on), whereas negative values correspond to a strengthening risk-off regime.

Table 6

Event effects on CAΔTRRAI

Type	Event	CAΔTRRAI [-1,+1]	t-stat. [-1,+1]	CAΔTRRAI [-3,+3]	t-stat. [-3,+3]	CAΔTRRAI [-10,+10]	t-stat. [-10,+10]
E10	Global financial crisis	0.367	0.433	-1.545	-1.194	-0.915	-0.408
E3	Constitutional Referendum	0.111	0.164	0.009	0.008	0.317	0.177
E6	Syrian civil war	0.030	0.052	-0.675	-0.779	-0.416	-0.277
E4	Gezi Park protests	-6.105	-1.427	-0.007	-0.001	-1.923	-0.170
E3	17–25 December episode	-10.238	-1.668	-9.085	-0.969	-6.396	-0.394
E5	Suruç attack	0.158	0.277	0.217	0.199	-0.595	-0.308
E5	Ankara Train Station attack	-0.720	-1.308	-0.092	-0.110	-0.258	-0.178
E6	Downing of Russian jet	-0.972	-1.748	-0.322	-0.379	-0.570	-0.383
E5	Atatürk Airport attack	-0.302	-0.625	-0.134	-0.182	-0.777	-0.611
E5	Coup attempt (15 July)	-17.968	-2.985	-11.635	-1.266	-1.778	-0.112
E3	Presidential system referendum	0.418	0.069	0.400	0.043	0.660	0.041
E1	2018 FX and debt crisis	-2.372	-3.730	1.079	1.111	1.553	0.930
E6	Operation Peace Spring	-0.271	-0.593	-0.551	-0.794	-1.161	-0.957
E10	COVID-19	-9.664	-1.338	-24.572	-2.228	-22.105	-2.378
E6	Eastern Mediterranean tensions	0.055	0.106	-0.032	-0.041	-0.115	-0.087
E7	Thodex crisis	0.021	0.038	-0.072	-0.086	-0.559	-0.375
E7	Introduction of FPD	-3.702	-5.162	0.347	0.317	-0.586	-0.308
E8	Inflation exceeds 80%	4.493	0.297	-1.673	-0.072	-4.287	-0.107
E9	Kahramanmaraş earthquakes	-20.875	-1.928	2.859	0.173	-0.010	-0.005
E2	Presidential election (runoff)	0.408	0.473	-0.221	-0.167	-1.508	-0.660
E7	Rate-hike cycle (first hike)	0.350	0.407	-0.442	-0.337	0.477	0.210
E7	Policy rate hike (continuation)	0.740	0.836	-0.327	-0.242	-0.072	-0.031
E2	2024 local elections	-5.248	-0.569	9.405	0.668	-1.064	-0.044
E7	Termination of FPD	-2.620	-0.426	-1.018	-0.108	-7.023	-0.431

FPD is FX-protected deposits

Note. Created by the author using microsoft excel and stata statistical softwares.

The results in Table 6 indicate that TRRAI responds strongly—and predominantly negatively—to events characterized by elevated uncertainty. Among global shocks, the COVID-19 episode shows a risk-appetite deterioration that is already negative in the short window and deepens substantially in the medium and wide windows, indicating that global uncertainty rapidly intensifies risk-off behavior in Türkiye's financial markets. For the global financial crisis, the short-window estimate is mildly positive, but it turns negative in the medium and wide windows, suggesting that the inferred response is sensitive to window length and that the initial-day movement may have been offset by subsequent adjustment through other channels.

For episodes dominated by domestic uncertainty and institutional disruptions, TRRAI produces rapid and sharp declines, confirming the composite index's sensitivity to country-specific risk shocks. The Gezi Park episode yields a pronounced short-window risk-off response, while the 17–25 December episode remains negative across short, medium, and wide windows, consistent with the interpretation that perceived institutional–political crisis

compresses risk appetite. Similarly, the coup attempt is associated with a very large negative short-window response and continued negative effects in the medium window, indicating that domestic security shocks can strengthen the risk-off regime sharply, likely through the FX and sovereign risk premium channels.

By contrast, political-process events such as the constitutional referendum and the presidential system referendum are associated with positive or limited CAΔTRRAI values in short windows. This pattern is consistent with the notion that such events may, in some cases, be interpreted as uncertainty-resolving and thus supportive of risk appetite. The modest magnitude, however, suggests that the effect of political developments on risk appetite may depend on the event's surprise component and concurrent macro-financial conditions.

Results around macro-financial crises and policy interventions indicate that risk appetite is sensitive to the time horizon. For the 2018 FX and debt crisis and the introduction of FX-protected deposits, the short-window reaction is risk-off, but partial recovery signs emerge in the medium window, suggesting that markets may rebalance risk perceptions after the initial shock through adjustment and policy-response channels. For the inflation-threshold episode, the short-window response is positive, whereas medium and wide windows weaken or turn negative, implying that macro-data threshold news may initially be interpreted through a policy-response-expectations channel, but can subsequently be dominated by concerns about the growth–inflation path.

Finally, for several events, negative short-run CAΔTRRAI values attenuate or reverse in medium and wide windows, indicating that risk appetite can react quickly yet transiently to shocks in Türkiye's markets. Overall, Table 6 shows that TRRAI captures both global and domestic uncertainty shocks and that risk appetite dynamics are event-specific, horizon-dependent, and asymmetric. These findings provide empirical support for using TRRAI as a reliable risk-perception proxy in subsequent econometric analysis.

Table 7 jointly reports, for the [−1, +1] window, the events that generate the strongest short-run risk-off effects according to TRRAI and the contemporaneous BIST-100 response. This comparison is intended to reveal the extent to which sharp declines in risk appetite translate into equity-market pricing and whether the direction and magnitude of the two measures align in the short run.

Table 7

Strongest risk-off events based on TRRAI and the BIST-100 response ([-1, +1])

Type	Event	BIST CAR [-1,+1] (%)	t-stat. (BIST)	TRRAICA Δ[-1,+1]	t-stat. (TRRAI)
E9	Kahramanmaraş earthquakes	-6.671	-2.172	-20.875	-1.928
E5	Coup attempt (15 July)	-8.532	-2.921	-17.968	-2.985
E3	17–25 December episode	-6.414	-2.439	-10.238	-1.668
E10	COVID-19	-5.379	-2.361	-9.664	-1.338
E4	Gezi Park protests	-4.688	-2.450	-6.105	-1.427
E2	2024 local elections	2.079	0.575	-5.248	-0.569
E7	Introduction of FX-protected deposits	-5.663	-1.908	-3.702	-5.162
E7	Termination of FX-protected deposits	1.667	0.582	-2.620	-0.426
E1	2018 FX and debt crisis	-4.018	-1.934	-2.372	-3.730
E6	Downing of Russian jet	-6.755	-2.982	-0.972	-1.748

Note. Created by the author using microsoft excel statistical software.

The results in Table 7 indicate that events associated with pronounced declines in TRRAI largely coincide with economically and statistically meaningful negative cumulative abnormal returns in the BIST-100. For the Kahramanmaraş earthquakes and the coup attempt—two episodes producing the sharpest short-window deteriorations in risk appetite—the steep declines in TRRAI occur simultaneously with strongly negative BIST CARs, indicating rapid transmission of the risk-off regime into equity pricing. Similarly, for high-uncertainty events such as the 17–25 December episode, the Gezi Park protests, and the first COVID-19 case announcement, deteriorations in risk appetite align with sizable negative repricing in equity returns. Taken together, these patterns suggest that a worsening of risk perceptions as captured by TRRAI is, in the short run, associated with a clear flight-to-safety response in Türkiye's equity market.

At the same time, some events exhibit a non-aligned responses between TRRAI and BIST-100. The 2024 local elections and the termination of FX-protected deposits illustrate cases in which TRRAI signals risk-off in the short window while BIST CARs are positive. This pattern suggests that equity pricing may reflect not only contemporaneous risk perceptions but also the perceived policy direction and normalization expectations, the extent of pre-event pricing, and concurrent expectation channels. Thus, while TRRAI and BIST-100 often move in the same direction, the equity market may price certain events through a different expectation set.

Table 7 also highlights that a policy-regime step such as the introduction of FX-protected deposits generates a strong risk-off coincidence in both TRRAI and BIST-100, consistent with the view that markets may interpret unconventional policy interventions as heightened uncertainty signals in the short run. By contrast, within the same policy category, the termination of the instrument is associated with a positive BIST response, indicating that the content and context of a policy action (e.g., perceived normalization and improved predictability) can reverse the direction of equity pricing. This distinction is consistent with the broader finding that effects within E7-type events can be asymmetric at the event level.

Overall, Table 7 shows that TRRAI-identified risk-off regimes largely coincide with negative cumulative abnormal returns in the BIST-100, supporting TRRAI as a reliable composite gauge of short-run risk perceptions in Türkiye's financial markets. Nevertheless, the presence of events for which TRRAI is negative while BIST CARs are positive or muted indicates that equity-market responses can vary with event characteristics, policy expectations, and the degree of pre-event pricing. These findings imply that sharp deteriorations in risk appetite typically transmit quickly and negatively to equities, while certain policy transitions may generate more complex pricing patterns, underscoring TRRAI's usefulness as a complementary and powerful analytical tool.

Having examined event-level reactions for the strongest risk-off episodes identified by TRRAI, we next deepen the analysis at the event-type level. To this end, Table 8 groups similar events and reports mean and median BIST CAR and TRRAI CAΔ values across alternative windows ([-1, +1], [-3, +3], and [-10, +10]), enabling a comparative assessment of short- and longer-horizon reactions by event type.

Table 8

Mean/median effect sizes by event type

		BIST100					
Type	**Events**	**CAR (-1,+1) Mean(%)**	**CAR (-1,+1) Med.(%)**	**CAR (-3,+3) Mean(%)**	**CAR (-3,+3) Med.(%)**	**CAR (-10,+10) Mean(%)**	**CAR (-10,+10) Med.(%)**
E1	1	-4.018	-4.018	-3.505	-3.505	-1.403	-1.403
E2	2	5.403	5.403	5.056	5.056	2.780	2.780
E3	3	-2.136	-0.946	-2.261	-0.329	-1.824	3.133
E4	1	-4.688	-4.688	-6.915	-6.915	-17.854	-17.854
E5	4	-3.768	-2.787	-3.090	-2.919	-1.752	-2.804
E6	4	0.438	0.705	-2.040	-2.111	-5.399	-7.679
E7	5	1.096	1.667	-0.886	0.208	2.010	-2.642
E8	1	7.094	7.094	8.639	8.639	5.269	5.269
E9	1	-6.671	-6.671	-2.851	-2.851	-14.864	-14.864
E10	2	-7.193	-7.193	-17.727	-17.727	-9.404	-9.404
				TRRAI			
E1	1	-2.372	-2.372	1.079	1.079	1.553	1.553
E2	2	-2.420	-2.420	4.592	4.592	-1.286	-1.286
E3	3	-3.236	0.111	-2.892	0.009	-1.806	0.317
E4	1	-6.105	-6.105	-0.007	-0.007	-1.923	-1.923
E5	4	-4.708	-0.511	-2.911	-0.113	-0.852	-0.686
E6	4	-0.290	-0.121	-0.395	-0.436	-0.566	-0.493
E7	5	-1.042	0.021	-0.302	-0.327	-1.553	-0.559
E8	1	4.493	4.493	-1.673	-1.673	-4.287	-4.287
E9	1	-20.875	-20.875	2.859	2.859	-0.010	-0.010
E10	2	-4.649	-4.649	-13.058	-13.058	-11.510	-11.510

Note. Created by the author using microsoft excel statistical software.

The evidence in Table 8 shows that responses of both BIST-100 and TRRAI vary markedly by event type, in both magnitude and persistence. For global shocks (E10), the BIST response becomes substantially more negative as the window widens, indicating that the global risk-off regime is not confined to the event day but unfolds over several days. Consistent with this pattern, TRRAI exhibits strongly negative values in medium and wide windows, suggesting that global shocks generate persistent flight-to-safety dynamics across Turkish assets. By contrast, for the domestic financial crisis category (E1), BIST CAR remains negative but

attenuates in wider windows, while TRRAI turns positive in the medium and wide windows, pointing to a potential role for post-shock adjustment and policy-response channels.

For domestic political and institutional categories (E3) and the mass-protest category (E4), the evidence is heterogeneous. The protest event (E4) produces strong negative responses in the short window for both BIST and TRRAI, consistent with a broad-based deterioration in risk perceptions across channels. Within E3, negative mean effects coexist with less negative or even positive medians in wider windows, indicating substantial cross-event dispersion and a potential influence of outliers. For TRRAI, the contrast between a negative mean and a near-zero/positive median in the short window similarly suggests that institutional episodes vary considerably in how sharply they deteriorate risk appetite.

Security and terror events (E5) generate clearly negative average BIST CARs in short and medium windows, consistent with accelerated flight-to-safety in equities following security shocks. The attenuation of the negative effect in the wide window suggests partial market adjustment or recovery dynamics. TRRAI also shows negative short-window means in E5, confirming an initial deterioration in risk perceptions, while the less negative medians point to heterogeneity across episodes.

For geopolitical events (E6), BIST responses become more negative as the window widens: the short-window mean is slightly positive, whereas medium and wide windows turn distinctly negative, consistent with the idea that risk-premium and uncertainty channels may be priced more gradually over several days. TRRAI shows small negative effects in the short window and a strengthening of negative effects in medium and wide windows, suggesting that geopolitical tensions feed into risk perceptions in a more cumulative manner.

For policy regime and economic-policy events (E7), the divergence between means and medians is pronounced, highlighting strong event-level asymmetry. In BIST, the mean is positive in the short window, turns negative in the medium window, and returns to positive in the wide window, while the median remains positive in the short window, suggesting that some policy steps can be interpreted as uncertainty-reducing and thus priced favorably. In TRRAI, the negative short-window mean alongside a near-zero/positive median indicates that E7 is not a uniformly risk-off category. This pattern is consistent with the notion that the surprise component of policy actions, the communication framework, and the extent of pre-event pricing play a critical role in shaping the direction of market reactions.

Finally, despite being based on a limited number of events, E8 and E9 display noteworthy patterns. For E8, BIST CAR is strongly positive across all windows, whereas TRRAI is positive in the short window but turns negative in the medium and wide windows, suggesting that macro-threshold announcements may initially be interpreted via a policy-response-expectations channel but can later be dominated by concerns about the inflation–growth trade-off and the risk premium. For E9, BIST exhibits large and persistent negative effects, while TRRAI shows an extremely sharp short-window decline followed by rapid attenuation, implying that while disasters can leave deep imprints on equity pricing, the composite risk-appetite measure may rebalance over time as different components adjust.

Overall, Table 8 indicates that BIST responses are closely related to event type, scope, and the extent to which pricing unfolds over time, whereas TRRAI produces stronger and more persistent risk-off signals especially for global shocks and high-uncertainty domestic disruptions. These results underscore that risk perceptions in Türkiye's financial markets differ

meaningfully by event type and that the composite risk-appetite approach (TRRAI) offers an informative framework for capturing market dynamics in systemic episodes.

Having summarized event-level co-movements in Table 7 and type-level mean/median magnitudes in Table 8, we next consider through which channels the risk-on/risk-off responses are transmitted. Because TRRAI combines sovereign credit risk (CDS), the exchange rate (USD/TRY), equity returns (BIST), and volatility, identifying the dominant component(s) is informative. To this end, Table 9 decomposes the short-window ([-1, +1]) TRRAI response into component contributions by event type and reports the dominant component and its share.

Table 9

Component contributions to TRRAI by event type: dominant channel and share

Event type	Dominant component	Dominant share (%)	2nd component	2nd share (%)
E1	CDS	50.558	USD/TRY	38.689
E2	CDS	37.058	BIST (equity)	27.522
E3	BIST (equity)	60.793	CDS	17.434
E4	BIST (equity)	63.676	USD/TRY	23.826
E5	BIST (equity)	37.549	CDS	29.272
E6	BIST (equity)	54.590	CDS	23.974
E7	BIST (equity)	56.858	USD/TRY	20.854
E8	CDS	42.327	BIST (equity)	33.340
E9	BIST (equity)	70.568	CDS	26.026
E10	Volatility	40.738	BIST (equity)	21.899

TRRAI is constructed by z-score standardizing the components—CDS, USD/TRY, BIST returns (with sign reversed), and BIST volatility—and combining them with equal weights (1/4). For the [-1, +1] event window, component contributions are based on the weighted cumulative sum of each component's abnormal change (AC) over the same window. "Share (%)" is computed, by event type, as the ratio of the absolute magnitude of a component's contribution to the sum of absolute magnitudes across components (i.e., it emphasizes magnitude rather than direction).
Note. Created by the author using microsoft excel statistical software.

Table 9 decomposes the short-window ([-1, +1]) TRRAI response into component contributions at the event-type level, highlighting the empirical relevance of TRRAI's multi-component structure. The results indicate that the dominant transmission channel varies systematically across event types. For domestic financial stress/crisis events (E1) and elections/political power dynamics (E2), the dominance of the CDS component suggests that risk perceptions in these episodes are priced primarily through the sovereign risk premium. By contrast, for institutional/constitutional disruptions (E3), protests and social unrest (E4), security and terror shocks (E5), and policy–macroprudential/regulatory shocks (E7), the equity component (BIST) accounts for the largest share, implying that short-run risk appetite responses are frequently transmitted through equity pricing.For natural disasters (E9), the particularly high equity share indicates that initial repricing tends to concentrate in the domestic equity channel. In global shock episodes (E10), the prominence of the volatility component points to a regime in which risk-off dynamics operate mainly through volatility, with risk perceptions becoming most visible via the uncertainty/volatility channel.

Overall, Table 9 strengthens the empirical rationale for a composite approach: risk perceptions can be transmitted through different market segments depending on the nature of the shock—CDS and FX in some cases, equities in others, and volatility during global systemic episodes. Accordingly, relying on a single proxy is likely to miss important aspects of the transmission mechanism. The multi-component TRRAI framework offers a more encompassing lens for interpreting both the co-movements and divergences documented in Table 7 and the type-level effect magnitudes summarized in Table 8.

In the next step, we present event-time plots (t = −10, …, +10) of average responses by event type to trace the dynamic pattern of adjustment around the event date. These figures visualize the timing of reactions (immediate versus delayed) in both BIST abnormal returns and TRRAI changes, thereby complementing the window-based evidence by revealing the underlying response dynamics.

Figure 1

Event-time Average Responses for Type E1

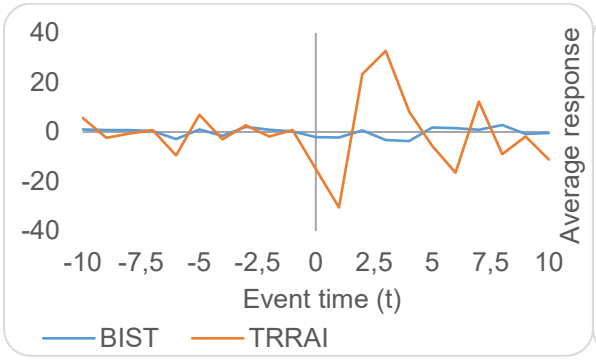


Figure 2

Event-time Average Responses for Type E2

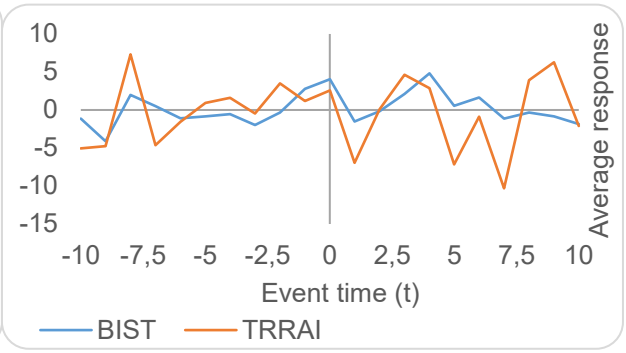


Figure 3

Event-time Average Responses for Type E3

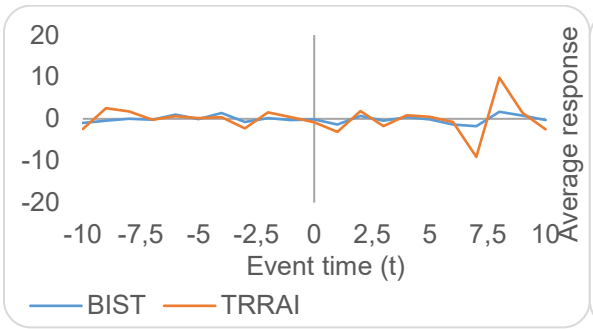


Figure 4

Event-time Average Responses for type E4

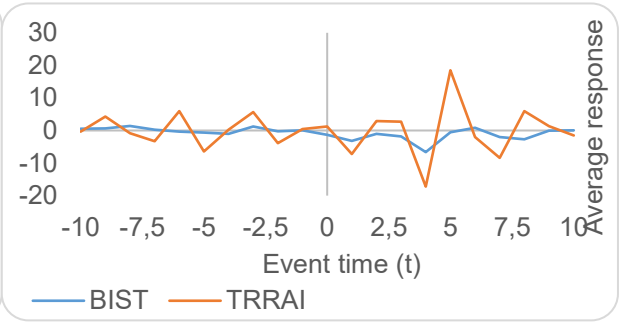


Figure 5

Event-time Average Responses for Type E5

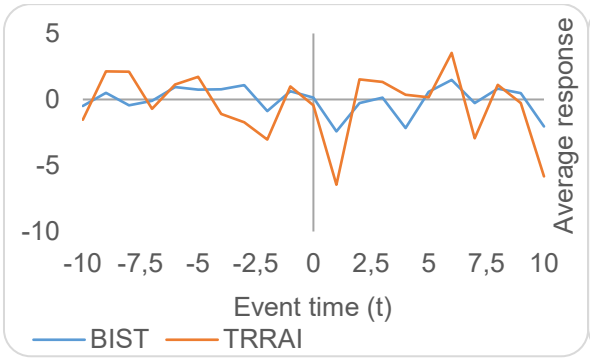


Figure 6

Event-time Average Responses for Type E6

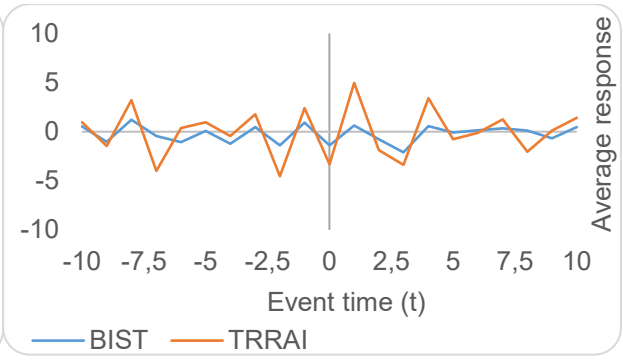


Figure 7

Event-time average responses for type E7

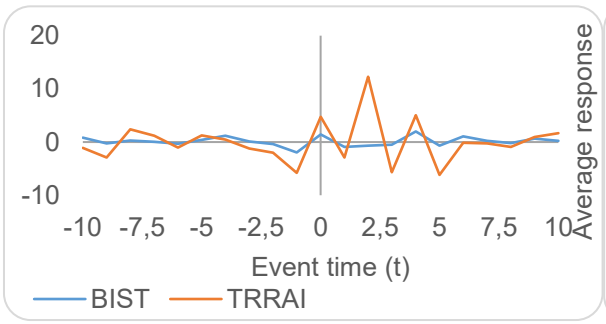


Figure 8

Event-time average responses for type E8

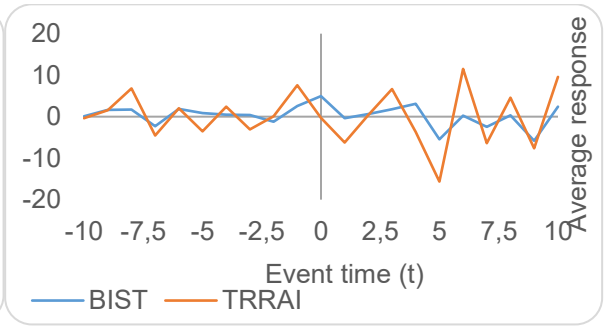


Figure 9

Event-time average responses for type E9

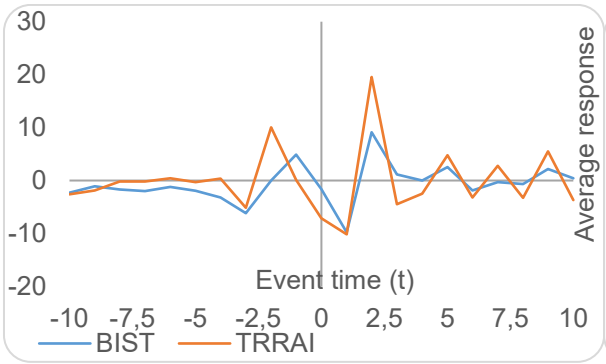
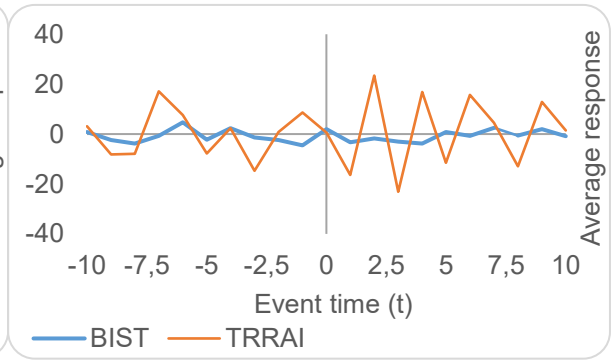


Figure 10

Event-time average responses for type E10



Note. Figures (1-10) are created by the author using microsoft excel statistical software.

The figures complement the window-based evidence reported in Tables 7 and 8 by adding a dynamic perspective. While Table 8 summarizes mean/median effect magnitudes by

event type, the event-time profiles reveal when these effects concentrate around the event date, whether responses are immediate or delayed, and the extent to which impacts persist. In this sense, the figures make transparent which specific days drive differences across windows (e.g., [-1, +1] versus [-10, +10]) by visualizing the accumulation of responses over event time. Moreover, even when mean/median values appear similar for a given type in Table 8, the event-time curves help distinguish between a "one-off shock" and a "cumulative repricing" pattern. The possibility that TRRAI and BIST curves exhibit different timing within the same type also provides dynamic support for the dominant-channel heterogeneity documented in Table 9 (CDS/FX versus equities versus volatility).

For global shocks (E10), the TRRAI response displays higher-amplitude and more persistent post-event fluctuations, indicating that risk-off regimes associated with global episodes tend to unfold over several days rather than being confined to the event day. This pattern is consistent with Table 8, where the adverse effects for E10 become more pronounced as the window widens, and it also aligns with Table 9, which identifies the volatility component as dominant for this type. Taken together, the evidence suggests that, in global shock episodes, risk perceptions in Turkish assets are transmitted most visibly through the volatility regime.

For natural disasters (E9) and the types associated with domestic security and institutional or social disruptions—particularly E5, E3, and E4—the responses are not uniformly concentrated exactly at $t = 0$. E9 exhibits a sharp movement around $t = 0$ followed by a pronounced rebound in the immediate post-event days, indicating a strong short-run shock with rapid subsequent adjustment. E5 shows relatively more "near-event" concentration (around $t \approx 0$ to $t \approx +1$), whereas E3 and E4 display more delayed and oscillatory post-event dynamics (with the larger swings emerging several days after the event), rather than a purely synchronized $t = 0$ response. The stronger deviations close to the event date in both BIST mean abnormal returns and TRRAI profiles (especially for E9 and E5) imply rapid transmission of deteriorating risk perceptions into equity pricing in the short run. This visual evidence is consistent with Table 7, where many of the sharpest short-window declines in TRRAI coincide with negative BIST CARs in the [-1, +1] window.

By contrast, policy and macroprudential/regulatory shocks (E7) exhibit a more asymmetric pattern. The TRRAI curve can oscillate sharply in both directions after the event, suggesting fast updating of risk perceptions in response to policy actions. At the same time, the BIST response can reflect additional channels—such as "normalization" or improved predictability—leading to pricing in the opposite direction in certain episodes. When interpreted jointly with Table 7 (where some policy transitions show negative TRRAI alongside positive BIST CAR) and Table 9 (which highlights channel heterogeneity), the event-time evidence indicates that E7 should not be expected to generate a uniform, one-directional response.

Finally, for elections (E2) and geopolitical events (E6), the response profiles tend to be lower-amplitude and more oscillatory, suggesting that market reactions in these categories may be more heterogeneous and context-dependent. This provides visual support for the window sensitivity and mean/median divergences observed for these types in Table 8.

Overall, the figures complement the evidence in Tables 7–8 along three dimensions: (i) they identify the timing of responses (immediate versus delayed), (ii) they show whether short-run deviations evolve into persistent effects, and (iii) they strengthen the mechanism-based discussion of why TRRAI and BIST responses may diverge for some event types. Accordingly,

when used jointly with the window-based summary tables, the event-time plots improve interpretability and enhance the persuasiveness of the empirical findings.

Discussion

Taken together, the results indicate that risk-off responses in Turkish markets do not operate through a single channel; rather, the transmission mechanism reflects an event-specific "mix" of channels. While the BIST-100 results (Table 5) capture the direct equity-pricing response around events, TRRAI (Table 6) provides a broader measure of the common risk-perception component generated simultaneously across sovereign credit risk, the exchange rate, and volatility. Accordingly, the evidence suggests that a single-channel interpretation—such as "a negative BIST implies a risk-off episode"—can be incomplete, because in some periods risk perceptions are transmitted more strongly through CDS, FX, or volatility than through equities.

At the event level (Tables 5 and 6), high-uncertainty episodes tend to produce co-movement across the two measures: risk-off signals in TRRAI frequently coincide with negative CARs for BIST. This alignment is especially pronounced for shocks related to domestic security, institutional disruptions, social unrest, and natural disasters. Table 7 makes this co-movement particularly clear in the short window [-1, +1] by showing that many of the events associated with the sharpest declines in TRRAI are also accompanied by statistically meaningful negative BIST responses. This pattern strengthens the inference that deteriorations in risk perceptions transmit rapidly to equity prices, implying that risk-off regimes often trigger swift short-run price adjustments.

At the same time, the non-aligned cases documented in Table 7—most notably for certain policy transitions and some election-related events—indicate that a risk-off signal does not always translate into an equity response of the same sign. In several of these episodes, TRRAI is negative in the short window while BIST generates a positive CAR, suggesting that markets may price two forces simultaneously: a deterioration in broad risk perceptions and an expectations update about policy direction and/or normalization that supports equities. Rather than implying inconsistency, this pattern points to channel-specific transmission: broad risk perceptions can worsen through CDS, FX, or volatility even when equities reprice positively due to perceived policy-path improvement or uncertainty resolution. Accordingly, focusing exclusively on equity-market reactions may understate risk-off dynamics when non-equity channels dominate the adjustment.

Statistical significance should also be interpreted with caution. Some event windows display economically meaningful effects with weaker t-statistics, partly because event-study inference is constrained by limited statistical power when certain categories contain only a small number of events. For this reason, the findings are evaluated jointly in terms of effect magnitude, window-length sensitivity (attenuation versus accumulation), and cross-measure consistency between BIST AR–CAR and CAΔTRRAI, rather than on the basis of any single test statistic alone.

Type-level summaries (Table 8) highlight that this heterogeneity has systematic components. For global shocks (E10), the fact that negative effects become more pronounced as the event window widens for both BIST and TRRAI implies that risk-off behavior is not confined to the event day but reflects a repricing process unfolding over several days. By contrast, for some categories—particularly policy/regulation and geopolitics—the larger mean–median gaps indicate that reactions are more heterogeneous across events, consistent with context-dependent pricing rather than a uniform response. This observation clarifies a central

implication of the study: beyond the magnitude of the average response, the dispersion and stability of responses also vary meaningfully by event type.

In this regard, the channel decomposition in Table 9 helps clarify the underlying mechanism. The dominance of the CDS component in types such as E1–E2 indicates that risk perceptions in these episodes are transmitted primarily through the sovereign risk premium. In contrast, the prominence of the equity component in types such as E3–E7 and E9 suggests that short-run risk-appetite responses often become most visible through equity pricing. For global shocks, the dominance of the volatility component points to a "volatility/stress regime" in which risk perceptions are most clearly reflected through heightened uncertainty. This component structure provides a coherent explanation for the co-movements and cross-measure differences documented in Table 7: when TRRAI declines mainly due to CDS/FX/volatility channels, it becomes more likely that BIST will not respond with the same sign or magnitude.

The event-time figures complement these findings by adding a time dimension: for some types, responses sharpen around t = 0, whereas for others effects accumulate in the days after the event and persist with oscillatory dynamics. This visual evidence links the window-sensitivity in Table 8 to a delayed or extended repricing mechanism. Conversely, in types characterized by sharp breaks followed by rapid stabilization, the effects are concentrated in the short window. Hence, the figures reinforce the interpretation that similar mean/median effects can arise through distinct dynamics (an immediate shock versus cumulative repricing), strengthening the overall interpretation.

Overall, the study shows that risk-off events in Turkish markets typically generate fast and adverse reactions, but the direction, magnitude, and persistence of these effects differ systematically by event type. BIST-based evidence provides a clear view of equity-market repricing, while TRRAI captures the joint risk-perception component across credit, FX, equity, and volatility, offering a more holistic perspective. Considering both measures jointly allows one to distinguish between regimes characterized by synchronized risk-off behavior and regimes characterized by channel heterogeneity, thereby underscoring the empirical necessity of reading risk appetite through multiple market segments in economies such as Türkiye that are highly sensitive to the country risk premium and the exchange-rate channel.

Conclusion

This study examines the effects of global and domestic risk-off events on Türkiye's equity market within a holistic framework that measures risk perceptions using a composite index. The evidence shows that Türkiye's equity market responds strongly—and predominantly negatively—to abrupt increases in uncertainty and perceived risk, particularly over short horizons. This finding implies that investors' flight-to-safety tendencies are rapidly incorporated into prices and that market reactions occur with limited delay.

The results further indicate that market responses differ markedly by event characteristics. Financial and macroeconomic crises generate larger and more persistent negative effects than other event types, and these episodes tend to coincide with pronounced deteriorations in global risk sentiment. This pattern suggests that Turkish markets are more fragile in the presence of systemic risks and move in close coordination with shifts in global risk appetite during such periods. Type-level summaries show that, especially for global shocks (E10), adverse effects become more pronounced as the event window expands for both BIST and TRRAI, consistent with a repricing process that unfolds over several days. The event-time

profiles corroborate this mechanism by showing that for some types the response concentrates around t = 0, whereas for others it accumulates in post-event days and becomes more persistent.

By contrast, political developments, election episodes, and certain geopolitical events display more heterogeneous responses in both sign and magnitude. In some cases, negative effects are limited or transient, and in others markets appear to adjust quickly. These patterns suggest either partial pre-pricing of political/geopolitical risks or an investor perception that such uncertainties are comparatively more manageable. The presence of episodes in which TRRAI signals short-run risk-off while BIST produces a positive CAR indicates that equity prices may simultaneously reflect expectation management and perceived policy direction/normalization. This pattern underscores that reading risk perceptions through a single market segment can be insufficient and that TRRAI's multi-component structure plays a complementary role in capturing such differences.

When considered jointly, the behavior of the risk-perception index and equity-market responses points to a tight link between local market dynamics and global investor sentiment during risk-off regimes. In periods when uncertainty becomes global, pricing reactions are sharper and more synchronized; in more locally driven episodes, effects appear more episodic and sometimes less persistent. This reinforces the well-established sensitivity of emerging markets to global capital flows and international risk sentiment. The fact that many of the sharpest TRRAI declines coincide with negative short-window BIST CARs indicates rapid and often aligned transmission of deteriorating risk perceptions into equities. Importantly, the component decomposition shows that this co-movement does not necessarily arise through the same channel: CDS/FX can dominate in some types, equities in others, and volatility in global episodes.

Overall, the evidence suggests that risk-off events induce strong short-run price adjustments in Türkiye's equity market, while medium- and longer-horizon effects depend on the event type. Finally, we note that inference can be constrained by limited statistical power in categories with few events. Accordingly, conclusions are drawn by jointly considering economic magnitudes, window sensitivity, and the consistency between the equity-based AR–CAR evidence and the multi-market CAΔTRRAI evidence.

From an investor perspective, the results highlight the importance of portfolio diversification and event-based risk management; for policymakers, they underscore the need to monitor sources of uncertainty closely to safeguard financial stability. More broadly, the study contributes to the literature by demonstrating that risk-perception-based composite indicators provide a powerful tool for analyzing financial fragility in emerging markets. This contribution is strengthened by the joint use of BIST (the pricing channel) and TRRAI (the common risk component across multiple market segments), which enables the identification of both the direction of risk-off regimes and the composition of their transmission mechanisms.

The findings are consistent with the view that uncertainty shocks are priced through the risk-premium channel. In particular, for domestically driven uncertainty sources—such as institutional disruptions, protests, and security shocks—the ability of TRRAI to generate stronger risk-off signals via the CDS and exchange-rate components indicates that equities alone may not fully capture the underlying shift in risk perceptions. This reinforces the importance of measuring risk appetite through multiple market segments in economies such as Türkiye, where dollarization and sovereign risk-premium sensitivity are high. Policy-regime episodes, especially within E7, further illustrate that equity repricing and broader multi-market risk perceptions may adjust through different channels and on different horizons. This

observation is consistent with the type-level heterogeneity reported in Table 8 and with the timing differences visible in the event-time profiles.

These conclusion-level implications are consistent with related literatures in ways that mirror our empirical findings. First, our evidence that BIST and TRRAI can diverge for the same episodes—and that component contributions vary by event type—suggests that single-market proxies may be insufficient and that multi-market composite measures can provide complementary information (Illing & Liu, 2006; Holló et al., 2012). Second, our heterogeneous responses around political and geopolitical episodes, together with patterns consistent with partial pre-pricing in some cases, align with evidence that policy uncertainty and geopolitical risk are important channels through which uncertainty is priced in financial markets (Baker et al., 2016; Caldara & Iacoviello, 2022). Third, where domestically driven episodes involve social unrest/protests and security-related shocks, our results indicate pronounced short-run risk-off signals—at times captured more clearly in TRRAI than in equities alone—consistent with findings that protest-related instability can be associated with adverse market responses in emerging economies (Barrett et al., 2021; Barrett et al., 2024). Finally, our component decomposition shows that CDS and exchange-rate channels can dominate the aggregate risk-off signal in several event types, which is consistent with broader evidence that sovereign CDS–exchange rate spillovers are time-varying and can strengthen during "market fear" conditions (Feng et al., 2021) and with Türkiye-specific evidence of volatility connectedness across Turkish CDS spreads and FX-related markets that varies across volatility states and major events (Gök et al., 2023). In addition, the stronger and more synchronized responses observed during global shock episodes are consistent with Türkiye-focused evidence that foreign portfolio flows and global risk aversion are closely linked and can materially shape local market dynamics (Kartal et al., 2022; Güneş et al., 2024).

References

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives, 21*(2), 129–152. https://doi.org/10.1257/jep.21.2.129

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics, 131*(4), 1593–1636. https://doi.org/10.1093/qje/qjw024

Barrett, P., Bondar, M., Chen, S., Chivakul, M., & Igan, D. (2021, March 19). *Pricing protest: The response of financial markets to social unrest* (IMF Working Paper No. 2021/079). International Monetary Fund. https://doi.org/10.5089/9781513572765.001

Barrett, P., Bondar, M., Chen, S., Chivakul, M., & Igan, D. (2024). Pricing protest: The response of financial markets to social unrest. *Review of Finance*, *28*(4), 1419–1450. https://doi.org/10.1093/rof/rfae008

Bekaert, G., & Hoerova, M. (2014). The VIX, the variance premium and stock market volatility. *Journal of Econometrics, 183*(2), 181–192. https://doi.org/10.10.1016/j.jeconom.2014.05.008

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science, 2*(1), 1–8. https://doi.org/10.1016/j.jocs.2010.12.007

Brown, S. J., & Warner, J. B. (1985). Using daily stock returns: The case of event studies. *Journal of Financial Economics, 14*(1), 3–31.https://doi.org/10.1016/0304-405X(85)90042-X

Caldara, D., & Iacoviello, M. (2022). Measuring geopolitical risk. *American Economic Review, 112*(4), 1194–1225. https://doi.org/10.1257/aer.20191823

Eden, L., Miller, S. R., Khan, S., Weiner, R., & Li, D. (2022). The event study in international business research: Opportunities, challenges, and practical solutions. *Journal of International Business Studies*, *53*(5), 803–817. https://doi.org/10.1057/s41267-022-00509-7

Ekinci, A. (2013). Financial stress index for Turkey. *Doğuş Üniversitesi Dergisi, 14*(2), 213–229.

Engle, R. F., & Campos-Martins, J. (2023). Measuring and hedging global common volatility. *Journal of Financial Economics*, *147*(1), 221–242. https://doi.org/10.1016/j.jfineco.2022.10.009

Feng, C., Sun, J., Liu, S., & Li, J. (2021). Spillovers between sovereign CDS and exchange rate markets: The role of market fear. *The North American Journal of Economics and Finance*, 55, Article 101308. https://doi.org/10.1016/j.najef.2020.101308

Gök, R., Bouri, E., & Gemici, E. (2023). Volatility spillovers between sovereign CDS and futures markets in various volatility states: Evidence from an emerging economy around the pandemic. *Research in International Business and Finance, 66*, Article 102023. https://doi.org/10.1016/j.ribaf.2023.102023

Güneş, D., Özkan, İ., & Erden, L. (2024). Investor sentiment and foreign portfolio flows: Evidence from Türkiye. *Central Bank Review*, *24*(1), 100147. https://doi.org/10.1016/j.cbrev.2024.100147

Holló, D., Kremer, M., & Lo Duca, M. (2012, March). *CISS—A composite indicator of systemic stress in the financial system* (ECB Working Paper No. 1426). European Central Bank. https://www.ecb.europa.eu › scpwps › ecbwp1426

Illing, M., & Liu, Y. (2006). Measuring financial stress in a developed country: An application to Canada. *Journal of Financial Stability, 2*(3), 243–265. https://doi.org/10.1016/j.jfs.2006.06.002

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–291. https://doi.org/10.1142/9789814417358_0006

Kartal, M. T., Ertuğrul, H. M., & Ulussever, T. (2022). The impacts of foreign portfolio flows and monetary policy responses on stock markets by considering COVID-19 pandemic: Evidence from Turkey. *Borsa Istanbul Review, 22*(1), 12–19. https://doi.org/10.1016/j.bir.2021.06.003

Kuttner, K. N. (2001). Monetary policy surprises and interest rates: Evidence from the Fed funds futures market. *Journal of Monetary Economics, 47*(3), 523–544. https://doi.org/10.1016/S0304-3932(01)00055-1

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics, 47*(1), 13–37. https://doi.org/10.2307/1924119

MacKinlay, A. C. (1997). Event studies in economics and finance. *Journal of Economic Literature, 35*(1), 13–39. https://www.jstor.org/stable/2729691

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance, 7*(1), 77–91.

Ranco, G., Aleksovski, D., Caldarelli, G., Grčar, M., & Mozetič, I. (2015). The effects of Twitter sentiment on stock price returns. *PLoS ONE, 10*(9), e0138441. https://doi.org/10.1371/journal.pone.0138441

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory, 13*(3), 341–360. https://doi.org/10.1016/0022-0531(76)90046-6

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance, 19*(3), 425–442. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

Shiller, R. J. (2015). *Irrational exuberance* (3rd ed.). Princeton University Press.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance, 62*(3), 1139–1168. https://doi.org/10.1111/j.1540-6261.2007.01232.x

Appendix

Appendix 1:

Table A

Event timing classification and event-day (t = 0) definition

Calendar Event Date	Event Title	Event Type	Event Day Used (t = 0)	Shift (Trading Days)	Timing Classification
2008-09-15	Global financial crisis (post-Lehman shock)	E10	2008-09-15	0	Trading-hours-or-same-day
2010-09-12	Constitutional Referendum	E3	2010-09-13	1	After-close-or-non-trading-day
2011-03-15	Syrian civil war / onset of uprising	E6	2011-03-15	0	Trading-hours-or-same-day
2013-05-28	Gezi Park protests (onset)	E4	2013-05-28	0	Trading-hours-or-same-day
2013-12-17	17–25 December episode (first wave)	E3	2013-12-17	0	Trading-hours-or-same-day
2015-07-20	Suruç attack	E5	2015-07-20	0	Trading-hours-or-same-day
2015-10-10	Ankara Train Station attack	E5	2015-10-12	1	After-close-or-non-trading-day
2015-11-24	Downing of Russian jet	E6	2015-11-24	0	Trading-hours-or-same-day
2016-06-28	Atatürk Airport attack	E5	2016-06-28	0	Trading-hours-or-same-day
2016-07-15	Coup attempt	E5	2016-07-15	0	Trading-hours-or-same-day
2017-04-16	Presidential system referendum	E3	2017-04-17	1	After-close-or-non-trading-day
2018-08-10	2018 FX and debt crisis (critical wave)	E1	2018-08-10	0	Trading-hours-or-same-day
2019-10-09	Operation Peace Spring	E6	2019-10-09	0	Trading-hours-or-same-day
2020-03-11	COVID-19 (first case announcement in Türkiye)	E10	2020-03-11	0	Trading-hours-or-same-day
2020-08-10	Eastern Mediterranean tensions (NAVTEX / Oruç Reis phase)	E6	2020-08-10	0	Trading-hours-or-same-day
2021-04-22	Thodex crisis (withdrawals halted / collapse)	E7	2021-04-22	0	Trading-hours-or-same-day
2021-12-20	Introduction of FX-protected deposits	E7	2021-12-20	0	Trading-hours-or-same-day
2022-09-05	Inflation exceeds 80% (August CPI release)	E8	2022-09-05	0	Trading-hours-or-same-day
2023-02-06	Kahramanmaraş earthquakes	E9	2023-02-06	0	Trading-hours-or-same-day
2023-05-28	Presidential election (runoff)	E2	2023-05-29	1	After-close-or-non-trading-day
2023-06-22	Start of monetary tightening cycle	E7	2023-06-22	0	Trading-hours-or-same-day

Calendar Event Date	Event Title	Event Type	Event Day Used (t = 0)	Shift (Trading Days)	Timing Classification
2023-07-20	Policy rate hike (continuation)	E7	2023-07-20	0	Trading-hours-or-same-day
2024-03-31	Local elections	E2	2024-04-01	1	After-close-or-non-trading-day
2025-08-23	Termination of FX-protected deposits	E7	2025-08-25	1	After-close-or-non-trading-day

Event day (t = 0) is defined as the first trading day on which the event could be incorporated into market prices. If the calendar event date falls on a non-trading day or if the first observable pricing occurs on the next trading day, the event day is shifted accordingly. The "Shift (Trading Days)" column reports the number of trading days between the calendar event date and the event day used in the analysis.

Note. Created by the author

Information About the Article/Makale Hakkında Bilgiler

The Ethical Rules for Research and Publication / Araştırma ve Yayın Etiği

The author declared that the ethical rules for research and publication followed while preparing the article.

Yazar makale hazırlanırken araştırma ve yayın etiğine uyulduğunu beyan etmiştir.

Conflict of Interests/ Çıkar Çatışması

The author have no conflict of interest to declare.

Yazar çıkar çatışması bildirmemiştir.

Grant Support/ Finansal Destek

The author declared that this study has received no financial support.

Yazar bu çalışma için finansal destek almadığını beyan etmiştir.

Author Contributions/ Yazar Katkıları

The draft process of the manuscript/ Taslağın Hazırlanma Süreci A.Ö, Data Collection/Verilerin Toplanması A.Ö, Writing the Manuscript/ Makalenin Yazılması A.Ö, Submit, Revision and Resubmit Process/ Başvuru, Düzeltme ve Yeniden Başvuru Süreci A.Ö.

Journal of Applied and Theoretical Social Sciences

JATSS, 2026; 8(1), 78-104

First Submission:05.01.2026
Revised Submission After Review:16.03.2026
Accepted For Publication:24.03.2026
Available Online Since:25.03.2026

Research Article

The "Eco-Chameleon" Effect in Fast Fashion: Assessing Consumer Deception Risks via a Multi-Criteria Approach

Mahmut Selami Akın[a] & Serkan Eti[b]

Abstract

Introduction: Sustainability has emerged as a central concern in marketing science. In the fast fashion industry, however, it is complicated by widespread greenwashing practices. Although existing literature primarily evaluates greenwashing through linguistic ambiguities and corporate motivations, a deeper understanding of its perceptual dynamics remains necessary. The present study integrates marketing analytics with greenwashing tactics to investigate the mechanisms underlying manipulative environmental claims and their effects on consumer deception.

Method: A quantitative and updated multi-criteria approach was employed using the Simple Weight Calculation (SIWEC) method. Seven distinct greenwashing tactics were evaluated by marketing experts and ranked through SIWEC, systematically assessing their relative influence on consumer perception and the associated risks of consumer deception.

Results or Findings: The results indicate that "False Labeling" and "Visual Manipulation" are the most influential tactics used to mislead consumers. The analysis demonstrates that, instead of relying exclusively on text-based claims, companies are increasingly employing non-verbal, nature-themed visual fabrications, referred to as "Eco-Chameleon" strategies, to manipulate consumer perceptions.

Discussion or Conclusion: In contrast to traditional research that primarily emphasizes linguistic deception, the study highlights the influential manipulative role of visual imitation on the consumer decision-making process.The findings have substantial implications for marketing practitioners and policymakers, indicating that regulatory frameworks should be updated to explicitly address visual semiotics in addition to reported sustainability claims to enhance market transparency and protect consumer welfare.

Keywords: consumer behavior, greenwashing, fast fashion, sustainability, marketing ethics, SIWEC method.

JEL Codes: M30, M31, M39

[a] Asist. Prof. Dr., Istanbul Medipol University, Vocational School of Social Sciences, Marketing and Advertising Department, Istanbul/Türkiye, selamiakin@medipol.edu.tr, ORCID ID: https://orcid.org/0000-0002-8175-4618 (Corresponding Author)

[b] Assoc. Prof.Dr., Istanbul Medipol University, Vocational School, Computer Programming Department, Istanbul/Türkiye, seti@medipol.edu.tr, ORCID ID: https://orcid.org/0000-0002-4791-4091

Journal of Applied and Theoretical Social Sciences

JATSS, 2026; 8(1), 78-104

İlk Başvuru: 05.01.2026
Düzeltilmiş Makalenin Alınışı: 16.03.2026
Yayın İçin Kabul Tarihi: 24.03.2026
Online Yayın Tarihi: 25.03.2026

Araştırma Makalesi

Hızlı Modada 'Eko-Bukalemun' Etkisi: Tüketici Aldatma Risklerinin Çok Kriterli Bir Yaklaşımla Değerlendirilmesi

Mahmut Selami Akın[a] & Serkan Eti[b]

Öz

Giriş: Sürdürülebilirlik, pazarlama biliminde temel bir odak noktası olarak öne çıkmaktadır. Ancak hızlı moda endüstrisinde bu durum, yaygın yeşil aklama (greenwashing) uygulamaları nedeniyle karmaşık bir hal almaktadır. Mevcut literatür yeşil aklamayı ağırlıklı olarak dilsel belirsizlikler ve kurumsal motivasyonlar üzerinden değerlendirse de, bu kavramın algısal dinamiklerinin daha derinlemesine anlaşılmasına halen ihtiyaç duyulmaktadır. Bu çalışma, manipülatif çevresel iddiaların altında yatan mekanizmaları ve bunların tüketicinin aldatılması üzerindeki etkilerini incelemek amacıyla pazarlama analitiği ile yeşil aklama taktiklerini entegre etmektedir.

Yöntem: Basit Ağırlık Hesaplama (SIWEC) yöntemi kullanılarak nicel ve güncel çok kriterli bir yaklaşım benimsenmiştir. Yedi farklı yeşil aklama taktiği pazarlama uzmanları tarafından değerlendirilmiş ve SIWEC aracılığıyla sıralanarak; bu taktiklerin tüketici algısı üzerindeki göreceli etkileri ile barındırdıkları tüketiciyi aldatma riskleri sistematik olarak ölçülmüştür.

Sonuçlar ya da Bulgular: Sonuçlar, "Sahte Etiketleme" ve "Görsel Manipülasyon"un tüketicileri yanıltmak için kullanılan en etkili taktikler olduğunu göstermektedir. Analiz, şirketlerin yalnızca metin tabanlı iddialara dayanmak yerine, tüketici algılarını manipüle etmek için giderek artan bir oranda "Eko-Bukalemun" stratejileri olarak adlandırılan sözsüz ve doğa temalı görsel kurgulara başvurduğunu ortaya koymaktadır.

Tartışma ya da Yapılan Çıkarımlar: Araştırma, ağırlıklı olarak dilsel belirsizliğe odaklanan geleneksel çalışmaların aksine, görsel kurgunun tüketicilerin karar verme süreçleri üzerindeki manipülatif rolüne dikkat çekmiştir. Bulgular ise pazar şeffaflığını yeniden sağlamak ve tüketici refahını korumak için yasal düzenlemelerin görsel göstergeleri de kapsaması gerektiğini önererek sosyal politika ve pazarlama etiği alanlarına önemli katkılar sunmaktadır.

Anahtar Kelimeler: tüketici davranışı, yeşil aklama, hızlı moda, sürdürülebilirlik, pazarlama etiği, SIWEC yöntemi.

JEL Kodları: M30, M31, M39.

[a] Dr. Öğr. Üye., İstanbul Medipol Üniversitesi, Sosyal Bilimler Meslek Yüksekokulu, Pazarlama ve Reklamcılık Bölümü, İstanbul/Türkiye, selamiakin@medipol.edu.tr, ORCID ID: https://orcid.org/0000-0002-8175-4618 (Sorumlu Yazar)

[b] Doç. Dr., İstanbul Medipol Üniversitesi, Meslek Yüksekokulu, Bilgisayar Programcılığı Bölümü, İstanbul/Türkiye, seti@medipol.edu.tr, ORCID ID: https://orcid.org/0000-0002-4791-4091

Introduction

The global fashion industry constitutes one of the most dominant sectors in the world economy, with the apparel market revenue reaching approximately 1.79 trillion USD in 2024 and projected to exceed 2.26 trillion USD by 2030. Within this expansive landscape, the fast fashion segment has demonstrated remarkable momentum, valued at approximately $148 billion USD in 2024, showing a rapid increase to 150.82 billion USD. It is expected to grow at a compound annual growth rate (CAGR) of over 10% in the coming decade, driven largely by consumption trends in the Asia-Pacific and North American regions. This industrial expansion is mirrored by a massive surge in production volume, which has effectively doubled since 2000, now exceeding 100 billion garments annually worldwide (Tcholakova, 2025; Thuy, 2025). However, parallel to this mass production, the sustainable fashion market is witnessing an unprecedented rise. Valued at approximately 7.9 billion USD in 2024, this niche market is projected to expand at a CAGR of 23.5%, reaching nearly 33 billion USD by 2033, reflecting a structural shift in consumer paradigms toward ethical consumption (Tcholakova, 2024). Shein dominates the fast fashion market in the United States with a 50% share, followed by H&M (16%) and Zara (13%) (Cardona, 2025).

Despite its economic contributions, the fashion industry, particularly fast fashion, is widely seen as a major contributor to global environmental degradation. The sector uses excessive water, causes chemical pollution, and generates large amounts of textile waste. It creates a paradox between industrial growth and ecological preservation (Mousavi et al., 2024). Meanwhile, consumer environmental awareness is rising. Modern consumers now scrutinize the ecological footprint of their purchases and increasingly support brands that share their environmental values. The growing sustainability awareness pressures firms to disclose their environmental performance. As a result, "greenness" has become a significant competitive advantage in the crowded marketplace (Badhwar et al., 2024). However, consumers who aim for sustainable consumption often face complex informational barriers (Testa et al., 2018). In fashion, these barriers are worsened by deceptive marketing that exploits consumers' inability to verify environmental claims.

The lucrative nature of the green market has incentivized some firms to manipulate consumer demand through deceptive marketing strategies, a phenomenon known as greenwashing. Grounded in Signaling Theory, greenwashing can be understood as a strategic abuse of information asymmetry, where firms transmit false or exaggerated positive signals about their environmental performance to consumers who lack the expertise to verify these claims (Santos et al., 2024b). In the context of fast fashion, where supply chains are complex and opaque, brands frequently exploit this asymmetry to gain unwarranted legitimacy. By presenting products as more environmentally friendly than they actually are, firms aim to capitalize on the green trend without incurring the actual costs of sustainable production, thereby misleading stakeholders and distorting market competition (Alizadeh et al., 2024).

As greenwashing tactics become more sophisticated, fast fashion companies are increasingly adopting an “Eco-Chameleon” strategy. This approach involves manipulating semiotic codes, including fake certifications, ambiguous eco-friendly terminology, and nature-inspired aesthetics, to circumvent consumer skepticism and superficially integrate into the sustainable market. Analogous to the historical evolution of discursive strategies in the fossil fuel industry that delayed energy transitions (Gentile & Gupta, 2025), fast fashion brands may employ these deceptive visual and linguistic cues to preserve market dominance without implementing substantive environmental improvements.

Moreover, the landscape of greenwashing has evolved from simple false claims to more sophisticated and covert tactics, challenging the consumer's ability to distinguish between genuine and fake sustainability. Brands utilize a diverse array of strategies, ranging from vague terminology (e.g., conscious, eco-friendly) and visual manipulation (e.g., using green themes or nature imagery) to hiding trade-offs and price-based deception (Ende et al., 2023; Mathew & Spinelli, 2025). According to Perceived Risk Theory, this proliferation of misleading tactics significantly increases the perceived risk and confusion among consumers, fostering a deep-seated skepticism toward all green claims (Lu et al., 2022). As consumers struggle to decipher the validity of environmental assertions, their trust in the industry erodes, leading to a phenomenon where even genuinely sustainable brands face scrutiny and skepticism (Nguyen et al., 2019).

Although the existing literature has extensively documented the existence of greenwashing and its general negative impact on consumer trust and brand equity (Munir & Mohan, 2022; Xiao et al., 2022), there remains a notable gap regarding the comparative risk analysis of specific greenwashing tactics in the fashion industry. Most prior studies have focused on the binary presence of greenwashing or its broad classification, yet there is a lack of quantitative research that prioritizes these tactics based on their potential to deceive consumers. Specifically, it is not sufficiently clear which specific tactic, such as false labeling versus vague claims, poses the highest risk of consumer deception in the fast fashion context. Understanding the hierarchy of these risks is crucial for developing targeted interventions. Therefore, multi-criteria decision-making (MCDM) approaches should be highly effective in uncovering the greenwashing implementations causing consumer deception.

Recent studies have increasingly adopted quantitative and multi-criteria decision-making MCDM methods to dissect the complex mechanisms of greenwashing across various high-impact industries. Daga et al. (2025) utilized TISM and MICMAC analyses to model the interaction of factors discouraging ESG greenwashing, highlighting the systemic relationships between regulatory pressures and internal corporate drivers. In the construction sector, Li et al. (2025) employed agent-based modeling integrating new institutionalism theory to demonstrate how government regulation and public scrutiny synergistically curb greenwashing behaviors, while an earlier study by Li et al. (2023) used the Grey-DEMATEL method to map the causal relationships of these influencing factors, identifying environmental regulation as a critical determinant. Focusing on corporate governance in the financial sector, Poiriazi et al. (2025) integrated Entropy Weight and TOPSIS methods to rank the influence of board characteristics on greenwashing mitigation, finding that specific board attributes significantly affect transparency.

Moreover, Sklavos et al. (2025) assessed the eco-efficiency of European financial institutions using a hybrid DEA-CRITIC-TOPSIS approach, establishing a performance ranking based on green accounting and ESG indicators. Furthermore, within the context of digital transformation, Xu et al. (2025) applied a multi-period difference-in-differences (DID) model to reveal that digital-driven mergers and acquisitions significantly reduce corporate greenwashing among Chinese listed firms by enhancing governance mechanisms. Despite the growing application of MCDM techniques to evaluate greenwashing drivers in the finance and construction sectors, there remains a significant scarcity of research applying these robust decision-making tools to the fast fashion industry, specifically to prioritize consumer deception tactics. This study addresses this gap by employing the Simple Weight Calculation (SIWEC) method to establish a risk hierarchy of specific greenwashing tactics, thereby offering a novel quantitative perspective on consumer deception in the fashion context.

Expanding upon previous methodologies, the present study applies the Simple Weight Calculation (SIWEC) method, a robust MCDM approach, to consumer deception tactics, thereby complementing existing research and redirecting the focus from corporate causalities to market outcomes. By ranking these tactics according to their deceptive impact, the study offers a clearer understanding of the role of greenwashing on consumer decision-making and market dynamics. Consequently, the study aims to fill the gap by prioritizing greenwashing tactics according to the risk of consumer deception. To address this critical gap, the central research question of the study is below.

> RQ: How do different greenwashing tactics in the fast fashion industry compare in their potential to generate consumer deception risk, and which specific non-verbal or textual strategies present the greatest threat to sustainable consumption?

Furthermore, several criteria were defined in the research through a comprehensive literature review, enabling a novel quantitative ranking of greenwashing practices. The greenwashing criteria derived from the comprehensive literature review were subsequently evaluated and prioritized by a panel consisting of academic and sectoral experts. On the other hand, the findings should provide actionable insights for policymakers to regulate the most detrimental practices and for comprehending sustainable consumer behavior to improve green literacy, thereby supporting a more transparent and trustworthy sustainable fashion ecosystem. Additionally, for fast fashion brands, these insights offer strategic guidance to align marketing communications with authentic sustainability initiatives, reducing reputational risks and promoting consumer trust and long-term brand loyalty in a market characterized by increasing consumer skepticism. Ultimately, adopting research recommendations positions all stakeholders to drive meaningful progress toward a genuinely sustainable fashion industry.

Literature Review

Theoretical Framework

To understand the dynamics of greenwashing in the fast fashion industry and its impact on consumers, this study draws upon two complementary theories: Signaling Theory and Perceived Risk Theory. These theoretical lenses provide a robust basis for explaining why firms engage in deceptive practices and how consumers process and react to these misleading environmental claims.

Signalling Theory, initially developed within economics and finance (Spence, 1973) to clarify the resolution of information asymmetry in labour and financial markets, addresses challenges that arise when two parties, such as sellers and buyers, possess unequal information. The foundational theory has been widely adopted in marketing literature to explain how brands communicate in sustainable marketing strategy (Taoketao et al., 2018) and unobservable product attributes (Wang et al., 2020), including eco-labels (Sigurdsson et al., 2022) and environmental friendliness (Cao & Zhang, 2024). In the fashion industry, information asymmetry is particularly prevalent; brands have comprehensive knowledge of their supply chains and production processes, whereas consumers typically lack the expertise or resources to verify a product's environmental attributes prior to purchase (Alizadeh et al., 2024). To address this disparity, firms employ signals such as eco-labels, sustainability reports, and green marketing campaigns to communicate these otherwise unobservable quality attributes to buyers.

However, greenwashing emerges when these signals are manipulated. According to recent literature, dishonest firms in the fast fashion sector exploit information asymmetry by transmitting false or exaggerated signals to gain a competitive advantage without incurring the cost of genuine sustainable practices (Santos et al., 2024a). For instance, a conscious collection tag acts as a positive signal, suggesting environmental responsibility. When this signal does not correspond to the actual environmental performance, it transforms into a deceptive mechanism that distorts market efficiency and misleads stakeholders (Badhwar et al., 2024).

Within the context of manipulated signals, the study introduces and highlights the 'Eco-Chameleon' concept. The 'Eco-Chameleon' effect describes the highly adaptive and deceptive strategies that corporations use to superficially integrate into the sustainable market without implementing substantive operational changes. Analogous to a chameleon changing its colours to blend into its surroundings, fast fashion brands manipulate external semiotic codes (Ertürk, 2024), including visual aesthetics (Basso et al., 2023), packaging (Boncinelli et al., 2023), and marketing rhetoric (Siew et al., 2024), to create an illusion of ecological responsibility (Roos, 2025). This phenomenon extends beyond traditional false advertising; it constitutes a systemic adaptation in which firms continually refine deceptive cues to circumvent growing consumer skepticism and regulatory oversight, resulting in a significant disconnect between brand image and actual environmental performance (Gallas et al., 2025; Ha et al., 2022; Tu et al., 2024).

While Signaling Theory and the Eco-Chameleon effect explains the firm's behavior, Perceived Risk Theory elucidates the consumer's reaction. First introduced by Bauer (1960), this theory posits that consumer behavior is driven by the desire to minimize the uncertainty and potential negative consequences associated with a purchase. In the realm of sustainable fashion, greenwashing significantly amplifies consumers' perceived risk, specifically the risk of being deceived or making an unethical choice despite good intentions (Lu et al., 2022).

When consumers are exposed to inconsistent or vague green claims (e.g., eco-friendly without certification), their skepticism increases, and they perceive a higher risk of psychological loss and functional loss such as buying a product that does not perform as promised (Wang & Jung, 2025). Consequently, the prevalence of greenwashing tactics in the fast fashion industry not only erodes trust in specific brands but also creates a generalized green skepticism, making it difficult for genuinely sustainable brands to signal their legitimacy effectively (Akturan, 2018). Moreover, recent research emphasize that misleading environmental information does not merely cause confusion but triggers a sense of brand betrayal among consumers, directly leading to boycotts and reputational damage (Zhong & Kim, 2025). Therefore, identifying the specific tactics that create this deception is crucial for market stability. The present study utilizes these theories to categorize and prioritize specific greenwashing tactics based on the magnitude of the risk they pose to the consumer.

Taxonomy of Greenwashing Tactics

To systematically analyze the risk of consumer deception in the fast fashion industry, this study synthesizes the dispersed tactics identified in recent literature into two distinct cognitive categories: (1) *Claim and Evidence-Based Tactics*, which rely on explicit textual misinformation and unverifiable assertions; and (2) *Associative and Structural Manipulations*, which exploit psychological heuristics through visual cues and pricing strategies.

Consequently, seven specific greenwashing tactics emerged through the literature review. These were not determined arbitrarily; rather, they were systematically derived from a comprehensive review of the current sustainability and marketing literature. Additionally, they represent the most prevalent and impactful deceptive tactics specifically utilized within the fast

fashion industry. Other potential greenwashing criteria were excluded from the scope of this study to maintain a focused evaluation on the tactics that most directly manipulate consumer perception and visual semiotics in the fast fashion industry.

Claim and Evidence-Based Tactics

This category encompasses tactics where brands manipulate the direct information provided to consumers. These practices exploit the consumer's reliance on linguistic cues and certifications to assess environmental legitimacy.

- False Labeling: False labeling represents the most explicit and legally risky form of greenwashing. It involves the use of fake labels, counterfeit certificates, or misleading text that claims third-party endorsement where none exists. In the fast fashion sector, this often manifests as brands self-declaring a product as "Certified Organic" or "100% Recycled" without valid verification from recognized bodies such as the Global Organic Textile Standard (GOTS) (Mousavi et al., 2024). Wang and Jung (2025) argue that false labeling poses a severe cognitive risk because consumers utilize labels as primary heuristic cues to reduce decision-making effort; when these cues are fabricated, the fundamental trust mechanism is breached. Furthermore, recent literature highlights the need for technological interventions to address greenwashing and false labeling. The use of Digital Product Passports (DPP) and blockchain-based traceability is seen as essential to reducing information asymmetry in supply chains. Scholars argue these decentralized technologies can verify market signals, increase transparency, and help regain consumer trust in sustainability claims (Acciai & Pérez-Bou, 2025; Guo et al., 2020).
- Vague Claims: Unlike false labeling, which relies on falsifiable lies, vague claims utilize semantic ambiguity to create a positive impression without legal accountability. This tactic involves the use of broad, undefined, and non-regulated terms such as "eco-friendly," "conscious," "green," or "nature-inspired" (Badhwar et al., 2024; Simão et al., 2022). Since these terms lack standardized metrics in the textile industry, brands can interpret them subjectively. For instance, a product may be labeled sustainable simply for using slightly less water than the industry average, creating a gap between the literal meaning and the consumer's interpretation of environmentally harmless (Alizadeh et al., 2024).
- Lack of Proof: This criterion refers to environmental claims that cannot be substantiated by accessible evidence. According to Mousavi et al. (2024), while brands frequently emphasize transparency in their marketing rhetoric, they often fail to provide verifiable data, such as supplier lists, carbon footprint calculations, or valid certification codes (e.g., a QR code linking to a report), at the point of purchase. The absence of proof forces consumers to rely on blind trust. When claims are not backed by accessible verification, information asymmetry widens, leaving the consumer unable to validate the product's actual environmental impact (Santos et al., 2024a).

Associative and Structural Manipulations

This category includes tactics that do not necessarily rely on explicit verbal claims but rather exploit consumers' psychological heuristics and lack of technical knowledge through visual cues, pricing strategies, and selective disclosure.

- Visual and Atmospheric Manipulation (Green Sheen): Often termed "green sheen," this tactic involves the strategic use of non-verbal cues—such as nature-related imagery

(e.g., leaves, forests) and a specific color palette dominated by greens, browns, and soft earthy tones—to evoke a false perception of sustainability. Ende et al. (2023) demonstrate that fast fashion brands use these visual heuristics to trigger an automatic eco-friendly association in the consumer's mind. Even without explicit claims, this atmospheric illusion masks the synthetic nature of the garments, leading consumers to perceive the brand as environmentally responsible based solely on aesthetic codes (Alizadeh et al., 2024). Moreover, recent methodological discussions emphasize the need to move beyond traditional self-report measures to better understand the latent psychological impacts of visual manipulation and the 'Eco-Chameleon' effect. Researchers recommend incorporating neuroscientific methods, including neuromarketing techniques such as eye-tracking and electroencephalography (EEG), to objectively assess unconscious cognitive processing related to greenwashing. These advanced approaches are considered essential for elucidating how nature-inspired visual stimuli and green aesthetics may circumvent consumers' rational evaluation (Ćirović et al., 2024; Liu et al., 2023).

- Price-Based Perception Management (Price Deception): This tactic exploits the consumer heuristic that higher price implies higher ethical quality. In the fast fashion context, consumers often equate ultra-low prices with exploitation and poor environmental standards, conversely if higher-priced items (e.g., a Premium Sustainable Line) are genuinely eco-friendly (Ende et al., 2023). Brands manipulate this perception by artificially inflating the prices of their green collections to signal legitimacy. This price-based signaling creates a false sense of ethical consumption, where the consumer pays a premium not for verified environmental benefits, but for the alleviation of guilt (Wang & Jung, 2025).

- Hidden Trade-off (The Capsule Collection Paradox): Uniquely prevalent in the fashion industry, this tactic involves promoting a small, exclusive collection made from sustainable materials (e.g., organic cotton) as the face of the brand's identity, while the vast majority of production continues to rely on resource-intensive conventional materials (Munir & Mohan, 2022). By disproportionately advertising this niche capsule collection, brands create a "halo effect" that greenwashes the entire organization (Hora & Subramanian, 2019). This distracts stakeholders from the unsustainable reality of the brand's core business model (Santos et al., 2024a).

- Downsizing Negative Aspects: This tactic involves highlighting a minor positive environmental attribute while deliberately concealing significant negative impacts (Wang & Jung, 2025). For example, a brand may promote a garment made from recycled polyester to appeal to circular economy trends, while omitting the fact that the dyeing process releases toxic chemicals into water systems (Rashid & Arshad, 2023). By focusing attention on a single hero attribute, brands successfully divert scrutiny from their carbon-intensive supply chains.

Given the sophisticated nature of the mentioned greenwashing tactics, consumer behavior in sustainable fashion reveals a notable paradox within the generational context. Although current research demonstrates that Gen Z exhibits high environmental awareness and climate sensitivity, they frequently display a significant intention-behavior gap, primarily influenced by the attractiveness of low-cost, trend-driven fast fashion (Kang & Badal, 2025). This paradox indicates that younger consumers may be particularly susceptible to advanced greenwashing strategies, which complicates their capacity to make authentically sustainable purchasing decisions (Bytof & Ritch, 2023; Williams & Hodges, 2022).

The next section explains the methodology of the research.

Methodology

The present study utilizes a quantitative research design grounded in the Simple Weight Calculation (SIWEC) method, a multi-criteria decision-making (MCDM) approach developed by Puška et al. (2023). The application of the SIWEC method to consumer deception tactics shifts the analytical focus from corporate origins to market consequences. The primary aim of the methodology is to determine the relative importance weights of identified greenwashing tactics according to the risk of consumer deception. In contrast to previous research that has focused on internal firm dynamics or the existence of greenwashing, the mentioned multi-criteria approach offers a quantitative framework for assessing the hierarchical impact of various deceptive strategies on consumer decision-making in the marketplace.

The SIWEC Method

The SIWEC method was selected for this study due to its distinct advantages over traditional MCDM methods such as AHP or ANP. Unlike methods that require complex pairwise comparisons, which can lead to inconsistency and high cognitive burden for experts, SIWEC allows decision-makers to evaluate criteria directly using a defined scale. This approach ensures high consistency in results while effectively capturing expert judgements in environments characterized by subjective uncertainty (Puška et al., 2024). In addition, its use of standard deviation differentiates it from other methods. It allows experts to conduct a broader evaluation, thus leading to more realistic results. Unlike methods such as Analytic Hierarchy Process (AHP) and Analytic Network Process (ANP), SIWEC differs in this respect and also uses less data because it does not rely on pairwise comparisons in the evaluation process. While the number of items increases exponentially with increasing criteria in methods like AHP, ANP, and DEMATEL, it increases linearly in SIWEC (Özçalıcı et al., 2025). It positively impacts the process when there are many criteria (Puška et al., 2023; Pamucar et al., 2018). The method is particularly suitable for this study as it facilitates the prioritization of seven distinct greenwashing criteria without the mathematical complexity that often obscures practical insights.

Moreover, comparative studies in the literature highlight a significant limitation of traditional methods such as AHP and ANP. These methods are vulnerable to the rank reversal phenomenon and depend on enforced matrix consistency (Cinelli et al., 2020; Wang & Luo, 2009). When evaluating deceptive marketing practices, the relative eigenvector calculations in AHP frequently smooth out divergent expert opinions. This process aims to achieve an acceptable consistency ratio, yet it should result in misclassification of greenwashing risks. Evidence from the literature demonstrates that methods emphasizing absolute evaluation and direct variance, such as SIWEC, yield fundamentally different and more robust outcomes. Instead of imposing a relative trade-off between essential criteria such as 'False Labeling' and 'Visual Manipulation', SIWEC maintains genuine divergence in expert risk assessments. This approach ensures final rankings more accurately represent the complexities of real-world markets, rather than reflecting mathematical compromises imposed by matrix normalization (Munier & Hontoria, 2021).

The next section is about the formulation of methodology, including the Fermatean fuzzy SIWEC for calculating the importance of criteria.

Fermatean Fuzzy Sets

Under uncertainty, fuzzy logic is used to perform calculations with words. Within the scope of fuzzy logic, the membership and non-membership degrees of an object to a set, along with hesitancy, are considered. In Fermatean fuzzy sets, the third powers of membership (μ) and non-membership (ϑ) degrees are used. This definition offers a unique approach to expressing unreliable, imprecise, and ambiguous information in a fuzzy environment.

A Fermatean fuzzy set $(\tilde{F})$ in a set S is represented in Equation (1).

$$\tilde{F} = \{u, \mu_{\tilde{F}}(u), \vartheta_{\tilde{F}}(u) : u \in S\} \quad (1)$$

Where μ and ϑ are between zero and one. The condition in Fermatean fuzzy sets is shown in Equation (2).

$$0 \leq \mu_{\tilde{F}}(u)^3 + \vartheta_{\tilde{F}}(u)^3 \leq 1 \quad (2)$$

The hesitancy function is obtained Equation (3).

$$\pi_{\tilde{F}}(u) = \sqrt[3]{1 - (\mu_{\tilde{F}}(u)^3 + \vartheta_{\tilde{F}}(u)^3)} \quad (3)$$

Let $\tilde{F} = (\mu_{\tilde{F}}, \vartheta_{\tilde{F}})$ be a Fermatean fuzzy number. The score and accuracy values are defined in Equations (4) and (5), respectively.

$$S(\tilde{F}) = 1 + \mu_{\tilde{F}}^3 - \vartheta_{\tilde{F}}^3 \quad (4)$$

$$A(\tilde{F}) = \mu_{\tilde{F}}^3 + \vartheta_{\tilde{F}}^3 \quad (5)$$

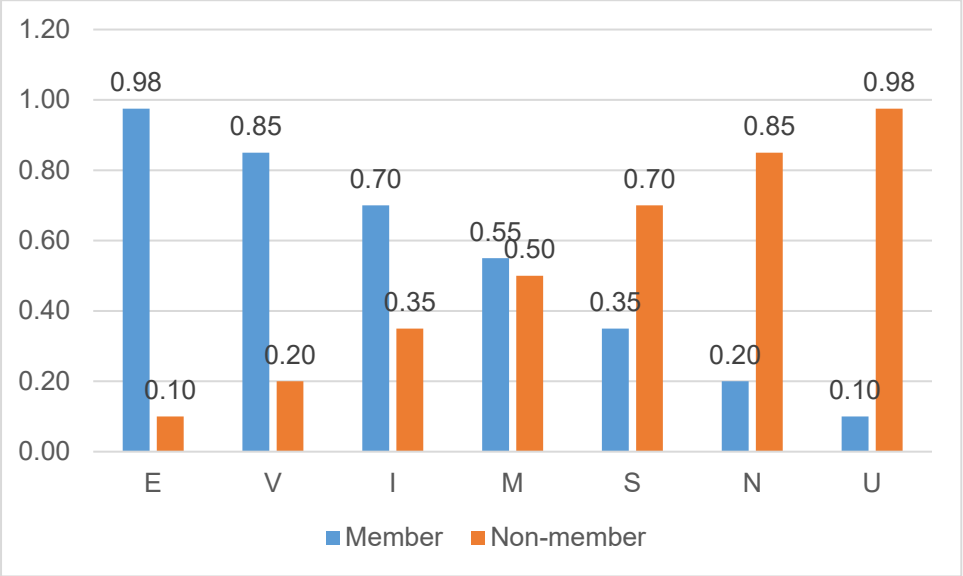
FF-SIWEC

The SIWEC method uses the standard deviation of expert ratings to determine the importance of criteria. This allows for more realistic and accurate ratings from experts. The calculation of the method using Fermatean fuzzy numbers is as follows.

Once n criteria are determined, e experts evaluate the importance of each criterion individually using linguistic variables in Figure 1.

Figure 1

Linguistic Variables



Note. Created by the authors using MS Excel.

Then these evaluations are converted into Fermatean fuzzy numbers to create the initial decision-making matrix shown in Equation (6).

$$\tilde{X} = [\tilde{x}_{ij}]_{e\times n} \tag{6}$$

Afterwards, the values are defuzzified using score function. Next, defuzzified values are normalized with Equation (7).

$$\mathfrak{n}_{ij} = \frac{S(\tilde{x}_{ij})}{\max S(\tilde{x}_{ij})} \tag{7}$$

The standard deviations of experts' normalized values are calculated via Equations (8) and (9).

$$\bar{\mathfrak{n}}_i = \frac{1}{n}\sum_{j=1}^{n} \mathfrak{n}_{ij} \tag{8}$$

$$\sigma_i = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\mathfrak{n}_{ij} - \bar{\mathfrak{n}}_j\right)^2} \tag{9}$$

Next, the expert-weighted normalized values are computed with the help of Equation (10).

$$p_{ij} = \pi_{ij}\sigma_i \tag{10}$$

Finally, the expert-weighted normalized values for each criterion are summed by Equation (11), and then, the weights of criteria are estimated using Equation (12).

$$s_j = \sum_{i=1}^{e} p_{ij} \tag{11}$$

$$w_j = \frac{s_j}{\sum_{j=1}^{n} s_j} \tag{12}$$

Data Collection

The data were collected directly by the researchers through a structured online expert consultation process using Microsoft Forms, without the involvement of a third-party research company. The data collection process was conducted between December 8, 2025, and December 31, 2025. To ensure validity and reliability, strict selection criteria were used: the panel consisted of academicians specializing in sustainable marketing, sustainable fashion, and consumer behavior, each with at least 10 years of sector experience. Rigorous panel criteria are essential for reliable qualitative-to-quantitative evaluations (Langfeldt, 2002). Initially, 17 academicians responded; 3 were excluded due to insufficient experience. Thus, data from 14 highly qualified experts were analyzed.

The experts were presented with the seven greenwashing criteria identified in the literature review: False Labeling (C1), Vague Claims (C2), Lack of Proof (C3), Visual Manipulation (C4), Price Deception (C5), Hidden Trade-off (C6), and Downsizing Negative Aspects (C7). They were asked to rank these criteria based on their potential to deceive consumers, using a 7-point scale where 1 represents extremely important and 7 represents extremely unimportant. To ensure conceptual clarity, a detailed definition table for each criterion was provided to the experts prior to the evaluation.

In the SIWEC method, capturing expert judgments directly through this scale ensures high consistency; unlike traditional MCDM models (e.g., AHP) that require Consistency Ratio (CR) calculations, SIWEC inherently handles evaluation variance by utilizing the standard deviation of the responses, thereby providing a reliable ranking without the need for additional validity coefficients.

Findings

This section includes the results of the expert evaluations of greenwashing tactics. Initially, the criteria are demonstrated in Table 1 with short codes.

Table 1

Criteria List

Criterion	Definition	Short Code
False Labeling	involving creating fake third-party certification marks—meaning counterfeit badges from organizations that usually review product compliance—or using self-assigned labels to mislead consumers into believing a product meets legal or regulatory requirements.	FSLBL
Vague Claims	using broad, ill-defined terminology, such as all-natural or eco-friendly, which lacks clear scientific parameters or legal definitions.	AMBST
Lack of Proof	including environmental claims made by firms that cannot be substantiated by accessible evidence or independent verification at the point of purchase.	LKEVD
Visual Manipulation	employing non-verbal cues, such as nature-inspired imagery and green color palettes, to psychologically influence consumers to perceive a product as sustainable, irrespective of its actual environmental impact.	VSMNP
Price Deception	including artificially inflating product prices to exploit the consumer heuristic that associates higher costs with ethical production standards.	PRBPR
Hidden Trade-off	emphasizing a single positive environmental feature, such as packaging labeled 'sustainable,' to obscure more significant negative impacts elsewhere in the supply chain (the processes and steps involved in making and delivering a product).	SBCCL
Downsizing Negative Aspects	involving deliberate efforts to minimize, omit, or obscure information about the harmful environmental footprint of the company's broader operations.	HDNGT

Note. Created by the authors using MS Excel.

The evaluations are collected from 14 academics with an average of 17 years of experience. The findings are presented in Table 2.

Table 2

Evaluations

	FSLBL	AMBST	LKEVD	VSMNP	PRBPR	SBCCL	HDNGT
Academician_1	E	S	M	E	E	M	E
Academician_2	V	I	M	E	E	E	I
Academician_3	V	M	S	E	E	V	V
Academician_4	E	E	V	E	E	E	E
Academician_5	E	V	S	I	M	E	E
Academician_6	E	I	M	E	S	S	I
Academician_7	E	E	E	E	E	E	E
Academician_8	V	I	M	I	I	V	V

Academician_9	E	V	E	V	V	V	E
Academician_10	E	E	V	E	E	E	E
Academician_11	I	V	V	E	S	V	E
Academician_12	E	E	V	E	E	E	E
Academician_13	E	S	S	I	V	V	I
Academician_14	E	E	E	E	E	E	E

E=Extremely Important, V=Very Important, I=Important, M=Moderately Important, S=Slightly Important, N=Not Important, U=Extremely Unimportant

Note. Created by the authors using MS Excel.

The fuzzy numbers corresponding to the linguistic expressions in Table 2 are derived using the methodology illustrated in Figure 2. Accordingly, the linguistic evaluations are transformed into fuzzy numbers as shown in Figure 2. The resulting fuzzy-to-numerical mixtures of the evaluations are presented in Table 3.

Table 3

Initial Decision-Making Matrix

	FSLBL	**AMBST**	**LKEVD**	**VSMNP**	**PRBPR**	**SBCCL**	**HDNGT**
Academician_1	(.975,.1)	(.35,.7)	(.55,.5)	(.975,.1)	(.975,.1)	(.55,.5)	(.975,.1)
Academician_2	(.85,.2)	(.7,.35)	(.55,.5)	(.975,.1)	(.975,.1)	(.975,.1)	(.7,.35)
Academician_3	(.85,.2)	(.55,.5)	(.35,.7)	(.975,.1)	(.975,.1)	(.85,.2)	(.85,.2)
Academician_4	(.975,.1)	(.975,.1)	(.85,.2)	(.975,.1)	(.975,.1)	(.975,.1)	(.975,.1)
Academician_5	(.975,.1)	(.85,.2)	(.35,.7)	(.7,.35)	(.55,.5)	(.975,.1)	(.975,.1)
Academician_6	(.975,.1)	(.7,.35)	(.55,.5)	(.975,.1)	(.35,.7)	(.35,.7)	(.7,.35)
Academician_7	(.975,.1)	(.975,.1)	(.975,.1)	(.975,.1)	(.975,.1)	(.975,.1)	(.975,.1)
Academician_8	(.85,.2)	(.7,.35)	(.55,.5)	(.7,.35)	(.7,.35)	(.85,.2)	(.85,.2)
Academician_9	(.975,.1)	(.85,.2)	(.975,.1)	(.85,.2)	(.85,.2)	(.85,.2)	(.975,.1)
Academician_10	(.975,.1)	(.975,.1)	(.85,.2)	(.975,.1)	(.975,.1)	(.975,.1)	(.975,.1)
Academician_11	(.7,.35)	(.85,.2)	(.85,.2)	(.975,.1)	(.35,.7)	(.85,.2)	(.975,.1)
Academician_12	(.975,.1)	(.975,.1)	(.85,.2)	(.975,.1)	(.975,.1)	(.975,.1)	(.975,.1)
Academician_13	(.975,.1)	(.35,.7)	(.35,.7)	(.7,.35)	(.85,.2)	(.85,.2)	(.7,.35)
Academician_14	(.975,.1)	(.975,.1)	(.975,.1)	(.975,.1)	(.975,.1)	(.975,.1)	(.975,.1)

Note. Created by the authors using MS Excel.

Afterwards, the initial decision-making values are defuzzified using score function. Next, defuzzified values are normalized with Equation (7). The normalized matrix is shown in Table 4.

Table 4

Normalized Matrix

	FSLBL	AMBST	LKEVD	VSMNP	PRBPR	SBCCL	HDNGT
Academician_1	1.000	.363	.541	1.000	1.000	.541	1.000
Academician_2	.834	.675	.541	1.000	1.000	1.000	.675
Academician_3	.834	.541	.363	1.000	1.000	.834	.834
Academician_4	1.000	1.000	.834	1.000	1.000	1.000	1.000
Academician_5	1.000	.834	.363	.675	.541	1.000	1.000
Academician_6	1.000	.675	.541	1.000	.363	.363	.675
Academician_7	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Academician_8	.834	.675	.541	.675	.675	.834	.834
Academician_9	1.000	.834	1.000	.834	.834	.834	1.000
Academician_10	1.000	1.000	.834	1.000	1.000	1.000	1.000
Academician_11	.675	.834	.834	1.000	.363	.834	1.000
Academician_12	1.000	1.000	.834	1.000	1.000	1.000	1.000
Academician_13	1.000	.363	.363	.675	.834	.834	.675
Academician_14	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Note. Created by the authors using MS Excel.

The standard deviations of experts' normalized values are calculated via Equations (8) and (9). The standard deviation values are equal to 0.262, 0.176, 0.219, 0.058, 0.236, 0.245, 0.000, 0.105, 0.082, 0.058, 0.203, 0.058, 0.203, 0.058, 0.224, and 0.000, respectively. Then, the expert-weighted normalized values are computed with the help of Equation (10). The expert-weighted normalized matrix is demonstrated in Table 5.

Table 5

Expert-Weighted Normalized Matrix

	FSLBL	AMBST	LKEVD	VSMNP	PRBPR	SBCCL	HDNGT
Academician_1	.262	.095	.142	.262	.262	.142	.262
Academician_2	.147	.119	.095	.176	.176	.176	.119
Academician_3	.183	.118	.080	.219	.219	.183	.183
Academician_4	.058	.058	.048	.058	.058	.058	.058
Academician_5	.236	.197	.086	.159	.128	.236	.236
Academician_6	.245	.166	.133	.245	.089	.089	.166
Academician_7	.000	.000	.000	.000	.000	.000	.000
Academician_8	.087	.071	.057	.071	.071	.087	.087
Academician_9	.082	.069	.082	.069	.069	.069	.082

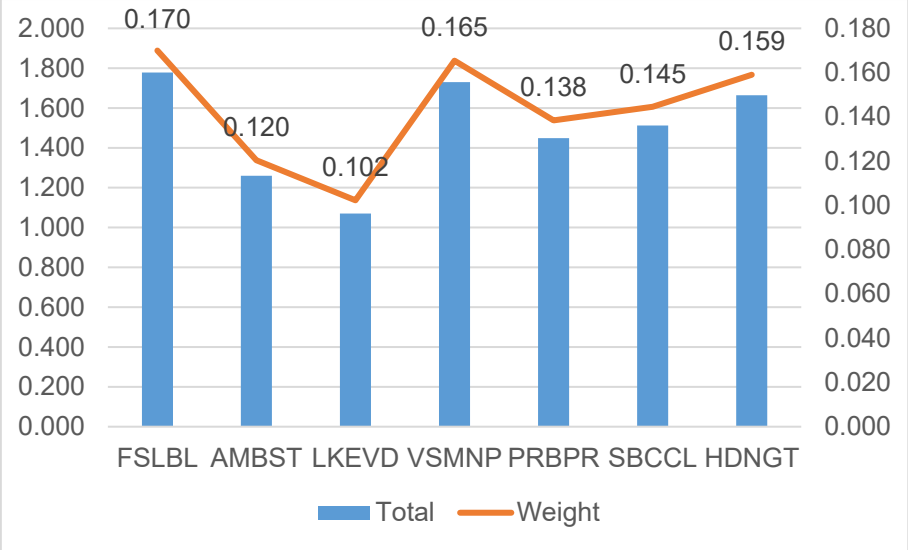
Academician_10	.058	.058	.048	.058	.058	.058	.058
Academician_11	.137	.170	.170	.203	.074	.170	.203
Academician_12	.058	.058	.048	.058	.058	.058	.058
Academician_13	.224	.081	.081	.151	.187	.187	.151
Academician_14	.000	.000	.000	.000	.000	.000	.000

Note. Created by the authors using MS Excel.

Consequently, the expert-weighted normalized values for each criterion are summed by Equation (11), and then, the weights of criteria are estimated using Equation (12). The results are visualized in Figure 1.

Figure 2

Total and Weight Values



Note. Created by the authors using MS Excel.

According to Figure 2, the most important criterion for greenwashing is False Labeling, with 0.17. The second important criterion is Visual Manipulation with 0.165. The third criterion, Downsizing Negative Aspects, has 0.159 importance weight. Then, the fourth one is Hidden Trade-off (0.145). The least important criteria are Price Deception (0.138), Vague Claims (0.120), and Lack of Proof (0.102).

Discussion

The study investigates deceptive mechanisms in the fast fashion industry by ranking greenwashing tactics based on consumer deception risk using the SIWEC method. Quantitative analysis reveals that "False Labeling" (w=0.17) and "Visual/Atmospheric Manipulation" (w=0.165) are the most significant determinants of deception risk. These results offer a refreshed outlook that contrasts with traditional research, which primarily focuses on textual ambiguity. Previous studies have often identified text-based 'Vague Claims' and insufficient information specificity as primary drivers of greenwashing perceptions and consumer skepticism (Musgrove et al., 2018; Orazi & Chan, 2020; Polonsky et al., 2025; Schmuck et al., 2018). Although Yang et al. (2020) also found vague statements to be the most common tactic

in general markets, the current findings indicate that, in the fast fashion sector, the most harmful tactics are those that leverage structural trust through labels and cognitive heuristics via visual cues.

The identification of "False Labeling" as the highest risk is consistent with the information asymmetry framework outlined by Hung and Chang (2024), who contend that consumers lack the technical expertise to verify environmental claims and must therefore depend on external signals. Eco-labels function as essential cognitive shortcuts for consumers navigating complex sustainability information, making their credibility critical (Aasmäe, 2021; Van der Ven, 2019). Recent empirical studies demonstrate that consumers rely extensively on these certification marks to mitigate perceived risks, which increases their susceptibility to deception when such labels are fabricated or unregulated (Kothiyal & Mehta, 2025). Fabrication of these signals results in a complete breach of trust. The finding is further substantiated by Busalim et al. (2025), who identified image barriers and specific doubts regarding sustainability labeling as primary drivers of consumer resistance to sustainable fashion. Therefore, false labeling constitutes not only a deceptive practice but also a structural barrier to the adoption of authentic eco-friendly products.

Kinnunen (2020) further supports the result by demonstrating that baseless claims result in green consumer confusion, which significantly reduces purchase intention. In contrast, Poiriazi et al. (2025) found that robust board governance and transparency mechanisms in the financial sector can mitigate such unethical behaviors. However, the fast fashion industry appears to have a governance gap that enables explicit falsification. The analysis indicates that the regulatory voids identified by Li et al. (2025) in the construction sector are also prevalent in the fashion industry, particularly in the form of unchecked false labeling. On the other hand, recent comprehensive reviews emphasize that although eco-labels are intended to bridge the information gap between producers and consumers, their effectiveness is severely compromised without strict validation mechanisms (Nakaishi & Chapman, 2024). Therefore, the high deception risk of false labeling identified in this study suggests a failure in the current policy tool function of these labels and highlights the need for standardized verification systems.

“Visual or Atmospheric Manipulation” (Green Sheen) is identified as a high-risk greenwashing tactic that exemplifies the 'Eco-Chameleon' strategy within the fast fashion sector. In line with Szabo and Webster (2021), who found that executional greenwashing significantly shapes consumer perception. The current findings elevate this tactic to a critical risk level, approaching that of explicit deception. Moreover, Boncinelli et al. (2023) demonstrated that green-colored packaging and nature-inspired aesthetics can distort market share by creating an unwarranted perception of sustainability. Similarly, Andreoli and Minciotti (2023) observed that explicit textual claims primarily elicit consumer skepticism, whereas visual manipulation often bypasses conscious awareness and avoids scrutiny, making it particularly challenging for consumers to detect deception (Fella & Bausa, 2024). This form of manipulation generates a psychological halo effect, as reported by Li et al. (2025) in the context of corporate reputation. Consumers frequently infer sustainability based solely on green aesthetics (Ende et al., 2023), which enables companies to obscure the synthetic nature of their products. The principal risk emerges when false labels are combined with visual manipulation, as eco-labels function as visual cues, supported by past studies. Lim et al. (2025) further argue that in online retail environments where consumer attention is limited, labels serve both as information sources and as shortcuts for sustainability assessment. Liu et al. (2025) also show that the source and specificity of claims influence purchase intent, reinforcing the view that unverified visual signals can be especially deceptive.

Methodologically, the research bridges the gap between consumer behavior studies and the expanding body of multi-criteria decision-making (MCDM) applications in sustainability. Recent studies have applied MCDM methods to evaluate greenwashing drivers in various sectors. For example, Daga et al. (2025) used TISM to model ESG barriers, while Sklavos et al. (2025) employed DEA-CRITIC-TOPSIS to rank eco-efficiency in banking. Building on these approaches, the present study applies the SIWEC method to consumer deception tactics, thereby complementing prior work and shifting the focus from corporate causes to market consequences. Furthermore, whereas Xu et al. (2025) demonstrated that digital transformation and mergers and acquisitions (M&As) curb greenwashing by enhancing transparency in Chinese firms, the findings of this study suggest that, in the absence of such structural digitization in fashion supply chains, consumers remain vulnerable to "Hidden Trade-off" tactics (Munir & Mohan, 2022), where capsule collections distract from the brand's overall carbon footprint.

Several practical implications for stakeholders can be made. For managers, the elevated risk associated with visual manipulation serves as a clear warning. Employing "nature-inspired" marketing for synthetic products is no longer a permissible "grey area" but constitutes a liability that provokes substantial green skepticism (Andreoli & Minciotti, 2023; Shojaei et al., 2024). Managers are required to ensure that their visual identity accurately reflects material reality to mitigate reputational harm linked to the "Sin of Fibbing" (Yang et al., 2020). For policymakers, these findings indicate that current regulations, which primarily address textual misinformation, are inadequate. As Pagán Martinez et al. (2020) noted, given the complexity of green consumer behavior, regulatory frameworks must expand to include visual semiotics by prohibiting the use of nature imagery on packaging for products that do not meet established sustainability criteria, thereby enhancing the reliability of market signals.

Conclusion and Insights for Future Research

The study examined the complex structure of consumer deception within the fast fashion industry by prioritizing greenwashing tactics based on the risk of consumer deception using the SIWEC method. Expert evaluations enabled the establishment of a quantitative risk hierarchy, which identified "False Labeling" and "Visual/Atmospheric Manipulation" as the most significant threats to market integrity. The findings indicate that greenwashing has evolved from linguistic ambiguity, such as vague claims, to more advanced forms of structural and visual falsification. The evidence confirms that deception has shifted from explicit denial to subtle mimicry through visual manipulation, supporting the perspective that corporate resistance to sustainability is becoming more implicit and psychological. This trend contributes to a significant breakdown in the reliability of marketplace signaling.

In response to recent calls in the literature to evaluate sustainability communications using neuroscientific methods rather than relying solely on self-reported data (Ćirović et al., 2024; Liu et al., 2023), future marketing research should prioritize consumer experiments that examine the cognitive effects of visual greenwashing. Given that visual manipulation represents a significant risk, scholars may employ neuromarketing techniques such as eye-tracking or EEG to investigate how green aesthetics, including color palettes, nature imagery, and rustic store designs, influence subconscious consumer decision-making. It is important to determine whether these cues diminish critical thinking and enable consumers to rationalize the purchase of unsustainable fast fashion despite environmental concerns.

Given the risks associated with False Labeling, future research should examine the potential of technological tools to restore consumer trust. Recent literature underscores the

importance of technological interventions in achieving operational transparency (Guo et al., 2020). Researchers could assess the impact of Digital Product Passports (DPPs) and blockchain-enabled traceability on consumer perceptions of greenwashing (Acciai & Pérez-Bou, 2025; Hina et al., 2024). Specifically, studies should evaluate whether scannable, immutable supply chain data, such as QR codes, reduces perceived deception more effectively than traditional third-party certifications, which are frequently viewed with suspicion due to the possibility of falsification.

A comparative analysis of generational and cultural cohorts represents a promising direction for future research. Recent studies highlight a significant 'consumption paradox' and a pronounced intention-behavior gap among Generation Z, who are both prominent climate advocates and the largest consumer segment for fast fashion (Kang & Badal, 2025). Although this study relied on expert consensus, susceptibility to deception may differ substantially between Generation Z and older cohorts. Future research should investigate whether climate-aware younger consumers demonstrate higher levels of green literacy, enabling them to identify the "Hidden Trade-off" tactics described in this study (Bytof & Ritch, 2023; Williams & Hodges, 2022), or whether their preference for fashionable, low-cost apparel increases their likelihood of overlooking deceptive signals due to cognitive dissonance reduction strategies.

For policymakers and non-governmental organizations (NGOs), the present study highlights the urgent need to broaden consumer protection regulations beyond textual accuracy. Visual Manipulation is a significant deception risk. Regulatory bodies should implement stricter guidelines regarding the semiotics of sustainability. These may include restricting the use of nature-inspired imagery and specific color palettes on packaging for products that do not meet established environmental standards. To address the critical issue of False Labeling, governments should expedite the shift from voluntary corporate disclosures to mandatory, standardized verification systems. For example, a government-backed Digital Product Passport including blockchain technology could be used (Hina et al., 2025). This approach would reduce the ambiguity of self-issued certifications. Accordingly, NGOs play a key role in auditing supply chain transparency. They are also essential in launching visual literacy campaigns to help consumers distinguish between aesthetic cues and true environmental claims.

Although the study offers significant insights, it faces several limitations that indicate avenues for future research. While the SIWEC method quantifies uncertainty effectively, it depends on expert judgment. Consequently, the risk hierarchy reflects theoretical assessments from 14 marketing academicians rather than empirical consumer behavior. Although these experts possessed extensive experience in sustainable marketing, the absence of marketing industry professionals or sector representatives limits the practical industry perspective of the evaluation. Therefore, future research should aim to include a more diverse decision-making panel, incorporating both academic and sectoral experts, to provide a more comprehensive assessment of greenwashing tactics. On the other hand, the sample size limits the generalizability to wider cultural or demographic groups, though sufficient for consistency in multi-criteria decision-making. Finally, the study focuses solely on the fast fashion industry. Thus, the prioritization of tactics such as False Labeling and Visual Manipulation may differ in sectors with stricter regulations, such as food, pharmaceuticals, or automotive.

References

Aasmäe, S. (2021). *Consumer awareness of greenwashing and the reliability of eco-labels* [Unpublished bachelor's thesis]. Tallinn University of Technology.

Acciai, E., & Pérez-Bou, S. (2025). Exploring the role of digital product passports for supply chain traceability in the "Made in Italy" textile sector. *Asia Pacific Journal of Marketing and Logistics*, 1-17. https://doi.org/10.1108/APJML-03-2025-0371

Akturan, U. (2018). How does greenwashing affect green branding equity and purchase intention? An empirical research. *Marketing Intelligence & Planning, 36*(7), 809-824. https://doi.org/10.1108/MIP-12-2017-0339

Alizadeh, L., Liscio, M. C., & Sospiro, P. (2024). The phenomenon of greenwashing in the fashion industry: A conceptual framework. *Sustainable Chemistry and Pharmacy, 37*, Article 101416. https://doi.org/10.1016/j.scp.2023.101416

Andreoli, T. P., & Minciotti, S. A. (2023). Attention and skepticism to greenwashing: Analysis of influence on consumer behavior. *Revista de Administração da UFSM*, *16*(4), e6. https://doi.org/10.5902/1983465973743

Badhwar, A., Islam, S., Tan, C. S. L., Panwar, T., Wigley, S., & Nayak, R. (2024). Unraveling green marketing and greenwashing: A systematic review in the context of the fashion and textiles industry. *Sustainability, 16*(7), 2738. https://doi.org/10.3390/su16072738

Basso, K., Pauli, J., Cerutti, P., Perin, M., Corte, V. F. D., & Moro, L. D. (2023). Greenwash, show your true colours: how verbal and visual messages influence consumers' perception?. *International Journal of Environment and Sustainable Development, 22*(2), 210-225. https://doi.org/10.1504/IJESD.2023.129891

Bauer, R. A. (1960). Consumer behavior as risk taking. In R. S. Hancock (Ed.), *Dynamic marketing for a changing world: Proceedings of the 43rd conference of the American Marketing Association* (June 15–17, Chicago, IL, United States) (pp. 389–398). American Marketing Association.

Boncinelli, F., Gerini, F., Piracci, G., Bellia, R., & Casini, L. (2023). Effect of executional greenwashing on market share of food products: An empirical study on green-coloured packaging. *Journal of Cleaner Production, 391*, Article 136258. https://doi.org/10.1016/j.jclepro.2023.136258

Busalim, A., Lynn, T., & Wood, C. M. (2025). Consumer resistance to sustainable fashion: Evidence from the USA and India. *Journal of Fashion Marketing and Management: An International Journal*, *29*(3), 371–394. https://doi.org/10.1108/JFMM-02-2024-0052

Bytof, M. P., & Ritch, E. L. (2023). Exploring the perceptions of climate-aware Generation Z towards fast-fashion's greenwashing for the climate crisis. In M. Brandstrup, L. P. Dana, D. Ryding, G. Vignali, & M. Caratù (Eds.), *The garment economy* (pp. 515–531). Springer. https://doi.org/10.1007/978-3-031-33302-6_24

Cao, C., & Zhang, X. (2024). What triggers consumers to purchase eco-friendly food? The impact of micro signals, macro signals and perceived value. *British Food Journal, 126*(5), 2204-2226. https://doi.org/10.1108/BFJ-08-2023-0697

Cardona, N. (2025). *Environmental impact of fast fashion statistics (2025)*. Uniform Market. Retrieved December 3, 2025, from https://www.uniformmarket.com/statistics/fast-fashion-statistics

Cinelli, M., Kadziński, M., Gonzalez, M., & Słowiński, R. (2020). How to support the application of multiple criteria decision analysis? Let us start with a comprehensive taxonomy. *Omega*, *96*, Article 102261. https://doi.org/10.1016/j.omega.2020.102261

Ćirović, M., Dimitriadis, N., Janić, M., Alevizou, P., & Dimitriadis, N. J. (2024). More than words: Rethinking sustainability communications through neuroscientific methods. *Journal of Consumer Behaviour, 23*(1), 15-30. https://doi.org/10.1002/cb.2125

Daga, S., Yadav, K., Singh, D., Pamucar, D., & Simic, V. (2025). Unveiling greenwashing: Analyzing the interaction of factors discouraging ESG greenwashing through TISM and MICMAC. *Journal of Environmental Management*, *380*, Article 124850. https://doi.org/10.1016/j.jenvman.2024.124850

Ende, L., Reinhard, M. A., & Göritz, L. (2023). Detecting greenwashing! The influence of product colour and product price on consumers' detection accuracy of faked bio-fashion. *Journal of Consumer Policy*, *46*(2), 155–189. https://doi.org/10.1007/s10603-023-09537-8

Ertürk, E. (2024). Consumer society and sustainability: A semiotic analysis of campaigns on ecological fashion thought. *Ordu Üniversitesi Sosyal Bilimler Enstitüsü Sosyal Bilimler Araştırmaları Dergisi, 15*(1), 227-255. https://doi.org/10.48146/odusobiad.1424243

Fella, S., & Bausa, E. (2024). Green or greenwashed? Examining consumers' ability to identify greenwashing. *Journal of Environmental Psychology, 95*, Article 102281. https://doi.org/10.1016/j.jenvp.2024.102281

Gallas, S., Chouchene, M. R., & Bouzgarrou, H. (2025). Greenwashing illusion and financial performance: The moderating role of environmental practices. *Business Strategy & Development, 8*(4), e70234. https://doi.org/10.1002/bsd2.70234

Gentile, G., & Gupta, J. (2025). Orchestrating the narrative: The role of fossil fuel companies in delaying the energy transition. *Renewable and Sustainable Energy Reviews*, *212*, Article 115359. https://doi.org/10.1016/j.rser.2025.11535

Guo, S., Sun, X., & Lam, H. K. (2020). Applications of blockchain technology in sustainable fashion supply chains: Operational transparency and environmental efforts. *IEEE Transactions on Engineering Management, 70*(4), 1312-1328. https://doi.org/10.1109/TEM.2020.3034216 .

Ha, M. T., Ngan, V. T. K., & Nguyen, P. N. (2022). Greenwash and green brand equity: The mediating role of green brand image, green satisfaction and green trust and the moderating role of information and knowledge. *Business Ethics, the Environment & Responsibility, 31*(4), 904-922. https://doi.org/10.1111/beer.12462

Hina, M., Islam, N., & Dhir, A. (2024). Blockchain for sustainable consumption: an affordance and consumer value-based view. *Internet Research, 34*(7), 215-250. https://doi.org/10.1108/INTR-07-2023-0523

Hina, M., Islam, N., & Luo, X. R. (2025). Towards sustainable consumption decision-making: Examining the interplay of blockchain transparency and information-seeking in

reducing product uncertainty. *Decision Support Systems*, *189*, Article 114370. . https://doi.org/10.1016/j.dss.2024.114370

Hora, M., & Subramanian, R. (2019). Relationship between positive environmental disclosures and environmental performance: An empirical investigation of the greenwashing sin of the hidden trade-off. *Journal of Industrial Ecology, 23*(4), 855-868. https://doi.org/10.1111/jiec.12823

Hung, C.-Z., & Chang, T.-W. (2024). Have I purchased the right product? Consumer behavior under corporate greenwash behavior. *Journal of Consumer Behaviour*, *23*(3), 1102–1113. https://doi.org/10.1002/cb.2263

Kang, J. H., & Badal, C. (2025). Gen Z college students' fashion consumption paradox: Sustainable vs. fast fashion. *Fashion and Textiles, 12*, Article 35. https://doi.org/10.1186/s40691-025-00442-x

Kinnunen, A. (2020). *The effects of greenwashing on consumer behavior: Study of the effects of greenwashing on green consumer confusion, green perceived risk and green trust* [Unpublished bachelor's thesis]. Aalto University

Kothiyal, M., & Mehta, S. (2025). Effectiveness of Certification Marks and Eco-labels in Reducing Greenwashing: A Comparative Study. *Journal of Intellectual Property Rights (JIPR), 30*(6), 787-797. https://doi.org/10.56042/jipr.v30i6.15643

Langfeldt, L. (2002). *Decision-making in expert panels evaluating research: Constraints, processes and bias* [Unpublished doctoral dissertation]. University of Oslo.

Li, X., Hao, J., Ding, Z., Yang, J., & Yao, Y. (2025). Unveiling the driving mechanism of greenwashing behaviors: An agent-based modeling approach integrating new institutionalism and halo effect theories. *Humanities & Social Sciences Communications*, *12*, Article 895. https://doi.org/10.1057/s41599-025-05121-2

Li, X., Li, J., He, J., Huang, Y., Liu, X., Dai, J., & Shen, Q. (2023). What are the key factors of enterprises' greenwashing behaviors under multi-agent interaction? A grey-DEMATEL analysis from Chinese construction materials enterprises. *Engineering, Construction and Architectural Management*. Advance online publication. https://doi.org/10.1108/ECAM-01-2023-0027

Lim, H., Kim, I., & Cho, M. (2025). Eco-labels as visual heuristics: Empowering online consumer agency in responsible fashion consumption. *Journal of Fashion Marketing and Management: An International Journal*. Advance online publication. 1-24. https://doi.org/10.1108/JFMM-08-2024-0316

Liu, M., Liang, D., Zhou, X., & Sun, J. (2025). Communication strategies for eco-friendly apparel: Experimental evidence from Generation Z consumers in China. *Journal of Fashion Marketing and Management: An International Journal*, *29*(6), 1130–1153. https://doi.org/10.1108/JFMM-07-2024-0254

Liu, Y., Zhao, R., Xiong, X., & Ren, X. (2023). A bibliometric analysis of consumer neuroscience towards sustainable consumption. *Behavioral Sciences, 13*(4), Article 298. https://doi.org/10.3390/bs13040298

Lu, X., Sheng, T., Zhou, X., Shen, C., & Fang, B. (2022). How does young consumers' greenwashing perception impact their green purchase intention in the fast fashion

industry? An analysis from the perspective of perceived risk theory. *Sustainability, 14*(20), 13473. https://doi.org/10.3390/su142013473

Mathew, M., & Spinelli, R. (2025). Decoding sustainable drivers: A systematic literature review on sustainability-induced consumer behaviour in the fast fashion industry. *Sustainable Production and Consumption*. 55,132–145 https://doi.org/10.1016/j.spc.2025.02.011

Mousavi, K., Kowsari, E., Ramakrishna, S., Chinnappan, A., & Gheibi, M. (2024). A comprehensive review of greenwashing in the textile industry (life cycle assessment, life cycle cost, and eco-labeling). *Environment, Development and Sustainability, 27*(9), 21737-21777. https://doi.org/10.1007/s10668-024-04508-6

Munier, N., & Hontoria, E. (2021). *Uses and limitations of the AHP method: A non-mathematical and rational analysis*. Springer.

Munir, S., & Mohan, V. (2022). Consumer perceptions of greenwashing: Lessons learned from the fashion sector in the UAE. *Asian Journal of Business Ethics*, *11*(1), 1–44. https://doi.org/10.1007/s13520-021-00140-z

Munir, S., & Mohan, V. (2022). Consumer perceptions of greenwashing: Lessons learned from the fashion sector in the UAE. *Asian Journal of Business Ethics*, *11*(1), 1–44. https://doi.org/10.1007/s13520-021-00140-z

Musgrove, C. C. F., Choi, P., & Chris Cox, K. (2018). Consumer perceptions of green marketing claims: An examination of the relationships with type of claim and corporate credibility. *Services Marketing Quarterly, 39*(4), 277-292. https://doi.org/10.1080/15332969.2018.1514794

Nakaishi, T., & Chapman, A. (2024). Eco-labels as a communication and policy tool: A comprehensive review of academic literature and global label initiatives. *Renewable and Sustainable Energy Reviews*, *202*, Article 114708. https://doi.org/10.1016/j.rser.2024.114708

Nguyen, T. T. H., Yang, Z., Nguyen, N., Johnson, L. W., & Cao, T. K. (2019). Greenwash and green purchase intention: The mediating role of green skepticism. *Sustainability, 11*(9), 2653. https://doi.org/10.3390/su11092653

Orazi, D. C., & Chan, E. Y. (2020). "They did not walk the green talk!:" How information specificity influences consumer evaluations of disconfirmed environmental claims. *Journal of Business Ethics, 163*(1), 107-123. https://doi.org/10.1007/s10551-018-4028-6

Özçalıcı, M., Ersoy, N., Van Duc, D., Bao, N. C., & Son, N. H. (2025). SIWEC-R: A rank-sensitive improvement to SIWEC methodology. *Teknomekanik, 8*(2), 183-210. https://doi.org/10.24036/teknomekanik.v8i2.40172

Pagán Martínez, M., Cremasco, C. P., Gabriel Filho, L. R. A., Braga Junior, S. S., Bednaski, A. V., Quevedo-Silva, F., ... Padgett, R. C. M. (2020). Fuzzy inference system to study the behavior of the green consumer facing the perception of greenwashing. *Journal of Cleaner Production*, *242*, Article 116064. https://doi.org/10.1016/j.jclepro.2019.116064

Pamucar, D., Stević, Ž., & Zavadskas, E. K. (2018). Integration of interval rough AHP and interval rough MABAC methods for evaluating university web pages. *Applied Soft Computing*, *67*, 141–163. https://doi.org/10.1016/j.asoc.2018.02.057

Poiriazi, E., Zournatzidou, G., & Konteos, G. (2025). Do board characteristics matter with greenwashing? An investigation in the financial sector with the integration of entropy weight and TOPSIS multicriteria decision-making methods. *Risks*, *13*(4), Article 64. https://doi.org/10.3390/risks13040064

Polonsky, M. J., Rotman, J., Weber, V., & Kumar, P. (2025). How meaningless and substantive green claims jointly determine product environmental perceptions. *International Journal of Advertising, 44*(3), 396-427. https://doi.org/10.1080/02650487.2024.2377503

Puška, A., Nedeljković, M., Pamučar, D., Božanić, D., & Simić, V. (2024). Application of the new simple weight calculation (SIWEC) method in the case study in the sales channels of agricultural products. *MethodsX, 13*, Article 102930. https://doi.org/10.1016/j.mex.2024.102930

Puška, A., Nedeljković, M., Prodanović, R., Vladisavljević, R., & Suzuki, K. (2023). Market assessment of pear varieties in Serbia using fuzzy CRADIS and SIWEC methods. *Agronomy*, *13*(2), Article 339. https://doi.org/10.3390/agronomy13020339

Rashid, A., & Arshad, M. Y. (2023). Integrating circular economy, SBTI, digital LCA, and ESG benchmarks for sustainable textile dyeing: A critical review of industrial textile practices. *Global NEST Journal, 25*(7), 39–51. https://doi.org/10.30955/gnj.005145

Roos, A. (2025). Greenwashing: a critical realist approach to the study of false ecological claims. *Journal of Critical Realism, 24*(2), 176-197. https://doi.org/10.1080/14767430.2025.2503634

Santos, C., Coelho, A., & Marques, A. (2024a). A systematic literature review on greenwashing and its relationship to stakeholders: state of art and future research agenda. *Management Review Quarterly, 74*(3), 1397-1421 https://doi.org/10.1007/s11301-023-00337-5 .

Santos, C., Coelho, A., & Marques, A. (2024b). The greenwashing effects on corporate reputation and brand hate, through environmental performance and green perceived risk. *Asia-Pacific Journal of Business Administration, 16*(3), 655-676. https://doi.org/10.1108/APJBA-05-2022-0216

Schmuck, D., Matthes, J., & Naderer, B. (2018). Misleading consumers with green advertising? An affect–reason–involvement account of greenwashing effects in environmental advertising. *Journal of Advertising, 47*(2), 127-145. https://doi.org/10.1080/00913367.2018.1452652

Shojaei, A. S., Barbosa, B., Oliveira, Z., & Coelho, A. M. R. (2024). Perceived greenwashing and its impact on eco-friendly product purchase. *Tourism & Management Studies*, *19*(4), 1–12. https://doi.org/10.18089/tms.20240201020

Siew, M. L. M., Ahmad, R. A. R., & Zainon, S. (2024). Greenwashing: Analysing the discrepancy between environmental rhetoric and corporate practices. *Academic Journal of Business and Social Sciences, 8*(1), 1-23

Sigurdsson, V., Larsen, N. M., Pálsdóttir, R. G., Folwarczny, M., Menon, R. V., & Fagerstrøm, A. (2022). Increasing the effectiveness of ecological food signaling: Comparing sustainability tags with eco-labels. *Journal of Business Research, 139*, 1099-1110. https://doi.org/10.1016/j.jbusres.2021.10.052

Simão, S. A. V., Rohden, S. F., & Pinto, D. C. (2022). Natural claims and sustainability: The role of perceived efficacy and sensorial expectations. *Sustainable Production and Consumption*, *34*, 505–517. https://doi.org/10.1016/j.spc.2022.09.026

Sklavos, G., Zournatzidou, G., Ragazou, K., & Sariannidis, N. (2025). Green accounting and ESG-driven eco-efficiency in European financial institutions: A two-stage DEA-CRITIC-TOPSIS evaluation. *PLoS One*, *20*(10), e0334882. https://doi.org/10.1371/journal.pone.0334882

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374. https://doi.org/10.2307/1882010

Szabo, S., & Webster, J. (2021). Perceived greenwashing: The effects of green marketing on environmental and product perceptions. *Journal of Business Ethics*, *171*, 719–739. https://doi.org/10.1007/s10551-020-04461-0

Taoketao, E., Feng, T., Song, Y., & Nie, Y. (2018). Does sustainability marketing strategy achieve payback profits? A signaling theory perspective. *Corporate Social Responsibility and Environmental Management*, *25*(6), 1039-1049.

Tcholakova, E. (2024). *Sustainable fashion worldwide - statistics & facts*. Statista. Retrieved December 3, 2025, from https://www.statista.com/topics/9543/sustainable-fashion-worldwide/#topicOverview

Tcholakova, E. (2025). *Fast fashion in the U.S. - statistics & facts*. Statista. Retrieved December 3, 2025, from https://www.statista.com/topics/13181/fast-fashion-in-the-us/#topicOverview

Testa, F., Boiral, O., & Iraldo, F. (2018). Internalization of environmental practices and institutional complexity: can stakeholders pressures encourage greenwashing?. *Journal of Business Ethics, 147*(2), 287-307. https://doi.org/10.1007/s10551-015-2960-2

Thuy, D. (2025). *Sustainable fashion consumption in the Asia-Pacific region - statistics & facts*. Statista. Retrieved December 3, 2025, from https://www.statista.com/topics/13707/sustainable-fashion-consumption-in-the-asia-pacific-region/

Tu, J. C., Cui, Y., Liu, L., & Yang, C. (2024). Perceived greenwashing and its impact on the green image of brands. *Sustainability, 16*(20), Article 9009. https://doi.org/10.3390/su16209009

van der Ven, H. (2019.). *Beyond greenwash: Explaining credibility in transnational eco-labeling*. Oxford University Press.

Wang, X., Guo, J., Wu, Y., & Liu, N. (2020). Emotion as signal of product quality: Its effect on purchase decision based on online customer reviews. *Internet Research, 30*(2), 463-485. https://doi.org/10.1108/INTR-09-2018-0415

Wang, Y. M., & Luo, Y. (2009). On rank reversal in decision analysis. *Mathematical and Computer Modelling*, *49*(5-6), 1221-1229. https://doi.org/10.1016/j.mcm.2008.06.019

Wang, Y., & Jung, S. (2025). The aftermath of greenwashing in the fashion industry: Chinese consumer responses to different types of greenwashing. *Asia Pacific Journal of Marketing and Logistics, 37*(9), 2669-2685. https://doi.org/10.1108/APJML-06-2024-0858

Williams, A., & Hodges, N. (2022). Adolescent Generation Z and sustainable and responsible fashion consumption: exploring the value-action gap. *Young Consumers, 23*(4), 651-666. https://doi.org/10.1108/YC-11-2021-1419

Xiao, Z., Wang, Y., & Guo, D. (2022). Will greenwashing result in brand avoidance? A moderated mediation model. *Sustainability, 14*(12), Article 7204. https://doi.org/10.3390/su14127204

Xu, J., Cai, Z., & Li, X. (2025). Unveiling greenwashing trends: The curious effect of digital-driven M&As among listed Chinese firms. *Research in International Business and Finance*, *80*, Article 103140. https://doi.org/10.1016/j.ribaf.2024.103140

Yang, Z., Nguyen, T. T. H., Nguyen, H. N., Nguyen, T. T. N., & Cao, T. T. (2020). Greenwashing behaviours: Causes, taxonomy and consequences based on a systematic literature review. *Journal of Business Economics and Management*, *21*(5), 1486–1507. https://doi.org/10.3846/jbem.2020.13225

Zhong, P., & Kim, C. (2025). When are global fashion companies socially irresponsible? Relations among misleading information, brand betrayal and saving face in consumer boycotts. *Journal of Fashion Marketing and Management: An International Journal*, *29*(6), 1085–1108. https://doi.org/10.1108/JFMM-08-2024-0337

Information About the Article/Makale Hakkında Bilgiler

The Ethical Rules for Research and Publication / Araştırma ve Yayın Etiği

The authors declared that the ethical rules for research and publication followed while preparing the article.

Yazarlar makale hazırlanırken araştırma ve yayın etiğine uyulduğunu beyan etmiştir.

Conflict of Interests/ Çıkar Çatışması

The authors have no conflict of interest to declare.

Yazarlar çıkar çatışması bildirmemiştir.

Grant Support/ Finansal Destek

The authors declared that this study has received no financial support.

Yazarlar bu çalışma için finansal destek almadığını beyan etmiştir.

Author Contributions/ Yazar Katkıları

The draft process of the manuscript/ Taslağın Hazırlanma Süreci M.S.A/S.E., Data Collection/Verilerin Toplanması M.S.A/S.E., Writing The Manuscript/ Makalenin Yazılması M.S.A/S.E., Submit, Revision and Resubmit Process/ Başvuru, Düzeltme ve Yeniden Başvuru Süreci M.S.A/S.E.

Journal of Applied and Theoretical Social Sciences

JATSS, 2026; 8(1), 105-136

First Submission:25.01.2026
Revised Submission After Review:10.03.2026
Accepted For Publication:24.03.2026
Available Online Since:25.03.2026

Research Article

Gender Differences and Regional Disparities in Youth NEET Status: Empirical Evidence for Türkiye

Hayriye Özgül Özkan Değirmenci[a]

Abstract

Introduction: This study measures youth NEET (not in employment, education, or training) in Türkiye and identifies its main correlates, with emphasis on gender gaps and regional inequality. The objective is to describe how human capital, household constraints, and place-based conditions shape NEET risk in 2022.

Method: Nationally representative 2022 Household Labor Force Survey microdata for ages 15–29 (N=123,834) is analyzed using binary logit models estimated for the full sample and separately by gender. Covariates include age (and squared), education, marital status, household size, child-under-six indicator, and NUTS-1 regions; results are reported as average marginal effects with household-clustered standard errors.

Results or Findings: The NEET rate is 26.6% overall, with a large gender gap (36.5% for women; 16.8% for men). NEET probability declines strongly with higher education, but gender differences remain across education levels. Regional disparities are substantial, with higher NEET prevalence in disadvantaged regions. Marriage and young-child presence are strongly associated with higher NEET probability among women, while associations for men are weaker and often opposite in sign.

Discussion or Conclusion: NEET in Türkiye reflects regional inequality, education gradients, and gendered household constraints that raise women's inactivity. Priorities include reducing early school leaving, improving school-to-work transitions, expanding childcare, and targeting activation and quality formal job creation in high-NEET regions. Future work should identify causal program impacts using longitudinal or policy-variation designs.

Keywords: NEET, youth, labor force survey, regional inequality, gender, education

JEL Codes: J13, J16, J21, R23, I26

[a] Research Assistant Dr., Muğla Sıtkı Koçman University, Faculty of Economics and Administrative Sciences, Department of Economics, Muğla/Türkiye, ozguldegirmenci@mu.edu.tr, ORCID ID: https://orcid.org/0000-0003-1728-7442 (Corresponding Author)

Journal of Applied and Theoretical Social Sciences

JATSS, 2026; 8(1), 105-136

İlk Başvuru: 25.01.2026
Düzeltilmiş Makalenin Alınışı: 10.03.2026
Yayın İçin Kabul Tarihi: 24.03.2026
Online Yayın Tarihi: 25.03.2026

Araştırma Makalesi

Gençlerin NEET Statüsünde Toplumsal Cinsiyet Farklılıkları ve Bölgesel Eşitsizlikler: Türkiye İçin Ampirik Kanıtlar

Hayriye Özgül Özkan Değirmenci[a]

Öz

Giriş: Bu çalışma, Türkiye'de genç NEET (istihdamda olmayan, eğitimde veya öğretimde bulunmayan) düzeyini ölçmekte ve 2022'de toplumsal cinsiyet farkları ile bölgesel eşitsizliklere odaklanarak temel belirleyicileri incelemektedir. Amaç, insan sermayesi, hane halkı kısıtları ve mekânsal koşulların NEET riskini nasıl şekillendirdiğini göstermektir.

Yöntem: 15–29 yaş grubuna ait ulusal temsiliyete sahip 2022 Hane halkı İşgücü Anketi (N=123.834) mikro verileri ikili logit modelleriyle, toplam örneklemde ve cinsiyete göre ayrı ayrı analiz edilmiştir. Açıklayıcı değişkenler yaş (ve karesi), eğitim, medeni durum, hane halkı büyüklüğü, hanede 6 yaş altı çocuk varlığı ve Düzey-1 İBBS bölgeleridir; sonuçlar hane düzeyinde kümelenmiş standart hatalarla ortalama marjinal etkiler olarak raporlanmıştır.

Sonuçlar ya da Bulgular: NEET oranı toplamda %26,6 olup cinsiyet farkı büyüktür (kadın %36,5; erkek %16,8). Eğitim yükseldikçe NEET olasılığı belirgin biçimde azalır; ancak cinsiyet farkları eğitim düzeyleri boyunca sürer. Dezavantajlı bölgelerde NEET daha yüksektir. Evli olmak ve hanede küçük çocuk bulunması kadınlarda NEET olasılığıyla güçlü biçimde pozitif ilişkiliyken, erkeklerde ilişkiler daha zayıf ve çoğu zaman ters işaretlidir.

Tartışma ya da Yapılan Çıkarımlar: Türkiye'de genç NEET, bölgesel eşitsizlik, eğitim gradyanı ve kadınların hareketsizliğini artıran toplumsal cinsiyete dayalı hanehalkı kısıtlarının bileşimini yansıtır. Öncelikler erken okul terkini azaltmak, eğitimden işe geçişi güçlendirmek, çocuk bakımını genişletmek ve yüksek-NEET bölgelerde etkinleştirmeyi nitelikli ve kayıtlı istihdam yaratımıyla birlikte hedeflemektir. Gelecek çalışmalar, uzunlamasına veriler veya politika varyasyonu tasarımlarıyla programların nedensel etkilerini belirlemelidir.

Anahtar Kelimeler: NEET, gençlik, işgücü araştırması, bölgesel eşitsizlik, toplumsal cinsiyet, eğitim

JEL Kodlar: J13, J16, J21, R23, I26

[a] Araştırma Görevlisi Dr., Muğla Sıtkı Koçman Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İktisat Bölümü, Muğla/Türkiye, ozguldegirmenci@mu.edu.tr, ORCID ID: https://orcid.org/0000-0003-1728-7442 (Sorumlu Yazar)

Introduction

In many middle-income economies, the transition from school to stable work has become longer and more uncertain, as young people increasingly experience extended spells of job search, intermittent employment, or inactivity before achieving sustained labor-market attachment (Choudhry & Pastore, 2023). In this context, the NEET concept, youth not in employment, education, or training, has become a central indicator for assessing social inclusion and the efficiency of school-to-work transitions because it captures both open unemployment and a broader set of young people who are detached from formal work and human-capital accumulation (Bynner & Parsons, 2002; Karma, 2025; Organisation for Economic Co-operation and Development (OECD), 2025b). At the same time, the NEET category is inherently heterogeneous, and its policy relevance depends on whether NEET status reflects short transition spells, discouragement, care constraints, or more persistent detachment (Furlong, 2006; Yates & Payne, 2006). This heterogeneity is especially salient in settings where inactivity, rather than unemployment, dominates youth disconnection.

Türkiye is a particularly important case for revisiting NEET dynamics because youth disengagement intersects with persistently low female labor-force participation and marked regional inequality. Prior work on Türkiye has consistently emphasized that NEET status is strongly gendered and closely linked to household formation and care responsibilities, suggesting that women's inactivity pathways are central to understanding aggregate NEET outcomes (Lüküslü & Çelik, 2022). These patterns have direct implications for both growth and equality. High and persistent NEET shares imply underutilized human capital and weaker long-run productivity, while gendered inactivity reinforces labor-market segmentation and unequal access to income and social protection.

Comparative evidence underscores the scale of Türkiye's challenge. For ages 15–29, Eurostat reports that the NEET rate averaged 11.0% in the European Union (EU) in 2024, ranging from about 5% in the Netherlands to about 19% in Romania, and notes an EU-level target to reduce NEET below 9% by 2030 (Eurostat, 2025). In contrast, the 2022 Household Labor Force Survey (HLFS) indicates that 27.5% of youth in Türkiye are NEET, with a very large gender disparity (38.6% for women versus 16.2% for men; Table 1). This gap is substantially larger than the EU-wide gender gap in 2024 (12.1% for young women versus 10.0% for young men), reinforcing the view that mechanisms beyond standard job-search frictions, particularly inactivity-related constraints, are likely to be central for Türkiye (Eurostat, 2025).

A time-series benchmark further highlights the divergence. For the EU, Eurostat shows that the NEET share for ages 15–29 declined steadily from 2014 to 2019, rose with the onset of the COVID-19 shock in 2020, and then fell again, reaching 11.0% in 2024, below pre-pandemic levels, indicating a recovery in youth labor-market attachment (Eurostat, 2025). Türkiye's 2022 level is therefore not only above the contemporary EU average but also far above the EU's medium-term trajectory and its policy target for 2030 (Eurostat, 2025).

For additional cross-country context, the World Bank's World Development Indicators publishes a harmonized NEET series for ages 15–24 (indicator SL.UEM.NEET.ZS), sourced from the International Labour Organization (ILOSTAT) labor force statistics framework. This series is useful for comparisons beyond Europe and for longer-run benchmarking, but it can differ from national survey tabulations because it applies internationally harmonized definitions, uses a different age band than 15–29, and may incorporate modelled estimates for some country-years (World Bank, 2025; ILOSTAT, 2025). These differences underscore the

importance of being explicit about age coverage and measurement when comparing Türkiye's NEET outcomes across sources.

OECD benchmarks point to a similarly stark picture. OECD reporting highlights that NEET shares can be decomposed into unemployed NEET and inactive NEET, which is crucial for diagnosing whether disconnection reflects labor-demand shortfalls or barriers to participation (OECD, 2025a). In Education at a Glance 2025, the OECD reports that the average NEET rate among 18–24-year-olds was 14% in 2024 and notes that the share exceeds 25% in only a small number of countries, including Türkiye; it further indicates that Türkiye's NEET composition is heavily weighted toward inactivity rather than unemployment (OECD, 2025a). This composition aligns with Türkiye-focused research emphasizing that youth disconnection is not solely a labor-demand problem but is also shaped by constraints on participation, especially for women (Lüküslü & Çelik, 2022).

Against this background, this manuscript asks three questions. First, how large is the NEET gender gap once observable characteristics are accounted for? Second, how does educational attainment relate to NEET probability for women and men? Third, to what extent do regional disparities shape NEET outcomes in Türkiye? To answer these questions, we estimate logit models of NEET status using nationally representative microdata and a set of covariates that capture life-cycle transitions, human capital, household structure, and spatial heterogeneity (age and age-squared, education, marital status, household size, the presence of a child under age six, and NUTS-1 region indicators). We report average marginal effects to present magnitudes in probability units and estimate models for the full sample and separately by gender, allowing correlates to differ systematically between women and men.

The remainder of the paper is organized as follows. The next section provides a comparative context situating Türkiye's NEET outcomes relative to benchmarks in the EU and the OECD. The paper then reviews the related literature on NEET in Türkiye, highlighting established findings and remaining gaps that motivate the empirical contribution. The subsequent section describes the HLFS 2022 microdata, sample construction, and the main variables used to operationalize NEET status and covariates. The empirical strategy section outlines the logit framework, explanation of covariate effects in a cross-sectional setting, and the use of average marginal effects, including gender-stratified estimates. The results section presents the baseline estimates and key heterogeneity patterns by gender, education, and region. The discussion section interprets the findings and develops policy implications with attention to gender-responsive and regionally targeted interventions. The final section concludes by summarizing the main contributions and outlining priorities for future research.

Comparative Context

EU monitoring indicates both progress over the past decade and substantial cross-country dispersion in NEET outcomes. In 2024, 11.0% of individuals aged 15–29 in the EU were neither in employment nor in education or training, but national rates varied widely, from around 5% in the Netherlands to about 19% in Romania, highlighting persistent structural heterogeneity in youth transitions across member states (Eurostat, 2025). Over the longer horizon, EU aggregates suggest improvement: the NEET share for ages 15–29 declined between 2014 and 2024, despite the temporary reversal associated with the COVID-19 shock, and the European Pillar of Social Rights sets an EU-level target of reducing the NEET rate below 9% by 2030 (Eurostat, 2025). This policy framing matters because it places NEET at the center of broader EU objectives on inclusion, labor-market integration, and the intergenerational transmission of disadvantage.

EU evidence also underscores that NEET is not gender-neutral and that age gradients remain pronounced. In 2024, the NEET rate among young women aged 15–29 was 12.1%, compared with 10.0% among young men, with women more likely to be outside the labor force rather than unemployed, an important compositional detail for interpreting policy needs (Eurostat, 2025). The age pattern is steep: NEET prevalence rises from the mid-single digits among those aged 15–19 to double-digit levels in the early and late twenties, reflecting the accumulation of transition risks as young people leave education and confront job-search and matching frictions, as well as potential household-formation constraints (Eurostat, 2025). These EU patterns provide a useful benchmark for Türkiye because they distinguish between unemployment-linked NEET and inactivity-linked NEET, which can imply very different policy levers.

OECD comparisons convey a similarly heterogeneous landscape and position Türkiye among the highest-NEET countries in the organization. In Education at a Glance 2025, the OECD reports that, on average, 14% of 18–24-year-olds across OECD countries were NEET in 2024, while the share exceeded 25% in only a small number of cases, including Colombia and Türkiye, as well as partner country South Africa (OECD, 2025a). Crucially, the OECD decomposition highlights that Türkiye's NEET profile is heavily weighted toward inactivity: among 18–24-year-olds, roughly 8% are unemployed NEET while around 24% are inactive (outside the labor force) (OECD, 2025a). This composition is consistent with an interpretation of NEET in Türkiye that extends beyond labor-demand shortfalls and instead emphasizes barriers to participation, particularly those related to gendered constraints and household roles.

Cross-source comparisons require careful attention to definitional choices and statistical construction. This paper's HLFS-based estimates focus on ages 15–29 to align with the EU's headline NEET indicator and to capture the period in which late transitions into stable work and household formation are most consequential. In contrast, the World Bank's World Development Indicators report a harmonized NEET series for ages 15–24 (SL.UEM.NEET.ZS), sourced from the ILOSTAT labor force statistics framework; the indicator is defined as the share of youth not in employment and not in education or formal training within the total youth population, and in some country-years it is reported as a modelled ILO estimate (World Bank, 2025; ILOSTAT, 2025). These differences in age coverage (15–24 versus 15–29), definitions, and estimation procedures can generate level differences relative to national tabulations, implying that benchmarking Türkiye internationally should be explicit about the underlying measurement framework.

Literature Review

Türkiye's NEET profile is distinctive in comparative perspective because a large share of youth disconnection takes the form of economic inactivity rather than unemployment. The NEET indicator, youth not in employment, education, or training, was developed to complement unemployment statistics by capturing young people who are outside both paid work and formal human-capital accumulation, including discouraged workers and those constrained from job search (OECD, 2025b; International Labour Office, 2020). The concept was initially developed to capture forms of social exclusion associated with delayed or fragile school-to-work transitions, beyond what youth unemployment alone measures (Bynner & Parsons, 2002; Redmond & McFadden, 2023). In Türkiye, this broader lens is essential: Gürsel and Tanrıverdi (2025) emphasizes that, unlike many European contexts where the youth unemployment rate exceeds the NEET rate, Türkiye exhibits the opposite pattern, implying that

NEET is driven not only by joblessness but also by early educational disengagement, gendered roles, and structural barriers in school-to-work transitions .

Eurostat's cross-country evidence also provides a benchmark for how NEET correlates with age, education, and place. In the EU, NEET incidence rises sharply with age (5.4% at ages 15–19, 12.8% at 20–24, and 14.6% at 25–29 in 2024) and is strongly graded by education (12.6% among those with low education versus 7.9% among those with tertiary education). NEET is also higher outside cities (10.0% in cities, 11.4% in towns/suburbs, and 12.3% in rural areas), and women are more likely to be NEET than men (12.1% versus 10.0% in 2024) (Eurostat, 2025). These stylized facts motivate the present study's emphasis on age profiles, education gradients, and regional disparities, while also highlighting that Türkiye's exceptionally large gender gap implies a different underlying composition of NEET, one that is likely to be more strongly linked to inactivity and care constraints.

A second recurring theme is heterogeneity within the NEET category. Turkish scholarship stresses that NEET aggregates multiple pathways, unemployed job-seekers, discouraged workers, family-care constrained youth, and those temporarily outside institutions, so that one-size-fits-all interventions are unlikely to address the full problem. This point is consistent with European evidence showing that the NEET population contains distinct subgroups with different needs and policy levers. Accordingly, identifying whether NEET reflects unemployment, inactivity, or temporary transitions is essential for both interpretation and targeting (Maguire, 2015). This heterogeneity has encouraged empirical work that distinguishes compositional differences (unemployed versus inactive NEET) and identifies the mechanisms that keep youth disconnected from both education and the labor market. In policy terms, the implication is that reducing NEET requires coordinated actions that span education retention, active labor market support, and social policies that lower participation constraints (United Nations Development Programme (UNDP) Türkiye, 2025).

Evidence on determinants in Türkiye

Recent econometric studies using the Household Labor Force Survey (HLFS) consistently identify education, age, marital status, and place of residence as core correlates of NEET outcomes. Earlier HLFS-based evidence for Türkiye similarly documents strong associations of NEET status with education and gender, highlighting the role of inactivity among young women (Kılıç, 2014). In a regional analysis based on pooled HLFS microdata, (Özdemir et al., 2023) show that demographic characteristics, including gender and marital status, remain statistically important after controlling for education and other observables, and that regional differences are themselves significant determinants of NEET status. Their descriptive evidence also points to a sharp education gradient, with particularly high NEET incidence among low-educated youth, underscoring the importance of preventing early exit from school and strengthening pathways from education into formal employment. Recent work further extends this line of analysis by documenting a robust role for education, demographic characteristics, and spatial differences in shaping NEET risks in Türkiye, reinforcing the importance of estimating gender-heterogeneous models with regional controls (Özdemir et al., 2024).

Newer contributions extend this determinants framework by placing the family at the center of NEET risk. Using HLFS data spanning 2010–2023, Elitaş (2025) finds that parental education, parental employment, and household income are strongly associated with youth NEET probability. A nuanced pattern emerges that medium levels of parental education can be protective, maternal education may exhibit stronger associations than paternal education, and

parental “decent work” status reduces NEET risk, highlighting how household economic security shapes young people’s ability to remain attached to education and the labor market. Complementing this approach, Akgül (2025) emphasizes that family dynamics and socio-economic conditions intersect with gender, implying that the same household circumstances can translate into different NEET risks for women and men.

Gendered mechanisms and the composition of NEET

Perhaps the most robust finding across the Türkiye NEET literature is the magnitude and structure of the gender gap. Qualitative and mixed-method work argues that treating NEET as a gender-neutral category can obscure distinct mechanisms that generate disconnection for women. Lüküslü and Çelik (2022) document that young women are disproportionately represented among NEET and that women’s NEET status is more likely to consist of inactivity rather than active job search, whereas men’s NEET status more often reflects unemployment. This compositional distinction is analytically important because it suggests that women’s NEET outcomes cannot be addressed solely through job creation or job-search assistance; they also reflect constraints on labor-force participation. Earlier qualitative evidence similarly documents a gendered form of youth disconnection in Türkiye linked to domestic roles and constrained opportunity sets for young women, reinforcing that inactivity pathways are central to the NEET phenomenon (Çelik & Lüküslü, 2011).

Gender-sensitive accounts connect NEET status to job quality, safety, and social norms. Lüküslü and Çelik (2022) report that many young women define “good work” as formal employment with social security and safe working conditions, and that precarious and informal jobs can be seen as unacceptable or risky. In such settings, withdrawal from the labor market may be a rational response to low expected returns and high non-monetary costs, reinforced by norms that prioritize domestic roles. Quantitative work similarly treats education as protective but recognizes that education alone may be insufficient if local labor markets do not offer family-compatible, decent jobs (Özdemir et al., 2024; UNDP Türkiye, 2025).

Marriage, childcare, and household constraints

Household formation and care burdens are repeatedly identified as pivotal to women’s NEET risk in Türkiye (Yiğit et al., 2023). In the pooled HLFS regional analysis, Özdemir et al. (2023) show that marriage and the presence of young children are associated with higher NEET probabilities for women but lower probabilities for men, consistent with the breadwinner–caregiver division of labor and asymmetric assignment of unpaid work (see also Değirmenci, 2023). Akgül (2025) similarly frames women’s NEET vulnerability as rooted in traditional family roles and caregiving responsibilities, while also emphasizing how socio-economic circumstances shape the feasibility of combining employment with family responsibilities.

Policy-oriented evidence triangulates these mechanisms and emphasizes that constraints are multi-dimensional. UNDP Türkiye’s needs assessment on NEET young women documents barriers such as limited affordable childcare, weak school-to-work links, restricted mobility, and safety concerns in some contexts, factors that can discourage young women’s participation in employment even when education levels rise (UNDP Türkiye, 2025). Building on this diagnosis, the UNDP Türkiye White Paper advocates a bundle of gender-responsive interventions, including strengthened care infrastructure, targeted active labor market policies, and stronger coordination across education, employment, and social services (UNDP Türkiye, 2025).

Regional inequality and place-based opportunity structures

A further consensus in the Turkish literature concerns the spatial concentration of NEET disadvantage. Spatial econometric evidence finds that regional disparities are systematic and persistent, supporting a place-based interpretation of NEET disadvantage (Karahasan, 2025). Beyond individual characteristics, regional labor demand, sectoral composition, and access to education and training opportunities shape the probability that youth become and remain NEET. Özdemir et al. (2023) provide direct evidence that regional differences are significant determinants of NEET status in Türkiye. BETAM's recent profiling similarly emphasizes that Türkiye's NEET burden is exceptionally high in European comparisons and that it is geographically concentrated, with pronounced disadvantages in less developed regions and particularly high rates among young women (Gürsel & Tanrıverdi, 2025).

Importantly, regional inequality is likely to interact with gendered constraints. Where local labor markets offer fewer formal jobs, where transport is limited, or where social norms are more restrictive, women's "acceptable opportunity set" may shrink, increasing the likelihood of inactivity. This interaction is explicitly noted in Akgül (2025), who argues that regional disparities exacerbate NEET vulnerability for women by intensifying the joint constraints of limited opportunities and traditional expectations. Relatedly, UNDP's policy work frames NEET among young women as a place-sensitive problem in which the availability of local services (childcare, career guidance, employment intermediation) can materially influence transitions into employment (UNDP Türkiye, 2025).

Positioning and contribution of the present study

Taken together, the literature points to a multi-layered explanation of NEET in Türkiye. Education and individual characteristics are important, but household constraints and gendered norms shape women's inactivity pathways, and regional opportunity structures condition the feasibility of transitions into work and training. Three gaps remain salient for the contemporary policy debate and motivate the present study.

First, much of the quantitative evidence relies on pooled multi-year data, which is valuable for documenting long-run regularities but can mask how the configuration of risks and constraints looks in a specific, post-pandemic year. The 2022 HLFS captures the Turkish labor market after the COVID-19 disruptions to educational enrolment and female participation, providing a timely snapshot that updated knowledge requires.

Second, even when gender differences are emphasized, prior studies do not always present a unified empirical account that simultaneously (i) quantifies gender heterogeneity using fully comparable specifications, (ii) expresses magnitudes in probability units directly suitable for policy interpretation, and (iii) embeds these estimates within an explicit regional framework aligned with NUTS-1 benchmarking. Reporting logit coefficients alone, without conversion to average marginal effects and predictive margins, might limit the policy interpretability of findings. This present study addresses this gap by reporting all results in probability-scale terms alongside the structural coefficients.

Third, while regional disparities in NEET have been documented, prior work has not simultaneously combined gender-stratified modelling with a full set of NUTS-1 region fixed effects and household constraint variables in a single, harmonized specification. Doing so is important because household variables and region indicators are potentially correlated: women

in less-developed regions face both stronger care norms and weaker formal job availability, so failing to control for both simultaneously can confound the estimated gendered effects.

Using nationally representative HLFS 2022 microdata, this study addresses these gaps by estimating harmonized logit models for the pooled sample and separately for women and men, and by reporting average marginal effects and predictive margins that translate covariate differences into percentage-point changes in NEET probability. Substantively, it brings the most policy-relevant household constraints into the core specification, marital status, and the presence of a child under age six, while also controlling for regional fixed effects to isolate gendered and household mechanisms from persistent place-based differences. By combining gender-stratified modelling with regionally grounded interpretation and probability-scale reporting, the paper provides an empirically transparent and policy-oriented reassessment of NEET in Türkiye in 2022, directly informing the design of regionally targeted and gender-responsive youth inclusion strategies.

Four empirical expectations follow directly from the theoretical and substantive mechanisms reviewed above and guide the interpretation of the regression results:

H1: (Education gradient): Higher completed educational attainment is expected to reduce NEET probability for both women and men, with the strongest protective effects at tertiary levels, reflecting the role of human capital in improving employability and school-to-work transition quality.

H2: (Marriage asymmetry): Marriage is expected to be positively associated with NEET probability among women but negatively associated among men, reflecting gender-differentiated household specialization in which marriage corresponds to labor market withdrawal for women and stronger breadwinner attachment for men.

H3: (Care constraint asymmetry): The presence of a child under age six in the household is expected to increase NEET probability among women, reflecting time-intensive care responsibilities, while its association among men is expected to be weaker or of opposite sign, consistent with asymmetric allocation of unpaid work.

H4: (Regional opportunity gradient): NUTS-1 region indicators corresponding to structurally disadvantaged areas, particularly those in Eastern and Southeastern Anatolia, are expected to carry positive and significant marginal effects relative to Istanbul, reflecting place-based deficits in formal labor demand, institutional capacity, and access to education and care infrastructure.

The empirical specification translates the theoretical mechanisms directly into estimable quantities. Human capital theory and school-to-work transition research motivate the inclusion of categorical education indicators, which capture the extent to which completed schooling lowers NEET risk. In addition to that, gendered labor supply theory and the household-specialization literature motivate the inclusion of marital status and the presence of a child under age six as proxies for the care and domestic responsibilities. These proxies are hypothesized to constrain women's participation disproportionately. Spatial inequality perspectives motivate NUTS-1 region fixed effects, which absorb persistent place-based differences in labor demand, institutional capacity, and social norms that would otherwise confound individual-level associations. By specifying the model in this way, the empirical analysis tests whether the theoretical mechanisms identified in the literature operate as expected in the Turkish context, and whether their magnitudes differ between women and men.

Data and Variables

The analysis uses the Turkish Household Labor Force Survey (HLFS) 2022 microdata provided by the Turkish Statistical Institute (TurkStat, 2022). The HLFS provides harmonized information on labor-market attachment, job-search behaviour, educational enrolment, and household composition, which allows NEET status to be constructed consistently and linked to a set of individual and contextual correlates within a single cross-section. The analytical sample is restricted to youth aged 15–29 and includes only observations with non-missing values for the dependent variable and the core covariates, yielding N = 123,834 individuals (62,413 women and 61,421 men; see Table 1). The dependent variable which is an indicator of NEET status is constructed from three HLFS survey items. First, employment status is determined using the reference-week employment question. The respondents who performed any paid work of at least one hour, or who held a job from which they were temporarily absent, are classified as employed. Second, educational enrolment is assessed using the question on current registration in a formal education or training program (including apprenticeships and publicly recognized vocational courses). Third, a residual job-search activity variable is used descriptively to distinguish unemployed NEET (those seeking work but not employed and not enrolled) from inactive NEET (those neither employed nor enrolled nor seeking work). An individual is assigned NEET = 1 if and only if they satisfy neither the employment condition nor the education/training enrolment condition at the time of the survey interview. Descriptively, 27.5% of youth are NEET (72.5% non-NEET), but the distribution is sharply gendered: 38.6% of young women are NEET compared to 16.2% of young men (Table 1). This large raw gap motivates modelling gender differences explicitly and, in the econometric analysis, estimating specifications separately by women and men to allow determinants to differ in magnitude and sign.

The HLFS provides person-level survey weights constructed by TurkStat to ensure national representativeness through post-stratification to population benchmarks. Following Wooldridge (2001), when the regression model includes the full set of variables used in survey stratification and post-stratification, the unweighted maximum likelihood estimator is consistent and asymptotically efficient, and weighting is not required for valid inference. As a transparency check, the weighted and unweighted NEET rates differ by at most 0.6 percentage points in the overall sample and by at most 0.9 percentage points within each gender sub-group, confirming that the estimation sample closely mirrors the target population.

The first set of covariates capture demographic and life-course transitions. Gender enters as a binary indicator (the sample is approximately balanced by sex; Table 1), and marital status is measured with married versus not married indicators. Marriage is empirically salient in the Türkiye context because it often coincides with changes in educational continuation, labor supply, and household responsibilities. In the HLFS youth sample, 23.0% are married overall, but marriage is substantially more prevalent among women (30.9%) than men (14.9%), implying that marital transitions are a plausible contributor to observed gender differences in NEET status (Table 1). Age is included as a continuous variable (mean 21.667) together with age squared to permit non-linearity over the youth life-course (Table 1). This quadratic specification is theoretically justified because NEET risk may rise during schooling-to-work transitions in the late teens and early twenties and then decline as labor-market attachment increases, and the age-squared term allows the data to reflect such turning points rather than imposing a linear profile.

A second core block of covariates measures education as a categorical indicator of human capital and school-to-work transition capacity. Education is coded into mutually exclusive categories: no schooling, primary, middle school, high school, vocational high school, vocational college, university, and master's/PhD. The education distribution is concentrated in middle school (41.3%), followed by high school (18.3%) and vocational high school (13.9%), while tertiary attainment remains more limited (university 10.7%, master/PhD 1.1%; Table 1). Importantly, women are not uniformly less educated than men in this youth cohort; women have higher shares in university (12.6% vs 8.8%) and vocational college (9.8% vs 7.5%), whereas men are more concentrated in middle school (44.0% vs 38.6%) and vocational high school (15.8% vs 12.1%) (Table 1). This compositional fact matters for interpretation because it implies that the higher NEET prevalence among women cannot be attributed solely to lower educational attainment; instead, it points to additional constraints, potentially household roles, care burdens, or local labor-market conditions, that can keep women out of employment and education/training even at higher schooling levels.

Household composition and constraints are proxied through household size and the presence of young children. Household size (mean 4.295) captures living arrangements that may influence NEET through resource pooling, dependency burdens, and norms around youth economic participation (Table 1). The expected relationship is ambiguous a priori; larger households can reduce immediate labor-supply pressure via shared resources or increase it via dependency demands, so it is treated as an empirical control rather than a variable with a predetermined sign. The presence of a child under age six in the household is included as a binary indicator (23.7% overall), with substantial gender differences in exposure: 29.8% for women versus 17.5% for men (Table 1). This variable is central for the Türkiye context because it serves as a strong proxy for time-intensive care needs and household responsibilities, which are often disproportionately borne by young women and can directly limit feasible participation in work or education/training; the descriptive gender gap in exposure to young children therefore anticipates gender-asymmetric NEET determinants in the multivariate analysis.

Finally, regional opportunity structures are captured through NUTS-1 (IBBS-1) region indicators: Istanbul, Western Marmara, Aegean, Eastern Marmara, Western Anatolia, Mediterranean, Central Anatolia, Western Black Sea, Eastern Black Sea, Northeastern Anatolia, Central Eastern Anatolia, and Southeastern Anatolia. These controls are necessary because NEET risk is not determined solely by individual endowments; it is also shaped by spatially uneven labor demand, sectoral composition, informality, institutional capacity, and access to education and care infrastructure. The sample distribution reflects population concentration across regions (Table 1), and modest gender differences in regional composition reinforce the importance of conditioning on region when interpreting gender gaps. Taken together, this variable set operationalizes a compact, theoretically coherent framework for NEET status: age and age squared capture life-course transitions; education captures human capital and employability; marital status and young-child presence capture household constraints that may be gendered; household size proxies living arrangements and resource contexts; and regional indicators capture spatial inequality in opportunity. This structure is aligned with the paper’s objective of explaining why NEET status in Türkiye is simultaneously gendered and geographically uneven, using nationally representative survey microdata.

Table 1

Descriptive Statistics of Variables

	Overall	Women	Men
Labor Market Status			
Neet	0.275	0.386	0.162
Non-Neet	0.725	0.614	0.838
Gender			
Women	0.504	1.000	0.000
Men	0.496	0.000	1.000
Marital Status			
Married	0.230	0.309	0.149
Not Married	0.770	0.691	0.851
Age	21.667	21.804	21.528
Age Squared	489.432	495.407	483.361
Household Size	4.295	4.350	4.238
Child under the age of 6	0.237	0.298	0.175
Education Level			
No Schooling	0.043	0.057	0.028
Primary	0.018	0.020	0.016
Middle School	0.413	0.386	0.440
High School	0.183	0.181	0.185
Voc. High School	0.139	0.121	0.158
Voc. College	0.087	0.098	0.075
University	0.107	0.126	0.088
Master's and PhD	0.011	0.011	0.010
Region			
Istanbul	0.079	0.077	0.082
Western Marmara	0.048	0.046	0.050
Aegean	0.105	0.101	0.109
Eastern Marmara	0.091	0.090	0.093
Western Anatolia	0.098	0.099	0.098
Mediterranean	0.111	0.113	0.109
Central Anatolia	0.069	0.070	0.069
Western Black Sea	0.070	0.070	0.071
Eastern Black Sea	0.040	0.039	0.040
Northeastern Anatolia	0.076	0.079	0.074
Centraleastern Anatolia	0.086	0.087	0.085
Southeastern Anatolia	0.125	0.130	0.120
Sample Size	123834	62413	61421

Note. Values are proportions. Calculated using HLFS 2022 Household Labor Force Survey microdata. Created by the author using STATA 19 statistical software.

Empirical Strategy

This study examines correlates of NEET status using cross-sectional microdata from the Turkish Household Labor Force Survey (HLFS) for 2022. The empirical approach treats NEET as a binary outcome and estimates how its likelihood varies with individual, household, and regional characteristics in a multivariate setting. Let $NEET_i \in \{0,1\}$ indicates whether individual i is classified as NEET (equal to 1 if not in employment, education, or training, and 0 otherwise).

To model this outcome, the paper estimates binary logit specifications of the form $\Pr(NEET_i = 1 \mid X_i) = \Lambda(X_i'\beta) = \exp(X_i'\beta)/[1 + \exp(X_i'\beta)]$, where $X_i$denotes the covariate vector and $\beta$is estimated by maximum likelihood. The logit functional form ensures predicted probabilities lie between zero and one and allows covariates to affect the outcome probability nonlinearly through the logistic index. To assess the sensitivity of the results to the choice of binary model functional form, the pooled and gender-stratified models are re-estimated using a probit specification; the results are reported in Appendix A and are so similar to the logit estimates in both sign and magnitude.

Models are estimated for the pooled sample and separately for women and men to examine gender heterogeneity. In the pooled model, gender enters directly as a regressor; in gender-stratified estimations, the gender indicator is omitted because it is constant within each subsample. Estimating separate models avoids imposing that the relationship between covariates and NEET status is identical for women and men, which is particularly relevant in Türkiye, where household roles and labor market attachment differ sharply by gender.

The covariate set is chosen to capture human capital, demographic constraints, household context, and spatial heterogeneity. Age is included in levels and squared to flexibly model the life-cycle profile of NEET risk, which may be non-monotonic over youth ages due to schooling completion, job search, and delayed transitions into stable employment. Educational attainment is incorporated through a set of categorical indicators based on completed education levels (with one category serving as the reference group), so coefficients capture differences in NEET propensity relative to that baseline.

Marital status is controlled by being married, motivated by the role of household formation and within-household specialization in shaping labor supply and training participation. Household characteristics include an indicator for the presence of at least one child under age six in the household and a measure of household size. The child-under-six indicator proxies early-childhood care responsibilities that can constrain participation in paid work and training, while household size captures broader household composition and potential support or care burdens that may correlate with disengagement.

To account for systematic geographic differences in opportunities and norms, the specification includes NUTS-1 (IBBS-1) region fixed effects (implemented as region indicators). These fixed effects absorb time-invariant regional heterogeneity in labor market conditions, education and training access, and social norms, thereby reducing confounding from omitted region-level factors correlated with individual and household characteristics. All regressors are entered jointly, so estimated associations are conditional on the full set of controls and region fixed effects.

For statistical inference, standard errors are clustered at the household level to allow for arbitrary within-household correlation in unobserved determinants of NEET status. This adjustment is important in household surveys because multiple individuals can face shared constraints and shocks, and ignoring within-household dependence can understate uncertainty.

Results are reported both as logit coefficients with clustered standard errors and as average marginal effects (AMEs) to facilitate interpretation in probability units. For continuous covariates, the marginal effect is $\partial \Pr(NEET_i = 1 \mid X_i)/\partial x_{ik} = \Lambda(X_i'\beta)[1 - \Lambda(X_i'\beta)]\beta_k$, and AMEs average these individual-level effects over the estimation sample; for indicator variables and categorical covariates such as education groups and regions, AMEs are computed as discrete changes in predicted probabilities relative to the reference category. Presenting AMEs

alongside coefficients makes the magnitude of associations directly interpretable as percentage-point differences in the probability of being NEET.

The average adjusted predictions (predictive margins) by education level are estimated separately for women and men. For each education category, the predicted NEET probability is computed for every individual in the relevant sub-sample by setting the education variable to that category while retaining all other covariates at their observed values; the displayed value is the mean of these individual-specific predictions across the sub-sample. The margins at representative values approach ensures that the plotted probabilities reflect the actual covariate composition of the estimation sample rather than an artificial reference profile. It is therefore interpretable as the average NEET probability that would be observed if all individuals in the sub-sample have the specified education level.

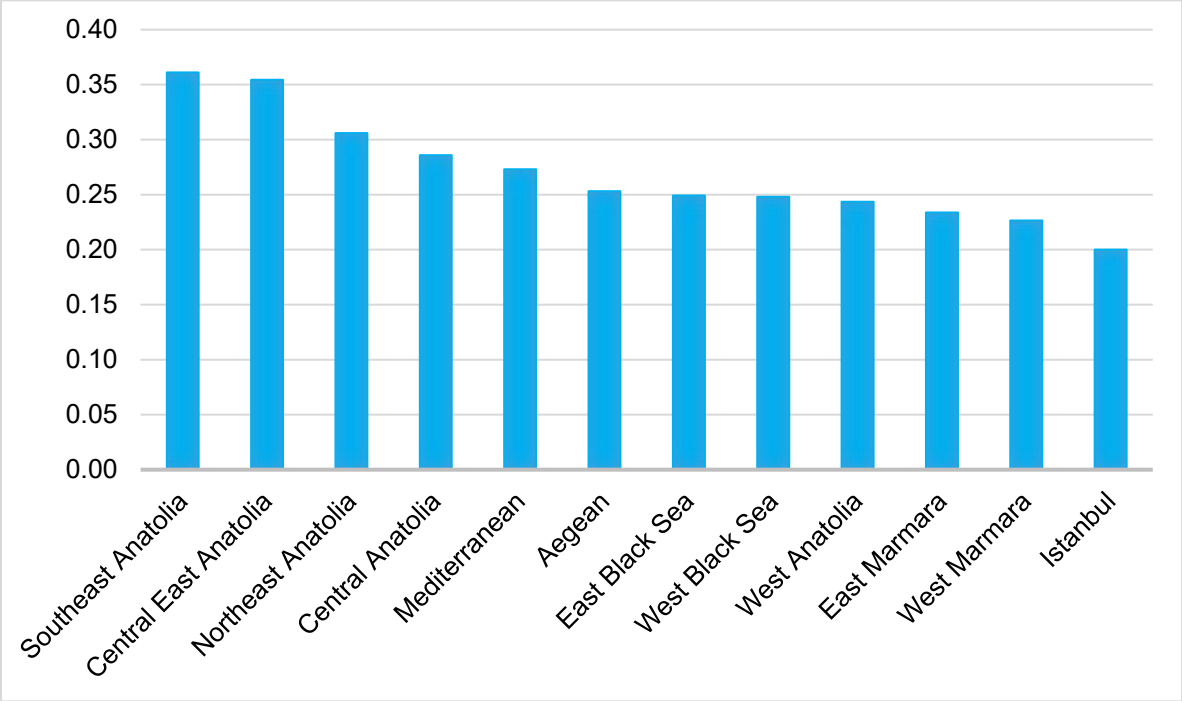
Because the analysis is cross-sectional, all estimates are interpreted as conditional correlations rather than causal effects, and it is important to be explicit about the boundaries of causal inference. Three sources of potential endogeneity are worth noting. First, marital status is likely jointly determined with labor market attachment: women with weaker expected earnings or lower attachment may both marry earlier and have lower employment rates, so the positive association between marriage and female NEET could reflect selection rather than a causal constraint imposed by marriage itself. Second, the presence of a child under age six may similarly be endogenous: households where a young woman is already out of employment or training may face weaker financial disincentives to having a child at an earlier stage of the life course. Third, educational attainment, while predetermined for most of the sample, may remain endogenous for younger individuals who have not yet exited school. In the absence of credible instruments, the estimated associations should not be read as causal impacts. The empirical strategy nonetheless provides a systematic and policy-relevant description of how NEET status is patterned across key characteristics in Türkiye, and whether these patterns differ between women and men. While this cannot substitute for casual evidence, it offers a descriptive foundation on which future causal research using longitudinal panels or quasi-experimental policy variation can build.

Results

The descriptive evidence indicates substantial heterogeneity in NEET status across Türkiye, both geographically and across demographic and socioeconomic groups. At the regional level, Figure 1 documents a clear dispersion in NEET prevalence across NUTS-1 (IBBS-1) regions: Istanbul stands out with the lowest NEET rate (roughly around one-fifth), whereas Southeast Anatolia and Central East Anatolia record the highest rates (in the mid–0.30 range), with the remaining regions typically lying in between. Figure 1, therefore, motivates the inclusion of NUTS-1 region fixed effects in the multivariate specifications to account for persistent spatial differences in labor market opportunities, institutional capacity, and local norms that may correlate with individual characteristics.

Figure 1

NEET rates by NUTS-1 region



Note. Calculated using HLFS 2022 Household Labor Force Survey microdata using STATA 19 statistical software.

Table 2 reports logit estimates for the pooled sample, presenting both coefficients with household-clustered standard errors and average marginal effects (AMEs) expressed in probability units. In the pooled model, the gender gap is large, precisely estimated, and economically meaningful. Women display a substantially higher probability of being NEET relative to men, even after conditioning on age, education, marital status, household characteristics, and region fixed effects. The corresponding AME implies a sizable percentage-point difference in NEET probability attributable to gender conditional on observables, which indicates that the observed gender gap cannot be explained solely by compositional differences in education, age structure, or region of residence.

The age profile is nonlinear and statistically significant, consistent with a youth transition process that is not monotonic. In Table 2, age enters positively while age-squared enters negatively, implying an inverted-U relationship between age and NEET probability. NEET risk rises over early youth ages and then flattens or declines at later ages. This pattern is present in the pooled model and persists in the gender-stratified models, but the curvature is steeper for women than for men. Substantively, the stronger age profile for women is consistent with the timing of school exit and labor market entry coinciding with household formation, which can interact with norms and constraints around women's market participation.

Education exhibits a strong gradient in Table 2, with higher completed schooling associated with markedly lower NEET probabilities relative to the omitted base category. Across most education levels, the estimated AMEs imply large reductions in the likelihood of NEET status, underscoring that educational attainment remains one of the most powerful correlates of youth engagement. However, the pattern is not purely monotonic when interpreted

alongside the predicted probabilities. Some intermediate education levels are associated with elevated risk relative to the highest education group, consistent with a difficult school-to-work transition segment in the distribution where youths have completed schooling but have not yet achieved stable labor market attachment.

Table 2

Determinants of NEET Status, Logit Model Estimates (Overall)

Variable	Coefficient (Std. Error)	Marginal Effect (Std. Error)
gender	1.0978***	0.1754***
	(0.0158)	(0.0023)
age	1.2528***	0.2002***
	(0.0223)	(0.0035)
age2	-0.0264***	-0.0042***
	(0.0005)	(0.0001)
primary school	-0.3791***	-0.0766***
	(0.0680)	(0.0137)
middle school	-1.7076***	-0.3237***
	(0.0430)	(0.0086)
high school	-1.4672***	-0.2845***
	(0.0442)	(0.0087)
vocational high school	-1.4560***	-0.2826***
	(0.0452)	(0.0088)
vocational collage	-1.6874***	-0.3205***
	(0.0467)	(0.0089)
university	-1.7353***	-0.3280***
	(0.0466)	(0.0088)
master's and PhD	-2.0808***	-0.3777***
	(0.0859)	(0.0130)
married	0.6344***	0.1014***
	(0.0226)	(0.0036)
western marmara	0.2147***	0.0343***
	(0.0443)	(0.0071)
aegean	0.4052***	0.0647***
	(0.0360)	(0.0057)
eastern marmara	0.2675***	0.0427***
	(0.0372)	(0.0059)
western anatolia	0.2851***	0.0456***
	(0.0362)	(0.0058)
mediterranean	0.5023***	0.0803***
	(0.0357)	(0.0057)
central anatolia	0.5541***	0.0885***
	(0.0385)	(0.0061)
western black sea	0.3851***	0.0615***
	(0.0395)	(0.0063)
eastern black sea	0.4586***	0.0733***
	(0.0468)	(0.0075)
northeastern anatolia	0.5294***	0.0846***
	(0.0392)	(0.0062)
centraleastern anatolia	0.7719***	0.1233***
	(0.0383)	(0.0061)
southeastern anatolia	0.7384***	0.1180***
	(0.0357)	(0.0057)
HH has at least one child (<6)	0.2850***	0.0455***
	(0.0212)	(0.0034)
HH size	0.0565***	0.0090***

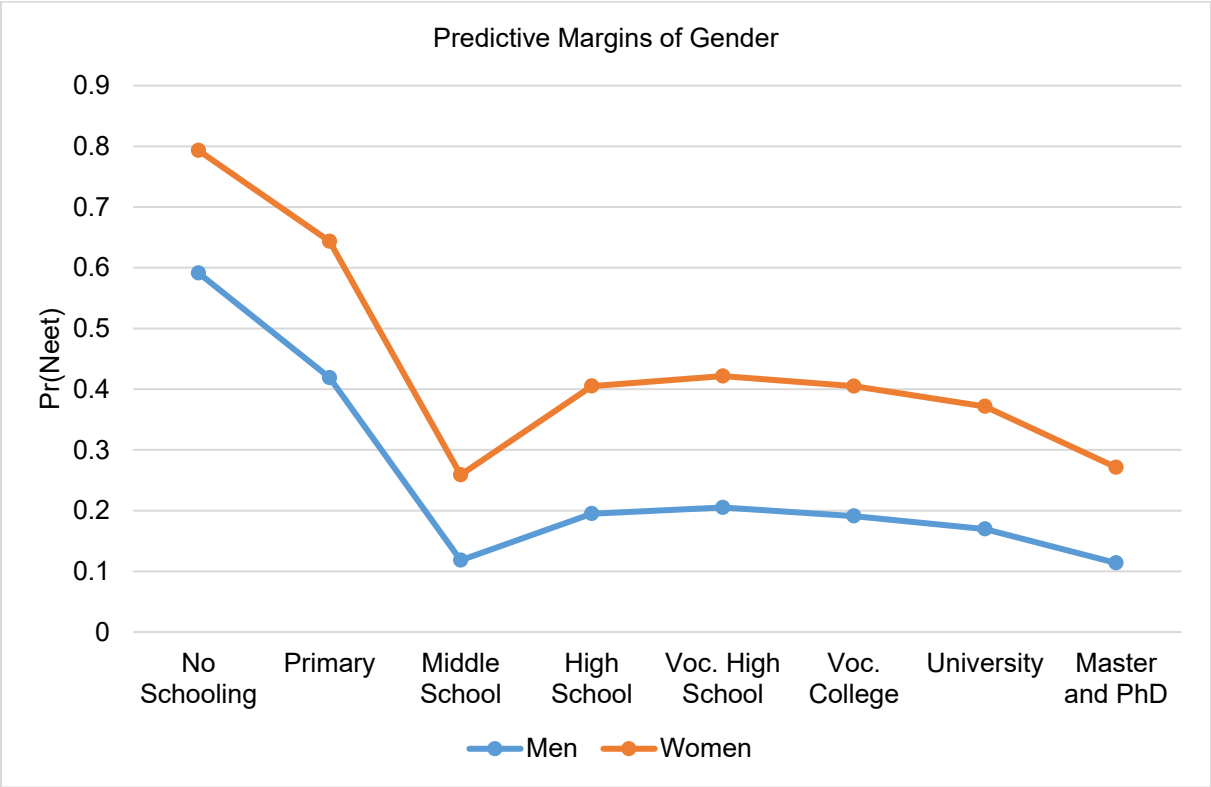
	(0.0051)	(0.0008)
Constant	-15.3641***	
	(0.2517)	

Significance levels *, ** and *** denote .10, .05, .01 respectively. N=123,834
Note. Created by the author using STATA 19 statistical software.

Figure 2 visualizes predicted probabilities of NEET status by education level separately for men and women and complements the regression evidence in two keyways. First, the plot shows that women have higher predicted NEET probabilities at every education level, reinforcing the finding from Table 2 that the gender gap persists conditional on covariates. Second, the predicted education profile suggests that NEET risk is especially high at the very bottom of the education distribution, falls sharply at middle schooling, rises through certain upper-secondary and vocational tracks, and then is lower again at tertiary levels. This nonlinearity is consistent with the coexistence of structural disadvantages at low schooling and transition-related frictions at intermediate schooling levels.

Figure 2

NEET and the gender–education gap



Note. Calculated using HLFS 2022 Household Labor Force Survey microdata using STATA 19 statistical software. Predicted probabilities are average adjusted predictions (predictive margins) from the gender-stratified logit models. For each education category, the NEET probability is computed for every individual in the sub-sample at that education level, holding all other covariates at their observed values; the displayed value is the sub-sample mean.

Marital status is strongly associated with NEET status, but the direction of the relationship differs sharply by gender. Table 3 shows that, conditional on controls, marriage is associated with a higher probability of NEET among women, while Table 4 indicates that

marriage is associated with a lower probability of NEET among men. This asymmetry is consistent with household specialization and gendered labor market attachment. Marriage is correlated with women's withdrawal from employment and training or reduced likelihood of entering them, whereas it is correlated with stronger labor force attachment among men. The gender-stratified results are therefore important because the pooled estimate alone would mask these opposite-signed relationships.

Household constraints, proxied by the presence of a child under age six and household size, further reinforce the gendered structure of NEET status. In Table 3, living in a household with a young child is associated with a significantly higher probability of being NEET for women, while in Table 4, the association is small and negative for men. This pattern aligns with a care-constraint channel in which childcare responsibilities reduce women's ability to participate in employment and training, while men's labor market attachment remains comparatively stable. Household size also carries information about household composition and potential care burdens or support networks, and its estimated association provides additional evidence that NEET status is shaped not only by individual characteristics but also by the household context in which youths make schooling and labor market decisions.

Regional fixed effects remain substantively important even after controlling for a rich set of individual and household covariates. Table 2 indicates that, relative to the omitted base region, many regions exhibit higher predicted probabilities of NEET status, and the largest regional differences remain meaningful in probability units. This pattern is consistent with the descriptive dispersion observed in Figure 1 and suggests that place-based factors, such as local labor demand, sectoral composition, and access to education and training opportunities, continue to shape NEET risk beyond individual education and household characteristics.

Taken together, the combined evidence from Table 2 and Figures 1–2 points to three core empirical messages about youth disconnection in Türkiye in 2022, each of which carries distinct implications for how NEET should be interpreted and what kinds of policy levers are likely to be effective. First, NEET status is characterized by a large and persistent gender gap that survives extensive controls for human capital, demographics, household structure, and NUTS-1 region fixed effects. The fact that the female–male difference remains sizeable in the fully specified model implies that Türkiye's NEET disadvantage for women cannot be reduced to compositional differences in observable characteristics such as education or regional residence. Instead, the persistence of the gap is consistent with a structural interpretation in which women face systematically tighter constraints on labor-market participation and training access, constraints that operate through norms, the organization of unpaid work, and the availability of "acceptable" jobs. Importantly, Figure 2 reinforces this point visually by showing that women's predicted NEET probabilities exceed men's across the entire education distribution. That pattern suggests that the gender gap is not confined to the lowest-skilled segment; it also extends into tertiary-educated groups, where one would otherwise expect education to translate more reliably into labor-market attachment. In substantive terms, this means that even if educational attainment rises, aggregate NEET outcomes will remain high unless the institutional and labor-market conditions mediate women's transitions, such as childcare availability, flexible but protected work, safe transport, and the availability of formal entry-level jobs, also improve. In this sense, the results are consistent with the broader framing in the paper's introduction and comparative context: Türkiye's NEET problem is not simply a youth unemployment issue but a gendered disengagement challenge in which inactivity pathways play a central role.

Second, education is strongly protective overall, yet the education–NEET relationship is shaped by transition dynamics in ways that caution against a purely linear "more schooling always and everywhere reduces NEET" narrative. The regression estimates show large probability reductions associated with higher completed education relative to the base category, indicating that human capital accumulation remains one of the most consequential correlates of NEET status. However, Figure 2 suggests that predicted NEET risk does not decline smoothly at every step of the education ladder. Instead, the profile is consistent with non-monotonicity at intermediate schooling levels, an empirical pattern that aligns with the idea that NEET status is partly a function of where individuals sit in the school-to-work transition process. Intermediate levels of schooling may coincide with a phase where young people have exited education but have not yet secured stable employment, making them especially exposed to matching frictions, limited work experience, or constrained local opportunities. This is particularly plausible in a context where entry-level formal jobs are scarce, where vocational and general tracks differ in their signaling value and labor-market relevance, and where transitions may depend heavily on local employer networks. The key implication is that policies focused only on raising attainment, while important, are not sufficient unless they are complemented by measures that improve the transition architecture, career guidance, work-based learning, employer engagement, and credible pathways from schooling into formal employment. Moreover, because Figure 2 shows that women remain at higher predicted NEET risk even at higher education levels, improving transitions is also inseparable from addressing gendered constraints that limit the conversion of education into employment for women.

Third, household formation variables, marriage, and the presence of a child under age six, emerge as central correlates that operate in sharply different ways for women and men, as documented in Table 3 and Table 4. This gender asymmetry is more than a technical detail; it provides direct evidence that the mechanisms behind NEET status differ by gender and that pooling women and men can conceal fundamentally different pathways into disconnection. Among women, marriage and young-child presence are associated with substantially higher NEET probabilities, consistent with a care-constraint channel in which unpaid domestic and childcare responsibilities reduce the feasibility of job search, training participation, and sustained employment, thereby pushing women toward inactivity within the NEET category. Among men, by contrast, the associations are weaker and often opposite in sign, consistent with a breadwinner-oriented pattern in which household formation correlates with stronger labor-market attachment. This asymmetry helps explain why the overall NEET problem in Türkiye is so strongly gendered and why conventional labor-market interventions targeted at unemployed youth will not necessarily reach the larger inactivity-driven component among women. It also reinforces the paper's policy logic: reducing NEET among women requires interventions that relax participation constraints, especially childcare and care support, alongside improvements in job availability and job quality.

A final cross-cutting message concerns geography. Figure 1 documents pronounced regional dispersion in NEET prevalence, and the regression results indicate that substantial geographic heterogeneity persists even after conditioning on education, demographics, and household characteristics. This persistence suggests that place-based opportunity structures, local labor demand, sectoral composition, the density of formal employment, and the accessibility of education and training services remain salient determinants of NEET risk rather than mere reflections of population composition. The coexistence of strong regional differences with the household and gender patterns highlighted above implies that spatial inequality and gendered constraints likely interact in regions where formal jobs are scarce and services are

thin, the effective opportunity set, particularly for young women, may be even narrower, raising the likelihood of inactivity. In sum, the combined evidence supports an interpretation of NEET in Türkiye as a multi-layered outcome produced by the interaction of (i) persistent gender disparities in participation, (ii) education-linked differences in transition vulnerability, and (iii) place-based inequality in opportunities and services. This integrated reading aligns with the comparative benchmarks discussed earlier and provides a coherent empirical foundation for the discussion section's emphasis on gender-responsive and regionally targeted policy packages that link transition support and activation to childcare provision and quality job creation.

Table 3

Determinants of NEET Status, Logit Model Estimates (Women)

Variable	Coefficient (Std. Error)	Marginal Effect (Std. Error)
age	1.3490***	0.2387***
	(0.0299)	(0.0051)
age2	-0.0281***	-0.0050***
	(0.0007)	(0.0001)
primary school	0.1397	0.0243
	(0.1114)	(0.0193)
middle school	-1.5890***	-0.3010***
	(0.0629)	(0.0113)
high school	-1.3868***	-0.2634***
	(0.0643)	(0.0115)
vocational high school	-1.3606***	-0.2585***
	(0.0662)	(0.0118)
vocational collage	-1.6848***	-0.3183***
	(0.0668)	(0.0118)
university	-1.8813***	-0.3528***
	(0.0662)	(0.0116)
master's and PhD	-2.2301***	-0.4099***
	(0.1133)	(0.0183)
married	1.1650***	0.2062***
	(0.0306)	(0.0052)
western marmara	0.2788***	0.0493***
	(0.0584)	(0.0103)
aegean	0.4043***	0.0715***
	(0.0472)	(0.0083)
eastern marmara	0.3058***	0.0541***
	(0.0487)	(0.0086)
western anatolia	0.4216***	0.0746***
	(0.0470)	(0.0083)
mediterranean	0.5463***	0.0967***
	(0.0465)	(0.0082)
central anatolia	0.7115***	0.1259***
	(0.0510)	(0.0090)
western black sea	0.4537***	0.0803***
	(0.0520)	(0.0092)
eastern black sea	0.3596***	0.0636***
	(0.0613)	(0.0108)
northeastern anatolia	0.5771***	0.1021***
	(0.0519)	(0.0092)
centraleastern anatolia	0.7991***	0.1414***
	(0.0501)	(0.0088)

southeastern anatolia	0.8297***	0.1468***
	(0.0466)	(0.0082)
HH has at least one child (<6)	0.4215***	0.0746***
	(0.0278)	(0.0049)
HH size	0.0697***	0.0123***
	(0.0073)	(0.0013)
Constant	-15.8974***	
	(0.3357)	
Observations	62413	

Significance levels *, ** and *** denote .10, .05, .01 respectively.
Note. Created by the author using STATA 19 statistical software.

Table 4

Determinants of NEET Status, Logit Model Estimates (Men)

Variable	**Coefficient (Std. Error)**	**Marginal Effect (Std. Error)**
age	0.9801***	0.1230***
	(0.0381)	(0.0048)
age2	-0.0209***	-0.0026***
	(0.0009)	(0.0001)
primary school	-0.8568***	-0.1845***
	(0.0952)	(0.0195)
middle school	-1.9608***	-0.3461***
	(0.0610)	(0.0132)
high school	-1.5433***	-0.2966***
	(0.0615)	(0.0133)
vocational high school	-1.5823***	-0.3018***
	(0.0626)	(0.0134)
vocational collage	-1.6459***	-0.3101***
	(0.0685)	(0.0137)
university	-1.3348***	-0.2667***
	(0.0662)	(0.0139)
master's and PhD	-1.6006***	-0.3042***
	(0.1287)	(0.0198)
married	-0.7141***	-0.0896***
	(0.0482)	(0.0060)
western marmara	0.1115	0.0140
	(0.0745)	(0.0093)
aegean	0.4236***	0.0532***
	(0.0581)	(0.0073)
eastern marmara	0.2248***	0.0282***
	(0.0612)	(0.0077)
western anatolia	0.1022*	0.0128*
	(0.0618)	(0.0078)
mediterranean	0.4736***	0.0594***
	(0.0580)	(0.0073)
central anatolia	0.3723***	0.0467***
	(0.0644)	(0.0081)
western black sea	0.3123***	0.0392***
	(0.0650)	(0.0081)
eastern black sea	0.6055***	0.0760***
	(0.0721)	(0.0090)
northeastern anatolia	0.5213***	0.0654***
	(0.0629)	(0.0079)
centraleastern anatolia	0.8262***	0.1037***
	(0.0596)	(0.0075)

southeastern anatolia	0.7489***	0.0940***
	(0.0571)	(0.0072)
HH has at least one child (<6)	-0.1360***	-0.0171***
	(0.0424)	(0.0053)
HH size	0.0654***	0.0082***
	(0.0076)	(0.0010)
Constant	-11.6845***	
	(0.4194)	
Observations	61421	

Significance levels *, ** and *** denote .10, .05, .01 respectively.
Note. Created by the author using STATA 19 statistical software.

Discussion and Policy Implications

The empirical results portray NEET status in Türkiye as a multidimensional form of youth exclusion shaped jointly by individual endowments, household constraints, and spatially differentiated opportunity structures. Three findings are particularly salient. First, a large gender gap persists even after conditioning on education, age, marital status, household size, the presence of young children, and region fixed effects. This indicates that the observed female disadvantage is not simply compositional; instead, it is consistent with gendered constraints and norms that limit women's participation in paid work and training. Second, household formation variables operate in strongly asymmetric ways by gender: marriage and the presence of a child under age six are associated with higher NEET probabilities among women but lower or near-zero associations among men. This pattern is consistent with specialization within households and unequal distribution of unpaid care, which in turn translates into labor market detachment for women rather than unemployment in the narrow sense. Third, substantial regional heterogeneity remains even after extensive controls, implying that local labor demand conditions, sectoral structures, and the availability or quality of education and training opportunities meaningfully shape NEET risk. Together, these findings suggest that NEET in Türkiye is not reducible to a single "youth unemployment" problem; it reflects a combination of demand-side constraints, uneven human capital accumulation and utilization, and gendered barriers to sustained attachment to the labor market and training systems.

A central implication of the results is that policy packages focused narrowly on job-search assistance, while potentially useful for unemployed NEET, will have limited reach for the segment of NEET youth whose disengagement is driven by inactivity. The gender-stratified estimates provide strong evidence that women's NEET status is closely linked to household roles and care responsibilities, particularly in households with young children (Danner et al., 2021). In such settings, activation policies that assume immediate job readiness may fail because the binding constraints are time availability, caregiving, mobility, and the ability to access reliable, affordable childcare. Accordingly, policies aimed at reducing inactivity NEET among young women should place childcare and early-childhood services at the center rather than at the periphery (Ilkkaracan et al., 2021). Expansion of affordable childcare, especially for children under age three, where care needs are most intensive, can directly relax constraints that prevent job search, training participation, and sustained employment (Ekiz Gökmen, 2022; Gultekin et al., 2025). Importantly, childcare expansion is most effective when paired with working conditions that make labor market re-entry feasible, including predictable schedules, part-time options with social protection, and access to formal employment rather than informal or unstable work arrangements that may not justify the fixed costs of childcare and commuting.

The divergence in the association of marriage with NEET status across genders also has policy relevance. The estimates are consistent with marriage being correlated with higher

female NEET probability and lower male NEET probability, which aligns with household specialization and traditional gender roles. In this setting, interventions that only improve employability without addressing intra-household allocation of time and expectations may produce limited gains. Gender-responsive activation should therefore include components that reduce "secondary earner penalties" and increase incentives for female employment, such as subsidies for childcare and commuting, targeted training stipends, and formal employment incentives designed to offset the opportunity cost of market work relative to home production (Ernst et al., 2025). Parallel efforts to encourage more equal sharing of care responsibilities, through public awareness, employer practices, and potentially parental leave designs that do not reinforce gender specialization, can complement labor-market tools (Kang & Youn, 2024). While these measures extend beyond conventional labor policy, the results underscore that NEET status for young women is often rooted in the organization of care and household responsibilities rather than in job-search intensity alone (Odoardi et al., 2023).

The education gradient is strong in the pooled model, but the predicted probabilities by education reveal a more nuanced pattern than a simple monotonic decline in NEET risk with schooling. Very low education is associated with extremely high NEET probabilities, consistent with cumulative disadvantage, early school leaving, and weak attachment to the formal labor market. This group calls for preventative strategies: reducing early school leaving, strengthening second-chance education, and providing modular, stackable training that can generate labor market value quickly. However, the predicted profiles also suggest that some intermediate education levels coincide with elevated NEET probabilities, consistent with frictions in the school-to-work transition. This implies that policies should not only raise attainment but also improve the transition architecture: high-quality vocational guidance, structured work-based learning, employer engagement in curriculum design, and placement services that actively broker entry into formal jobs (Kebede et al, 2024; Farran & Nunez, 2025). In practice, this points toward strengthening apprenticeships and dual-training style mechanisms, but with careful attention to quality and to avoiding the channeling of young women into low-wage, low-mobility occupations. Good transition systems do not simply place youths into any job; they should facilitate matches that build skills and generate progression paths.

A particularly important and policy-relevant implication is that the gender gap persists across all education levels, including among tertiary-educated groups. Persistent NEET risk among highly educated women indicates that raising education alone is insufficient if labor demand does not translate credentials into stable employment or if workplaces fail to accommodate constraints that disproportionately affect women (Lu & Li, 2025). This finding supports policies that improve school-to-work matching and early-career job quality: strengthened career services in universities and vocational colleges, expanded internship pipelines with enforceable standards, employer partnerships that provide credible pathways from study to formal employment, and active monitoring of discriminatory hiring or workplace practices. It also highlights the importance of labor market institutions that raise the returns to formal employment for young women, such as enforcement of equal treatment, expansion of formal-sector opportunities, and social protection arrangements that make part-time or flexible employment compatible with long-term earnings and career progression.

The regional stratification documented in the descriptive evidence and reinforced by the regression results implies that national one-size-fits-all policies are unlikely to be efficient. The magnitude and ordering of the NUTS-1 region coefficients in Tables 2-4 provide a direct empirical basis for regional differentiation, consistent with broader evidence on the role of

territorial opportunity structures in shaping youth labor market outcomes (Cefalo & Scandurra, 2021; Lindblad et al., 2025). The two highest-burden regions, Central Eastern Anatolia and Southeastern Anatolia, exhibit the largest positive marginal effects relative to Istanbul and share structural characteristics: below-average formal employment density, high informality rates, weaker transport and childcare infrastructure, and, in the case of Southeastern Anatolia, particularly high female NEET shares. In these regions, the binding constraint is not principally a skills deficit but rather a shortage of formal entry-level jobs and a near-absence of care infrastructure; accordingly, the policy priority is demand-side: time-limited wage subsidies contingent on formal registration and retention, employer incentives to create apprenticeships and entry-level positions, and co-investment in local childcare and transport facilities that lower the fixed cost of participating in the labor market. Regions with intermediate coefficients, including the Mediterranean, Central Anatolia, and the two Black Sea regions, typically combine moderate formal job scarcity with stronger educational enrolment; here, transition-support measures take precedence: career guidance, structured work-based learning, employer engagement in curriculum design, and placement services that broker entry into formal employment. Finally, regions with relatively lower NEET margins, Western Anatolia, Eastern Marmara, and Western Marmara, may benefit most from targeted refinements to existing active labor market program, focusing on job-search quality, credential recognition, and reducing barriers for young women specifically. Effective place-based policy ultimately requires coordination across agencies. Training ministries, employment services, municipalities, and employers need to operate within a shared regional strategy rather than in parallel silos, with outcomes monitored separately by gender and region to track whether gaps are actually narrowing.

Across all these policy domains, the results also suggest that evaluation and targeting matter. Because NEET is heterogeneous, mixing unemployed jobseekers and inactive youths, interventions should be tailored to the underlying constraint. A practical approach is to segment NEET youth into groups using the same microdata logic as the paper: inactive NEET (often care-constrained), unemployed NEET (job-search active), and potentially "discouraged" or "long-term detached" subgroups where mental health, discouragement, or low perceived returns might be binding. The gender-specific estimates in the study provide a clear basis for such segmentation: women with young children and married women represent high-priority groups for childcare-centered activation; men in high-NEET regions may benefit more from demand-side and matching interventions; and low-educated youths require early interventions that prevent prolonged detachment. In addition, the strong regional variation implies that targeting should combine individual eligibility with geographic prioritization, especially where local labor markets are structurally weak.

Finally, the empirical findings should be interpreted with appropriate caution. The analysis is cross-sectional and therefore describes conditional associations rather than causal effects; marriage and fertility decisions are likely jointly determined with labor market attachment, and some household variables may capture both constraints and preferences. Nevertheless, the systematic patterns, especially the gender asymmetries in the correlates of NEET and the persistence of regional differences after controls, provide credible guidance for policy design. The evidence supports a policy agenda that shifts from a narrow unemployment-centered approach toward an integrated strategy combining childcare and care infrastructure, gender-responsive activation, improved transition systems from education to work, and regionally tailored demand-side measures that increase the availability of formal entry-level jobs.

Conclusion

Using nationally representative HLFS 2022 microdata, this paper provides an updated empirical account of the magnitude and correlates of youth NEET status in Türkiye and highlights that NEET is best understood as a layered form of disengagement rather than a simple proxy for youth unemployment. The results show that NEET risk is structured simultaneously by spatial opportunity gaps and by gendered pathways into inactivity. Regional stratification is pronounced in the descriptive evidence and remains meaningful even after conditioning on a rich set of individual and household characteristics, underscoring that place-based factors, local labor demand, institutional capacity, and access to education and training, shape the probability of youth disengagement beyond personal endowments. At the same time, the analysis documents a large and persistent gender gap that survives controls for age, education, marital status, household composition, and region fixed effects, indicating that women's higher NEET probabilities are not reducible to observable compositional differences.

Education emerges as a powerful protective correlate, with higher educational attainment associated with substantially lower NEET probabilities, yet the education gradient does not eliminate gender disparities. Predicted probability patterns by education indicate that women face elevated NEET risks at every schooling level, including among tertiary-educated groups, suggesting that improvements in educational attainment alone are insufficient if school-to-work transitions remain weak and if labor markets do not convert credentials into stable, formal employment for young women. Moreover, the gender-stratified estimates highlight the importance of household formation and care constraints: marriage and the presence of a child under age six are associated with higher NEET probabilities among women but not among men, pointing to the central role of unpaid care responsibilities and gendered norms in shaping labor market detachment. In this sense, Türkiye's NEET challenge reflects the intersection of human capital formation, labor demand conditions, and household-level constraints that operate differently for women and men.

These findings imply that effective policy responses must move beyond generic employability interventions toward an integrated strategy that differentiates between NEET subgroups and targets the binding constraints they face. Regionally, the persistence of large differences suggests the need for geographically tailored packages that link skills development to local job creation and employer engagement, rather than expanding training supply in isolation from demand. For young women, particularly those in households with young children, the results support gender-responsive activation centered on childcare availability and affordability, flexible and protected work arrangements, and measures that reduce mobility and safety barriers that limit access to jobs and training. The persistence of elevated NEET probabilities among highly educated women further emphasizes the importance of improving matching and early-career job quality through stronger career guidance, work-based learning pipelines, and credible pathways into formal entry-level employment.

The paper's contribution is therefore twofold. Substantively, it clarifies that NEET in Türkiye is shaped by a combination of spatial inequality and gendered inactivity mechanisms, and that these dimensions remain central even after controlling for education and demographics. Empirically, it provides evidence in both coefficient and probability terms, allowing the magnitude of associations to be interpreted as meaningful changes in predicted NEET risk. While the analysis is cross-sectional and thus does not identify causal effects, particularly for variables intertwined with household decisions such as marriage and child presence, the patterns are sufficiently systematic to guide policy prioritization. Overall, reducing NEET in

Türkiye will likely require coordinated policies that jointly expand opportunity in lagging regions, strengthen education-to-work transitions, and remove structural barriers that keep young women outside employment and training, with childcare and the quality of entry-level jobs occupying a central role in any credible strategy.

Three directions stand out for future research. First, on causal identification, the conditional correlations documented here establish a credible descriptive foundation, but moving to causal estimates requires either longitudinal panel data that track individuals across NEET transitions or quasi-experimental designs that exploit policy variation. For example, geographic or cohort-based rollouts of childcare subsidies, regional employment incentive schemes, or compulsory schooling reforms can be the sources of exogenous variation in the key household and education variables. Second, on NEET heterogeneity, subsequent studies could apply latent class or finite mixture models to disaggregate the NEET population into subgroups such as care-constrained inactive youth, discouraged workers, and those in short-term transitional spells, and examine whether the correlates identified in the present analysis differ across these groups, thereby enabling more precisely targeted policy segmentation. Third, on dynamic analysis, panel data tracking individuals over multiple years would allow examination of NEET duration, and the long-run scarring effects of disconnection on earnings and employment quality. These are dimensions that are central to the long-run case for early intervention but simply cannot be addressed with cross-sectional survey microdata.

References

Akgül, S. (2025). Family matters: Exploring gendered pathways into NEET status in Türkiye. *Anadolu University Journal of Economics and Administrative Sciences*, *26*(3), 575-603. https://doi.org/10.53443/anadoluibfd.1676154

Bynner, J., & Parsons, S. (2002). Social exclusion and the transition from school to work: The case of young people not in education, employment, or training (NEET). *Journal of Vocational Behavior*, *60*(2), 289-309. https://doi.org/10.1006/jvbe.2001.1868

Cefalo, R., & Scandurra, R. (2021). Territorial disparities in youth labour market chances in Europe. *Regional Studies, Regional Science*, *8*(1), 228-238, https://doi.org/10.1080/21681376.2021.1925580

Choudhry, M. T., & Pastore, F. (2023). Determinants of school-to-work transition: Global outlook. *International Journal of Manpower*, *44*(6), 989-999. https://doi.org/10.1108/IJM-07-2023-0365

Çelik, K., & Lüküslü, D. (2011). Spotlighting a silent category of young females: The life experiences of "house girls" in Turkey. *Youth & Society*, *44*(1), 28-48. https://doi.org/10.1177/0044118X10391636

Danner, M., Guegnard, C., & Maguire, S. (2021). Understanding economic inactivity and NEET status among young women in the UK and France. *Journal of Education and Work*, *34*(7-8), 839-854. https://doi.org/10.1080/13639080.2021.1997017

Değirmenci, S. (2023). Female labor force participation in Türkiye: Revisiting the 1988-2008 period. *İktisat Politikası Araştırmaları Dergisi/Journal of Economic Policy Researches*, *10*(1), 213-244. https://doi.org/10.26650/JEPR1201276

Ekiz Gökmen, Ç. (2022). A new perspective on women's care burden and employment in Turkey. *New Perspectives on Turkey*, *66*, 11-34. https://doi.org/10.1017/npt.2021.21

Elitaş, T. (2025). Individual and family related determinants of NEET probability: An empirical study for Türkiye. *Anadolu University Journal of Economics and Administrative Sciences*, *26*(1), 538-567. https://doi.org/10.53443/anadoluibfd.1619992

Ernst, E., Merola, R., & Reljic, J. (2025). Fiscal policy instruments for inclusive labor markets: A review. *Journal of Economic Surveys*, *39*(5), 2020-2041. https://doi.org/10.1111/joes.12685

Eurostat. (2025). *Statistics on young people neither in employment nor in education or training*. Retrieved October 24, 2025, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Statistics_on_young_people_neither_in_employment_nor_in_education_or_training

Farran, I., & Nunez, I. (2025). Converging pathways: New approaches to integrate vocational education training and higher education. *Journal of Vocational Education & Training*, 77(4), 1147-1165. https://doi.org/10.1080/13636820.2024.2428769

Furlong, A. (2006). Not a very NEET solution: Representing problematic labour market transitions among early school-leavers. *Work, Employment and Society*, *20*(3), 553-569. https://doi.org/10.1177/0950017006067001

Gultekin, D., Ilkkaracan, I., Bayar, A. A., & Ozcanlı-Baran, P. (2025). Investing in the care services sectors for employment generation and gender-inclusive growth: The case of Turkey. *New Perspectives on Turkey*, *72*, 90-114. https://doi.org/10.1017/npt.2024.32

Gürsel, S., & Tanrıverdi, S. (2025, Ekim 9). *Genç kuşakta ne eğitimde ne istihdamda olanların çok boyutlu analizi* (Araştırma Notu, No.25/282). Bahçeşehir Üniversitesi Ekonomik ve Toplumsal Araştırmalar Merkezi

Ilkkaracan, I., Kim, K., Masterson, T., Memiş, E., & Zacharias, A. (2021). The impact of investing in social care on employment generation, time-, income-poverty by gender: A macro-micro policy simulation for Turkey. *World Development*, *144*, 105476. https://doi.org/10.1016/j.worlddev.2021.105476

International Labour Office. (2020, March 9). *Global employment trends for youth 2020: Technology and the future of jobs*. https://www.ilo.org/publications/major-publications/global-employment-trends-youth-2020-technology-and-future-jobs

International Labour Organization. (2025). *Youth not in employment, education or training (NEET)* [Indicator]. ILOSTAT. Retrieved September 24, 2025, from https://ilostat.ilo.org/data/snapshots/youth-neet-rate/

Kang, C., & Youn, M. (2024). Bridging the gender gap in NEET risk: Institutional influences on NEET status in different age groups. *Sage Open*, *14*(4), 1-15. https://doi.org/10.1177/21582440241294126

Karahasan, B. C. (2025). Unemployment polarization and club convergence in Türkiye. *Growth and Change*, *56*(1), e70027. https://doi.org/10.1111/grow.70027

Karma, E. (2025). NEET youth in central and eastern European countries: A panel model approach. *Journal of Youth Studies*, *28*(9), 1458-1475. https://doi.org/10.1080/13676261.2024.2392189

Kebede, A., Asgedom, A., & Asfaw, A. (2024) Is linking technical and vocational training with industry a bridge to employability? Lessons from a literature review for Ethiopia. *Cogent Education*, *11*(1), 2406721. https://doi.org/10.1080/2331186X.2024.2406721

Kılıç, Y. (2014). Young people in Turkey who are not in education, employment or training (NEET). *Education and Science*, *39*(175), 121-135. https://doi.org/10.15390/EB.2014.3094

Lindblad, V., Lund, R. L., Gaardsted, P. S., Hansen, L. E. M., Lauritzen, F. F., & Melgaard, D. (2025). Place matters: Understanding geographic influences on youth not in education, employment, or training-A scoping review. *Journal of Adolescence*, *97*(3), 620-633. https://doi.org/10.1002/jad.12461

Lu, Y., & Li, X. (2025). Lost in translation: Gender gap in education–occupation mismatch among highly educated workers. *Work and Occupations*, *0*(0). https://doi.org/10.1177/07308884251380099

Lüküslü, D., & Çelik, T. (2022). Gendering the NEET category: Young NEET women in Turkey. *Turkish Studies*, *23*(2), 200-222. https://doi.org/10.1080/14683849.2021.1882861

Maguire, S. (2015). NEET, unemployed, inactive or unknown – why does it matter? *Educational Research*, *57*(2), 121-132. https://doi.org/10.1080/00131881.2015.1030850

Odoardi, I., D'Ingiullo, D., & Quaglione, D. (2023) Gender disparities between young and adult NEETs: do we need a more refined policy approach? *Applied Economics*, *55*(56), 6685-6699. https://doi.org/10.1080/00036846.2022.2161991

Organisation for Economic Co-operation and Development. (2025a, September 9). *Education at a glance 2025: OECD indicators.* OECD Publishing, https://doi.org/10.1787/1c0d9c79-en

Organisation for Economic Co-operation and Development. (2025b). *Youth not in employment, education or training (NEET).* [Indicator]. Retrieved November 12, 2025, from https://doi.org/10.1787/72d1033a-en

Özdemir, D., Metin Özcan, K., & Üçdoğruk Birecikli, Ş. (2023). Youth not in employment, education, or training (NEET) in Turkey: A regional analysis. *Ege Academic Review*, *23*(3), 409-426. https://doi.org/10.21121/eab.1178984

Özdemir, D., Metin Özcan, K., & Üçdoğruk Birecikli, Ş. (2024). What makes Turkish youth NEET? *Journal of Youth Studies*, 1-19. https://doi.org/10.1080/13676261.2024.2446967

Redmond, P., & McFadden, C. (2023). Young people not in employment, education or training (NEET): Concepts, consequences and policy approaches. *The Economic and Social Review*, *54*(4), 285-327.

Turkish Statistical Institute. (2022). *Household labor force survey 2022* [Data set].https://www.tuik.gov.tr/Kurumsal/Mikro_Veri

United Nations Development Programme (UNDP), Türkiye. (2025, May 30). *White paper – policy recommendations for NEET young women in Türkiye* [White paper]. https://www.undp.org/turkiye/publications/white-paper-policy-recommendations-neet-young-women-turkiye

Wooldridge, J. M. (2001). Asymptotic properties of weighted M-estimators for standard stratified samples. *Econometric Theory*, 17(2), 451-470. https://doi.org/10.1017/S0266466601172045

World Bank. (2025). *Share of youth not in education, employment or training, total (% of youth population) (SL.UEM.NEET.ZS)* [Data set]. World Development Indicators (WDI). Retrieved October 13, 2025, from https://databank.worldbank.org/metadataglossary/world-development-indicators/series/SL.UEM.NEET.ZS

Yates, S., & Payne, M. (2006). Not so NEET? A critique of the use of 'NEET' in setting targets for interventions with young people. *Journal of Youth Studies*, *9*(3), 329-344. https://doi.org/10.1080/13676260600805671

Yiğit, B., Çakmak, B. Y., & Çakmak, E. E. (2023). NEET in Turkey: A typology including jobless youths, parental education and employment status as determinants. *Education and Training*, *65*(5), 731–748. https://doi.org/10.1108/ET-01-2023-0005

Appendix

Appendix 1: Robustness Check: Probit Specification

The baseline logit models are re-estimated using a probit link function as a check on the sensitivity of the results to the choice of binary outcome model. Average marginal effects from the probit are reported alongside the logit AMEs in Table A for the overall sample, women-only and men-only sub-samples respectively. The two sets of estimates are virtually identical in sign and magnitude across all covariates. The education gradient, the gender gap in the pooled model, the asymmetric associations of marriage and child-under-six presence across genders, and the ordering of the regional fixed effects are all fully preserved under the probit specification. These results confirm that the findings reported in the main text are not an artifact of the logistic functional form assumption.

Table A

Average Marginal Effects (AMEs) – Obtained from Probit Model Results

Variable	Overall (Std. Error)	Women (Std. Error)	Men (Std. Error)
gender	0.1697***		
	(0.0023)		
age	0.1983***	0.2365***	0.1239***
	(0.0034)	(0.0050)	(0.0050)
age2	-0.0042***	-0.0049***	-0.0026***
	(0.0001)	(0.0001)	(0.0001)
primary school	-0.0726***	0.0286	-0.1816***
	(0.0130)	(0.0179)	(0.0192)
middle school	-0.3237***	-0.2798***	-0.3462***
	(0.0082)	(0.0104)	(0.0129)
high school	-0.2794***	-0.2375***	-0.2958***
	(0.0082)	(0.0106)	(0.0131)
vocational high school	-0.2783***	-0.2318***	-0.3022***
	(0.0083)	(0.0011)	(0.0132)
vocational collage	-0.3168***	-0.2944***	-0.3127***
	(0.0085)	(0.0110)	(0.0136)
university	-0.3225***	-0.3303***	-0.2704***
	(0.0084)	(0.0108)	(0.0137)
master's and PhD	-0.3917***	-0.3864***	-0.3070***
	(0.0127)	(0.0175)	(0.0195)
married	0.1039***	0.2117***	-0.0855***
	(0.0037)	(0.0052)	(0.0057)
western marmara	0.0325***	0.0501***	0.0134
	(0.0070)	(0.0103)	(0.0090)
aegean	0.0639***	0.0720***	0.0507***
	(0.0057)	(0.0083)	(0.0071)
eastern marmara	0.0412***	0.0546***	0.0261***
	(0.0059)	(0.0086)	(0.0074)
western anatolia	0.0430***	0.0747***	0.0116
	(0.0058)	(0.0083)	(0.0074)
mediterranean	0.0782***	0.0969***	0.0562***
	(0.0057)	(0.0082)	(0.0071)
central anatolia	0.0856***	0.1258***	0.0436***
	(0.0061)	(0.0090)	(0.0078)
western black sea	0.0592***	0.0796***	0.0363***
	(0.0063)	(0.0092)	(0.0079)
eastern black sea	0.0709***	0.0619***	0.0716***
	(0.0074)	(0.0108)	(0.0089)
northeastern anatolia	0.0824***	0.1013***	0.0627***
	(0.0062)	(0.0091)	(0.0078)
centraleastern anatolia	0.1216***	0.1402***	0.1009***
	(0.0061)	(0.0088)	(0.0074)
southeastern anatolia	0.1176***	0.1487***	0.0925***
	(0.0057)	(0.0082)	(0.0070)
HH has at least one child (<6)	0.0458***	0.0741***	-0.0167***
	(0.0035)	(0.0050)	(0.0051)
HH size	0.0092***	0.0122***	0.0087***
	(0.0008)	(0.0013)	(0.0009)

Information About the Article/Makale Hakkında Bilgiler

The Ethical Rules for Research and Publication / Araştırma ve Yayın Etiği

The author declared that the ethical rules for research and publication followed while preparing the article.

Yazar makale hazırlanırken araştırma ve yayın etiğine uyulduğunu beyan etmiştir.

Conflict of Interests/ Çıkar Çatışması

The author have no conflict of interest to declare.

Yazar çıkar çatışması bildirmemiştir.

Grant Support/ Finansal Destek

The author declared that this study has received no financial support.

Yazar bu çalışma için finansal destek almadığını beyan etmiştir.

Author Contributions/ Yazar Katkıları

The draft process of the manuscript/ Taslağın Hazırlanma Süreci H.Ö.Ö.D, Data Collection/Verilerin Toplanması H.Ö.Ö.D, Writing the Manuscript/ Makalenin Yazılması H.Ö.Ö.D, Submit, Revision and Resubmit Process/ Başvuru, Düzeltme ve Yeniden Başvuru Süreci H.Ö.Ö.D.

Journal of Applied and Theoretical Social Sciences

JATSS, 2026; 8(1), 137-168

First Submission:22.01.2026
Revised Submission After Review:09.03.2026
Accepted For Publication:24.03.2026
Available Online Since:25.03.2026

Research Article

The Impact of TSRS and IFRS S on Climate Discourse: A Longitudinal Analysis of Türkiye and EU Samples

Onur Özevin[a]

Abstract

Introduction: This study examines the impact of the mandatory TSRS/IFRS S standards on the climate discourse of BIST and EU companies. It holds an original value in terms of empirically revealing, through a longitudinal and global comparative methodology, whether the regulations transform corporate narratives into financial materiality-oriented, evidence-based technical information, whether the standards discipline corporate discourse, and whether they close the "intent-action" gap in the discourse.

Method: A longitudinal and comparative content analysis was conducted on the Integrated Reports of 24 manufacturing companies from Türkiye and the EU for the 2022–2024 period. Using MAXQDA software, the study coded texts based on TSRS/IFRS S 2 pillars. To measure narrative quality, specific discursive variables were examined, including "Climate Action Orientation", "Greenwashing Risk", and "Linguistic Tone". Statistical differences between samples and periods were analyzed via Mann-Whitney U and Kruskal-Wallis tests.

Results or Findings: Findings reveal that while Turkish companies significantly increased reporting volume in 2024 ("active adaptation"), they relied heavily on "future-oriented commitments" and "vague expressions" rather than concrete "impact" data. Furthermore, Turkish reports exhibited an overly optimistic tone compared to the balanced, evidence-based tone of the EU sample.

Discussion or Conclusion: The study concludes that mandatory standards triggered "excessive signaling" and symbolic compliance in Türkiye, increasing reporting volumes without immediately closing the "intent-action" gap. Consequently, the surge in "commitments" was not matched by concrete "impact" data. Furthermore, while governance structures exhibit global isomorphism, significant divergences in linguistic tone and operational depth persist between the Turkish and EU samples.

Keywords: sustainability reporting, TSRS & IFRS S, climate discourse, greenwashing, impression management, signal theory

JEL Codes: M41, M48, M14, G3, Q56

[a] Asst. Prof. Dr., Bolu Abant İzzet Baysal University, Gerede Vocational School, Department of Administration and Organization, Bolu/Türkiye, onurozevin@ibu.edu.tr, ORCID ID: https://orcid.org/0000-0002-1347-5027 (Corresponding Author)

Journal of Applied and Theoretical Social Sciences

JATSS, 2026; 8(1), 137-168

İlk Başvuru: 22.01.2026
Düzeltilmiş Makalenin Alınışı: 09.03.2026
Yayın İçin Kabul Tarihi: 24.03.2026
Online Yayın Tarihi: 25.03.2026

Araştırma Makalesi

TSRS ve IFRS S'nin İklim Söylemi Üzerindeki Etkisi: Türkiye ve AB Örneklemlerinin Boylamsal Bir Analizi

Onur Özevin [a]

Öz

Giriş: Bu çalışma, zorunlu TSRS/IFRS S standartlarının BIST ve AB şirketlerinin iklim söylemleri üzerindeki etkisini incelemektedir. Düzenlemelerin, kurumsal anlatıları finansal maddilik odaklı, kanıt temelli teknik bilgiye dönüştürüp dönüştürmediği, standartların kurumsal söylemi disipline edip etmediğini ve söylemdeki "niyet-eylem" boşluğunu kapatıp kapatmadığını boylamsal ve küresel karşılaştırmalı bir metodolojiyle ampirik olarak ortaya koyması açısından özgün bir değer taşımaktadır.

Yöntem: Türkiye ve AB'den 24 imalat şirketinin 2022–2024 dönemi Entegre Raporları üzerinde boylamsal ve karşılaştırmalı bir içerik analizi yürütülmüştür. Çalışma, MAXQDA yazılımını kullanarak metinleri TSRS/IFRS S 2 sütunlarına göre kodlamıştır. Anlatı kalitesini ölçmek için "İklim Eylemi Yönelimi", "Yeşil Aklama Riski" ve "Dilsel Ton" dahil olmak üzere belirli söylemsel değişkenler incelenmiştir. Örneklemler ve dönemler arasındaki istatistiksel farklar Mann-Whitney U ve Kruskal-Wallis testleri aracılığıyla analiz edilmiştir.

Sonuçlar ya da Bulgular: Bulgular, Türk şirketlerinin 2024 yılında raporlama hacmini önemli ölçüde artırmasına (aktif uyum) rağmen, somut "etki" verilerinden ziyade ağırlıklı olarak "gelecek odaklı taahhütlere" ve "muğlak ifadelere" dayandığını ortaya koymaktadır. Ayrıca, Türk raporları, AB örnekleminin dengeli, kanıt temelli tonuna kıyasla aşırı iyimser bir ton sergilemiştir.

Tartışma ya da Yapılan Çıkarımlar: Çalışma, zorunlu standartların Türkiye'de 'aşırı sinyal vermeyi' ve sembolik uyumu tetiklediği, 'niyet-eylem' boşluğunu hemen kapatmadan raporlama hacimlerini artırdığı sonucuna varmaktadır. Sonuç olarak, 'taahhütlerdeki' artış somut 'etki' verileriyle eşleşmemiştir. Ayrıca, yönetişim yapıları küresel izomorfizm sergilerken, Türk ve AB örneklemleri arasında dilsel ton ve operasyonel derinlik açısından önemli ayrışmalar devam etmektedir.

Anahtar Kelimeler: sürdürülebilirlik raporlaması, TSRS & IFRS S, iklim söylemi, yeşil aklama, izlenim yönetimi, sinyal teorisi

JEL Kodlar: M41, M48, M14, G3, Q56

[a] Dr. Öğr. Üyesi, Bolu Abant İzzet Baysal Üniversitesi, Gerede Meslek Yüksekokulu, Yönetim ve Organizasyon Bölümü, Bolu/Türkiye, onurozevin@ibu.edu.tr , ORCID ID: https://orcid.org/0000-0002-1347-5027 (Sorumlu Yazar)

Introduction

In recent years, the deepening of environmental crises such as global warming, climate change, and the depletion of natural resources has made the concept of sustainability a strategic priority in the business world. Stakeholders demand that businesses disclose not only their financial results but also their environmental, social, and governance (ESG) performance transparently has triggered a fundamental transformation in the corporate reporting ecosystem (Albitar et al., 2020). The inadequacy of traditional financial reporting in reflecting a business's value-creation process and long-term risks has increased the need for reporting non-financial information (Bernardi & Stark, 2018; Özdemir, 2024). In this context, sustainability reporting has become a critical tool for companies to maintain their legitimacy and establish a trust-based relationship with their stakeholders (Rabaya & Saleh, 2022). However, as Ali et al. (2025) argue, while these disclosures aim to reduce information asymmetry, they also carry the risk of being used opportunistically by managers to mask unethical practices or earnings management, thereby complicating the trust dynamic.

Furthermore, the fact that sustainability reporting proceeded for a long time on a voluntary and non-standardized basis has brought about various credibility issues. Practices referred to in the literature as "greenwashing"—defined as the effort by businesses to create misleading or exaggerated positive impressions about their environmental performance—became widespread during this period (Tepeli & Büklü, 2025; Yu et al., 2020). Research shows that firms often resort to impression management tactics by using vague and optimistic language that emphasizes "commitment" rather than concrete performance data (Anıl et al., 2025; Ozsozgun Caliskan, Esen & Barkemeyer, 2021). Narrative sections, such as CEO letters, are particularly criticized as texts that managers use to strategically shape the corporate image but carry the risk of being "cheap talk" (Treepongkaruna et al., 2024; Al Amosh, 2025).

To address this information asymmetry and ensure comparability, international regulators have accelerated the transition to mandatory reporting standards. While the IFRS S1 and S2 standards published by the International Sustainability Standards Board (ISSB) aim to create a global baseline, the Public Oversight, Accounting and Auditing Standards Authority (KGK) in Türkiye published the Turkish Sustainability Reporting Standards (TSRS 1 and TSRS 2) in alignment with these standards, initiating the mandatory reporting period for certain entities as of January 1, 2024 (Özdemir, 2024). Regulatory pressures and mandatory reporting are expected to affect not only the amount of information disclosed by firms but also the tone and content of the language they use (Albitar et al., 2020; Fang et al., 2025). Developed as a response to this credibility crisis, the TSRS 2 standard structurally diverges from frameworks like GRI that adopt a "double materiality" approach; instead, it centers on an investor-oriented "financial materiality" philosophy that directly links sustainability risks to cash flows and corporate value. This paradigmatic shift aims to theoretically minimize greenwashing risk by transforming the corporate narrative from "storytelling" into an "evidence-based" structure connected to financial statements.

This study aims to analyze the impact of the transition from voluntary reporting to a TSRS 2-oriented mandatory reporting regime on firms' sustainability discourse in Türkiye. Within the framework of legitimacy and signaling theories, the research comparatively examines the strategic structure of corporate disclosures, the focus on climate action, potential greenwashing risks, and tonal/linguistic changes in reporting language. In current literature, debates continue regarding whether mandatory regulations reduce greenwashing (Todaro & Torelli, 2024). This longitudinal analysis, conducted through BIST companies and a global sample, aims to contribute to the relevant literature by revealing the extent to which legal

regulations discipline companies' sustainability narratives and whether the gap between "discourse" and "action" is closing. Due to integrated reporting being the practice showing the highest alignment with the financial materiality principle brought by TSRS 2 and establishing the connectivity between financial data and ESG performance, this study includes only firms that publish integrated reports. Based on this motivation, the study hypothesizes that the transition to mandatory sustainability reporting standards will make corporate climate discourse more action-oriented and concurrently reduce the discursive divergence between local and global reporting practices.

Conceptual Framework

The corporate reporting ecosystem is evolving from the presentation of purely financial data toward a holistic structure encompassing ESG performance, in line with the increasing information demands of stakeholders and changing market dynamics (Özdemir, 2024; Rossi & Candio, 2023). The inadequacy of traditional financial reporting in reflecting value-creation processes and long-term risks has made the reporting of non-financial information a critical necessity. However, this process, which was long conducted on a voluntary basis through various frameworks (e.g., GRI, SASB, TCFD), led to inconsistencies between reports and weakened comparability for investors (Fianko et al., 2025). To address this fragmented structure and establish a global baseline, the IFRS S1 and IFRS S2 standards developed by the ISSB initiated a new era aimed at integrating sustainability disclosures with financial statements (Fianko et al., 2025). In Türkiye, the KGK published TSRS 1 (General Requirements) and TSRS 2 (Climate-related Disclosures) to ensure full alignment with these international norms, launching the mandatory reporting and assurance process as of January 1, 2024, for entities meeting specific criteria (Başar & Selimoğlu, 2025; Özçelik, 2025). This section details the legal and conceptual infrastructure forming the basis of the research.

Theoretical Background: Institutional and Signaling Theories

This study is grounded in two complementary theoretical perspectives: Institutional Theory and Signaling Theory. Institutional theory posits that organizational behaviors and reporting practices are profoundly shaped by external institutional pressures, leading to coercive, mimetic, or normative isomorphism (DiMaggio & Powell, 1983). In the context of this study, the mandatory implementation of TSRS 2 and IFRS S2 represents a coercive institutional pressure. Firms are compelled to adapt their reporting narratives to comply with these global standards and maintain their social legitimacy.

Concurrently, Signaling Theory (Spence, 1973) provides a framework for understanding how firms communicate their unobservable quality to stakeholders in an environment characterized by information asymmetry. High-performing firms use sustainability disclosures as positive signals to differentiate themselves from competitors. However, during transitional regulatory periods, these signals can be manipulated for impression management. Firms may adopt an overly optimistic or vague tone—signaling commitment rather than actual impact—which increases the risk of greenwashing (Tepeli & Büklü, 2025; Li et al., 2025). The interplay of these two theories explains why firms might show structural isomorphism (institutional compliance) while simultaneously using strategic rhetoric (signaling) to manage stakeholder impressions.

Structure of TSRS 2 (Climate-related Disclosures) Standards

TSRS 2 Climate-related Disclosures, prepared by the KGK in full compliance with ISSB standards to ensure international validity, is a technical framework requiring entities to integrate

climate-related risks and opportunities into their financial reporting processes (Başar & Selimoğlu, 2025). This regulation, which entered into force on January 1, 2024, became mandatory for entities exceeding certain size thresholds (total assets of 500 million TL, annual net sales revenue of 1 billion TL, and 250 employees) (KGK, 2024). The standard is built upon TCFD recommendations and structures reporting around four core pillars: Governance, Strategy, Risk Management, and Metrics and Targets (Fianko et al., 2025; KGK TSRS 2, 2023).

1.Governance: This pillar covers the governance processes, controls, and procedures used by the entity to monitor and manage climate-related risks and opportunities (KGK, TSRS 2, 2023). The standard requires disclosure of the competencies, job descriptions, and briefing processes of the body or individuals responsible for overseeing climate risks. Current research on BIST companies shows that while the level of compliance under this heading is higher than others, disclosures regarding the development of managers climate competencies remain limited (Kaldırım & Kaldırım, 2025; Özcan & Akbaş, 2025).

2.Strategy: Demonstrating the entity's resilience to climate change, this section requires disclosure of the impacts of short, medium, and long-term "physical risks" and "transition risks" (KGK TSRS 2, 2023). Beyond merely identifying risks, entities are expected to report the current and projected financial effects of these risks on cash flows, access to finance, and cost of capital using scenario analyses (Özcan & Akbaş, 2025; KGK TSRS 2, 2023). However, even in carbon-intensive sectors like electricity and construction, it has been observed that entities have not yet reached full readiness regarding financial impact modeling and scenario analyses (Özcan & Akbaş, 2025).

3.Risk Management: This component focuses on the processes for identifying, assessing, and prioritizing climate-related risks. A critical point is disclosing whether these processes are integrated into the entity's overall enterprise risk management system (KGK TSRS 2, 2023; Özçelik, 2025). The objective is to demonstrate that climate risks are not treated as an isolated issue but within the entity's main risk profile.

4.Metrics and Targets: Quantitative data used by the entity to measure climate performance is included in this pillar. TSRS 2 mandates the disclosure of absolute gross greenhouse gas emissions across "Scope 1" (direct), "Scope 2" (indirect from purchased energy), and "Scope 3" (other indirect in the value chain) (KGK TSRS 2, 2023). Calculating Scope 3 emissions stands out as the area where entities face the most difficulty and data gaps, as it relies on supply chain data (Özcan & Akbaş, 2025). Furthermore, the standard requires reporting sector-specific metrics, such as the clinker ratio for the construction sector or carbon intensity for the energy sector (KGK TSRS 2, 2023).

The core philosophy of TSRS 2 diverges from frameworks like GRI that adopt a "double materiality" approach covering both financial and impact dimensions (Fianko et al., 2025), instead relying on the investor-oriented "financial materiality" principle (Başar & Selimoğlu, 2025; KGK TSRS 1, 2023). Rather than focusing on an entity's impact on the environment, the standard centers on the effects of climate change on the entity's financial position, such as cash flows, asset values, and borrowing costs (KGK, TSRS 1, 2023).

In this context, information is considered "material" if its omission or misstatement could reasonably be expected to influence the decisions of investors in providing resources to the entity (Özdemir, 2024; KGK TSRS 1, 2023). This approach aims to transform sustainability reporting from a public relations activity into an investor-oriented, auditable information set linked to financial statements (Özçelik, 2025; Tepeli & Büklü, 2025). Indeed, recent evidence supports this investor-focused shift, as Anwar et al. (2025) demonstrate that in regulated

reporting environments, financial analysts are highly sensitive to disclosure tone, where overly optimistic or vague environmental disclosures can lead to increased forecast dispersion and lower accuracy. However, de Villiers et al. (2024) argue that this financial capture of sustainability reporting prioritizes investor needs, potentially sidelining broader social and environmental impacts. Similarly, Carungu et al. (2025) emphasize that the ISSB's influence is shifting the global discourse from 'double materiality' to a narrower single materiality focus. Abhayawansa and Adams (2022) further critique this investor-centric perspective, suggesting that it may fail to adequately capture long-term sustainable development risks such as pandemics and climate change. Nonetheless, as Grewal and Serafeim (2020) note, mandatory disclosure regulations remain a primary mechanism to force firms to measure and manage these sustainability outcomes, even if the transition poses challenges.

The transition from voluntary to mandatory reporting is a process that fundamentally alters the nature and content of corporate disclosures. This section establishes the research hypotheses by examining the potential of regulatory standards to mitigate information asymmetry and how firms utilize sustainability narratives as strategic tools.

There is strong evidence in the literature that mandatory reporting regimes reduce information asymmetry. Studies on the mandatory implementation of Integrated Reporting (IR) in South Africa have shown that this transition reduces analyst forecast errors and that ESG disclosures are used more effectively in predicting financial performance. Mandatory standards discipline corporate discourse by compelling firms to shift from symbolic disclosures—where they only share "good news"—to the sharing of concrete and auditable data. However, the literature also discusses the risk that firms may lean more toward greenwashing to appear "compliant" if environmental regulations increase financing constraints.

Albitar et al. (2020) investigated the role of corporate governance mechanisms in the impact of ESG disclosures on firm performance for FTSE 350 companies. The findings indicated a positive relationship between ESG disclosures and firm performance, which strengthened further after the transition to integrated reporting. Kılıç and Kuzey (2018) examined the determinants of forward-looking disclosures during the transition to integrated reporting among BIST 100 companies. They found that firm size, board independence, and gender diversity positively influenced the level of forward-looking disclosure. Consistently, Afifa et al. (2025) provide empirical evidence from developing markets that robust corporate governance mechanisms, particularly board independence and CSR strategy, are dominant internal drivers enhancing ESG disclosure quality, especially when firms face high carbon emission pressures.

Ozsozgun Caliskan et al. (2021) analyzed impression management tactics used in the CEO letters of companies in the BIST Sustainability Index. The research revealed that CEOs utilized more limited and specific tactics (particularly self-promotion and exemplification) in sustainability reports compared to annual reports, while avoiding quantitative performance comparisons. Ayrancı (2025) examined the sustainability reports of Turkish aviation companies (THY and Pegasus) for the 2021-2023 period, investigating the strategic role of these reports in corporate image building. The findings showed that the reports were used not only as transparency tools but also as strategic communication instruments; THY focused on technical and operational performance, while Pegasus developed different discourse strategies centered on social equality and digitalization.

Kaldırım and Kaldırım (2025) researched the extent to which reports from BIST Sustainability 25 Index companies met the TSRS 1 standard. Their 2022 review determined

that entities achieved high compliance in "governance" (90%) and "risk management" (86%), but fell short in "metrics and targets" (44%) and financial impact disclosures. Özcan and Akbaş (2025) analyzed the readiness levels of BIST companies in the electrical facilities and construction materials sectors for the TSRS 2 standard. The study found that while companies integrated climate-related responsibilities into corporate structures, significant deficiencies remained in scenario analyses, financial impacts of climate risks, and Scope 3 emissions reporting, with compliance remaining at a "partially appropriate" level. Başar and Doğruel (2025) examined the sustainability reports of Turkish fast-fashion firms (Koton, Mavi, Sun Tekstil) through content analysis. Results indicated that 58.4% of reporting focused on environmental themes, while social (10.1%) and governance (31.4%) dimensions remained secondary.

In corporate reporting, "discourse" is recognized as a strategic tool for building and maintaining corporate legitimacy, rather than just information transfer. Impression management literature suggests that managers can consciously manipulate the tone, complexity, and focus of language used in reports to shape stakeholder perceptions. CEO letters, being among the most read sections of sustainability reports, are critical for reflecting the "tone at the top". Research indicates a strong relationship between the language used by CEOs in these texts and a firm's ESG performance and corporate governance structure.

Albitar et al. (2023) found that a higher number of independent board members reduced "overly optimistic" language in CEO letters, leading to more realistic or pessimistic language. Conversely, firms with high ESG scores tend to use a more positive tone. Anıl et al. (2025) argue that firm visibility influences language choice. CEOs of firms in sensitive sectors use more "formal and analytical" language to protect legitimacy, whereas managers in less visible sectors prefer more "assertive" language to emphasize leadership. Saglam et al. (2024), in their review of the transportation sector, emphasized that CEO letters often contain "motivation" elements but fall short regarding the concrete "impact" of sustainability efforts and firm "commitment," thereby increasing stakeholder skepticism.

Language complexity is also used as an impression management tactic. According to the "obfuscation" hypothesis, poorly performing firms make their reports more complex and difficult to read to hide negative outcomes. Conversely, Rakhavendra et al. (2023) state that linguistic features in CEO letters (emotion, tone, readability) correlate with ESG performance, and CEOs who emphasize "well-being" or "environmental impact" are perceived more positively. Furthermore, Yuan (2025) suggests that CEOs' "emotional orchestration" abilities are decisive in corporate sustainability communication.

It has been determined that while greenwashing was studied infrequently in the early 2000s, it has been researched intensively since 2020 (Acar Uğurlu, 2024).The long-debated concept of "greenwashing" has evolved into "ESG-washing," covering social and governance dimensions as well. This phenomenon is defined as firms' attempts to manipulate stakeholders by exaggerating sustainability performance or providing misleading information. One of the primary reasons for greenwashing is the lack of auditing and verification mechanisms. Özçelik (2025) notes that assurance auditing of sustainability reports is not yet widespread in Türkiye, which undermines report reliability. Mandatory standards like TSRS 1 and 2 are expected to fill this audit gap and reduce the risk of "cheap talk".

Başar and Selimoğlu (2025) point to derivatives of greenwashing such as "social washing" and "blue washing" (mimicking collaboration with the UN). Tepeli and Büklü (2025) empirically demonstrated through event analysis that while greenwashing may increase firm

value in the short term, it leads to long-term reputation and financial losses once disclosed. They emphasized that greenwashing damages investor confidence in the long run and that a lack of auditing facilitates this process. Deng et al. (2024) found in the Chinese market that "top management team stability" reduces greenwashing behavior, as stable teams focus on long-term reputation rather than short-term opportunism.

The development of sustainability reporting in Türkiye has been detailed through studies focused on Borsa Istanbul (BIST) companies. These studies show that while Turkish companies follow global trends, their reporting contains sectoral differences and specific deficiencies. With the publication of TSRS 1 and TSRS 2 standards by the KGK, studies examining the compliance processes of BIST companies have commenced.

Gültepe (2025) examined the reports of ready-to-wear apparel brands in Türkiye, finding that firms prioritize environmental themes (58%) significantly more than social (10%) and governance themes. Furthermore, these brands attempt to gain legitimacy by utilizing sustainability discourse as a public relations strategy within the scope of "brand activism." Gürsoy and Özuyar (2024) investigated the integration of UN Sustainable Development Goals (SDGs) among BIST Sustainability 25 Index companies, revealing that while firms focus on popular targets like "climate action," they place less emphasis on fundamental social goals such as "no poverty." This situation points to the practice of "selective disclosure" by these companies.

Particularly, TSRS 2 has weakened the "storytelling" aspect of reporting while strengthening its "evidence-based" dimension by compelling entities to disclose the metrics and scenario analyses used in pursuit of climate targets. Findings indicating that mandatory reporting regimes reduce analyst forecast errors confirm that standards distance firms from "selective disclosure" tactics, thereby bringing corporate discourse under accounting discipline.

Corporate sustainability reports serve as strategic communication tools used by firms to gain legitimacy and manage stakeholder perception (Tepeli & Büklü, 2025). Impression management literature emphasizes that the language utilized in these reports may serve to construct a positive corporate image rather than reflecting actual performance (Al Amosh, 2025; Todaro & Torelli, 2024). In global samples, it has been observed that as firm visibility increases, the language employed adopts a more "formal and analytical" tone (Anıl et al., 2025). Even global firms with high ESG scores have been found to present their achievements using "narcissistic rhetoric" and an exaggerated tone, a tendency that can be balanced by board diversity (Mahran & Elamer, 2025). Additionally, the literature frequently criticizes firms for generating "cheap talk" during voluntary reporting periods, utilizing an optimistic tone that obscures actual carbon emissions (Treepongkaruna et al., 2024; Kathan et al., 2025).

Despite this extensive body of literature, the vast majority of studies on sustainability reporting in Türkiye remain limited to descriptive content analyses focusing on a single period or checklists measuring compliance levels with specific standards. In particular, the evolutionary impact of the transition from a voluntary reporting regime to a TSRS-based mandatory regime on the linguistic fabric and strategic tone of corporate discourse has not yet been sufficiently explored empirically. With its longitudinal design covering the years 2022, 2023, and 2024, this study fills a significant gap in the literature by revealing how changes in reporting discipline transform the analytical structure of language, the dosage of uncertainty, and greenwashing risks within a time series. Furthermore, by addressing Turkish data through a comparative perspective with global reporting trends and international literature findings, the research discusses on a universal level whether legal mandates transform corporate narratives

from symbolic legitimacy tools into disciplined sets of technical information. In this regard, the study offers an original contribution to the accounting and auditing literature in Türkiye through its methodological depth and capacity for periodic and global comparison.

Current literature indicates that during voluntary reporting periods, firms utilize language that conveys "intent" and "commitment" but lacks concrete actions and results (Saglam, Solak-Fiskin & Akgul, 2024). However, mandatory standards such as TSRS 2 require firms to disclose not only their targets but also the progress made toward achieving them, the metrics used, and the financial impacts (KGK TSRS 2, 2023). Consequently, the audit pressure and standardized data demand brought by legal mandates are expected to transform the language used by firms from "future-tense promises" to "past and present-tense actions." In this context, the following hypotheses are formulated:

H1: The transition to the mandatory TSRS/IFRS S2 regime significantly increases the action-oriented and concrete nature of corporate climate discourse.

H2: The implementation of mandatory global reporting standards significantly reduces the discursive divergence between emerging markets (Türkiye) and developed markets (EU), leading to a global isomorphic convergence.

Method

This study is designed to examine the impact of TSRS 1 and TSRS 2, which came into effect in Türkiye on January 1, 2024, on climate change-oriented corporate narratives within the integrated reports of companies.

Research Design and Approach

Narrative sections in financial and non-financial reports are considered strategic areas where managers communicate company performance and strategies to stakeholders, while simultaneously being able to apply impression management tactics (Bassyouny et al., 2020; Ozsozgun Caliskan et al., 2021). In this context, the research adopts a mixed-methods approach that blends quantitative and qualitative data analysis techniques.

The research is based on qualitative content analysis and critical discourse analysis techniques within the qualitative research tradition. The study features a longitudinal design measuring the impact of the transition from the voluntary reporting period (2022–2023) to the mandatory reporting period (2024) on discourse, alongside a comparative design measuring alignment with global standards. Corporate climate narratives have been deciphered through the frameworks of Legitimacy Theory, Impression Management, and Greenwashing, as well as the "financial materiality" and "transparency" criteria introduced by IFRS S2.

Population and Sample Selection

The population of this study consists of publicly traded companies listed on the Borsa Istanbul BIST 100 Index and their equivalent global peers. From this population, the sample was determined through a purposive sampling method. For the sample selection, the criteria were established as having published integrated reports in the manufacturing sector within the BIST 100 for the years 2022, 2023, and 2024, and similarly, for global companies to be in the manufacturing sector and have published integrated reports during the same period. The reason for selecting the manufacturing sector in this study is that it has the highest environmental impact in terms of carbon emissions, waste management, and resource use, a situation that necessitates these businesses to provide more comprehensive and transparent sustainability reporting under standards like TSRS 2. Out of 31 companies in the manufacturing sector within

the BIST 100, only 10 published continuous integrated reports during the relevant years. Consequently, a total of 24 companies—10 from Türkiye and 14 from European Union (EU) countries—with different levels of regulatory maturity were included in the research to ensure sample diversity.

The study's data set was formed from integrated reports that combine financial and non-financial information under the principle of "connected information," rather than independent sustainability reports where these elements are presented in isolation. The primary reason for this preference is that integrated reporting directly links ESG performance to the entity's value-creation process and financial capital (Bernardi & Stark, 2018; Zhou, Simnett & Green, 2017) and is the reporting practice that shows the highest alignment with the "financial materiality" principle forming the architecture of TSRS (IFRS S1/S2) (Kaldırım & Kaldırım, 2025). Beyond being a mere external communication tool, integrated reports reflect the integrated thinking structure and strategic decision-making processes within the business (Albitar et al., 2020; Maniora, 2017), providing the most suitable empirical ground for measuring the impact of the transition to mandatory standards on corporate narratives.

Within the scope of the research, integrated reports from the final period before the standards (2022 and 2023) and the first implementation period after the standards (2024) were obtained from the KAP website and corporate web pages. The analysis focused entirely on the verbal content, excluding numerical tables of the reports. In particular, the focus of the coding process was formed by the four core pillars of IFRS S2: Governance, Strategy, Risk Management, and Metrics and Targets.

Content Analysis and Measurement of Variables

Data analysis was carried out using the qualitative data analysis software MAXQDA 2024. The analysis process includes "dictionary-based approach" and "tone analysis" methods widely accepted in the literature (Albitar et al., 2023; Saglam et al., 2024).

Thematic Classification and Word Lists

To measure compliance with the "Governance," "Strategy," "Risk Management," and "Metrics and Targets" pillars introduced by the TSRS 2 Climate-related Disclosures standard, a specialized codebook was developed utilizing studies in the literature (Fianko et al., 2025; Özcan & Akbaş, 2025) and the standard's own terminology (KGK, TSRS 2, 2023).

- Climate Action Orientation: To distinguish whether corporate discourses contain only "intent" (commitment) or concrete "action" (impact), the classification of Commitment, Impact, and Motive proposed by Saglam et al. (2024) was utilized.
- Greenwashing Risk: To detect greenwashing and ESG-washing risks, the density of vague, commonplace, and forward-looking expressions in the texts was measured (Tepeli & Büklü, 2025; Todaro & Torelli, 2024).

Tone Analysis

To measure the level of optimism and tone in the reports, the Loughran and McDonald (2011) dictionary, considered standard in finance and accounting literature, was used as a basis (Albitar et al., 2023; Rakhavendra, Bala & Mukherjee, 2023). This dictionary is customized for financial texts and includes categories for "positive," "negative," "uncertainty," and "certainty". The following formula was used in the calculation of Net Tone (Albitar et al., 2023; Henry, 2008):

$$Net\ Tone = \frac{\text{Positive Wordcount} - \text{Negative Wordcount}}{\text{Positive Wordcount} + \text{Negative Wordcount}}$$

This formula expresses the overall sentiment of the report with a coefficient ranging from -1 (completely negative) to +1 (completely positive).

Linguistic Style and Impression Management

To analyze the strategic dimension of the language used by companies (impression management), the "Formal-Analytical," "Assertive," and "Defensive" linguistic style profiles developed by Anıl et al. (2025) were utilized. In particular, the "Analytic Thinking" score was used to determine the extent to which the texts are data-driven and objective (the structure expected by mandatory reporting) or based on storytelling (a characteristic of voluntary reporting) (Anıl et al., 2025).

Data Analysis Process

The analysis process was conducted in three stages:

1.Data Pre-processing: Integrated reports in PDF format were uploaded into the software; conjunctions, prepositions, and non-meaningful words were cleaned.

2.Coding and Frequency Analysis: In the MAXdictio module of the MAXQDA program, texts were subjected to automated coding through a codebook developed based on the literature as described above; the frequencies and densities of keywords related to TSRS 2 were calculated.

3.Statistical Analysis: The qualitative data obtained were quantified, and Mann-Whitney U and Kruskal-Wallis tests were performed to test the significance of the difference between sample groups (Türkiye-EU) and between mandatory and voluntary periods.

Validity and Reliability

To eliminate human error and subjectivity during the analysis process, the computer-aided automated coding method (MAXdictio), which is closed to researcher intervention, was preferred. While coding reliability was ensured by the software's algorithmic stability, the internal validity of the research was established by creating the codebook based on standard word lists in the literature (Loughran & McDonald, 2011; Saglam et al., 2024) and TSRS 2 terminology.

The lengths (word counts) of sustainability reports can vary significantly depending on the size, sector, and reporting traditions of the companies. This creates a risk of bias in comparisons made via raw code frequencies, as "longer reports may contain more codes". To eliminate this risk and measure discourse density independently of report volume, the data were normalized.

The "proportioning to word count" method, widely used in the literature, was preferred for the normalization process. Accordingly, the raw code counts for each document were divided by the total word count of the respective document, and the results were multiplied by a coefficient of 10,000 words to provide a comparable scale. The process was carried out according to the following formula. This method statistically ensures that the findings reflect the thematic density in each narrative unit (every 10,000-word block), rather than just the physical length of the report.

Limitations of the Study

Several methodological limitations should be considered when evaluating the results of this study. First, due to the fact that integrated reporting practice in Türkiye is still in its developmental stage and the number of firms providing continuous data is limited, the research sample was restricted to 10 companies in the BIST 100 manufacturing sector and their 14 global peers. This situation may limit the generalization of the findings to sectors with different dynamics, such as services or finance. Second, the analysis focused solely on the narrative (verbal) content and linguistic tone of the reports, while the technical analysis of financial data in numerical tables was excluded. Finally, as 2024 marks the first implementation year of mandatory reporting (TSRS 2), this study reflects companies' initial strategic responses to new regulations and a "transition period" snapshot rather than a long-term trend.

Findings

Discursive Focus and Strategic Framing: Word Cloud Analyses

In order to present the linguistic texture and the general overview of the strategic focal points in the integrated reports of the companies within the scope of the research, the most frequently used words were extracted and visualized through word cloud analysis. A structural and rhetorical divergence between the two sample groups, namely Türkiye and the EU, is strikingly evident. These visuals, where word sizes are directly proportional to their frequency of use, reveal how companies frame climate change beyond purely quantitative data. It is important to note that the word clouds presented in this section are generated cumulatively from the entire 2022–2024 dataset. The primary objective of this visual analysis is to capture and compare the overarching macro-discourse and general strategic framing differences between the Turkish and EU samples across the whole transition period, rather than tracking year-by-year fluctuations. While the longitudinal evolution over the years is detailed in the subsequent statistical tables, these cumulative word clouds serve to highlight the fundamental divergence in corporate mindsets between the two regions.

Figure 1

Word Cloud Analysis of Integrated Reports of Companies in the Türkiye Sample (2022-2024)

Note. Word cloud of search results. This figure was created via MAXQDA where Arial font customization is not supported

Figure 1 illustrates the word cloud analysis for the Türkiye sample. According to the figure, it is observed that "corporate legitimacy" and "value creation" are placed at the center of the reporting language. The word "value," located at the focus of the word cloud, indicates that the companies in the sample present sustainability as a fundamental "corporate value proposition". The combination of this focus with "financial/corporate" terms can be interpreted as an attempt to present sustainability efforts within a more rational and analytical framework by associating them with financial performance.

The density of the words "management," "integrated," and "board" within the text indicates a tendency to frame sustainability issues directly as a matter of Governance and Strategy during the reporting process. This choice of terminology can be interpreted as an effort to present the phenomenon of sustainability within the framework of a high-level managerial discipline rather than merely as an operational output. The position of concepts themed around "future," "transformation," and "target" in the frequency distribution confirms the Commitment (2.1) code, which yielded significant results in statistical tests. The interaction of the word "value" with these future-oriented terms suggests that companies are sending signals to their stakeholders based on "potential value" to be created in the future, rather than current concrete performance.

The frequency of conceptually broad words in the list, such as "activity," "general," and "important," visualizes the density in the Vague/General Expressions (3.1) category. While these terms contribute to the volumetric expansion of the reports, they indicate that technical details sometimes remain secondary to corporate rhetoric. The fact that words like "energy" and "water" appear significantly more frequently than social metrics such as "social" and "women" shows that the focus of reporting in Türkiye still concentrates on environmental (E) dimensions, particularly on resource efficiency.

Figure 2

Word Cloud Analysis of Prominent Concepts in Integrated Reports within the EU Sample (2022-2024)

Note. Word cloud of search results. This figure was created via MAXQDA, where Arial font customization is not supported

The word cloud analysis of the sample representing EU companies is presented in Figure 2. Accordingly, the leading positions of the terms "management" and "board" indicate the direct weight of the Governance function within EU companies. This structure, supported by the concepts of "governance" and "executive," suggests that the reports are constructed not as a public relations activity, but as a strategic management process.

The most striking finding is that the term "financial" possesses a frequency significantly higher than sustainability-themed concepts. This situation may be an indicator that EU reporting is intertwined with Analytical Language and financial discipline, and that sustainability data is presented with the same level of rigor as financial statements. The frequent repetition of the word "risk" reveals the centrality of the Risk Management code within the reports of the EU sample.

Although the term "impact" (f: 4,872) is a central concept, the higher frequencies of "performance" and "value" indicate that EU reports focus on demonstrating measurable performance through Metrics and Targets and the tangible outputs created via this performance. The terms "energy," "emissions," "water," and "climate" signify that the thematic content of the reports is technical and data-driven rather than speculative.

The absolute dominance of the "Value" concept in Türkiye, contrasted with the "Financial/Management" concepts in the EU sample, provides a fundamental framework regarding the difference in reporting philosophies between the two regions. When word frequencies and hierarchical distributions are juxtaposed, a divergence is observed between the Turkish and EU samples based on "Corporate Legitimacy Strategy" and "Financial Audit Discipline."

Sentiment of Reporting Language: Net Sentiment Analysis

To empirically reveal the dominant rhetoric and differences in corporate mindset of integrated reports, a "Net Sentiment" analysis was conducted on the top 50 key concepts with the highest frequency in both samples. These concepts, representing the vast majority of the total word volume, were classified as "Positive/Assertive" or "Negative/Risk-Oriented" according to their semantic weights in sustainability literature and financial lexicons (Loughran & McDonald, 2011).

An examination of the frequency list for the Turkish sample reveals that the reporting language possesses a highly optimistic and opportunity-oriented tone. In the Turkish list, the dominance of positive and assertive concepts numerically outpaces negative/risk-oriented terms. The net sentiment calculation is provided below:

- **Positive Word Group:** Value, Management, Integrated, Sustainability, Material/Significant, Future-oriented, Transformation, etc. (Total Frequency: ~75,450)
- **Negative/Risk Group:** Risk, Waste, Compliance (within the context of constraints), etc. (Total Frequency: ~10,811)
- **Total Key Word Volume:** ~245,000

$$Net\ Tone = \frac{75.450 - 10.811}{75.450 + 10.811} = 0{,}26$$

The net sentiment value of 0.26 in the Turkish sample indicates that the reports possess a highly optimistic and opportunity-oriented narrative. Conversely, the reporting tone in the EU sample was found to be situated on a more balanced, analytical, and realistic ground. In the EU list, it is observed that the tone is stabilized specifically by the increase in financial discipline and risk disclosures. The net sentiment calculation for the EU sample is as follows:

- **Positive Word Group:** Value, Sustainability, Performance, New, Development, etc. (Total Frequency: ~42,150)
- **Negative/Risk Group:** Financial (technical constraint), Risk, Risks, Emissions, Climate (within the context of risk), etc. (Total Frequency: ~33,450)
- **Total Key Word Volume:** ~280,000

$$Net\ Tone = \frac{42.150 - 33.450}{42.150 + 33.450} = 0{,}03$$

The net sentiment value of 0.03 in the EU sample indicates that the reports are prepared in a near-neutral and realistic tone, reflecting a balanced disclosure of both achievements and risks.

In conclusion, while the reports in Türkiye follow a signaling strategy based on "the value created by the company" (Positive Tone), EU reports declare sustainability as a financial risk and management process (Balanced Tone). The fact that the net sentiment in the Turkish sample is radically more positive compared to the EU may indicate that local reporting practices are still in the "Impression Management" phase.

Descriptive Analyses on Coding Frequencies and Discourse Density

Cross-tabulations were generated using MAXQDA to demonstrate the changes and differences over the years based on the codes determined for the climate discourse of the

Turkish and EU samples. In light of longitudinal data covering the 2022–2024 period, Table 1 illustrates the evolution of sustainability discourses of companies in the Turkish and EU samples, as well as the behavioral divergences in the process of alignment with mandatory reporting standards (TSRS 2/IFRS S2).

Table 1

Comparative Analysis of Coding Frequencies for Turkish and EU Samples by Year (2022–2024)

Analyzed Theme (Code)	Türkiye					EU				
	2022	Dif.	2023	Dif.	2024	2022	Dif.	2023	Dif.	2024
1.1.Governance	3227	20%	4025	39%	6583	4893	9%	5353	6%	5709
1.2.Strategy	400	42%	690	64%	1904	469	24%	621	21%	787
1.3.Risk Management	406	19%	503	51%	1023	78	23%	101	17%	121
1.4.Metrics&Targets	1133	32%	1666	17%	1996	947	25%	1259	40%	2091
2.1.Commitment	2348	14%	2732	40%	4524	2535	8%	2766	14%	3205
2.2. Impact	513	27%	704	19%	868	1834	7%	1971	1%	1992
2.3.Motive	301	41%	514	16%	614	736	31%	1067	4%	1115
3.1.Vague/General	2745	22%	3540	47%	6729	2439	18%	2972	9%	3252
3.2.Future-oriented	699	36%	1085	18%	1325	850	24%	1121	9%	1227
4.1. Positive	2041	19%	2510	26%	3379	753	17%	910	-13%	802
4.2. Negative	1226	27%	1669	38%	2685	908	16%	1087	14%	1264
4.3. Uncertainty	123	34%	187	22%	239	324	6%	343	-17%	293
5.1. Analytic	1553	24%	2033	21%	2570	942	-2%	924	1%	933
5.2. Assertive	1038	16%	1239	21%	1569	312	10%	346	-28%	271
5.3.Defensive	147	-6%	139	29%	195	156	-5%	148	-4%	142
Total	17900	23%	23236	36%	36203	18176	13%	20989	%10	23204
N=	10		10		10	14		14		14

Note. The dataset was derived from the official 2022-2024 integrated reports of the sample companies and calculated by the author using MAXQDA statistical software.

When evaluated together, Table 1—which contains the longitudinal raw coding frequencies obtained within the scope of the research—and Table 2—which presents normalized data per 10,000 words to eliminate potential biases arising from report lengths—reveal distinct strategic and discursive divergences between the Türkiye and EU samples during the alignment process with mandatory reporting standards (TSRS 2/IFRS S2).

Upon examining the data in Table 1, it is observed that the Türkiye sample recorded a dramatic increase in coding volume in 2024, the year of transition to mandatory reporting, compared to previous years. Specifically, in the "Strategy" category, the TR sample nearly tripled its coding count from 690 in 2023 to 1,904 in 2024; during the same period, the increase in the EU sample remained more limited (from 621 to 787), falling significantly behind the

Türkiye sample. Similarly, under the "Risk Management" heading, the Türkiye sample reached a coding frequency of 1,023 in 2024, while it is noteworthy that the EU sample remained at the 121 level. These data suggest that Turkish companies have entered a process of "active adaptation" by increasing reporting volumes in the face of regulatory pressure, whereas EU companies exhibit a more established and concise reporting practice. Table 2, which presents the normalized data, statistically confirms the qualitative difference between the two regions in terms of the content-orientation of the reports.

The most striking finding of the analysis is the disproportion between the "Commitment" and "Impact" codes. According to Table 1, "Commitment" codes in the Türkiye sample for 2024 (4,524) are approximately five times higher than the "Impact" codes (868), which express concrete results. In the normalized data (Table 2), the "Commitment" density of the Türkiye sample (Median: 32.54) was found to be statistically significantly higher than that of the EU sample (Median: 21.40). Conversely, in the "Impact" category, the density of EU companies (Median: 13.31) reaches nearly double that of Turkish companies (Median: 7.26). These findings can be interpreted as Turkish companies constructing their reporting on future-oriented promises, while EU companies adopt an "evidence-based" approach focusing on past performance and tangible outputs.

According to Table 1 data, although the Türkiye sample produces a numerically high frequency (1,023) regarding "Risk Management," this situation must be evaluated alongside the excessive increase in the "Vague/General Expressions" category (2024: 6,729 codes). Normalized data in Table 2 show that the density of risk discourse in the Türkiye sample (Median: 5.27) is significantly higher than in the EU (Median: 0.51); however, the density of "Vagueness" (Median: 40.28) is also approximately twice the level of the EU (Median: 21.79).

A trend of global convergence is observed in the reporting of institutional structures. In Table 2, the normalized median values in the "Governance" category (TR: 44.68 vs. EU: 37.49) are close to each other, and the Mann-Whitney U test conducted within the scope of Table 2 found no statistically significant difference between these two groups. Furthermore, the fact that both groups achieved their highest coding volume under the "Governance" heading in Table 1 (TR: 6,583; EU: 5,709) confirms that during the TSRS 2 and IFRS S2 alignment process, companies prioritize structural and auditable mechanisms such as committee establishment and job descriptions.

In conclusion, when the data in Table 1 and Table 2 are interpreted together, it is understood that while enterprises in Türkiye increased their information volume during the transition to mandatory reporting, this increase is largely concentrated on the axes of "future promises" (40% increase) and "vague expressions" (47% increase). Compared to the "impact-oriented" (Median: 13.31) structure of the EU sample, it can be argued that the reporting practice in Türkiye is still in the phase of "discursive expansion" and has not yet fully transitioned to the stage of "operational deepening."

Table 2

Discursive Density Differences Between Türkiye and the EU via Normalized Data

Analyzed Theme (Code)	TR Median (Density*)	EU Median (Density*)	p-value	Decision (p<0.05)
1.1.Governance	44,68	37,49	0,0234	Significant
1.2.Strategy	5,81	4,03	0,0332	Significant
1.3.Risk Management	5,27	0,51	0,0000	Significant
1.4.Metrics&Targets	17,81	11,05	0,0100	Significant
2.1.Commitment	32,54	21,4	0,0002	Significant
2.2. Impact	7,26	13,31	0,0000	Significant
2.3.Motive	4,45	6,27	0,0438	Significant
3.1.Vague/General	40,28	21,79	0,0000	Significant
3.2.Future-oriented	10,36	8,45	0,0234	Significant
4.1. Positive	26,72	5,42	0,0000	Significant
4.2. Negative	16,31	7,44	0,0000	Significant
4.3. Uncertainty	1,56	2,22	0,1833	Not Significant
5.1. Analytic	20,39	7,02	0,0000	Significant
5.2. Assertive	11,37	2,2	0,0000	Significant
5.3. Defensive	1,18	1,06	0,2238	Not Significant

*Density: Refers to the average coding frequency normalized per 10,000 words.

Note. Created by the author using MAXQDA statistical software.

Statistical Testing of Discursive Divergence Between Türkiye and the EU (Mann-Whitney U)

In order to determine the differences between the discursive structures of the integrated reports belonging to the Türkiye and EU samples examined within the scope of the study, statistical tests were applied. Due to the non-normal distribution of the data ($p < 0.05$), the median differences between the groups were tested using a non-parametric method. The results of the Mann-Whitney U test (Table 3), conducted to statistically examine the discursive variations in the sustainability reports of the Turkish and EU companies constituting the research sample, reveal that there are both structural similarities and strategic divergences in the reporting practices of the two groups.

Table 3

Mann-Whitney U Test Results for Discursive Differences Between Türkiye and EU Samples

Analyzed Theme (Code)	TR Median (Density*)	EU Median (Density*)	p-value	Decision (p<0.05)
1.1. Governance	417	225	0,1509	Not Significant
1.2. Strategy	58,5	28,5	0,016	Significant
1.3. Risk Management	50	4	0,0001	Significant
1.4. Metrics & Targets	147	69,5	0,0365	Significant
2.1. Commitment	277,5	163	0,0027	Significant
2.2. Impact	63,5	65	0,29	Not Significant
2.3. Motive	38	43	0,7337	Not Significant
3.1. Vague / General	315	152	0,0003	Significant
3.2. Future-oriented	93	63	0,0249	Significant
4.1. Positive	262,5	43	0,0001	Significant
5.1. Analytic	207	51	0,0001	Significant
5.3. Defensive	13	6,5	0,0713	Not Significant

Note. Created by the author using MAXQDA statistical software.

According to the analysis results, no statistically significant difference was found between the Türkiye (Median: 417) and EU (Median: 225) samples in the "Governance" theme. This finding suggests that structural elements such as oversight mechanisms, committee structures, and declarations of responsibility in corporate sustainability reporting have standardized at a global level, and that Turkish companies exhibit "isomorphism" with international norms in this field. Conversely, in the "Risk Management" category, the Türkiye sample possesses a considerably higher discourse volume (Median: 50) compared to the EU sample (Median: 4), a difference that is found to be highly significant from a statistical perspective. Similarly, a significant difference in favor of Türkiye was identified in the "Strategy" theme. This situation can be interpreted as Turkish companies operating within emerging market dynamics placing greater discursive emphasis on risk and strategy to manage uncertainties and consolidate corporate legitimacy during the alignment process with new regulations such as TSRS 2.

Another striking finding of the analysis is that the Türkiye sample exhibits a statistically significantly higher usage of "Vague / General Expressions" compared to the EU. When evaluated alongside the fact that Türkiye also has a higher density in sharing numerical "Metrics and Targets," this suggests that while companies strive for transparency through intensive data and information sharing, the frequency of vague expressions leaves this communication open to being interpreted as carrying a risk of "greenwashing."

The results in Table 3 reveal statistically significant and structural differences between the Türkiye and EU samples in terms of the temporal orientation and emotional tone of the reporting language. Specifically, the divergence in "Future Orientation" and "Positive Narrative" codes indicates that Turkish companies may be utilizing reporting communication as a strategic tool for legitimacy. According to the analysis, the median value for the Türkiye

sample in the "Future Orientation" category (93) is statistically significantly higher than that of the EU sample (63). As stated in the literature, forward-looking statements are an important tool used by companies to gain stakeholder trust and increase market value in environments of uncertainty. However, when read in conjunction with the high frequency of the "Commitment" code in Table 3, it can be interpreted that Turkish companies attempt to substitute a lack of concrete actions and results with intensive future-oriented promises and vision statements. The lower future-orientation score exhibited by EU companies may indicate that their reporting is more balanced, focusing primarily on realized data and the current state of affairs.

One of the most profound findings in Table 3 is the significant difference in the "Positive Narrative" code. The intensity of positive word usage in the Türkiye sample (Median: 262.5) is approximately six times that of the EU sample (Median: 43), and this difference is statistically highly significant. This excessive positivity in the Türkiye sample can be evaluated as a reflection of the "self-laudatory" and success-exaggeration tactics noted by Ozsozgun Caliskan et al. (2021) as being prevalent among Turkish companies. The fact that EU companies utilize a much lower positive tone (43) and a relatively more balanced language indicates that their reports are prepared with an objective, cautious style in accordance with financial reporting discipline.

In conclusion, the data in Table 3 demonstrate that Turkish companies, in the process of transitioning to mandatory reporting standards, utilize a more intensive, future-oriented, and positive language compared to their EU counterparts; however, this discursive expansion is shaped by commitment and risk rhetoric rather than concrete impact disclosures.

Longitudinal Analysis of the Transition Process to Mandatory Reporting (Kruskal-Wallis)

The results of the Kruskal-Wallis test (Table 4), conducted to analyze the longitudinal shifts in the sustainability reporting practices of companies in the Turkish sample between 2022 and 2024, indicate that the transition to mandatory reporting has created statistically significant differences in specific areas of the companies' discursive structures. However, it also suggests that a structural transformation has not yet fully materialized in certain fundamental domains.

Table 4

Analysis of Discursive Change by Year in the Turkish Sample via Kruskal-Wallis Test (Longitudinal Analysis)

Analyzed Theme (Code)	2022 Med.	2023 Med.	2024 Med.	H	p-value	Decision (p<0,05)
1.1.Governance	314,5	404	558	3,601	0,1652	No Significant Difference
1.2.Strategy	33,5	58,5	84	5,419	0,0666	Marginally Significant
1.3.Risk Management	42,5	40,5	79,5	3,337	0,1885	No Significant Difference
1.4.Metrics & Targets	95	138	181	1,93	0,381	No Significant Difference
2.1.Commitment	246,5	229	352,5	6,216	0,0447	Significant Difference
2.2.Impact	51,5	71	96,5	3,99	0,136	No Significant Difference
2.3.Motive	26,5	41	63,5	4,415	0,11	No Significant Difference
3.1.Vague/General	279	270,5	622,5	5,399	0,0672	Marginally Significant
3.2.Future-oriented	72,5	104	151,5	5,709	0,0576	Marginally Significant
4.1.Positive	205	204	335	3,12	0,2101	No Significant Difference
4.2.Negative	101	122	273	4,788	0,0913	Marginally Significant
4.3.Uncertainty	15	15	19,5	1,458	0,4824	No Significant Difference
5.1.Analytic	191	205	252	2,248	0,3249	No Significant Difference
5.2.Assertive	103	121	130	0,792	0,673	No Significant Difference
5.3.Defensive	11,5	11,5	20	1,012	0,6028	No Significant Difference

Note. Created by the author using R statistical software.

According to the analysis results presented in Table 4, the most distinct statistical divergence across the years was observed in the "Commitment" category. The commitment rhetoric, which followed a flatter trajectory in 2022 (Median: 246.5) and 2023 (Median: 229), showed a significant leap in 2024—the year mandatory reporting came into effect—reaching a median value of 352.5. This finding suggests that Turkish companies, during the alignment process with legal regulations, tend to increase their future-oriented declarations of intent and strategic promises rather than reporting current performance. It can be inferred that companies attempt to manage the pressure brought by TSRS 2 standards by presenting a "future vision" through actions that have not yet materialized.

While not statistically significant at the 5% level, categories approaching significance at the 10% level provide important clues regarding the transformation in reporting trends. The rise of the "Strategy" median value from 33.5 in 2022 to 84 in 2024 may indicate an increasing effort by companies to link their business models with climate risks. Similarly, the doubling of the median value for future-oriented statements supports the notion that reports are shifting toward a structure focused on the construction of the future rather than the accounting of the past.

Another significant finding is the upward trend in the "Vague / General Expressions" category over the years. The rise of the median value from 270.5 in 2023 to 622.5 in 2024

suggests that while fulfilling increased disclosure obligations, companies resort more frequently to general expressions open to interpretation rather than clear and verifiable information. As noted by Ozsozgun Caliskan et al. (2021), this aligns with the tendency of firms in emerging markets to employ "obfuscation" or "defensive" impression management tactics in environments of uncertainty. This vagueness, accompanying the increased volume of commitments, may harbor a potential risk of "greenwashing" regarding the quality of the reports.

Despite the increase in commitment and strategy rhetoric, no statistically significant change was detected over the years in the "Impact" category, which represents concrete results. This large gap between commitment (Median: 352.5) and impact (Median: 96.5) in 2024 can be interpreted as Turkish companies currently being in a phase of "discursive expansion" in sustainability reporting, whereas the translation of these discourses into tangible outputs may take time. This temporal lag aligns with the diffusion dynamics described by Jung (2025), who posits that broad institutional concepts like ESG tend to exert a slower, albeit more sustained, influence on broader engagement compared to immediate, action-oriented terms.

In conclusion, Table 4 data indicate that during the transition to mandatory reporting (2024), companies in Türkiye reconstructed their reporting language around the axes of future promises and strategic intentions. However, this signifies a period of "active adaptation" supported by vague expressions, where concrete impact disclosures have not yet been able to accompany this increase.

In contrast to Türkiye, analyses conducted on the EU sample showed no statistically significant discursive change in any of the 15 categories examined during the 2022–2024 period ($p > 0.05$). The high stability of p-values in categories such as Governance ($p=0.9558$), Analytical Language ($p=0.9868$), and Impact ($p=0.9949$) in EU reports may serve as an indicator that reporting practices in this region have approached a stage of institutional maturity and saturation. The reporting languages of EU companies have achieved a stable structure within the framework of established standards and did not undergo a radical discursive transformation during the period analyzed.

Table 5

Analysis of Discursive Change by Year in the EU Sample via Kruskal-Wallis Test (Longitudinal Analysis)

Analyzed Theme (Code)	2022 Med.	2023 Med.	2024 Med.	H	p-value	Decision (p<0,05)
1.1.Governance	211,5	280,5	229,5	0,09	0,9558	No Significant Difference
1.2.Strategy	17	31,5	44	1,472	0,4791	No Significant Difference
1.3.Risk Management	3,5	3,5	4,5	0,311	0,8559	No Significant Difference
1.4.Metrics & Targets	66,5	69	127,5	1,983	0,3711	No Significant Difference
2.1.Commitment	101,5	170	167,5	0,979	0,613	No Significant Difference
2.2.Impact	71	82	64,5	0,01	0,9949	No Significant Difference
2.3.Motive	34	43	47,5	0,514	0,7734	No Significant Difference
3.1.Vague/General	135,5	153,5	166,5	0,557	0,7569	No Significant Difference
3.2.Future-oriented	54	69	68	1,405	0,4954	No Significant Difference
4.1.Positive	34,5	64,5	42,5	0,236	0,8888	No Significant Difference
4.2.Negative	26	76	54,5	0,92	0,6313	No Significant Difference
4.3.Uncertainty	9	13	9	0,574	0,7504	No Significant Difference
5.1.Analytic	44	51	52	0,027	0,9868	No Significant Difference
5.2.Assertive	17	20,5	14	0,48	0,7868	No Significant Difference
5.3.Defensive	6	7,5	6,5	0,065	0,9679	No Significant Difference

Note. Created by the author using R statistical software.

This fundamental divergence between the two groups can be explained through the lens of Signaling Theory and Institutional Theory, specifically regarding the differences in the "learning process" and "level of institutionalization." While the EU sample exhibits a stable professionalism, the dynamic discursive surge in the Turkish sample—which occasionally pushes the boundaries of statistical significance—suggests that local companies are in an "active adaptation" phase within the regulatory compliance process. Particularly in Türkiye, the significant increase in commitment-oriented rhetoric reflects a transition period strategy where strategic intentions are prioritized over concrete performance data (Impact).

Conclusion

This study analyzed the impact of mandatory sustainability reporting standards (TSRS 1 and TSRS 2), which came into effect as of January 1, 2024, on the climate-oriented corporate narratives of BIST-listed companies through a comparative and longitudinal perspective with an European Union sample. The research findings indicate that the transition to legal mandates has created a dramatic surge in the reporting volume and strategic discourse intensity of Turkish companies; however, this increase does not fully align with content quality or the dimension of financial materiality.

The analysis results show that in the transition year (2024), companies in the Turkish sample disproportionately increased their coding frequencies in technical headings such as

"Strategy" and "Risk Management" compared to the EU sample. This situation can be interpreted as an "excessive signaling" strategy employed by firms in emerging markets to maintain legitimacy in the face of new regulations. As stated in the literature, firms in developing countries tend to increase reporting volumes to meet the expectations of global stakeholders and project an image of a "good corporate citizen." However, the fact that the increase in risk management discourse (103%) in the Turkish sample is accompanied by a high density in the "Vague/General Expressions" category (TR Median: 315 vs. EU Median: 152) suggests that this compliance carries the risk of remaining "symbolic." This observation is consistent with Tumewang et al. (2025), who conclude that despite the proliferation of reporting, many firms remain in a “cherry-picking” state, disclosing non-material information while failing to provide the in-depth strategic disclosures required by the standards. This finding aligns with Li et al. (2025), who demonstrate that firms often use extensive strategic disclosures and high ESG scores as a form of impression management to divert attention from a lack of genuine decarbonization performance. Furthermore, the prevalence of vague expressions in the Turkish sample corresponds to the attributes of greenwashing identified by Spaniol et al. (2024), who define it as making unsubstantiated claims with deceptive intent to gain a competitive advantage. Furthermore, Maniora (2017) empirically challenges the assumed superiority of integrated reporting frameworks, finding that they can be negatively associated with ESG integration levels compared to stand-alone reporting, as the dominance of financial priorities may unintentionally marginalize non-financial disclosure. This finding is consistent with Özcan and Akbaş (2025), who noted that while BIST companies identify climate risks, they fall short in disclosing financial impacts and scenario analyses. Similarly, in the international literature, Cho et al. (2012) and Marquis et al. (2016) emphasize that as scrutiny over firms' environmental performance increases, they tend to engage in impression management through selective disclosure and ambiguous phrasing. Turkish companies appear to have chosen to mitigate regulatory pressure by "narrating" risk verbally rather than "managing" it through technical and financial data.

One of the most critical findings of the study is that while the discourse of Turkish companies is built upon "Future-Oriented Commitments," EU companies focus on "Past Impact and Performance." The fact that "Commitment" codes in the Turkish data are approximately five times higher than "Impact" codes may serve as an indicator that companies are attempting to manage sustainability through future promises rather than tangible outputs. This suggests that the problem of "misalignment between motive and impact," as noted by Saglam et al. (2024), has deepened specifically in the context of Türkiye. Furthermore, the finding by Kaldırım and Kaldırım (2025) that BIST 25 companies showed the lowest compliance (44%) under the "Metrics and Targets" heading supports the "lack of impact" identified in this study. This situation recalls the "correlation between signaling and greenwashing" posited by Mahoney et al. (2013); intensive commitments made without concrete performance data are generally accepted in the literature as indicators of greenwashing. Companies are eager to disclose targets (commitments) but struggle to report realized performance (impact). However, it should be acknowledged that the high reliance on vague and future-oriented commitments may not exclusively stem from strategic greenwashing. Given that 2024 is the initial transition year, this gap could also reflect a 'learning curve' and organizational capacity constraints, particularly in modeling complex financial impacts and gathering Scope 3 emissions data.

The research demonstrates that there is no statistically significant difference between Türkiye and the EU in the field of "Governance" ($p > 0.05$). This indicates that a global "isomorphism" has been achieved in structural elements such as committee establishment and job descriptions. However, when the tone of the discourse is examined, it is observed that

Turkish companies utilize a much more "Positive" and "Assertive" language compared to the EU. This finding contrasts with the European context described by Anwar et al. (2025), who observed that following mandatory regulation (EUNFRD), firms adopted a more cautious and less optimistic tone to avoid regulatory scrutiny. Furthermore, Bassyouny et al. (2020) suggest that such positive tone may be driven by specific CEO characteristics, although strong corporate governance mechanisms, such as independent boards, typically moderate this excessive optimism. This excessive optimism aligns with the findings of Ozsozgun Caliskan et al. (2021), who identified that Turkish CEOs frequently employ "self-laudatory" tactics in reports. This "marketing-oriented" language, which may overshadow financial performance or concrete climate action, brings with it the risk of "greenwashing"—a risk that Tepeli and Büklü (2025) point out as being prevalent during transition periods where audit mechanisms are not yet fully established.

When evaluated alongside the studies of Ayrancı (2025) on the aviation sector and Başar and Doğruel (2025) on the fast fashion industry, the results suggest that Turkish companies use sustainability reporting as a tool for strategic communication and image management; however, the depth of technical data (financial materiality as per TSRS 2) has not yet reached EU levels. While the high density of "Analytical" and "Financial" terminology in EU companies indicates that reporting is a discipline of investor relations, the focus of Turkish companies on concepts such as "Value," "Future," and "Vision" suggests that the process has not yet fully emerged from the corporate public relations phase. The Net Sentiment analysis conducted within the scope of this research reveals a discursive chasm between the Türkiye (0.26) and EU (0.03) samples. The high positive tone in Turkish reports can be considered a sign that local companies use sustainability as a 'corporate image management' tool and, in accordance with Signaling Theory, consistently send success-oriented messages to stakeholders.

Contrary to this activity in Türkiye, longitudinal analyses of the EU sample showed no statistically significant change over the years in any of the 15 categories examined. The stability of critical indicators such as "Impact" and "Analytical Language" in EU reports can be interpreted as the reporting practice in this market having already reached a level of "institutional maturity" prior to TSRS 2 (IFRS S2). As noted by Rossi and Candio (2023), integrated reporting and ESG disclosures in Europe have transformed into an established information set that reduces analyst forecast errors. Consequently, EU companies did not have to radically alter their reporting languages to comply with new standards; instead, they maintained their existing "evidence-based" structures.

The longitudinal difference between the two regions supports the finding of Gürsoy and Özuyar (2024) that BIST companies utilize Sustainable Development Goals (SDGs) "selectively" in their reporting showcases rather than integrating them into business processes. The upward trend of the "Metrics and Targets" code in Türkiye, while the "Impact" code remains statistically stagnant, can be interpreted as the process still being in the "discursive expansion" phase; the transition to the "operational deepening" phase—representing the reporting of concrete outputs—has not yet been achieved. Conversely, by maintaining its high impact-oriented base effect, the EU continues to move sustainability reporting beyond being a "letter of intent" and uses it as a tool for transparency and accountability, as highlighted by Al-Amosh (2025).

When the primary hypotheses of the study are evaluated in light of the research findings, H1, which predicted that the transition to mandatory reporting would transform the climate discourse into an “action-oriented” structure, was not supported. Longitudinal analyses revealed that although BIST companies significantly increased their “Commitment” rhetoric,

they failed to create a statistically significant change in the “Impact” category, which expresses concrete results. On the other hand, H2, which predicted that mandatory standards would reduce the gap between global and local reporting languages (convergence), was partially rejected. While analyses indicate that a global isomorphism has been achieved under the structural heading of 'Governance,' it is understood that statistical differences between the Türkiye and EU samples are maintained across other categories. This suggests that mandatory standards provide formal alignment but have not yet created full global integration in terms of discursive quality.

In conclusion, longitudinal analyses reveal that mandatory reporting (TSRS 2) triggers "excessive signaling" behavior in emerging markets like Türkiye, increasing report volumes without immediately bringing narrative quality to the desired impact-oriented level. As Kaldırım and Kaldırım (2025) state, the production of high-quality information supported by financial materiality will take time. The EU example demonstrates that beyond legal mandates, corporate culture and stakeholder pressure are primary elements stabilizing reporting quality. Therefore, to close the "intent-action gap" and ground the expanding corporate discourse on "green trust" (Başar & Selimoğlu, 2025) rather than "greenwashing", the activation of mandatory independent assurance audits covering discursive consistency is essential (Özçelik, 2025). It is recommended that future research tracks reports from 2025 onwards, when TSRS 2 will be fully implemented, to observe whether these "transition period" pains (vagueness, excessive commitment) diminish and to what extent financial impact analyses are reflected in the reports.

References

Abhayawansa, S., & Adams, C. (2022). Towards a conceptual framework for non-financial reporting inclusive of pandemic and climate risk reporting. *Meditari Accountancy Research*, *30*(3), 710-738. https://doi.org/10.1108/MEDAR-11-2020-1097

Acar Uğurlu, Y. (2024). 2000-2024 dönemi için yeşil aklama konusundaki küresel eğilimler. *Mali Çözüm (İSMMMO), 34*(186), 1787-1823.

Afifa, M. A., Nguyen, N. M., & Bui, D. V. (2025). Environmental, social and governance (ESG) disclosure quality in developing countries: Evidence from the ASEAN region. *Corporate Governance: The International Journal of Business in Society*, 25(7), 1728-1754. https://doi.org/10.1108/CG-02-2024-0102

Al Amosh, H. (2025). Exchange rate volatility and ESG performance: An international empirical analysis. *The Journal of Corporate Accounting & Finance, 36,* 14-28. https://doi.org/10.1002/jcaf.22774

Albitar, K., Hussainey, K., Kolade, N., & Gerged, A. (2020). ESG disclosure and firmperformance before and after IR: The moderating role of governance mechanisms. *International Journal of Accounting and Information Management, 28*(2), 1-21.

Albitar, K., Abdoush, T., & Hussainey, K. (2023). Do corporate governance mechanisms and ESG disclosure drive CSR narrative tones? *International Journal of Finance & Economics*, *28*(4), 3876-3890.

Ali, K., Arslan, H. M., Mubeen, M., Azeem, H., Zhen-Yu, Z., Yushi, J., & Miao, M. (2025). From reporting to responsibility: Investigating the influence of sustainability disclosure on earnings management. *Environment, Development and Sustainability*, 27, 30431–30464. https://doi.org/10.1007/s10668-024-04920-y

Anıl, B., Aerts, W., Roszkowska-Menkes, M., & van Caneghem, T. (2025). CEO talk under scrutiny: Visibility and rhetorical impression management in sustainability narratives. *Accounting, Auditing & Accountability Journal.* https://doi.org/10.1108/AAAJ-08-2024-7297

Anwar, M., Sinnewe, E., & Quayle, A. (2025). Environmental disclosure tone and analyst forecast behaviour: A study of the EU Nonfinancial Reporting Directive. *Accounting & Finance, 65*(4), 3574–3604. https://doi.org/10.1111/acfi.70059

Ayrancı, M. (2025). Sürdürülebilirlik raporlarının kurumsal imaj oluşturmaya etkisi: Türk hava yolları ve Pegasus havayollarının karşılaştırmalı analizi. *Aksaray Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 17*(4), 221-230. https://doi.org/10.52791/aksarayiibd.1667007

Bassyouny, H., Abdelfattah, T., & Tao, L. (2020). Beyond narrative disclosure tone: The upper echelons theory perspective. *International Review of Financial Analysis, 70*, 101499. https://doi.org/10.1016/j.irfa.2020.101499

Başar, Ü., & Doğruel, M. (2025). Identifying sustainability trends in the fast fashion industry with content analysis. *Istanbul Gelisim University Journal of Social Sciences, 12*(3), 1141-1161. https://doi.org/10.17336/igusbd.1712549

Başar, A. B., & Selimoğlu, S. K. (2025). Legal framework in Türkiye within the scope of transition from greenwashing to green trust. *Muhasebe ve Finansman Dergisi, 108*, 131-168. https://doi.org/10.25095/mufad.1763631

Bernardi, C., & Stark, A. W. (2018). Environmental, social and governance disclosure, integrated reporting, and the accuracy of analyst forecasts. The British Accounting Review, 50(1), 16-31. https://doi.org/10.1016/j.bar.2016.10.001

Carungu, J., Dimes, R., & Molinari, M. (2025). The influence of global sustainability reporting standards on the discourse within the EU. *Sustainability Accounting, Management and Policy Journal*. Advance online publication. https://doi.org/10.1108/SAMPJ-12-2024-1378

Cho, C. H., Michelon, G., & Patten, D. M. (2012). Impression management in sustainability reports: An empirical investigation of the use of graphs. *Accounting and the Public Interest*, *12*(1), 16-37. https://doi.org/10.2308/apin-10249

de Villiers, C., Dimes, R., La Torre, M., & Molinari, M. (2024). The International Sustainability Standards Board's (ISSB) past, present, and future: Critical reflections and a research agenda. *Pacific Accounting Review, 36*(2), 255–273. https://doi.org/10.1108/PAR-02-2024-0038

Deng, B., Peng, Z., Albitar, K., & Ji, L. (2024). Top management team stability and ESG greenwashing: Evidence from China. *Business Strategy and the Environment, 34*(1), 450-467. https://doi.org/10.1002/bse.3998

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American sociological review*, *48*(2), 147-160.

Fang, C., Wang, Z., & Zhao, L. (2025). Environmental regulations and the greenwashing of corporate ESG reports. *Economic Analysis and Policy, 87,* 1469-1481. https://doi.org/10.1016/j.eap.2025.07.030

Fianko, S. K., Amoah, N., Addo, A., Agyeman, K., Gbadago, F. Y., Adjaye-Gyamfi, O., Agbemava, E., Agropah, F., Zaglago, L., Atiase, D. D., & Nooni, I. K. (2025). IFRS S1 and S2 implementation readiness in emerging markets: A multi-dimensional assessment framework and market readiness index. *Journal of Financial Reporting and Accounting.* https://doi.org/10.1108/JFRA-12-2024-0921

Grewal, J., & Serafeim, G. (2020). Research on corporate sustainability: Review and directions for future research. *Foundations and Trends in Accounting*, *14*(2), 73-127. https://doi.org/10.1561/1400000061

Gültepe, M. (2025) Sürdürülebilirlik stratejilerinin halkla ilişkiler açısından incelenmesi: Hazır giyim markalarında marka aktivizmi yaklaşımları. *INIF E- Dergi, 10*(2), 953-975. https://doi.org/10.47107/inifedergi.1736543

Gürsoy, Ç., & Özuyar, P. (2024). BIST sürdürülebilirlik 25 endeksindeki şirketlerin sürdürülebilirlik konusuna yaklaşımları. *Anadolu Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 25*(3), 441-464. https://doi.org/10.53443/anadoluibfd.1440136

Henry, E. (2008). Are investors influenced by how earnings press releases are written? *Journal of Business Communication*, 45(4), 363-407. https://doi.org/10.1177/0021943608319388

Jung, S.-U. (2025). Tracing the Diffusion of Sustainability Discourse: Institutional Signals and Consumer Search Behavior in the United States. *Sustainability*, *17*(17), Article 7697. https://doi.org/10.3390/su17177697

Kaldırım, Y., & Kaldırım, Z. (2025). BİST sürdürülebilirlik 25 endeksindeki işletmelerin sürdürülebilirlik raporlarının TSRS S1'e uyumu. *Mali Çözüm (İSMMMO), 35*(192), 1995-2020.https://archive.ismmmo.org.tr/docs/malicozum/192malicozum/10.pdf

Kamu Gözetimi, Muhasebe ve Denetim Standartları Kurumu (KGK). (2023). TSRS 1 Sürdürülebilirlikle İlgili Finansal Bilgilerin Açıklanmasına İlişkin Genel Hükümler. https://www.kgk.gov.tr

Kamu Gözetimi, Muhasebe ve Denetim Standartları Kurumu (KGK). (2023). TSRS 2 İklimle İlgili Açıklamalar. https://www.kgk.gov.tr

Kamu Gözetimi, Muhasebe ve Denetim Standartları Kurumu (KGK). (2024). 2024 faaliyet raporu. https://www.kgk.gov.tr

Kathan, M. C., Utz, S., Dorfleitner, G., Eckberg, J., & Chmel, L. (2025). What you see is not what you get: ESG scores and greenwashing risk. *Finance Research Letters, 74,* Article 106710. https://doi.org/10.1016/j.frl.2024.106710

Kılıç, M., & Kuzey, C. (2018). Determinants of forward-looking disclosures in integrated reporting. *Managerial Auditing Journal, 33*(1), 115–144. https://doi.org/10.1108/MAJ-12-2016-1498

Li, J., Peng, X., Gao, J., & Tang, Q. (2025). Detecting greenwashing behaviour in decarbonization performance. *Accounting & Finance, 65*(4), 3739–3762. https://doi.org/10.1111/acfi.70088

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance, 66(*1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Mahoney, L. S., Thorne, L., Cecil, L., & LaGore, W. (2013). A research note on standalone corporate social responsibility reports: Signaling or greenwashing? *Critical perspectives on accounting*, *24*(4-5), 350-359 https://doi.org/10.1016/j.cpa.2012.09.008

Marquis, C., Toffel, M. W., & Zhou, Y. (2016). Scrutiny, norms, and selective disclosure: A global study of greenwashing. *Organization science*, *27*(2), 483-504. https://doi.org/10.1287/orsc.2015.1039

Mahran, K., & Elamer, A. A. (2025). Shaping ESG commitment through organizational psychological capital: The role of CEO power. *Business Strategy and the Environment, 34*(1), 690-707. https://doi.org/10.1002/bse.4007

Maniora, J. (2017). Is integrated reporting really the superior mechanism for the integration of ethics into the core business model? An empirical analysis. *Journal of Business Ethics, 140*(4). https://doi.org/10.1007/s10551-015-2874-z

Ozsozgun Caliskan, A., Esen, E., & Barkemeyer, R. (2021). Impression management tactics in the CEO statements of Turkish sustainability reports. *Business Ethics, the Environment & Responsibility, 30*(4), 485–506. https://doi.org/10.1111/beer.12374

Özcan, Ö. D., & Akbaş, H. E. (2025). BIST sürdürülebilirlik endeksindeki şirketlerin TSRS 2 iklimle ilgili açıklamalar standardına hazırlık seviyeleri: Elektrik tesisleri ve güç jeneratörleri ile inşaat malzemeleri sektörleri üzerine bir araştırma. *Uluslararası Sosyal Alan Araştırmaları Dergisi, 14*(1), 151-171.

Özçelik, M. (2025). Sürdürülebilirlik raporlarının zorunlu güvence denetimi: Bağımsız denetçilerden görüşler. *İşletme Araştırmaları Dergisi, 17*(2), 1773-1793. https://doi.org/10.20491/isarder.2025.2064

Özdemir, N. (2024). Güncel gelişmeler ekseninde sürdürülebilirlik raporlama sürecinin ve Türkiye sürdürülebilirlik raporlama standartlarının değerlendirilmesi. *Ekonomik ve Sosyal Araştırmalar Dergisi*, *20*(2), 444-458. https://izlik.org/JA63UJ59SW

Rabaya, A., & Saleh, N. M. (2022). The moderating effect of IR framework adoption on the relationship between environmental, social, and governance (ESG) disclosure and a firm's competitive advantage. *Environment Development and Sustainabilty*, *24*(1). https://doi.org/10.1007/s10668-021-01519-5

Rakhavendra, A. H., Bala, P. K., & Mukherjee, A. (2023). CEO communications and ESG performance: Deciphering the impact of corporate narratives. *International Journal of Business & Economics (IJBE), 8*(2). https://doi.org/10.58885/ijbe.v08i2.190.ar

Rossi, P., & Candio, P. (2023). The independent and moderating role of choice of non-financial reporting format on forecast accuracy and ESG disclosure. *Journal of Environmental Management, 345,* 118891. https://doi.org/10.1016/j.jenvman.2023.118891

Saglam, B. B., Solak-Fiskin, C., & Akgul, E. F. (2024). Conveying the sustainability message through CEO letters: An investigation on selected transportation companies. *World Review of Intermodal Transportation Research, 11*(4), 436–458. https://doi.org/10.1504/WRITR.2023.10063023

Spaniol, M. J., Danilova-Jensen, E., Nielsen, M., Rosdahl, C. G., & Schmidt, C. J. (2024). Defining greenwashing: A concept analysis. *Sustainability*, *16*(20), Article 9055. https://doi.org/10.3390/su16209055

Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics*, *87*, 354-374.

Tepeli, Y., & Büklü, E. (2025). Çevre dostu açıklamalar, yeşil pazarlama ve piyasa gerçekliği: Yeşil aklamanın finansal etkileri. *Uluslararası Yönetim Akademisi Dergisi, 8*(4), 1102-1114. https://doi.org/10.33712/mana.1831443

Todaro, N. M., & Torelli, R. (2024). From greenwashing to ESG-washing: A focus on the circular economy field. *Corporate Social Responsibility and Environmental Management.* https://doi.org/10.1002/csr.2786

Tumewang, Y., Ntim, C., & Haque, F. (2025). Task force on climate-related financial disclosures: A systematic literature review and future research agenda. *Business Strategy and the Environment, 34*(8), 10033–10058. https://doi.org/10.1002/bse.70112

Treepongkaruna, S., Yong, H. H. A., Thomsen, S., & Kyaw, K. (2024). Greenwashing, carbon emission, and ESG. *Business Strategy and the Environment, 33*(8), 8526–8539. https://doi.org/10.1002/bse.3929

Yu, W., Chavez, R., Feng, M., Wong, C. Y., & Fynes, B. (2020). Green human resource management and environmental cooperation: an ability-motivation-opportunity and contingency perspective. *International Journal of Production Economics, 219,* 224-235. https://doi.org/10.1016/j.ijpe.2019.06.013

Yuan, M., Wang, X., Suo, X., Lin, H., Yu, M., & Jiang, W. (2025). Affective orchestration at the helm! Unraveling CEO affects' effect on corporate ESG performance. *Business Ethics, the Environment & Responsibility*. Advance online publication. https://doi.org/10.1111/beer.12813

Zhou, S., Simnett, R., & Green, W. (2017). Does integrated reporting matter to the capital market? *Abacus*, *53*(1), 94-132. https://doi.org/10.1111/abac.12104

Information About the Article/Makale Hakkında Bilgiler

The Ethical Rules for Research and Publication / Araştırma ve Yayın Etiği

The author declared that the ethical rules for research and publication followed while preparing the article.

Yazar makale hazırlanırken araştırma ve yayın etiğine uyulduğunu beyan etmiştir.

Conflict of Interests/ Çıkar Çatışması

The author have no conflict of interest to declare.

Yazar çıkar çatışması bildirmemiştir.

Grant Support/ Finansal Destek

The author declared that this study has received no financial support.

Yazar bu çalışma için finansal destek almadığını beyan etmiştir.

Author Contributions/ Yazar Katkıları

The draft process of the manuscript/ Taslağın Hazırlanma Süreci O.Ö, Data Collection/Verilerin Toplanması O.Ö, Writing the Manuscript/ Makalenin Yazılması O.Ö, Submit, Revision and Resubmit Process/ Başvuru, Düzeltme ve Yeniden Başvuru Süreci O.Ö.

Journal of Applied and Theoretical Social Sciences

JATSS, 2026; 8(1), 169-188

First Submission:04.01.2026
Revised Submission After Review:16.03.2026
Accepted For Publication:24.03.2026
Available Online Since:25.03.2026

Research Article

The Limits of the Digital State and Individual Privacy: A Public Sphere Perspective

Ejder Çetin [a] & Mustafa Demirkol [b]

Abstract

Introduction: Mutual trust lies at the heart of the social contract; yet, in the neoliberal era, this bond is being tested by the expanding limits of the digital state. This study problematizes the transformation of privacy not merely as a technical security breach, but as a deeper erosion of democratic legitimacy. It seeks to understand whether the digital sphere is evolving into a liberating 'agora' or a surveillance mechanism that threatens the individual's status as a political subject.

Method: Adopting a qualitative lens, the research weaves together a descriptive analysis with a theoretical critique grounded in political philosophy. The tension between the promise of digital freedom and the reality of surveillance society is interpreted through the comparative perspectives of Hannah Arendt's public/private distinction and Jürgen Habermas's communicative rationality.

Results or Findings: The analysis reveals a stark paradox: The digital sphere has morphed into a "transparency regime" reminiscent of a panopticon rather than a democratic public sphere. It is observed that the state is shifting from a service provider to a proactive structure of predictive policing, where privacy is commodified and ethical boundaries are redrawn by algorithms.

Discussion or Conclusion: The study concludes that privacy is not a luxury to be granted by the state, but a constitutive "constitutional threshold" that defines human dignity. To sustain legitimacy, policy must shift from data-hoarding to a privacy by design approach. Reclaiming the digital sphere as a space of freedom requires strictly limiting state authority through ethical and legal barricades.

Keywords: public sphere, private space, digital space, privacy, surveillance society

JEL Codes: H83, K24, H11, O33, O38

[a] Independent Researcher Dr., Social Security Institution of the Republic of Türkiye İzmir Provincial Directorate, İzmir/Türkiye, ejdercet@gmail.com , https://orcid.org/ORCID ID: 0000-0002-7185-4885

[b] Res. Asst. Dr., Tekirdağ Namık Kemal University, Faculty of Economics and Administrative Sciences, Department of Political Science and Public Administration, Tekirdağ/Türkiye, mustafademirkol@nku.edu.tr, ORCID ID: https://orcid.org/0000-0002-9860-8836 (Corresponding Author)

Journal of Applied and Theoretical Social Sciences

JATSS, 2026; 8(1), 169-188

İlk Başvuru: 04.01.2026
Düzeltilmiş Makalenin Alınışı: 16.03.2026
Yayın İçin Kabul Tarihi: 24.03.2026
Online Yayın Tarihi: 25.03.2026

Araştırma Makalesi

Dijital Devletin Sınırları ve Bireysel Mahremiyet: Bir Kamusal Alan Perspektifi

Ejder Çetin [a] & Mustafa Demirkol [b]

Öz

Giriş: Karşılıklı güven toplumsal sözleşmenin özüdür; ancak neoliberal çağda genişleyen dijital devlet pratikleri bu hayati bağı tehdit etmektedir. Bu çalışma, mahremiyeti salt teknik bir güvenlik sorunu olarak değil, demokratik meşruiyetin bir krizi olarak sorunsallaştırmakta; dijital alanın özgürleştirici bir agora mı yoksa bireyi nesneleştiren bir "gözetim mekanizması" mı olduğunu sorgulamaktadır.

Yöntem: Nitel bir yaklaşım benimseyen araştırma, betimsel analizi siyaset felsefesi temelli teorik bir eleştiriyle harmanlamaktadır. Dijital özgürlük vaadi ile "gözetim toplumu" gerçekliği arasındaki gerilim, Hannah Arendt'in özel/kamusal alan ayrımı ve Jürgen Habermas'ın iletişimsel eylem kuramı üzerinden karşılaştırmalı olarak okunmuştur.

Sonuçlar ya da Bulgular: Analiz, dijital alanın demokratik bir tartışma zemininden ziyade, panoptikonu andıran bir "şeffaflık rejimine" dönüştüğünü ortaya koymaktadır. Devletin pasif bir hizmet sağlayıcıdan, yapay zeka destekli "öngörücü polislik" yapısına evrildiği; bu süreçte mahremiyetin metalaştığı ve etik sınırların algoritmalarca ihlal edildiği görülmektedir.

Tartışma ya da Yapılan Çıkarımlar: Çalışma, mahremiyetin devlet tarafından bahşedilen bir lüks değil, insan onurunu ve siyasal özne olma hakkını koruyan kurucu bir "anayasal eşik" olduğu sonucuna varmaktadır. Dijital meşruiyeti sürdürmek için veri istifçiliğinden vazgeçilmeli; "tasarımla mahremiyet" ilkesi ve bağımsız etik denetimlerle dijital alan yeniden bir özgürlük sahasına dönüştürülmelidir.

Anahtar Kelimeler: kamusal alan, özel alan, dijital alan, mahremiyet, gözetim toplumu

JEL Kodlar: H83, K24, H11, O33, O38

[a] Dr., Bağımsız Araştırmacı, T.C. Sosyal Güvenlik Kurumu İzmir İl Müdürlüğü, İzmir/Türkiye, ejdercet@gmail.com, ORCID ID: 0000-0002-7185-4885

[b] Arş. Gör. Dr., Tekirdağ Namık Kemal Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Siyaset Bilimi ve Kamu Yönetimi Bölümü, Tekirdağ/Türkiye, mustafademirkol@nku.edu.tr, ORCID ID: https://orcid.org/0000-0002-9860-8836 (Sorumlu Yazar)

Introduction

While seeking the conditions of the ideal "discursive public sphere" attained in the 18th century, Habermas critiqued mass society through the narrative of the bourgeois public sphere which began to deteriorate in the 19th century, yet he simultaneously established the principles for the possibility of a political public sphere in modern democracies. Arendt, on the other hand, offering a more pessimistic critique than Habermas with her analysis of the "Modern Age" dating back to the 17th century, presented a total critique of modernization by idealizing the specific forms of political organization in ancient Greek and Roman societies, where participatory democracy was experienced in the "real" sense.

A fresh look at old ideas about public debate emerges when we consider how online spaces now shape conversation - algorithms quietly shift who gets heard, how voices connect, what counts as reason (Margetts & Dorobantu, 2019). In this respect, the digital public sphere no longer merely extends earlier communicative arenas but restructures them through data-driven governance and platform power.

The research question of our study is as follows: "How do individual privacy spaces change due to digital state practices in the context of surveillance in the new public sphere and what can be said about this change in terms of the classical public sphere theory?" The use of artificial intelligence, data analysis and decision making algorithms in the digital state, which has become a global phenomenon, increasingly erases the borders between the state and the citizen and ambiguity in governance, security and individual freedoms arise (Lee et al., 2024). In our study, which is a qualitative study, descriptive analysis and literature review methods are used. In this context, within the scope of the "new surveillance" concept specific to the digital age, we comparatively and interpretively examine the works of Hannah Arendt and Jürgen Habermas, who are the leading names of the differences between the public and private spaces. Our study, which is an interdisciplinary theoretical study, aims to introduce a new approach to the privacy debate, which is reduced to a technical and legal level and limited to data security in developing countries such as Turkey. By incorporating key concepts in political philosophy, such as the agora, the panopticon, and rational communication, the study aims to reconceptualize privacy as an integral part of the crisis of trust that characterizes the digital state-individual relationship. See also Véliz (2021) and Margetts & Dorobantu (2019). In this respect, the argument of this study echoes recent normative discussions of privacy that depart from the traditional understanding of privacy as a right of the individual and opt for the understanding that privacy can also be a social good with democratic relevance. The latter perspective posits that in digital governance environments, privacy can serve to sustain the basic principle of trust among citizens, governments, technology providers, and third parties, which is indispensable for efficient and accountable governance in the digital age.

This study aims to conceptualize privacy as a foundational boundary of democratic legitimacy in the digital age by reinterpreting a debate primarily addressed at the legal and technical levels in the literature on digital government and privacy, within the framework of Arendt and Habermas' theories of the public sphere. This article makes three contributions to the on-going discussion in the literature on digital state practices and privacy. Firstly, it leaves the rather formal and technological focus on the discussion about digital state practices and privacy behind by considering privacy as a political and normative issue in the context of the theory of the public sphere. Secondly, it examines in the light of the theories of Hannah Arendt and Jürgen Habermas on the relation between the state and the public sphere, to what extent the digital sphere can be regarded as a space of civil society for democratic communication between citizens, in a world of digital surveillance technologies. Thirdly, it criticizes the prevailing

individual rights-based understanding of privacy in the literature and introduces a constitutional alternative based on a normative concept of privacy as a constitutional threshold limiting the democratic potential of the digital state. Methodologically, this article employs qualitative and theoretical analysis, drawing on conceptual analysis, in order to evaluate the theoretical framework of the public sphere in the context of digital state practices.

From “discriminatory designs” to the digital state of exception The digital state is no longer just a technological tool for public administration, but a new political and power relationship between the state and the citizen. Through the use of “discriminatory designs” disguised as “input neutral” and “technological” - which Benjamin (2019) describes as the digital architecture that, according to the norms of the public space defined by Arendt and Habermas, undermines it and turns the citizens into transparent beings and keeps the state obscure in the algorithms of the decisions. This study investigates the preservation of the boundaries of the public space in a data-driven model of governance, while linking the theories of the public space with the digital practices that are currently exercised and analysing the consequences of achieving algorithmic efficiency at the cost of the political effectiveness of citizenship according to Hannah Arendt.

Methodology and Analytical Framework

In this study, preferred research design for this study would thus be conceptual and normative rather than empirical or field-based. The focus here is less on frequency counts of digital surveillance and more so on the qualitative understanding about how fundamentally, through these practices of a digital state, that's being dramatically rearranged. By emphasizing the normative implications of these community-building practices, I explain how the 'digital limit' of both political and non-political communities substantively alters the conditions for (and limits to) democratic legitimacy.

I have created a dual-lens analytical framework grounded in the political philosophy of Hannah Arendt and Jürgen Habermas to meet this need for adequate scientific theory. Further, in this model, Arendt’s public/private dichotomy is not the mere historical concept it has frequently been represented as; rather, it serves a diagnostic purpose to assess if any state intervention online would be an internal structural ‘invasion’ of the protected private domain (see Arendt, 2023). It also helps us see whether the individual is being shrunk from a political subject to a data object.

Conversely, Habermas’s theory of communicative rationality serves as a normative standard against which this analysis can be measured. I then use this lens to evaluate if the digital public sphere engages anymore in a broad rational-critical debate or has devolved into an algorithmically mediated echo chamber. This top-down and bottom-up blend allows the study to go deeper than a descriptive narrative. It establishes a theoretical lens through which contemporary governance forms, like predictive policing and infrastructuralisation of large-scale data can be critically examined.

The findings of this research relate to the ideal understanding of the political, and not to the statistical level of the political. The unit of analysis of this research is the structural change in political subjectivity resulting from the exercise of algorithmic governance. By analyzing through the specific research prism the legal framework and the digital infrastructure of algorithmic governance, this research aspires to achieve an original contribution to the philosophical debates on digital governance. In this sense, the result of this philosophical research is not the expression of dogmatic theses, but the outcome of a meticulous philosophical inquiry.

Discussions on the Public Sphere

The historically qualified concept of the public sphere has continuously differentiated and transformed up to the present day. In Ancient Greece, politics was conducted in the common area, the public sphere, and the "polis" was this very public sphere where relationships were based on the interactions of equal citizens. These relationships emerged on the basis of freedom and equality, for the citizen (polites) was a member of the public sphere. In this period, public life was lived in agoras, which consisted of the assembly, the court (lexis), and common actions (praxis). The two activities constituting political life were "speech" and "action." The life of the citizen in the public sphere was the "good life", and to ensure this good life, the necessities of life had to be met in the private sphere; thus the private sphere was a precondition for ensuring the freedom of citizens in public life.

Aristotle approached the public sphere in the context of citizens engaging in politics collectively and its relationship with the private sphere. 17th and 18th-century philosophers like Locke, Hobbes, and Rousseau discussed the public sphere in the context of civil society, yet after these periods the distinction between public and social concepts became blurred and civil society was identified with political power. Kant defined the public sphere as the area formed by public opinion, while in the subsequent period civil society underwent a change to include family and economy. Marx used civil society in this sense and addressed its dialectical relationship with the state, whereas in the 20th century Gramsci examined the cultural and ideological aspects of civil society contributing to the current usage of this concept.

Arendt questions human life by sharply drawing the boundaries of the public and private spheres; the public sphere is the realm of political activities while the private sphere is that of non-political activities. Arendt attempts to remind humans of the experience related to the life of the "polis", suggesting that the public sphere can be seen as the place where differences emerge, meaning the subjects constituting the public sphere exist with their differences. According to Arendt's model wherever citizens have the opportunity to speak and act they possess the possibility to create a public sphere, a possibility that exists where individuals are in harmony and where there is equality and participation. In spaces devoid of harmony, trust, and participation, we cannot speak of a public sphere. For Arendt, the public sphere—understood as the space of discussion, speech, and action—constitutes the realm of politics, since it is through speech that individuals become political actors (2023). In Arendt's public sphere model humans exist through their actions, and the understanding of doing politics based on differences—that is, justifications based on race, ethnic structure, or religion—is not valid here. The individual exists here on their own behalf and gains their identity in this sphere. In the public sphere the citizen freely defends their thoughts depending on their power of representation. In this sphere where equality prevails people come face to face having the opportunity to convey their thoughts to one another and to evaluate a problem from different angles.

The private sphere, on the other hand, possesses the quality of secrecy/privacy. The way to escape and hide from the transparency and openness of publicity is to own private property. Secrecy/privacy expresses the distinction between the private sphere and the public sphere, or in other words, the distinction between what is to be hidden and what is to be shown publicly. The condition of slaves and women determined the existence of secrecy/privacy in the private sphere. Arendt (2023) notes that individuals such as laborers, who were bound by biological necessities, and women, who ensured the physical survival of the species, were historically relegated to the hidden realms of the private sphere.

The situation that most clearly expresses the distinction between the private and public spheres is the existence of things that need to be hidden versus things that need to be displayed openly. The private sphere is closed to human knowledge and in one aspect keeps the hidden secret, harboring the realm of birth and death. This is primarily because the fundamental transitions of human existence, specifically the origins and ends of life, remain beyond the scope of certain knowledge (Arendt, 2023).

While grounding her model, Arendt emphasizes that “labor” and “work” activities, which she approaches dialectically, belong to the private sphere. Every human activity expresses the place it is located; labor and work belong to the private sphere while action belongs to the public sphere. Labor expresses activities aimed at meeting biological needs to sustain one's life, and according to this model labor is a concept distant from the public sphere and can be evaluated as a concept of the private sphere. Work expresses the world of needed objects. These two concepts belonged to the world of women, slaves, and merchants who continued their lives and occupations within the boundaries of the private sphere in the Ancient Greek period. In the private sphere secrecy and privacy were in question, and the citizen who was the head of the household applied force and violence in this sphere depending on secrecy.

We see that Arendt places the “polis,” which existed in Ancient Greece and expressed the public sphere in another way, at the foundation of the public sphere model she developed. For Arendt the “polis” is a historical fact, and while re-evaluating the “polis” that emerged in a specific period in Ancient Greece she was aware that it could not be repeated in today's world. The importance of the “polis” for Arendt lies in it being an organization that allows everything seen as a problem to be discussed by citizens and highlights solidarity rather than violence. For this reason, she places the “polis” at the center of her theory. According to Arendt (2023), the essence of the polis resides not in its physical or geographical boundaries, but in the collective existence and interaction of its citizens.

Speech in the political sphere is a speech with a special quality, not just any speech. Speech in this sphere is conscious speech based on knowledge regarding society. In Ancient Greece, the political sphere was the place where citizens' common words and deeds were shared. The polis is conceptualized by Arendt (2023) as a dynamic organizational structure that emerges whenever individuals engage in collective action and shared discourse. Action, speech, and freedom are the powers that complete the public sphere; if one of these is missing one cannot speak of a public sphere. Man can achieve immortality through his life in the public sphere. For Arendt, freedom is the most valuable quality of human life, and the public sphere must be recreated for freedom to be lived and human equality to be achieved. Participatory democracy and the concept of the active citizen are at the center of the understanding developed by Arendt.

Whether one is talking about the disappearance of public space or merely the loss of his own right to the future tense, in essence, this is how power operates today. Yes, as Zuboff (2019) says, the massive collection of human behavior data provides a new type of power that works underground, avoiding human awareness and effectively eliminating uncertainty. These methods, when utilized by the digital state, are nothing like Habermas's ideal of a public sphere where rational debate took place. Instead, it is transformed into a 'behaviorist mold ' regime. Here things change. The state no longer gathers information about its citizens to manipulate them. It forces them to make decisions by influencing their choices through digital signals. All of this upends Arendt’s understanding of politics and freedom as a “public life” understood as “activity and surprise.” A new, technological order establishes a new form of rationality that, as long as it remains imperceptible and concealed behind technology, can avert and eliminate

the dynamics and unpredictability of the public life of citizens, by turning their capacity for action into administrative efficiency. Wanting to remain private is therefore not the central issue at stake in this respect; rather, what is at stake is the defense of our freedom to oppose those anonymous systems that are devouring our democratic reality and thereby reducing politics to a purely ordinary affair.

One of the criticisms of Arendt's theory of the public sphere is that it portrays political activity in an unrealistic light. In her effort to portray political activity as being central to the everyday lives of citizens and to show that political activity is carried out in a space where citizens are seen as equals, she neglects to consider the structural inequalities and exclusions that actually exist in modern public spaces. While her framework remains normatively powerful, it provides limited analytical tools for understanding mediated and asymmetrical forms of publicity.

The public sphere defined by Habermas is the liberal bourgeois public sphere. He questions the rise and fall of this sphere which is specific within its historicity and limited to a certain period. Habermas's idea of the public sphere is an area at the level of legal rights where state intervention and market relations are absent. Habermas paints a picture of publicity separate from the state and society; this public sphere expresses an area separate from the political relations of family, market, and state.

Habermas's concept of the public sphere leads to the questioning of a category belonging to bourgeois society, and in this form, he is in an effort to reconstruct the public sphere as a historical category. The public sphere is an area open to everyone without the limitations of any social environment. According to Habermas, the public sphere is a specific area located within civil society. In this area, the citizen is located in a specific sphere open to freely explaining and publishing their thoughts, using the right to organize, and communicating with each other about problems concerning everyone. In the public sphere, citizens have the opportunity to discuss equally, rationally, and critically. According to Habermas, this form of discussion is the basis of democracy, and through discussion, the neutralization of social power is in question.

Habermas evaluates the category of the bourgeois public sphere as specific to a certain period. The establishment of institutional relations between parliament, press, and political parties is an indicator that the public acquired political functions in the 18th century. However, the nature of the public's political function must be evaluated in connection with the development history of the bourgeois class or society. In other words, the political function of the public can only be understood holistically by examining the historical phase in which commodity exchange and social labor achieved independence from state authority (Habermas, 2021). This structural transformation becomes particularly visible in contemporary digital platforms, where publicity is no longer mediated by deliberation but increasingly shaped by algorithmic visibility.

The public community is divided into a minority producing reason in a non-public form and a mass formed by consumers. In the name of making economic understanding and thought dominant, a literary, political, and consumer public was created, and the publicity of communication tools was destroyed. Habermas (2021) argues that the reality constructed by mass media maintains a public character only in superficial terms. Mass media has emptied the privacy of the private sphere by making it transparent.

According to Habermas, it is possible to read the structures of social integration in public spheres. The political publicity of the institution of democracy has a meaning for the integration of society. Habermas (2019) suggests that in modern societies where citizens lack personal

familiarity, the establishment and continuity of a fragile partnership depend on the collective process of forming public opinion and will. In this sense, the state of a democracy can be understood by listening to the pulse of public opinion politically.

According to Habermas's understanding, the function of the public sphere is communication. Today, the public sphere directed by mass media such as the media and advertising sector will reveal its critical political function based on reason in the direction of rationalizing communication. The aim in rationalization should not only be the rationalization of political power but beyond that, it should respond to the practice of life through scientific and technical developments. The relationship between technique and democracy should be questioned, and the possibilities of using technique should be arranged in a beneficial way for people. Communication in the public sphere is an activity showing that people have become aware of themselves as social subjects. The conditions providing communication and the public sphere are the same; these conditions are truth, reality, honesty, and intelligibility. These conditions are also the conditions of discourse ethics. Discourse ethics aims to solve political and legal problems fairly.

Habermas identifies the “social state” model belonging to capitalism as the structural transformation of publicity. In this model, mass media have developed sovereignty over publicity. Public and private spheres have intertwined. Here, organizations not only maintain their sovereignty in goods and the market; they also possess political functions in social environments. This process leads to the re-feudalization of the public sphere. Today, the boundaries between the public and private spheres have deteriorated in favor of mass culture. Mass media directs what is public. According to Habermas, fundamental rights are the guarantee of political publicity. In the public sphere, communication is important, and Habermas's understanding of communication expresses an interaction where consensus regarding the developed action plans of the participants is achieved. Consensus is achieved through rational discussion. Rational action is revived on the basis of communication in the public sphere.

The public sphere model developed by Habermas is open to criticism in certain aspects today. Although the public sphere seems open to everyone in the thinker's model, in practice it is open only to citizens who are “mature subjects”. The ambiguity of “everyone” carries within itself. Workers, women, and those from different ethnic groups do not take place in the public sphere because they are not mature subjects. In the said model, inequalities are bracketed. In this context, it is necessary to reconsider the concept of the citizen and think about whether a public understanding that includes differences should be developed. In the model developed by Habermas, the public sphere is accepted as an area separate from social environments. Habermas's understanding of the public sphere provides the possibility of a critical questioning of developments in the media and political world today.

The Digital Sphere As A New Public Sphere

In the context of the digital state, the concept of a 'limit' refers not to a geographical line of sovereignty but to a multi-layered threshold with legal, political, and ethical dimensions. The legal limit marks where the state's authority to collect data ends under legitimate laws (such as KVKK, GDPR); the political limit defines where the state's security reflex must be halted to prevent the violation of individual liberties; and the ethical limit represents the inviolable domain where an individual's digital footprint must be deemed 'private life'. In the neoliberal era, the state tends to continuously expand these limits to the detriment of the individual often under the guise of 'security' and 'service efficiency' thus the 'limit' problematized in this study

is not the limit of the state's technological capacity but rather the limit of its democratic legitimacy.

Just as Habermas discussed newspapers as a tool for the formation of the bourgeois public sphere in modern society in the 19th century, it is now in question that the internet, in a world rapidly individualizing under the influence of neo-liberalism, is revitalizing a global public sphere. Habermas's concept of publicity is brought to the agenda again with the spread of the internet, which is the resultant of new information and communication technologies, to society. The digital sphere is seen as the third stage of the public sphere coming after the newspaper and radio-TV. The thought regarding the internet creating a public sphere or spheres stems from the technological potential of the internet. The technological potential of the internet, characterized as the "network of networks" based on connecting computer networks to each other, has led to the revival of the positive relationship established between information and democratization.

The internet is a digital sphere where everyone with access can express themselves freely through different tools (e-mail, chat programs, forums, websites, etc.); it provides communication in every direction from person to person, person to group, group to group, and group to persons allowing for the global and simultaneous sharing of information. The abundance of information, the multiplicity of options, and the possibility of interaction on the internet have reactivated dreams of "participatory democracy", "direct democracy", and the "agora".

The conceptualization of the digital sphere as a new public sphere requires a move beyond the assumption of technological neutrality. What Hildebrandt noted back (2018) still holds true – digital spaces now take shape through systems designed to act ahead of thought. These platforms block out the space where we can weigh competing ideas and perspectives against one another and they thereby undermine the notion that through the process of argumentation, ideas can be proven or strengthened. Through their unexamined design, social media undermines the ideals of deliberative democracy. We do not discuss or deliberate about the architecture of social media sites and therefore the architecture of these sites defines the nature of our deliberative practices. The social media space is not a democratic agora, a civic space, for the purpose of discussion and deliberation about politics. Rather it is a technological design that acts as a filter to our social and political discourse prior to that discourse ever taking place. When the digital state integrates its functions into these infrastructures, the 'limit' of state authority becomes invisible and automated. This transformation suggests that we are moving from a public sphere of discourse to a public sphere of 'management,' where the individual's role is no longer to participate in a rational-critical debate, but to be efficiently processed by algorithmic systems. Therefore, the digital public sphere must be re-evaluated not just as a tool for communication, but as a site of constitutional struggle where individual agency is at risk of being designed out of the system.

However, this dream of an agora was shattered by the realization that the digital sphere is not a neutral field for public deliberation but rather an algorithmically designed echo chamber (Maslej et al., 2025). Large-scale empirical mappings of AI systems demonstrate that engagement-optimizing algorithms systematically privilege affective polarization over rational-critical debate, thereby structurally undermining the Habermasian ideal of communicative rationality (Maslej et al., 2025). The rational-critical discussion ground idealized by Habermas (2021) has been replaced on digital platforms by a design aimed at fueling users' emotional reactions to generate more interaction (and thus, more data). In The Transparency Society, Byung-Chul Han contends that the digital age is not suffering from a communication crisis, but

from a crisis of privacy: our information is exposed and we are forced to provide explanations for our actions and personal details (Han, 2017). While the individual believes they are being liberated in the digital sphere, they are in fact imprisoning themselves in a digital panopticon through voluntary disclosure (Han, 2017). In this regard, the digital sphere has transformed into a market of consensual surveillance rather than a public square of freedom.

According to Timisi, authors emphasizing the internet-democracy relationship support their views around five main themes: easy access and sufficiency in information access, the internet's organization immune from authorities, unlimited freedom of expression on the internet and the technical difficulty of content control, increased political participation, and the expansion and globalization of civil society (Timisi, 2003). Revisiting the democratic potential of the internet, Fuchs (2023) suggests that the digital public sphere is currently characterized by a struggle between liberating deliberative spaces and institutionalized surveillance mechanisms. It is argued by liberal theorists that the internet, thought to lead to many positive transformations in society, will democratize society and apply direct democracy because it provides individuals with the information they want by overcoming national borders and time problems in the political sphere with its structure that multiplies information and is based on interaction, facilitates citizens' participation in politics "technically", creates a discussion environment through virtual forums, gives dissident voices and alternative groups in society the opportunity to communicate and express themselves at a much lower cost than dominant communication tools under capital control, and facilitates the organization of civil society.

The internet owes its designation as a free space from its early days to being relatively independent of market and state control compared to other communication environments. Yet this relative freedom has been narrowing for some time due to the coercion of both states and the market. In the first stage, to ensure a safe shopping environment suitable for the needs of the market, the internet environment began to be taken under control for secure money transfer and the prevention of internet piracy.

British sociologist Antony Giddens argues that the relationship between information and society precedes modern internet technologies; instead, he maintains that the origins of the nation-state are fundamentally rooted in information collection, suggesting the modern state has inherently been an information society from its inception (Giddens, 1985). Furthermore, Giddens contends that classical thinkers like Marx, Weber, and Durkheim overlooked the pivotal roles of violence and surveillance in the formation of modern society (Giddens, 1985). A highly organized nation-state must constantly collect information about citizens and share this data with citizens again to maintain its existence, and in this way, society will be able to develop a social self-perception by knowing itself (Webster, 2021). On the other hand, the state itself needs information about citizens incessantly to understand the limits of its sovereignty, collect taxes, and distribute resources, which requires "effective surveillance" (Webster, 2021). The need for surveillance is not only to ensure institutional continuity over a large population; another important feature of these structures is that a significant part of nation-states were established in wartime and developed based on defense needs, meaning war is one of the basic existence dynamics of the nation-state (Webster, 2021). Therefore, surveillance goes beyond an activity needed to complete its functions; nation-states try to collect information about the "enemy" incessantly citing threats that may come from outside.

Today, states' activities of collecting information about citizens are becoming increasingly complex. Collecting information with the justification of a threat is not only an activity directed outwards but also a situation concerning the nation-state's own citizens. It has been observed that activists, union leaders, and media workers can also be included in the scope

of security-justified surveillance (Webster, 2021). While some of this collected information is kept constantly secret, some of it—divorce rates, crime rates, education levels, etc.-is presented to citizens through the media. Thanks to this information citizens have knowledge about their own lifestyles and how to plan their futures, yet this informing allows states to manipulate citizens as much as inform them. Modern capitalist state relations carry such a complex nature that it is not possible for individuals to collect high rates of information without state-supported organizations. Politicians and bureaucrats holding state authority possess so much control over information that all modern nation-states face "danger of totalitarianism" (Webster, 2021).

Nevertheless, this surveillance mechanism established by the state and the market today is no longer limited to merely 'watching what is' (surveillance); the primary goal is predicting what will be (prediction). Recent systematic reviews indicate that predictive policing technologies not only reproduce existing social biases but also expand the discretionary power of the state by translating probabilistic risk assessments into pre-emptive interventions (Lee et al., 2024). In the digital age, the role of the state has evolved from simply keeping data as an archive record to a model of predictive policing that forecasts future risks, crimes, and social mobilities through artificial intelligence and big data analytics (Lee et al., 2024). In this new regime which Shoshana Zuboff defines as "Surveillance Capitalism", not only the individual's current actions but also their future "behavioral surplus" are processed as raw material and converted into a commercial or political prediction product (Zuboff, 2019). The fact that we are being violated privacy means no longer that the intimate is being revealed but that our preferences are being formed before we are aware of it. Information about our preferences is already at the disposal of algorithms before we decide. This is fundamentally different from the classic concept of surveillance. Tore Velde of the Swedish Defence Research Agency has labeled this phenomenon "anticipatory governance", i.e. a form of state governance that is aimed at citizens' future actions (Tallberg et al., 2023).

Today, with the significant contribution of internet-based technologies, the number of people who can be kept under surveillance can be quite high. Many internet-based technologies can go beyond entertainment or social relationship purposes and turn into tools for keeping users under surveillance. Some website can detect faces in photos uploaded to the page, compare them with thousands, perhaps hundreds of thousands of photos on the internet, and find other photos where the face in the photo appears. According to Morozov (2011), such technologies have been exploited by oppressive regimes, such as in Iran, to identify and target political activists. Similarly, the Indian government has utilized digital tracking to monitor Facebook users in the Kashmir region, employing automated alerts to summon suspected individuals to police stations (Morozov, 2011). With developing surveillance technologies, totalitarian tendencies are also emerging in democracies. With developing surveillance technologies, totalitarian tendencies are also emerging in democracies. According to Jean Claude Paye, legal regulations in America and England pave the way for e-mails to be inspected without any permission, and "general supervision has replaced punishing certain behaviors" (Paye, 2004, p.153). According to Paye, legal regulations paving the way for internet surveillance allow contents on the internet to be monitored before a criminal act occurs by emphasizing the terror threat and creating a "proactive" surveillance system with police logic (Paye, 2004). Although states' tendency to use the internet as a surveillance tool is a structural situation regarding the emergence of nation-states, the legal ground of this surveillance tendency has increased significantly especially with the terror attacks on New York and Washington on September 11, 2001.

The Privacy Problematic

It is possible to address privacy in different dimensions. Privacy as secrecy is the individual's desire to keep information about themselves and certain actions secret from everyone or some people they choose. Anonymous privacy is the individual's desire to perform certain behaviors (even if done in society) in a way that cannot be associated with the individual, that is, the desire to be anonymous. Autonomy privacy is the individual's preference for some of their behaviors to remain public or secret to depend on themselves (Teh, 2002). Susser further argues that in the digital age, this autonomy is under threat from "algorithmic manipulation," where invisible data architectures diminish the individual's capacity for self-governance (Susser et al., 2019).

The right to privacy against the state differs from the right to privacy against other people in two respects. First, although there are situations of violating another's privacy areas without their consent through theft, etc., among people, their relationships with each other are generally voluntary. The individual has control over the personal information they will disclose to another and possesses the power and measures to protect their privacy against other people. However, the state can perform forced interventions against individual privacy in a very wide area. The second difference is that the state has sovereignty based on applying physical force over citizens. It is possible for individuals to take some measures to protect their own privacy against third parties (they can protect themselves against other citizens, take measures like locks and burglar alarms); however, they are very powerless when it comes to protecting themselves and their personal information against the state. The only way they can resort to protect their personal information from the state is to try to hide it (Flaherty, 2000). As Tanczer (2020) observes, cybersecurity in the digital state must be redefined as a foundational pillar of the social contract; without it, the state's intervention in personal domains loses its democratic legitimacy. The effects of this difference on privacy exhibit changes according to the point of view. If the state is thought to be well-intentioned, wanting the good of its citizens, and observing the public interest, then personal privacy will make it difficult for public officials to do what is good. However, on the other hand, if it is thought that the state can behave against citizens, in that case, personal privacy against the state and its protection is definitely a good thing. Although it is accepted that privacy can be used in hiding personal crimes, the view that it should be protected against the state is quite widely accepted.

Uslaner (2004b), in a study where he also addressed the relationship between trust in the internet and privacy anxiety, argued that individuals who are generally prone to trust have low levels of concern regarding privacy in transactions they carry out over the internet, whereas those with low propensity to trust have higher concerns regarding privacy and security. While trust tendency is effective on privacy perception, privacy perception is also an effective concept on trust towards e-government applications, for problems and concerns regarding privacy are among the critical obstacles to the implementation of e-government (Cullen & Reilly, 2007). While trust is vital, Lember ve Crompvoets (2023) notes that the 'black-box' nature of modern public service algorithms creates a legitimacy crisis that technical security alone cannot solve, as citizens feel alienated from the decision-making processes. Ensuring information security at the highest possible degree is of vital importance for information and communication technologies. Information security relates to the ability of personally identifiable information to be transmitted and disclosed over the internet and insecure networks and is the key to privacy, which is one of the most important elements of the trust environment that has become indispensable for the information society. The requirements of e-government systems to be established to ensure information security and reduce privacy anxiety as much as possible should be determined while the system is still being designed (Kalloniatis et al., 2004). In this

respect, e-government applications should not only offer better service opportunities for public institutions but also provide citizens with the guarantee of holding data ownership. Citizens do not want their personal information shared with third parties. The amount of information in the hands of the state and the sensitivity of this information is an important issue in terms of security, and the theft of citizens' data by hackers or access to such information by third parties must be prevented.

This security threat is not just a technical cybersecurity issue, but also has to do with the legal limbo that stems from combining "hybrid" relations between the state and tech behemoths. With states sharing data more and more with the tech giants in order to perform public services (smart cities, digital health etc.), it blurs the line of who actually owns and has sovereignty over public data. In Turkey, for example, the Law on the Protection of Personal Data (No. 6698, 2016), commonly known as KVKK, and GDPR in Europe became "digital shields" to protect ethical cookery as individual against this data tsunami (Tanışık & Bal, 2024). Véliz (2021) goes on to claim that privacy is not only a personal preference, but also a collective power essential for the survival of democracy, updating Arendt's insistence upon political freedom in the digital age. But such legal writings may still prove to be insufficient to fully secure the privacy of the individual from the state insofar as differences in room for manoeuvre are much greater when it comes to data that are processed on public security or other grounds based on national interest than data from the private sector. At this juncture law is slow to move and data colonialism indemnifies itself with legal gaps.

These legal boundaries are drawn with varying approaches on a global scale. For instance, while the European Union (EU), through the General Data Protection Regulation (European Parliament & Council of the European Union, 2016), treats privacy as a fundamental human right—effectively telling both the state and the market to 'back off' when it comes to personal data—the United States (US) plays a different game altogether. There, you see a more market-driven, sectoral pragmatism at work. Especially in that foggy, post-9/11 atmosphere where 'national security' became the ultimate trump card, the boundaries of state digital surveillance were pushed to an almost unrecognizable extent under the USA PATRIOT Act (Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism Act of 2001). Turkey, on the other hand, appears to have adopted the EU model (Directive 95/46/EC) with its Law No. 6698 (KVKK); however, due to broad exceptions such as 'state security and public order' in Article 28, it exhibits a hybrid model that leaves a wider margin of maneuver for state authority in practice (KVKK, 2016). This situation causes individual privacy in Turkey's digital state applications to remain in a more fragile position against the concept of public interest.

In fact, privacy is one of the most emphasized subjects in computer and information ethics. Privacy is directly related to information and access to information, and the problem must be handled through the relationship between personal information and personal identity. Like all types of information, personal information should be respected, and privacy should be ensured in this area. Moving from the fact that humanity is constantly developing, protecting the individual's information privacy also means recognizing their freedom to develop (Stahl, 2008). Tallberg et al., (2023) notes that the lack of algorithmic transparency in today's digital ecosystem makes Habermasian rational deliberation impossible. In other words, privacy is also one of the key concepts regarding development. Uslaner (2004a) argues that trust in the internet and relationships established over the internet can be structured not over strategic trust, as he defines it, but rather over moral trust. There are many examples where trust in the internet yields good results and good intentions are abused, yet the way to make the progress of a technological development like the internet possible passes through trusting humanity rather than strategic

trust. However, in internet-based applications, information moves much faster than controllable limits.

In Data Sovereignty and Structural Privacy Threats, Lember and Crompvoets (2023) argue that the lack of control over personal data (which they call data sovereignty in relation to citizens) experienced by citizens due to the automation of data processing systems constitutes a legitimacy crisis. They further argue that this lack of control over personal data, which we experience acutely especially after learning about the weaknesses of public infrastructure in the age of digitalisation, is what we call data sovereignty of citizens, a term first coined by Cullen and Reilly (2007). For instance, after all state institutions went online through e-government portals, it became apparent that the state's information is very exposed to potential internal leaks or third party cyber attacks (STM Defense Technologies Engineering and Trade Inc.,2025). Maalla (2021) highlights the lack of transparency in the use of AI within the public administration system, which may convert the administration system into an instrument of social control, turning public administration into an administrative system that does not provide any positive public service to citizens. Hence, this feeling of being in the data flow without control constitutes a structural privacy threat which is not something that the individuals or society can do anything about. Rather, it is a constraint that needs to be overcome for the modernisation of the public administration through the digitalisation of the state and for the citizens to adapt to the new digital government tools. As shown in the studies of social media users in Turkey (Turan, 2024), citizens' attitudes towards social media and privacy and trust issues in this regard may change over time as citizens become more aware of being under surveillance. As Margetts and Dorobantu (2019) argue in relation to the digital state, citizens cannot be expected to trust the digital state by merely installing firewalls and using encryption. Complex accountability mechanisms need to be developed to account for the AI systems that monitor citizens.

At this very moment, our privacy is no longer a constitutional right that needs to be defended, but rather a constitutional barricade which signals the boundaries of a surveillance state in the digital age. As far as Habermas is concerned, the digital world should be the "liberum veto" of our time – meaning free and unencumbered spaces for discussion. Privacy is not an entry ticket for these spaces. Rather it is a precondition for having a digital space of communication that is free of surveillance. When an individual calls on the state to implement state-wide end-to-end-encryption through means of cryptography, anonymity and data sovereignty, in a state where the citizen is continuously under surveillance by highly efficient algorithm-based crime prevention technology – then this call is, in the modern age, characterised by the same features that Arendt identified in civil disobedience as a "space of resistance". Therefore, the normative solution in digital democracies lies in constitutionally guaranteeing privacy not as a favor granted by the state but as a founding and untouchable 'negative liberty' that limits state power. Otherwise, the digital state carries the risk of transforming not into a servant for its citizens but into a warden who records every move.

Wide powers given to governments using terms such as national security or public order can quietly erode personal privacy. In real situations, those justifications tend to act like open permission slips for gathering private information. Without outside review or regular checks, officials may stretch their reach far beyond what laws intend. Over time, the lines meant to contain digital surveillance grow fuzzy, almost meaningless. Slowly, a space forms where rules about data access are bent, ignored, reshaped – simply because they can be.

Nowadays governments lean heavily on hidden algorithmic tools when making choices. Not even lawyers or ordinary people get to peek inside how those tools reach conclusions.

Hildebrandt pointed out back in 2018 that not being able to follow the reasoning shuts down fairness. When a machine tags your name without explanation, fighting its call becomes pointless. When power goes unchecked, it weakens what keeps democracies honest. Laws alone cannot fix this gap. To stay an active participant in politics, not just raw material for tracking, people need stronger moral limits on government access – limits shaped through ongoing public argument. If that tension fades, mass monitoring by authorities could settle into everyday life without challenge.

Conclusion

That privacy is a human right is a universally accepted premise. Individuals want to be sure that information regarding themselves will be processed and treated fairly, but the state's collection, storage, and management of personal data regarding citizens render the issue of privacy increasingly critical. Indeed, empirical findings demonstrate that there are numerous cases where personal information stored in public institutions could not be protected, was abused, or was leaked externally during the e-government process which is still in its developmental stage. This situation erodes individuals' sense of trust in e-government and negatively affects their behavioral patterns and sentiments. Therefore, problems in the storage and use of personal information collected in public institutions must be identified, and procedural measures and specialized institutions must be established to prevent damage to the right to privacy.

This study, conducted on the limits of the digital state and individual privacy, goes beyond these necessities to reveal that the tension between the state's role as a "service provider" and its "surveillance" nature in the neoliberal era has eroded classical public sphere theories. The fundamental theoretical conclusion reached by the study is that the digital sphere has transformed into a transparency regime in which visibility is no longer a condition of democratic accountability but a mechanism of normalization and control. In this regime, privacy erosion operates not merely as a technical side effect of digitalization but as a structural transformation of the relationship between the state and the citizen. This conclusion resonates with recent arguments emphasizing that democratic legitimacy in digital states depends not on technological efficiency, but on the preservation of epistemic autonomy and meaningful consent within data-driven governance structures.

In this context, the following policy recommendations are of vital importance for digital state applications to maintain their democratic legitimacy:

1. Data Minimization and Purpose Limitation: Public institutions must abandon the logic of "hoarding data in case it is needed one day"; data collection processes must be strictly limited to the minimum level required by the service provided.
2. Independent Audit and Transparency: The state's data processing activities must be audited by autonomous boards supported by civil society and academia, independent of the executive branch; algorithmic decision-making processes (black boxes) must be made transparent.
3. Privacy by Design: When establishing e-government systems, privacy must be constructed not as a patch added later but as the architectural foundation of the system.

How should we conduct research into privacy in the future? The answer is: by thinking about what the surveillance technologies of tomorrow could mean for society and thus what we need to do today to counterbalance them. The future of privacy research should explore, for

example, the ethics of the use of artificial intelligence in predictive policing or how blockchain technology could help give citizens control of their personal data. Privacy will be the civil rights battle of the digital age. Current literature on the effects of digitalisation generally describes privacy as a purely technological or security issue, while this article suggests that privacy should also be described as a political issue in the context of digitalisation. It uses the theories of Hannah Arendt and Jürgen Habermas on the public sphere as a theoretical basis and examines how privacy acts as a component of political subjectivity and the foundations of democratic politics in the digital age. In the context of the digital state in particular, a state that is increasingly governed through technology, it is no longer sufficient to look at whether a state has high-tech infrastructure or whether citizens have access to adequate online services. Instead the digital state must also be viewed in terms of ethics and politics. And the 'limit of the digital state' is not about reaching the limits of what is technically possible, but rather about the moment at which technology intrudes upon human existence as a human being, as a person, as something that is dignity-filled and therefore privacy-protected.

Privacy is about more than just technology. It is about politics and the political nature of the self. The possibility of a truly open digital life presupposes a fundamentally different understanding of the state and of state power. Thus, it is necessary to think about the technological means of realising epistemic autonomy — that is, the conditions under which one can constitute one's own knowledge and opinions independently of the structures of algorithmic processing or the statistical predictions of models. Stronger ethical and legal norms are needed in order to prevent the surveillance state from becoming the standard form of governance. Protecting private space is no secondary or third-order question of politics; it is rather a constitutional one, one that is fundamental to the very idea of democracy.

References

Arendt, H. (2023). *İnsanlık Durumu*, (B. S. Şener.Çev.). İletişim Yayınları.

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the New Jim Code*. Polity Press.

Cullen, R., & Reilly, P. (2007). Information privacy and trust in government: A citizen-based perspective from New Zealand. In *Proceedings of the 40th Annual Hawaii International Conference on System Sciences (HICSS'07)* (pp. 109–119). IEEE. https://doi.org/10.1109/HICSS.2007.271

European Parliament and the Council of the European Union. (2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation).* Official Journal of the European Union, L119, 1–88. Retrived January, 01, 2026, from https://eur-lex.europa.eu/eli/reg/2016/679/oj

Flaherty, D. (2000). Privacy impact assessments: An essential tool for data protection. *Privacy Law & Policy Reporter*, *7*(5), 85–90.

Fuchs, C. (2023). *Digital democracy and the digital public sphere*. Routledge.

Giddens, A. (1985). *The nation-state and violence: Volume two of a contemporary critique of historical materialism*. University of California Press.

Habermas, J. (2019). *Doğalcılık ve din arasında felsefi denemeler*, (A. Nalbant, Çev.). Yapı Kredi Yayınları.

Habermas, J. (2021). *Kamusallığın yapısal dönüşümü*. (T. Bora & M. Sancar, Çev.). İletişim Yayınları.

Han, B.-C. (2017). *Şeffaflık toplumu* (H. Barışcan, Çev.). Metis Yayınları

Hildebrandt, M. (2018). Algorithmic regulation and the rule of law. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376*(2128), Article 20170355. https://doi.org/10.1098/rsta.2017.0355

Kalloniatis, C., Kavakli, E., & Gritzalis, S. (2004). Security requirements engineering for e-government applications: Analysis of current frameworks. In R. Traunmüller (Ed.), *Electronic government (EGOV 2004)* (*Lecture Notes in Computer Science*, Vol. 3183, pp. 66–71). Springer. https://doi.org/10.1007/978-3-540-30078-6_11

Lee, Y., Bradford, B., & Posch, K. (2024). The effectiveness of big data-driven predictive policing: Systematic review. *Justice Evaluation Journal*, *7*(2), 127–160. https://doi.org/10.1080/24751979.2024.2371781

Lember, V., & Crompvoets, J. (2023). Public services and management in the digital age. In T. Bovaird & E. Loeffler (Eds.), *Public management and governance* (4th ed., pp. 141–152). Routledge. https://doi.org/10.4324/9781003282839-13

Maalla, H. A. (2021). Artificial intelligence in public sector: A review for government leaders about AI integration into government administrations. *International Journal of Academic Research in Economics and Management Sciences*, *10*(4), 27–39. https://doi.org/10.6007/IJAREMS/v10-i4/11911

Margetts, H., & Dorobantu, C. (2019). Rethink government with AI. *Nature*, 568 (7751), 163-165. https://doi.org/10.1038/d41586-019-01099-5

Maslej, N., Fattorini, L., Perrault, R., Gil, Y., Parli, V., Kariuki, N., Capstick, E., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Wald, R., Walsh, T., Hamrah, A., Santarlasci, L., … Oak, S. (2025). *The AI Index 2025 annual report*. Stanford University, Institute for Human-Centered AI. https://doi.org/10.48550/arXiv.2504.07139

Morozov, E. (2011). *The net delusion: Dark side of internet freedom*. PublicAffairs.

Paye, J. C. (2004). Terörle mücadele ve özel hayatın denetimi. *Conatus Çeviri Dergisi*, *1*(1), 153–165.

Resmi Gazete. (2016, 7 Nisan). 6698 sayılı *Kişisel Verilerin Korunması Kanunu (KVKK)*

Stahl, B. C. (2008). Discourses on information ethics: The claim to universality. *Ethics and Information Technology*, *10*(2-3), 97–108. https://doi.org/10.1007/s10676-008-9171-9

STM Savunma Teknolojileri Mühendislik ve Ticaret A.Ş. (2025, Ocak-Mart*). Siber tehdit durum raporu*. STM. https://thinktech.stm.com.tr/tr/siber-tehdit-durum-raporu-ocak-mart-2025

Susser, D., Roessler, B., & Nissenbaum, H. F. (2019). Online manipulation: Hidden influences in a digital world. *Georgetown Law Technology Review*, *4*(1), 1–45. https://doi.org/10.2139/ssrn.3306006

Tallberg, J., Erman, E., Furendal, M., Geith, J., Klamberg, M., & Lundgren, M. (2023). The global governance of artificial intelligence: Next steps for empirical and normative research. *International Studies Review*, *25*(3), Article viad040. https://doi.org/10.1093/isr/viad040

Tanczer, L. M. (2020). 50 shades of hacking: How IT and cybersecurity industry actors perceive good, bad, and former hackers. *Contemporary Security Policy*, *41*(1), 108–128. https://doi.org/10.1080/13523260.2019.1669336

Tanışık, S., & Bal, S. (2024). Dijital mahremiyet ve kurumsal sorumluluk: Kişisel verilerin korunmasında iletişim teknolojilerinin kamusal rolü. *Yeni Medya*, (16), 268–285. https://doi.org/10.55609/yenimedya.1424182

Teh, J. (2002). Privacy wars in cyberspace: An examination of the legal and business tensions in information privacy. *Yale Journal of Law & Technology*, 4, 1–96. https://yjolt.org/volume/4

Timisi, N. (2003). *Yeni iletişim teknolojileri ve demokrasi*. Dost Kitabevi Yayınları.

Turan, Y. (2024). Sosyal medyada gözetim ve mahremiyet algısı: Instagram kullanıcıları üzerine bir inceleme. *Akdeniz Üniversitesi İletişim Fakültesi Dergisi*, (47), 146–169. https://doi.org/10.31123/akil.1427283

USA PATRIOT Act, Pub. L. No. 107–56, 115 Stat. 272 (2001). https://www.congress.gov/107/plaws/publ56/PLAW-107publ56.pdf

Uslaner, E. M. (2004a). Trust online, trust offline. *Communications of the ACM*, *47*(4), 28–29. https://doi.org/10.1145/975817.975838

Uslaner, E. M. (2004b). Trust, civic engagement, and the internet. *Political Communication*, 21(2), 223–242. https://doi.org/10.1080/10584600490443895

Véliz, C. (2021). *Privacy is power: Why and how you should take back control of your data*. Penguin Random House.

Webster, F. (2021). *Theories of the information society* (4th ed.). Routledge. https://doi.org/10.4324/9781315867854

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.

Information About the Article/Makale Hakkında Bilgiler

The Ethical Rules for Research and Publication / Araştırma ve Yayın Etiği

The authors declared that the ethical rules for research and publication followed while preparing the article.

Yazarlar makale hazırlanırken araştırma ve yayın etiğine uyulduğunu beyan etmiştir.

Conflict of Interests/ Çıkar Çatışması

The authors have no conflict of interest to declare.

Yazarlar çıkar çatışması bildirmemiştir.

Grant Support/ Finansal Destek

The authors declared that this study has received no financial support.

Yazarlar bu çalışma için finansal destek almadığını beyan etmiştir.

Author Contributions/ Yazar Katkıları

The draft process of the manuscript/ Taslağın Hazırlanma Süreci E.Ç./M.D., Writing the Manuscript/ Makalenin Yazılması E.Ç./M.D., Submit, Revision and Resubmit Process/ Başvuru, Düzeltme ve Yeniden Başvuru Süreci M.D.

Journal of Applied and Theoretical Social Sciences

JATSS, 2026; 8(1), 189-212

First Submission:04.01.2026
Revised Submission After Review:05.03.2026
Accepted For Publication:24.03.2026
Available Online Since:25.03.2026

Research Article

The Power of Words for Credit Risk Classification: Sentiment Analysis As An Alternative Credit Rating Tool

Yunus Emre Akdoğan[a]

Abstract

Introduction: This study examines whether qualitative narratives in credit rating reports can be systematically structured using machine learning and to what extent indicators derived from these narratives carry information for predicting national credit ratings. The research is significant because it operationalizes narrative-based assessments, central to rating processes yet difficult to integrate into quantitative models, making them measurable and analytically tractable.

Method: This study analyzes publicly available institutional credit rating reports. Reports statements are first classified as strengths or constraints using a transformer-based model and then transformed into narrative indicators. Using only these features, national credit ratings are predicted via multi-class classification, with firm-level train–test splitting to avoid data leakage. Performance is assessed using accuracy, precision, recall, and F1-score.

Results or Findings: The strength–constraint classifier achieved high performance (accuracy = 0.977; F1 = 0.981). For rating prediction, accuracy was 0.517, with macro-F1 = 0.403 and weighted-F1 = 0.500. Class-level results were more stable for AA and BBB, but limited for AAA and BB.

Discussion or Conclusion: The findings indicate that indicators derived from report narratives carry a meaningful signal about credit ratings, yet this signal is not sufficient on its own. The proposed approach can be used for rapid summarization of newly released reports, tracking changes in the Strengths–Constraints balance, and supporting scalable credit surveillance and early-warning systems. Future research is encouraged to adopt approaches that integrate narrative indicators with financial and macroeconomic variables.

Keywords: credit rating reports, credit rating prediction, transformer-based classification, financial natural language processing

JEL Codes: G24, G17, C45, C55

[a] Assist. Prof. Dr., Yozgat Bozok University, Faculty of Economics and Administrative Sciences, Department of Business Administration, Yozgat/Türkiye, emre.akdogan@bozok.edu.tr, ORCID ID: https://orcid.org/0000-0002-1761-2869, (Corresponding Author)

Journal of Applied and Theoretical Social Sciences

JATSS, 2026; 8(1), 189-212

İlk Başvuru: 04.01.2026
Düzeltilmiş Makalenin Alınışı: 05.03.2026
Yayın İçin Kabul Tarihi: 24.03.2026
Online Yayın Tarihi: 25.03.2026

Araştırma Makalesi

Kredi Riski Sınıflandırmasında Kelimelerin Gücü: Alternatif Kredi Derecelendirme Aracı Olarak Sentiment Analizi

Yunus Emre Akdoğan[a]

Öz

Giriş: Bu çalışma, kredi derecelendirme raporlarında yer alan nitel anlatıların makine öğrenmesi ile sistematik biçimde yapılandırılıp yapılandırılamayacağını ve bu anlatıdan türetilen göstergelerin ulusal kredi notunu öngörmede ne ölçüde bilgi taşıdığını incelemektedir. Araştırma, derecelendirme süreçlerinde merkezi rol oynayan ancak çoğu zaman nicel modellere doğrudan entegre edilemeyen anlatı temelli değerlendirmeleri ölçülebilir ve analitik olarak işlenebilir hale getirmesi bakımından önemlidir.

Yöntem: Bu çalışma, kamuya açık kurumsal kredi derecelendirme raporlarını incelemektedir. Raporlardaki ifadeler transformer tabanlı bir modelle güçlü yönler ve kısıtlar olarak sınıflandırılarak anlatı göstergelerine dönüştürülmüştür. Yalnızca bu özellikler kullanılarak ulusal kredi notu çok sınıflı bir problem olarak tahmin edilmiş, veri sızıntısını önlemek için firma bazlı ayrım uygulanmış ve performans accuracy, precision, recall ve F1-skoru ile değerlendirilmiştir.

Sonuçlar ya da Bulgular: Güçlü yönler-kısıtlar sınıflandırması yüksek performans üretmiştir (accuracy=0,977; F1=0,981). Ulusal kredi notu tahmininde ise örneklem dışı doğruluk 0,517; macro-F1 0,403; weighted-F1 0,500 olarak elde edilmiştir. Sınıf bazında AA (F1=0,605) ve BBB (F1=0,591) sınıflarında görece daha istikrarlı sonuçlar görülürken, AAA'da düşük duyarlılık ve BB'de düşük destek nedeniyle zayıf performans dikkat çekmektedir.

Tartışma ya da Yapılan Çıkarımlar: Bulgular, rapor anlatılarından türetilen göstergelerin kredi notuna ilişkin anlamlı ancak tek başına yeterli olmayan bir sinyal taşıdığını göstermektedir. Önerilen yaklaşım, yeni yayımlanan raporların hızlı özetlenmesi, güçlü yönler-kısıtlar dengesindeki değişimlerin izlenmesi ve ölçeklenebilir kredi gözetimi/erken uyarı sistemleri için kullanılabilir. Gelecek araştırmaların, anlatı göstergelerini finansal ve makro değişkenlerle birleştirmesi gibi yöntemler önerilmektedir.

Anahtar Kelimeler: kredi derecelendirme raporları, kredi notu tahmini, transformer tabanlı sınıflandırma, finansal doğal dil işleme

JEL Kodlar: G24, G17, C45, C5

[a] Dr. Öğr. Üyesi, Yozgat Bozok Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İşletme Bölümü, Yozgat/Türkiye, emre.akdogan@bozok.edu.tr, ORCID ID: https://orcid.org/0000-0002-1761-2869, (Sorumlu Yazar)

Introduction

Accurate and timely measurement of credit risk is a fundamental requirement for financial institutions, investors, and regulators. In recent years, the credit risk literature has developed a broad set of "credit scoring/classification" techniques, ranging from traditional statistical approaches to machine learning methods (Baesens et al., 2003; De Oliveira & Basso, 2025; Pai et al., 2015; Siham et al., 2021; Teles et al., 2021; Thomas, 2000; Yobas et al., 2000). As the scale of credit markets has expanded, the fact that even small improvements in predictive accuracy can generate substantial economic gains through pricing, provisioning, capital allocation, and risk management has increased interest in benchmarking alternative algorithms. Indeed, numerous comparative studies show that modern machine learning approaches often outperform conventional baseline models across various credit scoring settings (Darwish, 2025; Lessmann et al., 2015; Machado et al., 2025; Machado & Karray, 2022; Marqués et al., 2012; Paleologo et al., 2010; Tripathi et al., 2019; Tsai, 2014; Xu et al., 2025).

However, a key element of real-world credit assessments, namely the qualitative narratives that accompany credit decisions, remains relatively underutilized in empirical modeling. In practice, credit risk decisions are communicated not only through a score or rating but also through textual explanations that summarize the rationale behind the assessment. Credit rating reports, in particular, often articulate the basis of the evaluation under headings such as *strengths* and *constraints*. Despite their central role in how stakeholders interpret credit risk, predictive models still predominantly rely on structured financial ratios and other tabular variables.

In corporate credit rating settings, this gap is particularly pronounced. Rating reports routinely articulate the main factors underlying the assigned grade as rating drivers, typically organized under sections resembling *strengths* and *constraints*. These narratives are a core part of how the rating decision is justified and communicated to stakeholders. However, turning such text into scalable, quantitative inputs is difficult because the content is unstructured, varies across reports, and relies on domain-specific phrasing. Recent advances in natural language processing, especially transformer-based architectures, make it increasingly feasible to parse these narratives systematically and derive measurable indicators from them. Even so, fully integrated studies that convert rating narratives into structured variables and evaluate their explanatory and predictive value for rating outcomes under leakage-robust designs remain relatively scarce.

This study focuses on the question: "Can qualitative descriptions in credit rating reports be systematically transformed into structured indicators through machine learning, and can these indicators provide statistically significant information in predicting national credit ratings?" In this context, the proposed two-stage framework aims to contribute to addressing this gap by systematically analyzing these narratives. In the first stage (Task 1), statements in rating reports are classified into *strengths* and *constraints* using a fine-tuned transformer-based classifier. In the second stage (Task 2), these outputs are aggregated at the report level to construct narrative-based indicators, such as the counts and balance of *strengths* and *constraints* as well as composite indices, and national credit rating classes are then predicted using a multi-class model based solely on these text-derived features. To reflect realistic deployment conditions and prevent information leakage, the train–test split is performed at the firm level, ensuring that reports from the same company do not appear in both the training and test sets.

The empirical findings draw a sharp distinction between a task that is linguistically learnable and a prediction task necessitating a more comprehensive information set.

Specifically, while the classification of *strengths* and *constraints* achieves near-perfect performance (Accuracy≈0.98; F1≈0.98), utilizing narrative-derived indicators alone for rating class prediction results in only moderate success (Accuracy≈0.52; macro−F1≈0.40), with significant performance drops observed in sparse or extreme classes. These results suggest that while rating narratives provide meaningful insights, they offer an incomplete signal regarding the final credit rating. Consequently, although text-based models excel at the scalable extraction and summarization of key credit drivers, they cannot fully replace quantitative fundamentals and the broader data environment used in the formal rating process.

This study contributes to the literature in three main ways. Methodologically, it employs machine learning to systematically transform qualitative credit rating report narratives into measurable indicators, thereby bridging the gap between narrative-based assessments and quantitative modeling, and establishing a reproducible and auditable feature space by modeling rating narratives at the granular level as *strengths* and *constraints* and synthesizing them into report-level indicators. From a data-driven perspective, the study integrates unstructured text data directly from rating reports into empirical analysis and predictive models, using a firm-level, leakage-robust evaluation framework to delineate the predictive boundaries of text-derived signals and identify conditions under which narrative-based modeling is insufficient. Empirically, by applying the framework to a national credit rating sample, the study tests the predictive power of sentiment-oriented indicators such as *strengths* and *constraints*, providing novel and concrete evidence, and offering a practical foundation for scalable credit monitoring through rapid processing of new reports and quantification of shifts in credit quality, while highlighting the importance of integrating complementary quantitative variables for robust early-warning analytics.

In the remainder of the paper, Section 2 reviews the literature on credit scoring and risk analytics. Section 3 describes the dataset, the models, the performance metrics used in the study and discusses the empirical findings. The final section concludes the paper by highlighting implications for credit monitoring and outlining directions for future research.

Literature Review

Credit risk assessment and rating/score prediction have long been central topics in finance and risk management, traditionally addressed with statistical classification models based on firm- and borrower-level quantitative indicators. Over the past two decades, rapid growth in computational power and data availability has shifted this literature toward machine learning approaches that can capture nonlinear relationships, complex interactions, and heterogeneous patterns in default and creditworthiness signals. Accordingly, a substantial body of research has benchmarked a wide range of algorithms, including logistic regression, decision trees, support vector machines, and neural networks, while a parallel stream has emphasized performance gains through ensemble learning, feature engineering, and robust evaluation under class imbalance and asymmetric error costs. More recently, the field has begun to extend beyond purely numerical predictors by incorporating unstructured text, such as disclosure narratives and analyst or rating-agency communications, to quantify qualitative information embedded in credit-related documents. Building on this broader trajectory, the following section reviews key methodological advances in credit scoring and related predictive frameworks, with particular attention to ensemble strategies and the emerging role of text-based signals.

In this context, ensemble learning methods have attracted particular attention for their potential to improve generalization performance and reduce classification errors. In their study,

Wang et al. (2011) compared the effectiveness of Bagging, Boosting, and Stacking ensemble methods using four base learners (Logistic Regression Analysis, Decision Tree, Artificial Neural Network, and Support Vector Machine) for credit scoring. The study revealed that the ensemble methods notably improved the performance of individual base learners. Specifically, Bagging outperformed Boosting across all credit datasets, and both Stacking and Bagging showed superior performance in terms of Decision Tree, average accuracy, Type I error, and Type II error.

Liu & Schumann (2005) conducted an empirical study to investigate how various algorithms, including "ReliefF", "Correlation-based", "Consistency-based", and "Wrapper", can enhance the performance of credit scoring models in terms of model simplicity, speed, and accuracy. They used real data sets and applied four classification algorithms: "model tree (M5)", "back-propagation multilayer perceptron", "logistic regression", and "k-nearest neighbor". The study findings revealed that the consistency-based and wrapper feature selection methods outperformed the other two methods. The learning curves of the consistency-based and wrapper methods displayed lower error rates in the early range (from 0 to approximately 33 features), indicating that their models achieved lower error rates earlier than the other two methods. Upon comparing the algorithms, it was noted that the k-nearest neighbor algorithm exhibited improvements in model accuracy, while no improvement was observed for the other algorithms. Consequently, the researchers concluded that reducing the number of features enhances classification accuracy, reduces training time, and simplifies the final models.

Teles et al. (2020) conducted a study to compare the effectiveness of fuzzy sets and neural network-based decision trees in credit scoring to estimate the recovered value using a dataset of 1890 borrowers. The study concluded that both models allow uncertainty modeling in credit scoring. Specifically, the study found that fuzzy logic is more adept at modeling uncertainty, while the decision tree model is better suited for addressing the problem at hand.

Trivedi (2020) utilized German credit data in their study to evaluate credit risk and focused on developing a credit scoring prediction model using artificial intelligence. The research included a comparative analysis of evaluation metrics to identify the most effective combination of machine learning classifier and feature selection technique. The study compared Bayesian, Naïve Bayes, SVM, C5.0, and Random Forest methods as machine learning classifiers, as well as Chi-Square, Information-gain, and Gain-Ratio as feature selection algorithms. Ultimately, the analysis concluded that the combination of Random Forest and Chi-Square is the most optimal pair for creating credit scoring models.

In their study, Dumitrescu et al. (2022) developed a penalized logistic tree regression (PLTR) model, which combines decision trees with logistic regression. The model was designed to address the lack of interpretability in many existing machine-learning algorithms used for credit scoring. Through Monte Carlo simulations and the application of the PLTR model to four real credit scoring datasets, they demonstrated its interpretability and predictive power. Their findings show that the PLTR effectively captures non-linear effects in credit scoring data while remaining interpretable. Empirical evidence indicates that the PLTR significantly improves credit risk prediction compared to logistic regression and competes closely with the random forest method. Furthermore, they assessed misclassification costs using expected maximum profit analysis and concluded that the PLTR method considerably reduces misclassification costs.

In a recent study, Bucker et al. (2022) sought to address the limitations of traditional credit scoring methods, which often result in higher reserves or increased loan defaults. Their

goal was to introduce a transparent and auditable machine learning model that not only delivers accurate risk estimation but also ensures understandability. The study emphasized the importance of transparency, auditability, and explainability in machine learning models for credit scoring. Interestingly, the researchers found that the baseline scorecard performed remarkably well compared to advanced machine learning techniques such as Gradient Boosting or Support Vector Machines. They also predicted that with the anticipated increase in the use of transactional and external data sources, credit scoring data will become more complex, underscoring the growing significance of attribute engineering.

In a study conducted by Gambacorta et al. (2024), the effectiveness of credit-scoring models based on machine learning techniques was compared to that of traditional loss and default models. The study utilized proprietary transaction-level data from a leading fintech company in China, spanning the period between May and September 2017. The performance of different models in predicting losses and defaults during both normal economic conditions and economic shocks was evaluated. Specifically, the study analyzed the impact of a regulatory policy change on shadow banking in China, which resulted in reduced lending and deteriorating credit conditions. The findings revealed that the machine learning-based model, incorporating non-traditional data, demonstrated superior predictive performance for losses and defaults relative to traditional models in response to a negative shock to aggregate loan supply.

In their study, Tripathi et al. (2020) underscored the importance of accurately distinguishing between "creditworthy" and "non-creditworthy" categories to enhance the performance of credit scoring models, especially for the unreliable group, thus affecting the profitability of financial institutions. Their investigation aimed to assess the influence of various combinations of feature selection and classification approaches. They applied nine feature selection methods and various classification approaches to datasets such as Australian, Bank-marketing, Bankruptcy, Japanese, German-categorical, German-numerical, and Taiwanese. Their findings indicated that TDNN (Time Delay Neural Network) and RF showcased the best and second-best performance across most of the credit scoring datasets. Furthermore, the study revealed that the Correlation Coefficients-based Feature Ranking (CFS) approach consistently delivered superior results across most of the classification algorithms.

In their study aiming to develop a model for evaluating personal loans based on big data from the lending club dataset, Wu & Pan (2021) employed the pdC-RF algorithm to optimize data feature correlation and reduce the dimensionality of personal loan data from 145 to 22 dimensions. They evaluated the dataset using random forest, support vector machine, and logistic regression models, and utilized weight of evidence (WOE) coding to measure the probability of a grouping in the features being predicted as a negative example. Following the model performance comparison, they concluded that logistic regression is more suitable for the personal loan assessment model.

In their study, Gunnarsson et al. (2021) sought to examine the effectiveness of deep learning algorithms for credit scoring. They developed and compared two deep learning architectures, namely a multilayer perceptron network and a deep belief network, with two traditional and two ensemble methods for credit scoring. The performance of these models was then evaluated using various credit scoring datasets and performance metrics. Furthermore, the researchers compared Bayesian statistical testing procedures with frequentist non-parametric testing procedures, which are conventionally considered best practices in credit scoring. The traditional methods considered Logistic Regression and Decision Tree, while the ensemble learning methods comprised Random Forest and XGBoost. For deep learning, the models used were Multilayer Perceptron Networks and Deep Belief Networks. The analysis of these

different classification algorithms for credit scoring led to two principal findings. Firstly, the XGBoost algorithm, an ensemble method, exhibited the best performance for credit scoring among all the methods considered. Secondly, it was found that deep neural networks did not outperform their shallower counterparts and were considerably more computationally intensive to construct. Consequently, based on this comparison, the researchers suggested that deep learning algorithms may not be the most suitable models for credit scoring and recommended the XGBoost method as the preferred approach for credit scoring activities due to its superior classification performance.

Boughaci & Alkhawaldeh (2020) undertook a study to compare various machine learning techniques using datasets from financial institutions in six different countries and the "Give Me Some Credit" dataset. The objective was to identify the most effective methods for credit scoring and insolvency prediction. Their findings showcased machine learning methods capable of distinguishing between suitable and unsuitable applicants or companies by generating scores for them. However, they determined that no single method consistently outperformed the others across all datasets. Furthermore, they noted significant variations in the performance of the methods across different datasets. For instance, the Bayesian net method proved most effective for the German and Give Me Some Credit datasets, while the LogitBoost method outperformed others for the Polish and Australian datasets, the AdaBoost method for the Japanese dataset, and the Random Forest method for the Taiwan dataset. In the Indian Qualitative Bankruptcy dataset, nearly all methods demonstrated comparable performance due to the unique nature of the data.

In their study, Radovanovic & Haas (2023) aimed to enhance traditional bankruptcy prediction models by integrating the socio-economic consequences of their forecasts alongside the prediction of the bankruptcy event itself. They leveraged a substantial real-world dataset comprising over 190,000 company-year observations of listed North American companies between 1985 and 2020. The study compared the performance of various machine learning algorithms, including logistic regression, support vector machines, random forest, and neural networks, with traditional statistical models like linear discriminant analysis. The results consistently demonstrated the superior performance of the machine learning models over the traditional statistical models. The study emphasized the importance of considering the financial and socio-economic implications when selecting the most appropriate bankruptcy forecasting model, as even minor disparities in model performance can have significant ramifications in these domains.

A study conducted by Toudas et al. (2024) focused on analyzing bankruptcy prediction within the construction industry in Greece during an economic crisis. The research utilized financial data from construction companies listed on the Greek Stock Exchange in Athens to compare the performance of three distinct models for predicting corporate insolvency, namely various iterations of the Altman Model, the Ohlson Model, and the Zmijewski Model. The objective was to assess the predictability of these models and to ascertain their efficacy in predicting bankruptcy for insolvent companies. The results indicated that both the original Altman model and its revised versions exhibited low overall predictability for the three years leading up to bankruptcy.

In 2023, Máté et al. conducted a study to determine the most effective method for predicting business failure in non-financial firms in Pakistan. They employed various machine learning models on 36 financial ratios to address questions regarding model selection and the best ratios to use. The study highlighted several financial ratios, such as return on assets, operating return on assets, debt coverage ratio, asset turnover, earnings per share, debt/asset

ratio, cash return on assets, and quick ratio, as valuable for predicting bankruptcy. The findings showed that decision tree, AdaBoost, and gradient boosting models performed exceptionally well with 100% accuracy, while SVM and logistic regression models demonstrated flexibility in feature selection, achieving accuracy rates between 89% and 99%. Conversely, the Naive Bayes model performed inadequately, yielding an accuracy range of 58% to 70%.

In 2023, Gajdosikova & Valaskova endeavored to develop a model for predicting bankruptcy based on the financial data of 3,783 Slovak enterprises in the manufacturing and construction sectors in 2020 and 2021. Utilizing multiple discriminant analyses, they identified key financial indicators for the prediction. The results emphasized the self-financing ratio as the most accurate variable in the model. Ultimately, the model, developed through multiple discriminant analyses, exhibited an impressive overall discriminant accuracy of 93%.

In 2022, a study by Muñoz-Izquierdo et al. explored the impact of specific sections within extended audit reports on corporate credit ratings. The researchers employed four machine learning techniques - C4.5 decision tree, PART algorithm rule classifier, rough set methodology, and logistic regression. Their findings demonstrated that accurately identifying key audit matters within the report enabled the evaluation of credit scores with 74% accuracy, using the rules provided by the PART algorithm. Additionally, internal and external key audit matters influenced a company's credit rating disclosure. The study also highlighted the similar predictive power of rule induction classifiers. Interestingly, when combining audit data with accounting ratios, the predictive accuracy of the model increased to 84%, surpassing that of the existing literature.

Issa et al. (2024) examined the probability of bankruptcy across 20 financial sector institutions by analyzing indicators such as liquidity, profitability, debt composition, and operational efficiency derived from financial statements spanning 2000 to 2018. These metrics were juxtaposed with regulatory standards and evaluated for their low, medium, or high-risk implications, culminating in an overarching risk assessment. Additionally, the model incorporates an algorithm that bolsters the dependability of the risk evaluation by accounting for excessive debt levels. The findings underscore that excessive debt adversely affects profitability, leading to lower stock returns and a greater likelihood of bankruptcy. Furthermore, excessive debt and leverage at major financial institutions pose a risk of systemic risk, potentially triggering a domino effect across the global financial system. These insights carry practical implications for investors and stakeholders, furnishing enlightening perspectives to aid decision-making processes, especially during periods of economic volatility.

Slapnik and Lončarski (2023) examine the determinants of sovereign credit ratings by applying textual analysis to credit rating reports. The study uses sentiment and subjectivity scores extracted from S&P, Fitch, and Moody’s reports, representing, respectively, the general perception of a country and the qualitative judgment of the rating committee. The findings indicate that soft information (objectively unobservable factors affecting a country’s debt repayment capacity) and bias proxies provide significant additional insight in predicting sovereign credit ratings. Differences were observed between emerging and advanced markets, with emerging markets exhibiting a higher reliance on the rating committee’s qualitative judgment. Moreover, the general tone of reports changed after the 2008 global financial crisis, although subjectivity did not show statistically significant variation. The study highlights that incorporating textual analysis and soft information can enhance the understanding of sovereign credit ratings and underscores the importance of qualitative judgment alongside traditional economic indicators.

The study by Fei, Gu, Yang, and Zhou (2015) examined the effectiveness of social media opinions in predicting the future credit risk of businesses. Traditional financial statement-based logit model results were used as a benchmark. Posts and comments from two social media platforms heavily used by financial investors in China were analyzed using text mining. Opinions of financial analysts were also included in the evaluation. The findings revealed that opinions obtained from social media were more successful than analyst opinions in predicting credit risk and contained meaningful value-related information. This study contributes to the literature by demonstrating that big data and textual analysis techniques can serve as alternatives to traditional methods in credit risk assessment.

In their study, Gül, Kabak, and Topcu (2018) proposed a multi-criteria approach integrating social media data into the credit rating process. In addition to traditional financial and non-financial indicators, data obtained from Twitter was processed using sentiment analysis, criterion weights were determined through pairwise comparisons, and company performance was evaluated using a cumulative belief rating approach. This method provided analysts with a more flexible and interpretable assessment by presenting credit ratings as a risk distribution rather than a single definitive value. The model, applied to 64 companies, revealed that social media data significantly contributed to creditworthiness assessments, but that credit ratings generally tended to decrease when social media data were taken into account.

The study by Chen and Chen (2022) examined the use of social media-derived big data in predicting corporate credit ratings. As an alternative to traditional financial statements, corporate governance, and macroeconomic indicators, they proposed integrating public perception of companies on social media into the credit rating process. In this context, a credit rating prediction process based on social media big data was designed, different machine learning techniques were developed, and the model's application and performance were evaluated. The findings revealed that predictions based on social media data provided higher accuracy compared to traditional indicators, and the K-Nearest Neighbor (KNN) algorithm, in particular, showed superior performance compared to other methods. The study makes a significant contribution to the literature by demonstrating that alternative data sources can be effectively used in credit rating processes within the FinTech context.

In Aksoy's (2020) study, the credit ratings of 11 insurance companies operating in Türkiye's non-life insurance sector with high market share were predicted using machine learning methods. Analyses were performed using Artificial Neural Networks (ANN), k-Nearest Neighbors, and Naive Bayes algorithms with financial statement data from the 2009–2019 period. Prediction performance was evaluated with 10-fold cross-validation. The results showed that ANN achieved a classification success rate of 98.55%, KNN 95.65%, and Naive Bayes 85.51%. Highlighting that these models are effective prediction tools that can be used by insurance companies, lenders, reinsurance companies, and regulatory bodies.

In the study conducted by Doğan, Büyükkör, and Atan (2022), the credit ratings of 1881 companies operating in three different sectors in Türkiye were estimated using machine learning and modern statistical methods. The study observed that logistic regression, support vector machines, Random Forest, and XGBoost algorithms achieved higher classification accuracy, sensitivity, specificity, and precision than decision tree and k-nearest neighbor methods. Furthermore, sector-based analyses were found to significantly affect the performance of credit rating prediction. This study demonstrates that machine learning methods can be used as a reliable and transparent tool in corporate credit ratings in Türkiye.

Sugozu, Verberi, and Yasar (2025) examine the credit risk of participation and traditional banks in Türkiye using machine learning methods. In this study, credit risk models were developed using CatBoost, XGBoost, Random Forest, and LightGBM algorithms with data from 33 traditional and 6 participation banks for the period 2009–2022, and the effects of variables were analyzed using the Tree SHAP method. The findings show that credit risk is higher in participation banks compared to traditional banks, competition increases credit risk, and loan size and profitability play a significant role in risk. Economic growth, however, reduces risk. The study suggests that participation banks should increase their economies of scale and reduce their risks through special regulations. This research makes a significant contribution to the literature by examining the impact of competition on credit risk in Türkiye's dual banking system using interpretable machine learning.

Method

This study adopts a quantitative text analytics design based on sentiment analysis and supervised machine learning. Sentiment analysis is a natural language processing (NLP) and text mining approach that aims to systematically identify subjective expressions in textual data and classify them according to their emotional orientation (positive, negative, neutral; like, dislike) and/or emotional intensity (Saberi & Saad, 2017). One of the pioneering works in the field, Bo Pang and Lillian Lee (2008), define sentiment analysis as the process of automatically extracting and classifying opinions, attitudes, and evaluations. Bing Liu (2012) offers a more comprehensive framework, defining sentiment analysis as the extraction of individuals' opinions and emotions about a particular entity, topic, or event from textual data. In the current literature, this field has evolved into multi-level (sentence, document, orientation/feature-based) analyses that more accurately capture contextual meaning through deep learning, particularly transformer-based models (Jacob Devlin et al., 2019). In this context, thanks to transfer learning, language models previously trained on large-scale datasets can be adapted to more limited, domain-specific datasets. Furthermore, Large Language Models (LLMs) capture contextual meaning, implicit sentiment tones, and complex discourse structures with higher accuracy, thus increasing both the accuracy and generalizability of sentiment analysis (Prottasha et al., 2022).

The dataset consists of corporate credit rating reports that include narratives under *strengths* and *constraints*, along with firm identifiers, sector information, and assigned national rating categories. In this context, the classification of expressions in the textual content as *strengths* and *constraints* is theoretically based on a polarity-based sentiment analysis approach. Within this framework, the "strengths" category represents positive evaluations and elements that strengthen institutional capacity, while the "constraints" category points to negative orientations, limiting factors, and risk indicators. Therefore, the classification performed can be conceptualized as a sentiment analysis application that systematically separates subjective evaluations in texts by sentiment orientation.

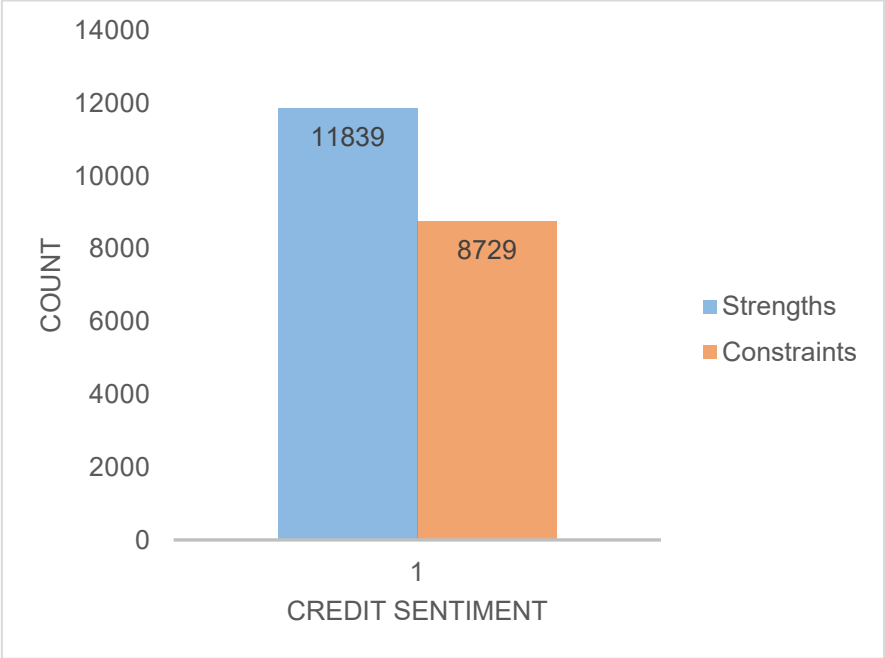
In the first stage, narratives are pre-processed (e.g., normalization, punctuation cleaning, and whitespace standardization) and used to fine-tune a transformer-based classifier to distinguish *strengths* from *constraints*. In the second stage, the model outputs are aggregated to construct report-level indicators, which are then used to predict the national credit rating using a multi-class classification model under a company-level train–test split to prevent information leakage across firms. Because the study relies solely on publicly available rating reports and does not require any data collection from human participants, ethical committee approval was not required.

Data Set

In the study, we used corporate credit rating reports prepared by credit rating agencies as the dataset. These reports generally contain two main types of information. The first consists of textual sections that describe the strengths and weaknesses of firms. In other words, the sections of credit rating reports that express firms' strengths and weaknesses provide a dataset labeled as *strengths* and *constraints* by experts. Within this context, 20,568 labeled observations were obtained from credit rating reports. As shown in Figure 1, 11,839 of these observations correspond to *strengths*, while 8,729 correspond to *constraints*.

Figure 1

Label distribution in the 20k expert-annotated credit sentiment datasets



Note. This figure was created by the author via Python.

The second type of data obtained from the reports consists of the credit ratings assigned to firms as a result of the evaluation process. As is well known, credit ratings are an assessment method based on a firm's ability to meet its financial obligations. They provide a forward-looking projection of the firm's potential to fulfill its financial liabilities in a timely manner.

Credit ratings provide a standardized, transparent language that facilitates global comparability, enabling investors to assess the likelihood that a firm or issuer will repay its debt obligations in full and on time. The rating scale used to evaluate firms and issuances is expressed through the categories "AAA" to "BBB" (Investment Grade), "BB" to "C" (Speculative), and "D" (Default). The use of plus (+) and minus (−) signs allows for further differentiation within the AA to CCC categories.

These categories do not imply that any security is recommended or approved for investment purposes. Ratings in the "Investment Grade" category indicate relatively low to moderate credit risk, whereas those in the "Speculative" category signal higher levels of credit risk. Finally, "Default Event" ratings indicate that a default event has partially or fully occurred.

Evaluation metrics

Model performance was evaluated using standard multi-class classification metrics (Powers, 2011; Sokolova & Lapalme, 2009) reported at both the class and overall levels. Precision measures the proportion of observations predicted as a given class that are correctly classified, while recall measures the proportion of true observations of that class that are correctly identified. Here, TP, FP, and FN denote true positives, false positives, and false negatives, respectively. Accordingly, precision and recall are defined as:

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

The F1-score (Bishop, 2006) is the harmonic mean of precision and recall and provides a single measure that balances these two aspects, particularly useful when there is a trade-off between false positives and false negatives:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{3}$$

Support denotes the number of test observations belonging to each class (n_k) and is critical for interpreting class-wise results, as metrics computed on very small supports can be unstable. At the aggregate level, accuracy is reported as the fraction of all test observations that are correctly classified:

$$Accuracy = \frac{\sum_{k=1}^{K} TP_k}{N} \tag{4}$$

where N is the total number of test observations and K is the number of classes. In addition, macro-averaged precision/recall/F1 are computed as the unweighted mean across classes, treating each class equally and thus reflecting performance under class imbalance:

$$Macro - M = \frac{1}{K} \sum_{k=1}^{K} M_k \tag{5}$$

By contrast, weighted-averaged metrics weight each class by its support, providing an overall summary that is more influenced by majority classes:

$$\text{Weighted} - \text{M} = \sum_{k=1}^{K} \frac{n_k}{N} M_k \tag{6}$$

Together, these metrics offer a comprehensive view of predictive performance, highlighting both overall correctness and class-specific discrimination.

Implementation Details

Creating sentimental strengths-constraints classifier models

In addition to finance-domain pre-trained models, we also employ other pre-trained large language models with the aim of developing classifier models that detect sentiment in the textual sources of credit rating reports within the *strengths–constraints* framework. Overall, the results indicate strong model performance, with accuracy exceeding 97%. These results remain

consistent across different hyperparameter settings, reinforcing the superiority of pre-trained models.

As shown in Table 1, the performance metrics demonstrate the suitability of the BERT and DistilBERT models for *strengths–constraints*–based sentiment analysis required in credit rating assessment. In the table, the accuracies of the environmental, social, and governance models are reported together with their standard deviations (std.) in parentheses.

Table 1

Strengths–Constraints Classification Results

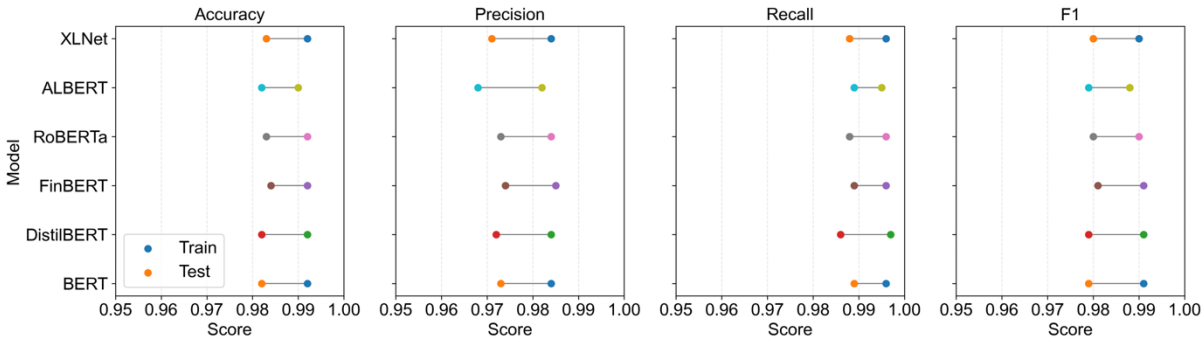
Model	checkpoint	Learning rate	Batch size	epochs	Accuracy		Precision		Recall		F1	
					Train	Test	Train	Test	Train	Test	Train	Test
BERT	bert-base-uncased	0.00002	8	3	0.992	0.982	0.984	0.973	0.996	0.984	0.990	0.978
BERT	bert-base-uncased	0.00003	16	4	0.992	0.982	0.984	0.969	0.997	0.989	0.991	0.979
RoBERTa	roberta-base	0.00002	8	3	0.991	0.984	0.983	0.974	0.996	0.988	0.989	0.981
RoBERTa	roberta-base	0.00003	16	4	0.992	0.983	0.984	0.973	0.996	0.986	0.990	0.980
DistilBERT	distilbert-base-uncased	0.00002	8	3	0.992	0.981	0.985	0.969	0.996	0.986	0.990	0.977
DistilBERT	distilbert-base-uncased	0.00003	16	4	0.992	0.982	0.984	0.972	0.997	0.986	0.991	0.979
ALBERT	albert-base-v2	0.00002	8	3	0.990	0.982	0.982	0.968	0.995	0.989	0.988	0.979
ALBERT	albert-base-v2	0.00003	16	4	0.989	0.973	0.981	0.963	0.993	0.975	0.987	0.969
XLNet	xlnet-base-cased	0.00002	8	3	0.991	0.982	0.983	0.972	0.996	0.986	0.989	0.979
XLNet	xlnet-base-cased	0.00003	16	4	0.992	0.983	0.984	0.971	0.996	0.988	0.990	0.980
FinBERT	ProsusAI/finbert	0.00002	8	3	0.991	0.983	0.983	0.972	0.996	0.989	0.990	0.981
FinBERT	ProsusAI/finbert	0.00003	16	4	0.992	0.984	0.985	0.974	0.996	0.988	0.991	0.981

Note. Created by the authors using Python.

Table 1 reports the performance of different Transformer-based language models on the *strengths–constraints* binary classification task. For BERT, RoBERTa, DistilBERT, ALBERT, XLNet, and FinBERT, test accuracy rates exceed 97%, while test F1 scores are above 0.97.

Figure 2

Train vs. Test Performance Across Models and Metrics

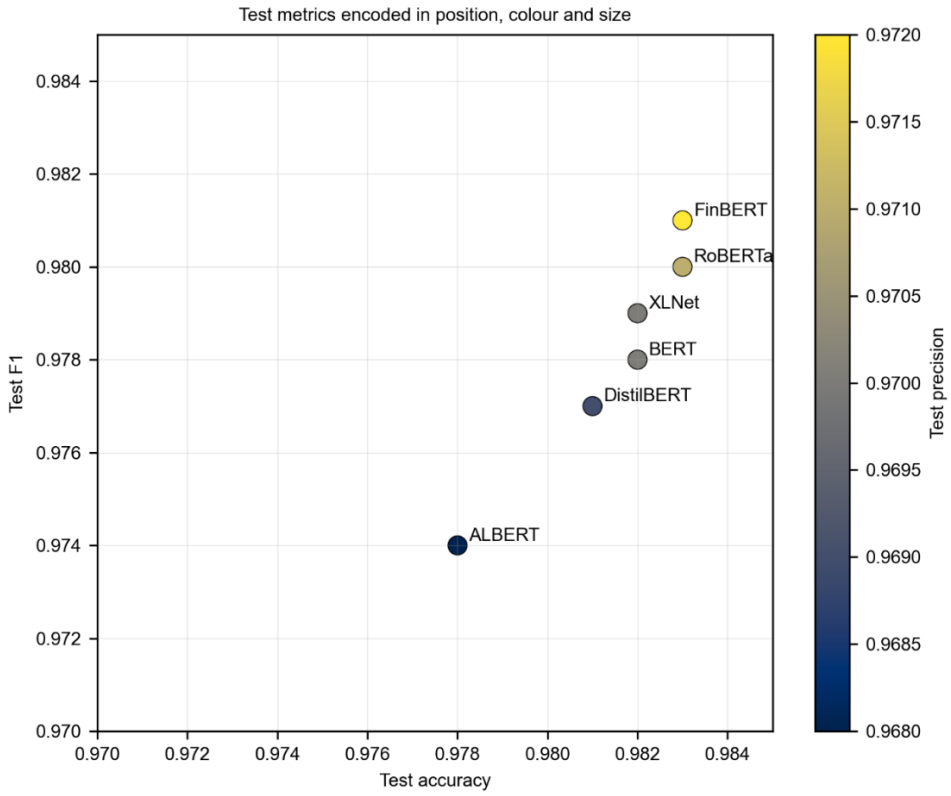


Note. This figure was created by the author via Python, where Arial font customization is not supported.

As shown in Figure 2, the fact that the gap between training and test metrics remains below 0.01 across all models indicates strong generalization capability. The highest test F1 score, 0.981, is achieved by the RoBERTa-base and FinBERT configurations. The results suggest that FinBERT's pre-training on financial text provides a limited but observable performance advantage in the context of credit rating reports.

Figure 3

Performance Results of Language Models

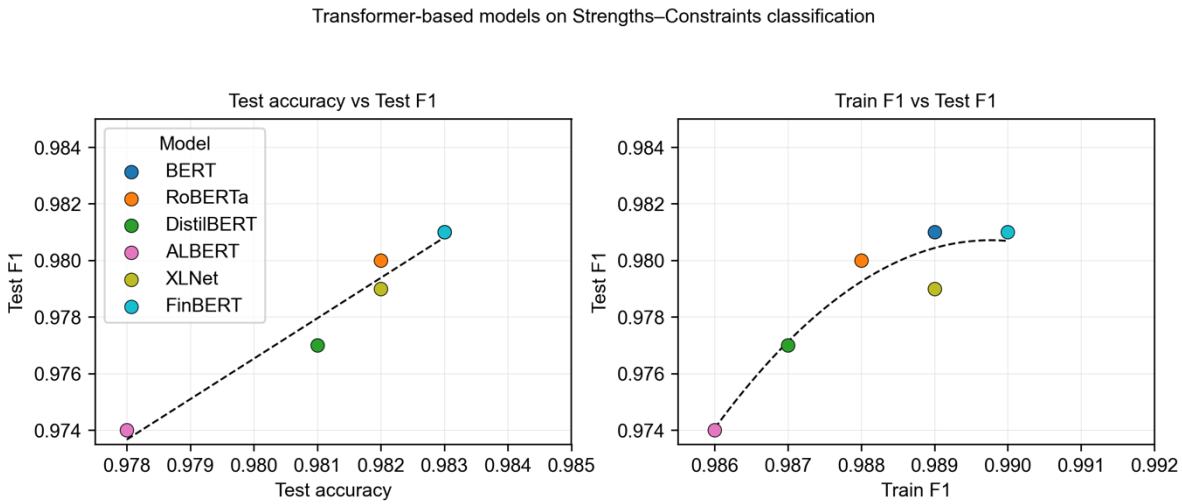


Note. This figure was created by the author via Python, where Arial font customization is not supported.

Nevertheless, the relatively small difference between general-purpose BERT-derived models and the domain-specific FinBERT indicates that the *strengths* and *constraints* sections are linguistically well separated and that the task is similarly tractable across different Transformer architectures. Moreover, the ability of the lighter DistilBERT architecture to produce results very close to those of larger models can be considered an important finding in terms of the trade-off between computational cost and performance at the deployment stage (Figure 3).

Figure 4

Transormer-Based Models on Strengths-Constraints Classification



Note. This figure was created by the author via Python, where Arial font customization is not supported.

As shown in Figure 4, test accuracy and test F1 scores (left) and train–test F1 scores (right) for different hyperparameter configurations of Transformer-based models in the *strengths–constraints* binary classification task. Each point represents a specific model–hyperparameter combination. The orange lines indicate smoothed curves obtained using a second-degree polynomial. The axes are zoomed into the 0.97–0.985 range. Therefore, differences across models are very small in absolute terms and are shown solely for relative comparison purposes.

From strengths-constraints to credit rating classifier models

Two different approaches are adopted to develop a model that predicts firms' credit ratings based on reports issued by credit rating agencies. The first approach classifies texts into *strengths* and *constraints*, then derives a sentiment score from this classification and uses that score to predict firms' credit ratings.

In the second approach, *strengths–constraints* texts are converted into vectors using Transformer libraries, thereby reducing all information to a single vector representation and embedding each text in a semantic vector space. In other words, if *strengths* are considered as positive aspects and *constraints* as negative aspects, these two types of text complement each other and together form a unified representation of a firm's overall financial condition. By providing *strengths* and *constraints* texts jointly to the model, it is ensured that the model learns the context of both sides within the same embedding space.

Table 2

Rating Scale

Investment Level	Rating Group	Credit Quality	Remarks
Investment Grade	AAA	Highest Credit Quality	Represents the minimal expected risk of default. Assigned only in exceptional circumstances where the entity demonstrates a very strong ability to meet its financial obligations, with this capacity being highly resilient to foreseeable events.
	AA	Very High Credit Quality	Reflects a very low expected level of default risk. It is assigned in exceptional situations where the entity demonstrates a strong ability to meet its financial obligations, and this ability is highly resilient to foreseeable events.
	A	High Credit Quality	Indicates a low expected level of default risk. The entity is considered to have a strong ability to meet its financial obligations. However, this ability may be more susceptible to negative business or economic developments compared with higher-rated entities.
	BBB	Good Credit Quality	Indicates that the expected default risk is presently low. The entity's capacity to meet financial obligations is considered adequateh. However, unfavorable business or economic conditions are more likely to compromise its solvency.
Speculative	BB	Speculative	Indicates a higher susceptibility to default risk, especially if business or economic conditions deteriorate over time. Nevertheless, the entity retains commercial or financial flexibility that helps ensure fulfillment of its financial obligations.
	B	Highly Speculative	Indicates a significant default risk, with only a narrow margin of safety. Although financial obligations are currently being met, the entity's ability to fulfill them remains vulnerable to adverse changes in business and economic conditions.

Note. Adapted by the author from "Credit Rating" (https://www.jcrer.com.tr/en/methodology/notations/credit-rating).

Table 2 presents a structured classification of credit ratings, linking investment levels, rating groups, and corresponding credit quality to the expected default risk and the entity's capacity to meet financial obligations. Investment-grade ratings, ranging from AAA to BBB, reflect minimal to low default risk, indicating a very strong to adequate ability to fulfill financial commitments, with resilience varying by rating. AAA and AA ratings represent the highest and very high credit quality, assigned in exceptional circumstances in which the entity demonstrates a strong, highly resilient capacity to meet its obligations. A and BBB ratings denote high to good credit quality, with the ability to meet obligations considered strong to adequate, though more susceptible to adverse business or economic developments compared to higher-rated entities. Speculative ratings, including BB and B, indicate elevated to significant default risk, where financial obligations may still be met but are increasingly vulnerable to deteriorating business and economic conditions (Category B was combined with Category BB in the analysis because there were very few observations in category B.). The table highlights not only the expected probability of default but also the degree of financial flexibility and resilience associated with each rating category, providing a comprehensive framework for evaluating credit quality across different investment levels.

Table 3

Credit Score Classification Results

Credit Class	Precision	Recall	F1-score	Support
AAA	0.4000	0.1081	0.1702	37
AA	0.5669	0.6486	0.6050	111
A	0.4202	0.4808	0.4484	104
BBB	0.5851	0.5978	0.5914	92
BB	0.5000	0.1250	0.2000	8
accuracy			**0.5170**	352
macro avg	0.4944	0.3921	0.4030	352
weighted avg	0.5093	0.5170	0.5003	352

Note. Created by the authors using Python.

The results indicate that narrative-based classification and credit rating prediction involve different levels of difficulty. In the *strength–constraint* stage (Task 1), the transformer model converged with low training loss and achieved strong test performance. In contrast, in Task 2, where the national credit rating was predicted using only report-level narrative features derived from Task1, performance was more limited (accuracy = 0.517; macro-F1 = 0.403; weighted-F1 = 0.500; n = 352). A class-wise evaluation shows relatively more stable performance for the AA (F1 = 0.605; recall = 0.649) and BBB (F1 = 0.591; recall = 0.598) categories, while the model attains moderate success for A (F1 = 0.448). By comparison, the markedly low recall for AAA (recall = 0.108; F1 = 0.170) suggests substantial difficulty in capturing the highest rating category. Similarly, due to the very small support for BB (support = 8), the corresponding metrics are unstable (recall = 0.125; F1 = 0.200). Overall, these findings suggest that indicators derived from report narratives contain meaningful signals about assigned ratings. However, when used alone, predictive power remains constrained, particularly for extreme classes and under pronounced class imbalance.

Discussion

This study examines the predictive performance of transformer-based language models for credit score classification. All evaluated models (BERT, RoBERTa, DistilBERT, ALBERT, XLNet, and FinBERT) showed very high performance on both training and test accuracy in the *strengths–constraints* classification task, achieving F1 scores above 0.97. This demonstrates that transformer architectures can effectively capture the necessary textual and financial information to identify *strengths* and *constraints* in corporate credit assessments, confirming the potential of deep learning approaches in financial risk assessment.

These findings support the work of Chen et al. (2022) and Fei et al. (2015), which showed that big data from social media and text analysis techniques can improve accuracy in credit risk prediction. In this study, transformer-based models were applied to structured and semi-structured textual data rather than social media data, and similarly high predictive performance was achieved. This situation parallels the general findings in the literature that

alternative textual data can improve the accuracy of credit rating processes (Gül, Kabak, & Topcu, 2018; Slapnik & Lončarski, 2023).

Unlike *strengths-constraints* classification, credit score classification proved to be a more challenging task. The weighted average accuracy was 0.517, and significant differences were observed across credit classes (e.g., AAA F1 score: 0.17, BBB F1 score: 0.59). This reflects the difficulty of distinguishing between high-grade and low-frequency classes. In this context, the low number of observations in some classes (e.g., 8 samples for the BB class) may have negatively impacted model performance.

Methodologically, FinBERT consistently demonstrated high performance in *strengths-constraints* tasks. This is consistent with the FinTech literature supporting the value of using domain-specific pre-trained models in financial NLP applications (Sugozu et al., 2025). The small performance differences between BERT, RoBERTa, and DistilBERT demonstrate that basic transformer models can generalize well, but domain adaptation and fine-tuning are critical to maximizing prediction accuracy. Overall, these results offer several important insights. First, transformer-based models can accurately predict performance-related constraints by extracting strong representations from financial texts. Second, credit rating predictions are sensitive to overclasses and limited sample sizes. Third, integration with text analyses from social media and credit reports (Slapnik & Lončarski, 2023; Fei et al., 2015) offers a promising avenue for improving predictive models.

Conclusion

This study makes a significant contribution to credit analytics by demonstrating that qualitative narratives in credit rating reports, which constitute the most interpretation-intensive component of the rating process, can be transformed into consistent quantitative information and integrated into an end-to-end modeling pipeline. First, the near-ceiling performance of the *strength–constraint* classifier provides strong evidence that rating reports contain stable and learnable linguistic patterns. This, in turn, enables the scalable and automated extraction of positive and negative rating drivers. More broadly, the findings point to a practical mechanism that standardizes the interpretation of narratives, reduces reliance on manual reading, and produces auditable summaries of credit rationales across firms and sectors.

Second, the rating prediction analysis delivers a clear empirical insight into the informational value of narratives. Even when the model is restricted to report-level features extracted from text, it can meaningfully distinguish among the more frequently observed rating categories, suggesting that qualitative assessments contain a non-trivial signal related to the final rating decision. At the same time, the weaker results for rare and extreme categories, most notably the difficulty in identifying AAA and the unreliability of BB estimates given its very small sample size, offer an important diagnostic implication. These patterns indicate where narrative information alone is insufficient and where factors such as class imbalance, threshold dynamics, and the inclusion of complementary non-textual inputs become crucial. Accordingly, the study clarifies the practical limits of text-only credit modeling rather than implying that narratives can fully substitute for the broader information set used in rating assignments.

From an applied standpoint, the proposed framework provides a scalable foundation for continuous credit monitoring. It can be deployed to automatically parse newly released rating reports, quantify shifts in the balance of *strengths* and *constraints*, and generate early-warning indicators for analysts, investors, and risk managers. Conceptually, the study bridges qualitative credit reasoning and quantitative prediction by operationalizing narrative content into transparent indicators that can be validated, compared over time, and integrated with other

information sources. Future research can build on this foundation by combining narratives with richer quantitative fundamentals and macro signals, adopting ordinal/hierarchical and cost-sensitive learning to better reflect the rating scale, and testing robustness across time periods and rating agencies to strengthen real-world generalizability.

References

Aksoy, B. (2020). Sigorta şirketlerinin derecelendirilmesinde makine öğrenmesi yöntemleri tahmin performansının karşılaştırılması: Türkiye örneği. *Akademik Araştırmalar ve Çalışmalar Dergisi (AKAD)*, 12(23), 579-597. https://doi.org/10.20990/kilisiibfakademik.710863

Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J., & Vanthienen, J. (2003). Benchmarking state-of-the-art classification algorithms for credit scoring. *Journal of the Operational Research Society*, *54*(6), 627–635. https://doi.org/10.1057/palgrave.jors.2601545

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Boughaci, D., & Alkhawaldeh, A. A. K. (2020). Appropriate machine learning techniques for credit scoring and bankruptcy prediction in banking and finance: A comparative study. *Risk and Decision Analysis*, *8*(1–2), 15–24. https://doi.org/10.3233/RDA-180051

Bucker, M., Szepannek, G., Gosiewska, A., & Biecek, P. (2022). Transparency, auditability, and explainability of machine learning models in credit scoring. *Journal of the Operational Research Society, 73*(1), 70–90. https://doi.org/10.1080/01605682.2021.1922098

Chen, Y.-J., & Chen, Y.-M. (2022). Forecasting corporate credit ratings using big data from social media. *Expert Systems with Applications, 207*, Article 118042. https://doi.org/10.1016/j.eswa.2022.118042

Darwish, J. A. (2025). Optimization and prediction of corporate credit rating through advanced feature selection based on AI and deep learning. *Alexandria Engineering Journal*, *127*, 586–594. https://doi.org/10.1016/j.aej.2025.05.043

De Oliveira, N. A., & Basso, L. F. C. (2025). Advancing credit rating prediction: The role of machine learning in corporate credit rating assessment. *Risks*, *13*(6), 116. https://doi.org/10.3390/risks13060116

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In J. Burstein, C. Doran, & T. Solorio (Eds.), *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long and Short Papers)* (pp. 4171–4186). Association for Computational Linguistics. https://doi.org/10.18653/v1/N19-1423

Doğan, S., Büyükkör, Y., & Atan, M. (2022). A comparative study of corporate credit ratings prediction with machine learning. *Operations Research and Decisions, 32*(1), 25–47. https://doi.org/10.37190/ord220102

Dumitrescu, E., Hué, S., Hurlin, C., & Tokpavi, S. (2022). Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects. *European Journal of Operational Research*, *297*(3), 1178–1192. https://doi.org/10.1016/j.ejor.2021.06.053

Fei, W., Gu, J., Yang, Y., & Zhou, Z. (2015). Credit risk evaluation based on social media. *Procedia Computer Science, 55*, 725–731. https://doi.org/10.1016/j.procs.2015.07.165

Gajdosikova, D., & Valaskova, K. (2023). Bankruptcy prediction model development and its implications on financial performance in Slovakia. *Economics and Culture*, *20*(1), 30–42. https://doi.org/10.2478/jec-2023-0003

Gambacorta, L., Huang, Y., Qiu, H., & Wang, J. (2024). How do machine learning and non-traditional data affect credit scoring? New evidence from a Chinese fintech firm. *Journal of Financial Stability*, *73*, Article 101284. https://doi.org/10.1016/j.jfs.2024.101284

Gunnarsson, B. R., vanden Broucke, S., Baesens, B., Óskarsdóttir, M., & Lemahieu, W. (2021). Deep learning for credit scoring: Do or don't? *European Journal of Operational Research*, *295*(1), 292–305. https://doi.org/10.1016/j.ejor.2021.03.006

Gül, S., Kabak, Ö., & Topcu, I. (2018). A multiple criteria credit rating approach utilizing social media data. *Data & Knowledge Engineering, 116*, 80–99. https://doi.org/10.1016/j.datak.2018.05.005

Issa, S., Bizel, G., Jagannathan, S. K., & Gollapalli, S. S. C. (2024). A comprehensive approach to bankruptcy risk evaluation in the financial industry. *Journal of Risk and Financial Management*, *17*(1), 41. https://doi.org/10.3390/jrfm17010041

Lessmann, S., Baesens, B., Seow, H.-V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research*, *247*(1), 124–136. https://doi.org/10.1016/j.ejor.2015.05.030

Liu, B. (2012). *Sentiment analysis and opinion mining*. Morgan & Claypool Publishers.

Liu, Y., & Schumann, M. (2005). Data mining feature selection for credit scoring models. *The Journal of the Operational Research Society*, *56*(9), 1099–1108. https://doi.org/10.1057/palgrave.jors.2601976

Machado, M. R., & Karray, S. (2022). Assessing credit risk of commercial customers using hybrid machine learning algorithms. *Expert Systems with Applications*, *200*, Article 116889. https://doi.org/10.1016/j.eswa.2022.116889

Machado, M. R., Chen, D. T., & Osterrieder, J. R. (2025). An analytical approach to credit risk assessment using machine learning models. *Decision Analytics Journal*, *16*, Article 100605. https://doi.org/10.1016/j.dajour.2025.100605

Marqués, A. I., García, V., & Sánchez, J. S. (2012). Two-level classifier ensembles for credit risk assessment. *Expert Systems with Applications*, *39*(12), 10916–10922. https://doi.org/10.1016/j.eswa.2012.03.033

Máté, D., Raza, H., & Ahmad, I. (2023). Comparative analysis of machine learning models for bankruptcy prediction in the context of Pakistani companies. *Risks*, *11*(10), Article 176. https://doi.org/10.3390/risks11100176

Muñoz-Izquierdo, N., Segovia-Vargas, M. J., Camacho-Miñano, M.-M., & Pérez-Pérez, Y. (2022). Machine learning in corporate credit rating assessment using the expanded audit report. *Machine Learning*, *111*(11), 4183–4215. https://doi.org/10.1007/s10994-022-06226-4

Pai, P.-F., Tan, Y.-S., & Hsu, M.-F. (2015). Credit rating analysis by the decision-tree support vector machine with ensemble strategies. *International Journal of Fuzzy Systems*, *17*(4), 521–530. https://doi.org/10.1007/s40815-015-0063-y

Paleologo, G., Elisseeff, A., & Antonini, G. (2010). Subagging for credit scoring models. *European Journal of Operational Research*, *201*(2), 490–499. https://doi.org/10.1016/j.ejor.2009.03.008

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, *2*(1–2), 1–135. https://doi.org/10.1561/1500000001

Powers, D. M. W. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies, 2*(1), 37–63. https://doi.org/10.9735/2229-3981

Prottasha, N. J., Sami, A. A., Kowsher, M., Murad, S. A., Bairagi, A. K., Masud, M., & Baz, M. (2022). Transfer learning for sentiment analysis using BERT based supervised fine-tuning. *Sensors*, *22*(11), 4157. https://doi.org/10.3390/s22114157

Radovanovic, J., & Haas, C. (2023). The evaluation of bankruptcy prediction models based on socio-economic costs. *Expert Systems with Applications*, *227*, Article 120275. https://doi.org/10.1016/j.eswa.2023.120275

Saberi, B., & Saad, S. (2017). Sentiment analysis or opinion mining: A review. .*International Journal of Advanced Science Engineering Information Technology*, *7*(5), 1660-1666.

Siham, A., Sekkate, S., & Adib, A. (2021). Feature selection based on machine learning for credit scoring: An evaluation of filter and embedded methods. In Z. H. Kilimci, T. Yildirim, V. Piuri, I. Czarnowski, D. Camacho, Y. Manolopoulos, & S. Solak (Eds.), *Proceedings of the 2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA 2021)* (pp. 1–6). IEEE. https://doi.org/10.1109/INISTA52262.2021.9548410

Slapnik, U., & Lončarski, I. (2023). Understanding sovereign credit ratings: Text-based evidence from the credit rating reports. *Journal of International Financial Markets, Institutions and Money, 88*, Article 101838. https://doi.org/10.1016/j.intfin.2023.101838

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, *45*(4), 427–437. https://doi.org/10.1016/j.ipm.2009.03.002

Sugozu, I. H., Verberi, C., & Yasar, S. (2025). Machine learning approaches to credit risk: Evaluating Turkish participation and conventional banks. *Borsa Istanbul Review*, *25*(3), 497-512. https://doi.org/10.1016/j.bir.2025.02.001

Teles, G., Rodrigues, J. J. P. C., Rabêlo, R. A. L., & Kozlov, S. A. (2021). Comparative study of support vector machines and random forests machine learning algorithms on credit operation. *Software: Practice and Experience*, *51*(12), 2492–2500. https://doi.org/10.1002/spe.2842

Teles, G., Rodrigues, J. J. P. C., Saleem, K., Kozlov, S., & Rabêlo, R. A. L. (2020). Machine learning and decision support system on credit scoring. *Neural Computing and Applications*, *32*(14), 9809–9826. https://doi.org/10.1007/s00521-019-04537-7

Thomas, L. C. (2000). A survey of credit and behavioural scoring: Forecasting financial risk of lending to consumers. *International Journal of Forecasting*, *16*(2), 149–172. https://doi.org/10.1016/S0169-2070(00)00034-0

Toudas, K., Archontakis, S., & Boufounou, P. (2024). Corporate bankruptcy prediction models: A comparative study for the construction sector in Greece. *Computation*, *12*(1), 9. https://doi.org/10.3390/computation12010009

Tripathi, D., Edla, D. R., Cheruku, R., & Kuppili, V. (2019). A novel hybrid credit scoring model based on ensemble feature selection and multilayer ensemble classification. *Computational Intelligence*, *35*(2), 371–394. https://doi.org/10.1111/coin.12200

Tripathi, D., Edla, D. R., Kuppili, V., & Bablani, A. (2020). Evolutionary extreme learning machine with novel activation function for credit scoring. *Engineering Applications of Artificial Intelligence*, *96*, Article 103980. https://doi.org/10.1016/j.engappai.2020.103980

Trivedi, S. K. (2020). A study on credit scoring modeling with different feature selection and machine learning approaches. *Technology in Society*, *63*, Article 101413. https://doi.org/10.1016/j.techsoc.2020.101413

Tsai, C.-F. (2014). Combining cluster analysis with classifier ensembles to predict financial distress. *Information Fusion*, *16*, 46–58. https://doi.org/10.1016/j.inffus.2011.12.001

Wang, G., Hao, J., Ma, J., & Jiang, H. (2011). A comparative assessment of ensemble learning for credit scoring. *Expert Systems with Applications*, *38*(1), 223–230. https://doi.org/10.1016/j.eswa.2010.06.048

Wu, Y., & Pan, Y. (2021). Application analysis of credit scoring of financial institutions based on machine learning model. *Complexity*, *2021*(1), Article 9222617. https://doi.org/10.1155/2021/9222617

Xu, Y., Chen, Y., Sun, L., & Chen, Y. (2025). Corporate credit scoring method based on unlabeled data and multi-source data. *Decision Support Systems*, *198*, Article 114543. https://doi.org/10.1016/j.dss.2025.114543

Yobas, M. B., Crook, J. N., & Ross, P. (2000). Credit scoring using neural and evolutionary techniques. *IMA Journal of Management Mathematics*, *11*(2), 111-125. https://doi.org/10.1093/imaman/11.2.11

Information About the Article/Makale Hakkında Bilgiler

The Ethical Rules for Research and Publication / Araştırma ve Yayın Etiği

The author declared that the ethical rules for research and publication followed while preparing the article.

Yazar makale hazırlanırken araştırma ve yayın etiğine uyulduğunu beyan etmiştir.

Conflict of Interests/ Çıkar Çatışması

The author have no conflict of interest to declare.

Yazar çıkar çatışması bildirmemiştir.

Grant Support/ Finansal Destek

The author declared that this study has received no financial support.

Yazar bu çalışma için finansal destek almadığını beyan etmiştir.

Author Contributions/ Yazar Katkıları

The draft process of the manuscript/ Taslağın Hazırlanma Süreci Y.E.A, Data Collection/Verilerin Toplanması Y.E.A, Writing the Manuscript/ Makalenin Yazılması Y.E.A, Submit, Revision and Resubmit Process/ Başvuru, Düzeltme ve Yeniden Başvuru Süreci Y.E.A.