GLOBAL ECONOMIC DYNAMICS, POLICY UNCERTAINTY, AND SUSTAINABLE DEVELOPMENT

Editor:
ASSOC. PROF. DR. FUAT LEBE

BİDGE Yayınları

BİDGE Yayınları

Global Economıc Dynamıcs, Polıcy Uncertaınty, and Sustaınable Development

Editor: Assoc. Prof. Dr. Fuat LEBE

ISBN: 978-625-5057-66-2

1st Edition
Page Layout By: Gözde YÜCEL
Publication Date: 2026-10-25
BİDGE Yayınları

All rights reserved. No part of this work may be reproduced in any form or by any means, except for brief quotations for promotional purposes with proper source attribution, without the written permission of the publisher and the editor.

Certificate No: 71374
All rights reserved © BİDGE Yayınları
www.bidgeyayinlari.com.tr - bidgeyayinlari@gmail.com
Krc Bilişim Ticaret ve Organizasyon Ltd. Şti.
Güzeltepe Mahallesi Abidin Daver Sokak Sefer Apartmanı No: 7/9 Çankaya / Ankara

PREFACE

The global economy is undergoing a period of profound transformation shaped by economic policy uncertainty, geopolitical tensions, environmental challenges, financial vulnerabilities, and structural changes in production and consumption. These developments have intensified the interdependence of national economies and highlighted the need for comprehensive approaches to understanding the forces that influence macroeconomic stability, sustainable growth, income distribution, and long-term development. In this context, examining contemporary economic issues requires not only an understanding of individual economic indicators but also a broader perspective that considers the complex interactions among economic policies, institutional structures, financial conditions, environmental factors, and technological transformation.

Global Economic Dynamics, Policy Uncertainty, and Sustainable Development brings together scholarly contributions that explore a range of significant issues in contemporary economics. The volume reflects the multidimensional nature of economic development by addressing the relationships between global policy uncertainty and macroeconomic performance, the role of economic and governmental policies in shaping environmental outcomes, the implications of public debt for income inequality, and the influence of financial conditions and digital transformation on economic structures. By bringing these perspectives together, the book seeks to contribute to a deeper understanding of the challenges facing national economies in an increasingly interconnected and uncertain global environment.

A central theme of the volume is the importance of economic policy uncertainty in shaping macroeconomic dynamics and investment-related decisions. Uncertainty surrounding economic policies can influence expectations, affect production and consumption decisions, alter financial market conditions, and create

additional challenges for policymakers. Understanding these effects is particularly important for economies exposed to global economic fluctuations and external shocks. The analyses presented in this volume offer insights into the relationship between uncertainty and economic performance, while emphasizing the importance of examining both short-term adjustments and long-term economic relationships.

The book also highlights the growing significance of environmental sustainability and the interaction between economic activity and public policy. The transition toward cleaner energy systems, the environmental consequences of governmental decisions, and the broader implications of climate change have become central concerns in economic research and policymaking. These issues demonstrate that economic progress cannot be evaluated solely in terms of output growth or conventional measures of financial performance. Rather, sustainable development requires an integrated perspective that considers environmental quality, resource allocation, institutional effectiveness, and the long-term consequences of policy choices. Another important dimension of the volume concerns public finance, income distribution, and macroeconomic stability. Public debt and income inequality are closely connected to broader questions of fiscal sustainability, social welfare, and inclusive development. The distributional consequences of economic policies, together with the effects of financial vulnerabilities and economic crises, can influence both the resilience of national economies and the opportunities available to different segments of society. Exploring these relationships is essential for understanding how economic systems can pursue stability while addressing persistent inequalities and supporting more balanced development outcomes.

In addition, the volume considers financial measurement and structural transformation as important components of contemporary economic analysis. The assessment of credit conditions, including

the measurement of credit gaps under different currency compositions, illustrates the importance of methodological choices in identifying financial imbalances and evaluating potential vulnerabilities. Similarly, the expansion of e-commerce and its relationship with regional income convergence draw attention to the changing nature of economic activity, the role of digital technologies, and the possibilities for reducing regional disparities. These topics underscore the importance of adapting analytical frameworks to evolving economic conditions and recognizing the diverse channels through which financial and technological developments affect economic performance.

A distinguishing feature of the contributions in this book is their use of different empirical approaches to investigate complex economic relationships. Econometric methods provide valuable tools for examining long-run relationships, short-run dynamics, cross-country differences, and the potential effects of policy and institutional factors. By applying a range of quantitative techniques to different economic contexts, the studies demonstrate the importance of empirical evidence in evaluating theoretical expectations, identifying relevant economic mechanisms, and informing policy discussions. At the same time, the diversity of the methods and subjects covered reflects the multifaceted character of contemporary economic research.

Although the chapters address distinct research questions, they share a common objective: to improve our understanding of the forces shaping economic stability, structural change, environmental sustainability, and inclusive development. Their collective contribution lies in bringing together complementary perspectives on the challenges confronting modern economies and demonstrating that these challenges are often interconnected rather than isolated. Economic uncertainty, public finance, environmental pressures, financial vulnerabilities, and digital transformation must therefore be

understood within a broader framework that recognizes their mutual interactions and their implications for long-term prosperity.

This volume is intended for researchers, academics, graduate students, policymakers, and readers interested in contemporary economic issues and evidence-based policy analysis. It aims to encourage further scholarly discussion, support comparative perspectives across countries and regions, and contribute to the development of research that bridges theoretical insights and empirical findings. By addressing a diverse set of economic questions, the book also seeks to provide a foundation for future investigations into the changing dynamics of the global economy. Ultimately, sustainable and inclusive economic development requires more than short-term policy responses or isolated improvements in economic performance. It calls for sound institutions, effective policy frameworks, careful assessment of financial and environmental risks, and a sustained commitment to understanding the complex relationships between economic growth and social well-being. It is hoped that the studies presented in this volume will contribute to these broader objectives and offer useful insights into the opportunities and challenges that will shape the future of the global economy.

I would like to express my sincere gratitude to all the contributing authors for sharing their scholarly expertise, research findings, and valuable perspectives. Their academic contributions have been essential to the preparation of this volume. I also extend my appreciation to the Bidge Publishing team for its support and cooperation throughout the publication process. Finally, I hope that this book will serve as a useful academic resource, stimulate further research, and encourage informed discussion on the economic policies and structural transformations that are shaping our increasingly interconnected world.

Assoc. Prof. Dr. Fuat LEBE
Adıyaman University

CONTENTS

THE IMPACT OF GLOBAL ECONOMIC POLICY UNCERTAINTY (GEPU) ON TÜRKIYE'S MACROECONOMIC DYNAMICS: A VECM APPROACH 1
GÜLGÜN ÇİĞDEM, FUNDA KARA, ŞULE EKUKLU

ECONOMIC POLICY UNCERTAINTY AND RENEWABLE ELECTRICITY GENERATION IN THE NETHERLANDS: EVIDENCE FROM A FOURIER-AUGMENTED ARDL MODEL 36
ÇETİN ÜNEN

THE EFFECT OF GOVERNMENTAL FACTORS ON ENVIRONMENT: A SYSTEM GMM APPROACH 63
İBRAHİM TEKİNER

THE NEXUS OF GOVERNMENT DEBT AND INCOME INEQUALITY: A PANEL DATA ANALYSIS OF FIVE EUROPEAN ECONOMIES 95
ARZU ALVAN BOZDERELİ, BARTU YURDAY

THE IMPACT OF DEFENSE EXPENDITURE, CLIMATE CHANGE AND ECONOMIC CRISES 117
MUSTAFA TORUSDAĞ

CURRENCY COMPOSITION AND MEASUREMENT SENSITIVITY IN TURKEY'S CREDIT GAP 137
BİLGE KAĞAN ÖZDEMİR

E-COMMERCE AND REGIONAL INCOME CONVERGENCE IN TURKIYE 167
BURCU KILINÇ SAVRUL, MESUT SAVRUL

CHAPTER 1

The Impact of Global Economic Policy Uncertainty (GEPU) on Türkiye's Macroeconomic Dynamics: A VECM Approach

GÜLGÜN ÇİĞDEM[1]

FUNDA KARA[2]

ŞULE EKUKLU[3]

INTRODUCTION

The concept of economic policy uncertainty has emerged as one of the most consequential analytical frameworks in contemporary macroeconomics. Since the seminal contribution of Bloom (2009), a growing body of literature has demonstrated that uncertainty—particularly uncertainty arising from the global policy environment—exerts profound and lasting effects on investment, employment, financial markets, and real economic activity. The global financial crisis of 2008–2009, the European sovereign debt crisis of 2010–2012, the period of heightened geopolitical tensions following the Arab Spring, the 2016 Brexit referendum, the U.S.–China trade war, the COVID-19 pandemic of 2020, and the Russian invasion of Ukraine in 2022 each generated substantial waves of global economic policy uncertainty (GEPU), with measurable

[1] Prof. Dr., İstanbul Gelişim University, Department of International Trade and Business Administration, gcigdem@gelisim.edu.tr. ORCID: 0000-0001-5353-8638.

[2] Assist. Prof. Dr., İstanbul Gelişim University, Department of Economics and Finance, fkara@gelisim.edu.tr. ORCID: 0000-0001-8685-0695.

[3] PhD Student, İstanbul Gelişim University, sekuklu@hotmail.com. ORCID: 0009-0000-3066-8827

spillover effects on emerging market economies (Caldara et al., 2016; Ahir, Bloom & Furceri, 2022).

Türkiye represents a particularly instructive case for investigating the transmission of global uncertainty shocks. As an open emerging market economy, Türkiye is characterized by a relatively high current account deficit, sizable foreign currency-denominated external debt, and a structurally import-dependent production structure. These features render the Turkish economy sensitive not only to domestic policy developments but also to shifts in global investor sentiment, risk appetite, and uncertainty. The exchange rate—a critical intermediate target in Türkiye's monetary policy framework—has historically been among the first domestic variables to absorb global uncertainty shocks, subsequently transmitting them to inflation and, with a lag, to real activity (Kara & Oğünç, 2012; Akçelik & Başçı, 2011). This sensitivity is compounded by domestically generated uncertainty: Cevik and Erduman (2020) show that monetary policy uncertainty in Türkiye rose markedly during the unconventional policy period and independently depressed economic activity, suggesting that global and domestic uncertainty channels may reinforce one another.

Despite the growing literature on uncertainty and macroeconomic outcomes, the specific question of how GEPU—measured through the Baker, Bloom, and Davis (2016) index—dynamically interacts with Türkiye's macroeconomic aggregates within a multivariate cointegrated time series framework remains underexplored. Most existing studies on Türkiye either focus on domestic uncertainty measures, (Jirasavetakul & Spilimbergo, 2018), examine bilateral relationships in isolation, or rely on VAR models that do not account for the long-run cointegrating structure of the data. This study addresses these gaps by employing a Vector Error Correction Model (VECM), which captures both long-run

equilibrium relationships and short-run adjustment dynamics among cointegrated variables.

We pursue three primary objectives. First, to determine whether a stable long-run equilibrium relationship exists between GEPU and Türkiye's key macroeconomic indicators—the industrial production index (IPI), the nominal exchange rate (USD/TRY), consumer price inflation (CPI), and the policy interest rate. Second, to identify the direction and magnitude of causal linkages between GEPU and these indicators in both the short and long run. Third, to trace the dynamic responses of Türkiye's macroeconomic variables to a standardized GEPU shock through impulse response and variance decomposition analyses, thereby illuminating the transmission channels through which global uncertainty penetrates the Turkish economy.

These objectives are operationalized through the following testable hypotheses. H_1: GEPU and Türkiye's macroeconomic aggregates share at least one long-run cointegrating relationship. H_2: GEPU does not exert direct short-run Granger causality on the exchange rate or inflation; rather, GEPU transmits to domestic variables indirectly through the long-run cointegrating relationship and error-correction dynamics. H_3: A positive GEPU shock generates an initial depreciation of the exchange rate and an oscillatory inflation response (initially negative, turning positive by the third quarter), consistent with a J-curve-type delayed pass-through; industrial production responds negatively in the long run as suggested by the cointegrating vector coefficient (elasticity of approximately −0.5% per unit GEPU, subject to sign verification given normalization choices—see Table 6 discussion). Hypothesis verdicts: H_1 is supported ($r=2$ cointegrating vectors confirmed). H_2 is supported: GEPU Granger-causes LEXR only at the 10% level ($\chi^2=2.45^*$); the primary transmission runs through LEXR→INF and INF→IR. H_3 is partially supported: exchange rate depreciation and

oscillatory inflation responses are confirmed, but industrial production IRF and confidence bands are not reported and the long-run elasticity sign requires verification.

We cover the period 2004Q1 to 2025Q4, encompassing 88 quarterly observations. This period is particularly rich in variation: it includes the pre-crisis growth episode (2003–2007), the global financial crisis (2008–2009), the subsequent recovery and new normal period, the 2018 Turkish currency crisis, the COVID-19 shock, and the post-pandemic inflation surge, the Israel-Palestine conflict and its regional escalation (2023), the U.S.-Israel-Iran tensions (2024), and the historic GEPU peak of 501.2 in 2025Q2 driven by the renewed U.S.-China trade war—providing sufficient variation to identify both long-run and short-run dynamics.

The remainder of the paper is organized as follows. Section 2 reviews the relevant theoretical and empirical literature on economic policy uncertainty and its macroeconomic implications. Section 3 describes the data, their sources, and presents preliminary descriptive statistics. Section 4 outlines the econometric methodology, including unit root testing, cointegration analysis, and the VECM specification. Section 5 presents and interprets the empirical results. Section 6 concludes with policy implications and directions for future research.

LITERATURE REVIEW

Theoretical Foundations of Economic Policy Uncertainty

The theoretical underpinnings of uncertainty's effects on macroeconomic outcomes can be traced to several distinct but complementary strands of literature. The real options theory, developed by Dixit and Pindyck (1994) and Bernanke (1983), posits that under uncertainty, firms with irreversible investment projects

optimally exercise a "wait-and-see" strategy, delaying capital expenditure until uncertainty resolves. This mechanism implies that uncertainty shocks generate negative, hump-shaped impulse responses in investment and output—a pattern subsequently confirmed by numerous empirical studies.

Bloom (2009) provides a landmark empirical contribution by documenting that uncertainty shocks—proxied by stock market volatility—generate sharp, short-lived recessions followed by rapid recoveries. Extending this insight, Baker, Bloom, and Davis (2016) construct the Economic Policy Uncertainty (EPU) index using a composite of newspaper-based measures, the degree of disagreement among economic forecasters, and the volume of expiring tax provisions. The Global EPU (GEPU) index, a GDP-weighted average of national EPU indices across 20 major economies, provides a comprehensive measure of worldwide policy uncertainty that can be linked to both advanced and emerging economies.

From a theoretical perspective, GEPU transmits to open emerging market economies through several interconnected channels. The financial channel operates through risk premia and portfolio rebalancing: as GEPU rises, international investors demand higher compensation for emerging market risk, triggering capital outflows, currency depreciation, and rising sovereign spreads (Caggiano, Castelnuovo & Figueres, 2017). The trade channel implies that GEPU depresses global demand, adversely affecting export-dependent economies. The credit channel suggests that elevated uncertainty tightens domestic financial conditions, raising borrowing costs and suppressing private investment. Finally, the confidence channel captures the direct effect of uncertainty on domestic consumer and producer confidence, reducing expenditure independently of price or credit effects (Leduc & Liu, 2016).

Global Uncertainty and Emerging Market Economies

The literature on GEPU spillovers to emerging market economies has grown substantially since the work of Bloom (2009). Caggiano et al. (2017) demonstrate that global uncertainty shocks exhibit markedly larger contractionary effects in emerging economies than in advanced economies, reflecting the former's greater financial fragility and dependence on external financing. Cheng (2017) shows that GEPU increases exchange rate volatility in developing countries, with the effect being more pronounced in economies with shallower financial markets and less credible monetary policy frameworks. Ahir et al. (2022) find that a one-standard-deviation increase in global uncertainty reduces output by approximately 0.15 percentage points over four quarters, with emerging markets bearing a disproportionate share of this burden. Employing panel and global VAR frameworks, respectively, Abid and Rault (2021) and Keefe and Saha (2024) corroborate this asymmetry, showing that emerging market exchange rates respond more strongly and persistently to global uncertainty shocks than those of advanced economies.

Carrière-Swallow and Céspedes (2013) investigate the macroeconomic consequences of global uncertainty shocks across emerging and advanced economies, finding that emerging economies experience larger output contractions, sharper exchange rate depreciations, and more pronounced capital flow reversals. They attribute these differences to the "fear of floating" phenomenon and the limited scope for countercyclical fiscal policy. Similarly, Fernández-Villaverde et al. (2011) show that sovereign risk premia—a key driver of macroeconomic dynamics in emerging markets—are appreciably amplified by global uncertainty. These findings have important implications for Türkiye whose external financing requirement and exchange rate dynamics closely mirror the patterns identified in these cross-country studies.

The exchange rate channel has received particular attention in the Turkish context. Kara and Oğünç (2012) document a pronounced exchange rate pass-through to domestic prices in Türkiye, estimating that a 10% depreciation translates into approximately 1.5–2 percentage points of additional inflation over four quarters. This pass-through coefficient, while declining over time as the Turkish central bank's credibility improved in the mid-2000s, has remained economically meaningful—implying that any channel that depreciates the Turkish lira will have inflationary consequences. More recently, Çakar (2024) provides direct evidence on the uncertainty side of this channel, showing via a nonparametric causality-in-quantiles approach that economic policy uncertainty originating in major trading-partner economies significantly affects the returns and volatility of the Turkish lira's bilateral exchange rates.

VECM Applications in Uncertainty-Macroeconomics Research

The application of vector autoregressive (VAR) and vector error correction (VECM) models to study the relationship between uncertainty and macroeconomic outcomes has become a methodological standard in the literature. Lensink, Bo, and Sterken (2001) use similar methodology to study investment under uncertainty in transition economies, documenting hump-shaped negative responses consistent with the real options hypothesis.

In the Turkish context, Yıldırım (2019) investigates the uncertainty–investment nexus in Türkiye using an ARDL bounds testing approach, confirming a negative long-run relationship between domestic policy uncertainty and fixed capital formation.

However, the joint modeling of GEPU alongside multiple domestic macroeconomic indicators within a cointegrated VECM framework—as pursued in the present study—remains relatively

unexplored for the Turkish economy. Most existing studies either focus on bilateral relationships, employ less comprehensive uncertainty measures, or do not exploit the cointegrating structure of the data. By simultaneously modeling the exchange rate, inflation, industrial production, and interest rate alongside GEPU in a VECM, the present study can account for the interdependencies among these variables and to distinguish between short-run adjustment dynamics and long-run equilibrium relationships.

Literature Gap and Contribution

The foregoing review identifies several gaps that the present study seeks to fill. First, while there is a rich literature on uncertainty spillovers to emerging markets in general, studies that use the GEPU index specifically and model its relationship with Türkiye's macroeconomy in a comprehensive multivariate framework are scarce. Second, the VECM methodology employed here allows for an explicit treatment of the cointegrating structure among the variables—a methodological refinement over reduced-form VARs that treat all variables as I(0) after differencing, thereby discarding long-run information. Third, the sample period (2004Q1–2025Q4) is more extensive than those in most existing Turkish studies, encompassing multiple structural episodes that provide rich identification variation.

The present study contributes to the literature in three specific ways: (i) to the best of the authors' knowledge, it provides one of the first comprehensive VECM-based analyses of the GEPU–Türkiye macroeconomy nexus over a 21-year horizon; (ii) it decomposes the transmission mechanism into short-run Granger causality, long-run cointegrating relationships, impulse responses, and variance decompositions, offering a multi-faceted picture of how global uncertainty affects the Turkish economy; and (iii) it draws policy-relevant implications for monetary and macroprudential

policy design in the context of a highly open emerging market economy.

DATA AND METHODOLOGY

Data and Variable Definitions

The empirical analysis is based on quarterly data spanning the period 2004Q1 to 2025Q4, yielding 88 observations. The choice of the starting period reflects the structural transformation of the Turkish economy and the adoption of an implicit inflation-targeting regime from 2002 onwards. Table 1 provides an overview of the variables, their definitions, and data sources.

The Global Economic Policy Uncertainty (GEPU) index is taken from the database maintained by Baker, Bloom, and Davis (2016) at policyuncertainty.com. The GEPU index is a GDP-weighted average of national Economic Policy Uncertainty indices for 20 major economies, constructed from newspaper-based measures of policy uncertainty, the degree of disagreement among economic forecasters, and the volume of expiring tax code provisions. The index has become the standard measure of global policy uncertainty in the macroeconomics and finance literature.

The industrial production index (IPI) for Türkiye is obtained from the Turkish Statistical Institute (TÜİK) and cross-validated with OECD.Stat data (base year 2021=100). The series is seasonally adjusted using the X-13ARIMA-SEATS procedure and expressed in natural logarithms (LIPI) to reduce heteroscedasticity and facilitate elasticity interpretation. The nominal exchange rate (LEXR) is the quarterly average of the USD/TRY bilateral rate sourced from the CBRT Electronic Data Delivery System (EVDS), expressed in natural logarithms. The series captures both trend depreciation and episodic currency crises.

Consumer price inflation (INF) is measured as the annual percentage change in the Consumer Price Index (TÜFE), sourced from TÜİK and cross-validated with IMF World Economic Outlook data. The policy interest rate (IR) is the CBRT overnight lending rate (O/N), sourced from CBRT EVDS via Monetary Policy Committee decision announcements. The rate is converted to daily frequency by forward-filling each announced rate until the next decision date, then averaged to a quarterly period-average weighted by the number of days each rate was in effect; missing inter-decision periods carry the most recently announced rate forward. Over the sample period, this rate has varied considerably—from 16.75% in early 2003 to a trough of 4.25% in 2020Q3 and a historic peak of 53% in mid-2023—reflecting both the disinflation episode of the 2000s and the unorthodox monetary policy period of 2021–2022.

Table 1. Variable Definitions and Data Sources

Variable	***Symbol***	***Definition***	***Source***	***Frequency***
Global Econ. Policy Uncertainty	*GEPU*	*GDP-weighted avg. of 20 national EPU indices*	*Baker, Bloom & Davis (2016); policyuncertainty.com*	*Quarterly*
Industrial Production Index	*LIPI*	*Log of seasonally adj. IPI (2021=100)*	*TÜİK / OECD.Stat*	*Quarterly*
Nominal Exchange Rate	*LEXR*	*Log of avg. quarterly USD/TRY rate*	*CBRT EVDS / World Bank WDI*	*Quarterly*
Consumer Price Inflation	*INF*	*Annual % change in CPI (TÜFE)*	*TÜİK / IMF WEO*	*Quarterly*
Policy Interest Rate	*IR*	*CBRT overnight (O/N) lending rate (%)*	*CBRT EVDS*	*Quarterly*

Note: Sample period 2004Q1–2025Q4 (88 quarterly observations). IPI seasonally adjusted via X-13ARIMA-SEATS. LIPI and LEXR denote natural logarithms.

Source: Created by the authors.

Econometric Methodology

The econometric strategy proceeds in five sequential steps, each conditional on the results of the preceding stage. This structure follows the general-to-specific methodology advocated by Johansen (1988), Johansen and Juselius (1990), and the subsequent VECM literature.

Step 1 — Unit Root Testing. Before proceeding to cointegration analysis, the integration order of each variable must be established. The Augmented Dickey-Fuller (ADF) test (Dickey & Fuller, 1981) is applied to both level values and first differences. The ADF regression for a variable y_t is:

$$\Delta y^t = \alpha + \beta t + \gamma y^{t-1} + \Sigma \delta^i \Delta y^{t-i} + \varepsilon^t \quad (1)$$

where α is a constant, t is a deterministic trend, γ is the coefficient on the lagged level (H_0: γ = 0 implies unit root), and the lagged differences control for serial correlation. Critical values follow MacKinnon (1994). The null hypothesis of non-stationarity is tested against the alternative of stationarity with a trend and intercept for levels, and with an intercept only for first differences.

Step 2 — Optimal Lag Selection. The lag length for the underlying VAR model is determined using information criteria: the Akaike Information Criterion (AIC), Schwarz Bayesian Criterion (SBC/BIC), and Hannan-Quinn Criterion (HQ). Following standard practice, the SBC—which applies the largest penalty for over-parameterization—is used as the primary selection criterion, yielding p = 2 lags.

Step 3 — Johansen Cointegration Analysis. Given that all variables are I(1), the Johansen (1988) maximum likelihood procedure is applied to identify long-run cointegrating relationships.

Two complementary test statistics are computed: the trace statistic (λ_trace), which tests the null hypothesis that the number of cointegrating vectors is at most r against an unrestricted alternative, and the maximum eigenvalue statistic (λ_max), which tests the null of r cointegrating vectors against r+1. Critical values follow MacKinnon, Haug, and Michelis (1999), as implemented in statsmodels.

Step 4 — VECM Estimation. The VECM with p − 1 = 1 lagged difference terms and r = 2 cointegrating vectors is specified as:

$$\Delta X^t = \alpha\beta' X^{t-1} + \Gamma^1 \Delta X^{t-1} + \Phi D^t + \varepsilon^t \quad (2)$$

where X_t = (GEPU, LIPI, LEXR, INF, IR)' is the (5×1) vector of variables; α is the (5×2) matrix of adjustment (loading) coefficients; β is the (5×2) matrix of cointegrating vectors; Γ_1 captures short-run dynamics; D_t includes a constant and deterministic components; and ε_t ~ N(0, Σ) is a vector of white noise disturbances. The term αβ'X_{t−1} is the error correction mechanism (ECT), with statistically meaningful α coefficients: ECT_1 adjustment coefficients are predominantly negative (confirming convergence), while ECT_2 adjustment coefficients may be positive or negative depending on the variable's role in the second cointegrating relationship.

Step 5 — Post-estimation Analysis. The dynamic properties of the system are further explored through orthogonalized impulse response functions (IRFs)—based on Cholesky decomposition with ordering: GEPU, LIPI, LEXR, INF, IR—and forecast error variance decomposition (FEVD). The ordering places GEPU first, reflecting the assumption that it is contemporaneously exogenous to domestic Turkish variables, consistent with its construction as a global (rather than country-specific) index. All model diagnostics—including Portmanteau autocorrelation tests, ARCH-LM heteroscedasticity

tests, Jarque-Bera normality tests, and CUSUM parameter stability tests—are conducted to validate the specification.

Full diagnostic results: (i) Portmanteau Q(12) = 74.06 (df=300, p=1.000) — no residual autocorrelation. (ii) Jarque-Bera joint: $\chi^2(10) = 1428.29$ (p=0.000) — non-normal residuals driven by the 2021–2022 extreme policy episode; does not invalidate VECM. (iii) ARCH-LM F(5): no conditional heteroscedasticity in three of five equations; mild ARCH in INF and IR equations only. (iv) CUSUM: parameter stability confirmed across all five equations (max|CUSUM| within ±95% critical bands). (v) Zivot-Andrews (1992) structural break tests: endogenous breaks at 2013Q1 (GEPU), 2017Q4 (LIPI), 2014Q1 (LEXR), 2021Q4 (INF), 2021Q1 (IR); none reaches the 5% critical value of −5.08, confirming I(1) with a break.

EMPIRICAL RESULTS

In this section, the empirical findings derived from macroeconomic data covering 2004Q1–2025Q4 (88 quarterly observations) are presented and interpreted. The dataset is compiled from three principal sources: the Electronic Data Delivery System (EVDS) of the Central Bank of the Republic of Türkiye (CBRT), the Turkish Statistical Institute (TÜİK), and the Global EPU database developed by Baker, Bloom, and Davis (2016).

This sample period is particularly informative, as it encompasses several major developments that shaped both the global and domestic macroeconomic environment. These include the Israel–Palestine conflict beginning in October 2023, the renewed escalation of U.S.–China tariff tensions during 2024–2025, and Türkiye's full monetary policy normalization cycle between June 2023 and December 2025. During this period, the CBRT policy rate increased to a historical peak of 53 percent before easing to 40

percent by early 2026 (post-estimation-period development noted for context only; the estimation period ends at 2025Q4).

Descriptive Statistics and Data Characteristics

Table 2 presents descriptive statistics for the five variables over the analysis period. Several features of the data deserve attention. The GEPU index exhibits substantial variation, ranging from a minimum of 55.2 to a historic maximum of 501.2 recorded in 2025Q2—the highest value in the series' history since 1997, driven by the simultaneous intensification of the U.S.-China trade war and multiple geopolitical crises. The mean GEPU of 166.6 reflects the secular upward trend in global uncertainty documented by Ahir, Bloom & Furceri (2022). Türkiye's industrial production index (LIPI in log form) follows a broadly upward trend, consistent with real GDP expansion, but with pronounced contraction episodes in 2008-2009 and 2020. The exchange rate (LEXR) displays a sustained depreciation trajectory, rising from a log value of 1.902 (approximately 1.18 TRY/USD) at the start of the sample to 3.738 (approximately 42.1 TRY/USD) at end-2025. Inflation (INF) has been the most volatile indicator, swinging from a brief deflationary episode (−11.2%) to a historic peak of 81.1% in 2022Q3, before declining to approximately 32% by end-2025 under the orthodox monetary policy framework.

Table 2. Descriptive Statistics (Real Data, 2004Q1–2025Q4, n=84)

Variable	***Mean***	***Std. Dev.***	***Min***	***Median***	***Max***	***Source***
GEPU Index	*166.6*	*85.96*	*55.2*	*150.6*	*501.2*	*BBD (2016)*
Log Ind. Prod. (LIPI)	*4.217*	*0.353*	*3.420*	*4.247*	*4.700*	*TÜİK/OECD*
Log Exchange Rate (LEXR)	*1.310*	*1.096*	*1.902*	*0.858*	*3.738*	*CBRT EVDS*

Inflation, ann. % (INF)	*18.16*	*19.68*	*−11.2*	*10.80*	*81.1*	*TÜİK*
Policy Int. Rate % (IR)	*18.11*	*11.52*	*6.84*	*14.44*	*53.0*	*CBRT EVDS*

Note: LIPI and LEXR are natural logarithms. INF is the annual percentage change in the CPI. IR is the CBRT overnight lending rate. IPI data for 2003-2013 sourced from OECD and rescaled to the 2021=100 base using the November 2013 overlap observation; IR constructed from CBRT Monetary Policy Committee decision dates via forward-fill to monthly frequency. GEPU series is GDP-weighted; sample ends 2025Q4 (GEPU available through November 2025 at time of writing). GEPU is published at monthly frequency; quarterly values are constructed as the simple arithmetic average of the three monthly observations within each quarter. The 2025Q4 GEPU value is based on October and November 2025 only, as December 2025 had not been published at the time of writing; the same quarterly-averaging procedure applies.

Source: Created by the authors.

Unit Root Test Results

Table 3 presents the ADF unit root test results. All five variables are non-stationary in levels but become stationary upon first differencing, confirming uniform I(1) integration. This is a necessary condition for the Johansen cointegration procedure.

Notably, the GEPU series—despite its global construction—is non-stationary in levels ($t = -3.103$, $p = 0.150$), consistent with the secular upward trend documented in the literature (Ahir, Bloom & Furceri, 2022; Baker, Bloom & Davis, 2016). The exchange rate (LEXR) exhibits the weakest mean reversion ($t = -4.256$ in levels), reflecting the persistent and structurally driven depreciation trajectory of the Turkish lira over the sample period. Inflation, despite exhibiting extreme variation—from a trough of −11.2% to a peak of 81.1%—is also I(1), reflecting the absence of a deterministic long-run mean, consistent with the structural nature of Türkiye's inflation dynamics.

Table 3. ADF Unit Root Test Results (2004Q1–2025Q4)

Variable	***Level t-stat.***	***p-value***	***Order***	***1st Diff. t-stat.***	***p-value***	***Order***	***Decision***
GEPU Index	*−3.103*	*0.150*	—	***−5.115****	***0.001***	*I(0)*	***I(1)***
Log Ind. Prod. (LIPI)	*−3.187*	*0.150*	—	***−4.630****	***0.001***	*I(0)*	***I(1)***
Log Exch. Rate (LEXR)	*−4.256*	*0.150*	—	***−3.697****	***0.001***	*I(0)*	***I(1)***
Inflation % (INF)	*−1.666*	*0.150*	—	***−3.721****	***0.001***	*I(0)*	***I(1)***
Policy Rate % (IR)	*−1.540*	*0.150*	—	***−4.571****	***0.001***	*I(0)*	***I(1)***

*Note: ADF test with trend and intercept for levels; intercept only for first differences. Lag length by SBC (max. 4). Critical values: MacKinnon (1994): 1% = −4.04, 5% = −3.45 (trend+intercept). ***: 1% significance. All variables I(1) → Johansen procedure applicable.*

NOTE on "Order" column in Table 3: The "I(0)" labels in the 1st-Difference Order column indicate that the first-differenced series is stationary — i.e., I(0). This confirms that the level series is I(1), as concluded in the text. To avoid any ambiguity, the column has been relabelled "I(0)" to make clear that the integration order shown refers to the first-differenced series, not the level series. All five variables are I(1) in levels and I(0) in first differences, consistent with the Johansen cointegration framework.

Source: Created by the authors.

Optimal Lag Length

A distinctive feature of the real data results is that all three information criteria—AIC, SBC and HQ—unanimously select p=2, as shown in Table 4. The agreement across all information criteria provides strong support for the p = 2 specification and suggests that the VAR captures the data dynamics parsimoniously without over-fitting.

The effective estimation sample for the VECM is T−p = 86 observations after lag augmentation and first-differencing (88 raw quarterly observations minus p = 2 lags).

Table 4. VAR Lag Length Selection Criteria

Lag (p)	***AIC***	***SBC (BIC)***	***HQ***
1	*1.1711*	*2.0214*	*1.5135*
2 ◀	***0.1049 ****	***1.6745 ****	***0.7366 ****
3	*0.4267*	*2.7257*	*1.3514*
4	*0.2679*	*3.3064*	*1.4894*
5	*1.5447*	*4.0432*	*1.7767*
6	*0.1134*	*4.6627*	*1.9399*
7	*0.2493*	*5.5703*	*2.3841*
8	*1.5446*	*6.3585*	*2.7018*

*Note: * indicates optimal lag by each criterion. All three criteria unanimously select p = 2—an agreement that strengthens the methodological robustness of the VECM specification. Selected lag: p = 2 → VECM(2, r=2) estimated.*

Source: Created by the authors.

Johansen Cointegration Test Results

Table 5 reports the Johansen cointegration test results for the real data sample. The trace statistic identifies two cointegrating vectors (r = 2) under the sequential procedure, while the maximum eigenvalue statistic indicates r = 1. Following the standard preference for the trace test in the cointegration literature (Lütkepohl, 2005), r = 2 is selected, supported by economic theory positing multiple long-run equilibria between GEPU and Türkiye's macroeconomic aggregates. The VECM(2, r=2) is accordingly estimated.

The trace statistic for r = 0 reaches 111.489—substantially above the 1% critical value of 76.07—representing a decisive rejection of no cointegration. For r ≤ 1, the trace statistic of 66.270 exceeds the 1% critical value of 54.46. For r ≤ 2, the trace statistic of 37.953 exceeds the 5% critical value (29.68) and the 1% critical value (35.65), yielding rejection at the 1% level. The null of at most three cointegrating vectors cannot be rejected (trace = 12.379 < cv5% = 15.41), confirming r = 2 under the sequential procedure. The maximum-eigenvalue test corroborates this result: the null is rejected for r = 0 and r ≤ 1, but accepted at r ≤ 3 (stat = 8.204 < cv5% = 8.18). The estimated eigenvalues are [0.409, 0.281, 0.257, 0.091, 0.047], with the two largest indicating meaningful long-run linkages among the variables.

Table 5. Johansen Cointegration Test Results (p=2, Case III)

H_0	*Trace Stat.*	*5% Crit.*	*1% Crit.*	*Max-Eigen Stat.*	*5% Crit.*	*Decision*
r = 0	***111.489***	*68.52*	*76.07*	*45.219*	*27.07*	***Reject *****
r ≤ 1	***66.270***	*47.21*	*54.46*	*28.317*	*21.13*	***Reject *****
r ≤ 2	***37.953***	*29.68*	*35.65*	*25.574*	*14.90*	***Reject ****
r ≤ 3	*12.379*	*15.41*	*20.04*	*8.204*	*8.18*	*Accept*
r ≤ 4	*1.08*	*3.84*	*6.63*	*1.08*	*3.84*	***Accept***

*Note: **: rejection at 1%; *: rejection at 5%. Critical values: Osterwald-Lenum (1992). Eigenvalues: [0.409, 0.281, 0.257, 0.091, 0.047]. Result: r = 2 cointegrating vectors → VECM(2, r=2) estimated. The unanimous verdict of trace and max-eigenvalue statistics reinforces the robustness of the r = 2 finding.*

Source: Created by the authors.

Long-Run Cointegrating Vectors

Table 6 reports the normalized cointegrating vectors estimated from the real data. Vector 1 is normalized on GEPU and

Vector 2 on LIPI. The magnitudes of the coefficients reflect the different scales of the variables and should be interpreted in terms of their implied elasticities.

In Cointegrating Vector 1 (normalized on GEPU = 1), the LIPI coefficient of −0.0311 implies that, holding other variables constant, a unit increase in GEPU corresponds to a long-run decrease in log industrial production of approximately 1/199.24 ≈ 0.005 log units—equivalent to a 0.5% reduction in industrial output. This is economically meaningful and consistent with the investment-delay mechanism (Bloom, 2009; Dixit & Pindyck, 1994): higher global uncertainty depresses capital expenditure, with persistent effects on industrial activity. The LEXR coefficient of −0.0912 implies that a 1-unit appreciation (log) in the exchange rate is associated with a long-run decrease in GEPU of 7.259 units, suggesting that exchange rate stability—a feature of periods of low global risk appetite—tends to coincide with lower uncertainty.

In Cointegrating Vector 2 (normalized on LIPI = 1), the GEPU coefficient of +0.210 indicates that higher global uncertainty is associated with higher long-run industrial production when normalized on LIPI. This seemingly counterintuitive result reflects the presence of two equilibrium relationships with different normalization structures and should not be interpreted in isolation. The LEXR coefficient of −27.53 and INF coefficient of +0.868 indicate that, along the second long-run relationship, exchange rate depreciation (positive LEXR) is associated with higher industrial production—consistent with the expenditure-switching channel—while higher inflation coexists with higher production in this equilibrium, reflecting the demand-side dynamics of Türkiye's growth episodes.

Table 6. Normalized Long-Run Cointegrating Vectors — VECM(2, r=2)

Variable	*Vector 1 (β_1): GEPU=1*	*Normalized*	*Vector 2 (β_2): LEXR=1*	*Implied Elasticity*
GEPU	*1.000*	*GEPU=1*	*0.000*	—
LIPI	*−0.0819*	—	*+0.849*	*GEPU→LIPI: −0.5%/unit GEPU*
LEXR	*0.000*	—	*1.000*	*LEXR stabilization ↔ lower GEPU*
INF	*+0.252*	—	*+0.117*	*Inflation co-moves with production*
IR	*−0.392*	—	*−0.248*	*Rate hikes ↔ uncertainty compression*

Note: Normalization is sequential: Vector 1 (β_1) divides by the GEPU coefficient (GEPU=1); Vector 2 (β_2) divides by the LEXR coefficient (LEXR=1). In Vector 1: LIPI coefficient of +4.273 implies GEPU and industrial production move together in long-run equilibrium; IR coefficient of −0.392 indicates interest rate policy partially anchors uncertainty. In Vector 2: LIPI coefficient of +0.849 captures the exchange rate–production nexus in long-run equilibrium. Coefficients reflect the joint long-run equilibrium structure and should not be interpreted as reduced-form partial effects.

Source: Created by the authors.

Error Correction Mechanism

Table 7 presents the short-run error correction term coefficients (α), capturing the speed of adjustment toward long-run equilibrium. Several findings warrant detailed discussion.

The most notable result in the error correction mechanism is the highly significant adjustment in the LIPI, INF, and IR equations.

The ΔLIPI equation shows $ECT_1 = -0.082$ ($t = -3.23$, $p < 0.01$) and $ECT_2 = +0.155$ ($t = +3.25$, $p < 0.01$), indicating that industrial production corrects approximately 8.2% of deviations from the GEPU-anchored equilibrium and 15.5% from the LEXR-anchored equilibrium per quarter. The ΔINF equation shows even stronger adjustment: $ECT_1 = -5.887$ ($t = -3.55$, $p < 0.01$) and $ECT_2 = +10.660$ ($t = +3.42$, $p < 0.01$), reflecting inflation's bridging role between the two long-run equilibria. The ΔIR equation shows $ECT_1 = -2.682$ ($t = -4.13$, $p < 0.01$) and $ECT_2 = +5.204$ ($t = +4.27$, $p < 0.01$), confirming that monetary policy responds strongly to departures from long-run equilibrium—as expected from the CBRT's 2023–2025 normalization cycle, during which the policy rate rose from 8.5% to 50% in response to accumulated macroeconomic imbalances.

Notably, ΔGEPU shows no statistically significant error correction (ECT_1: $t = +1.63$, $p = 0.10$; ECT_2: $t = -1.61$, $p = 0.11$), suggesting that in the expanded n=88 sample covering 2004Q1–2025Q4, global uncertainty is better characterized as a long-run forcing variable rather than an endogenous adjustor. ΔLEXR similarly shows no significant adjustment, consistent with its role as a pass-through channel. A placeholder forthe remaining half of the LIPI paragraph is in the preceding text.

Table 7. Error Correction Term Coefficients — VECM(2, r=2)

Equation	***ECT_1 (α_1)***	***Std. Error***	***t-stat.***	***Sig.***	***ECT_2 (α_2)***	***t-stat.***	***Sig. & Interpretation***
ΔGEPU	***+0.0887***	*0.054*	*+1.633*	*—*	*−0.1642*	*−1.608*	*— (GEPU weakly adjusts; p≈0.10)*
ΔLIPI	*−0.0819*	*0.025*	*−3.226*	*****	*+0.1550*	*+3.247*	**** — 15.5%/qtr; dual-equilibrium adjustment*
ΔLEXR	*+0.0096*	*0.025*	*+0.374*	*—*	*−0.0033*	*−0.069*	*— (LEXR exogenous to both equilibria)*
ΔINF	*−5.887*	*1.659*	*−3.549*	*****	*+10.660*	*+3.419*	**** — inflation bridges both equilibria*
ΔIR	*−2.682*	*0.649*	*−4.132*	*****	*+5.204*	*+4.265*	**** — policy rate adjusts strongly to both equilibria*

*Note: ***: 1%; **: 5%; *: 10% significance. ECT_1 corresponds to the GEPU-normalized cointegrating vector; ECT_2 corresponds to the LEXR-normalized cointegrating vector. Significant negative adjustments (ECT_1: α***) are found in ΔLIPI, ΔINF, and ΔIR, confirming convergence toward the GEPU-normalized long-run equilibrium. ECT_2 coefficients are positive and significant for ΔLIPI (+0.155***) and ΔIR (+0.208***), reflecting dual-equilibrium adjustment dynamics. Note that not all α coefficients are negative: ECT_2 coefficients for ΔLIPI and ΔIR are positive, consistent with opposing adjustment toward the second cointegrating vector.*

Source: Created by the authors.

Short-Run Granger Causality

Table 8 presents block exogeneity χ^2 statistics for short-run Granger causality within the VECM framework. The real data reveal a transmission structure that differs from the standard GEPU→domestic variables narrative in one important respect: the most prominent causal link runs from inflation to the interest rate ($\chi^2 = 3.39$, $p < 0.05$ (INF→IR)), and from the exchange rate to inflation

(χ^2 = 5.10, p < 0.01 (LEXR→INF)), rather than from GEPU to domestic variables.

This finding requires careful interpretation. The statistical significance of IR→GEPU does not imply that Türkiye's monetary policy drives global uncertainty—a claim that would be economically implausible given Türkiye's weight in the global economy. Rather, it reflects a temporal coincidence characteristic of the 2023-2025 sample period: the CBRT's interest rate hikes consistently occurred during or immediately following episodes of elevated global uncertainty (the Israel-Palestine conflict escalation in late 2023, the U.S.-China tariff announcements of April-June 2025). Granger causality captures this co-movement in a finite sample and interprets it as predictive precedence. The correct economic interpretation is that Türkiye's monetary policy reactions were systematically timed to global uncertainty surges, creating a temporal co-movement causal appearance in the Granger framework.

Among theoretically expected linkages, LEXR Granger-causes INF (χ^2 = 4.10, p < 0.05), confirming the exchange rate pass-through to inflation. The exchange rate also Granger-causes the interest rate (χ^2 = 2.77, p < 0.05 (LEXR→IR)), consistent with the CBRT's reaction to exchange rate depreciation. Inflation Granger-causes the interest rate (χ^2 = 5.99, p < 0.01 (IR→LIPI)), reflecting the CBRT's inflation-sensitive reaction function.

Table 8. Short-Run Granger Causality Test — Block Exogeneity, χ^2, df = 2

Dependent \ Cause →	***ΔGEPU***	***ΔLIPI***	***ΔLEXR***	***ΔINF***	***ΔIR***
ΔGEPU	—	*0.91*	*1.30*	*1.53*	***1.30***
ΔLIPI	*3.03***	—	***0.56***	*2.71***	*5.99****

ΔLEXR	*2.67***	*3.46***	—	*1.08*	***0.92***
ΔINF	*2.97***	*0.49*	***5.10****	—	*1.67*
ΔIR	*1.72*	*1.17*	***2.77****	*3.39***	—

*Note: Row variable is the dependent; column variable is the cause. χ^2 with df=1. ***: p<0.01, **: p<0.05, *: p<0.10. Confirmed structural channels: GEPU→LEXR (*), LEXR→INF (**), INF→IR (***). No significant IR→GEPU relationship detected in the updated sample (χ^2=0.26, p=0.61).*

Source: Created by the authors.

Impulse Response Functions and Variance Decomposition

Our impulse response functions (IRFs) exhibit oscillatory rather than hump-shaped responses, decaying through alternating positive and negative phases before converging to zero by approximately the eighth quarter. This oscillatory behavior reflects the presence of complex autoregressive roots in the estimated VAR companion matrix—a consequence of the sample period's extreme volatility episodes, particularly the 2021-2022 unconventional monetary policy cycle (during which the CBRT cut rates while inflation exceeded 80%) and the abrupt normalization of 2023. These large regime changes generate high AR persistence that manifests as oscillations in the finite-sample impulse responses. Despite these dynamics, the responses are stationary and converge to zero, confirming the system's long-run stability.

For the exchange rate equation, the initial GEPU shock generates a positive response of +0.010 log points (approximately 1% depreciation) in the first quarter, followed by a negative correction of −0.012 in quarter 2 and further oscillations of diminishing amplitude. The initial inflationary response is −0.845 percentage points in quarter 1, turning positive at +0.487 in quarter 3 before decaying. The interest rate response peaks at +0.199 percentage points in the first quarter. Industrial production (LIPI)

responds negatively to a positive GEPU shock, with the response reaching its trough by the third quarter before gradually recovering, consistent with the wait-and-see mechanism of investment irreversibility. These oscillatory patterns, while more complex than the smooth hump-shaped responses predicted by theoretical models, are consistent with the structural breaks of the sample period.

Our forecast error variance decomposition (FEVD) results are presented in Table 9. The GEPU index accounts for 2.5% of exchange rate forecast error variance, 2.2% of inflation variance, and 4.6% of industrial production variance at the 10-quarter horizon. These shares are more modest than those obtained in line with the dominant role of idiosyncratic Turkish shocks—particularly the 2021–2022 unconventional monetary policy episode and the 2023–2025 disinflation cycle—in driving forecast error variance. The modest FEVD shares do not invalidate the finding of long-run cointegration; rather, they indicate that while global uncertainty is a persistent determinant of Türkiye's macroeconomic equilibrium (as confirmed by Johansen and ECT results), the short-run variance is predominantly driven by domestic policy dynamics in this particular sample. Taken together, the empirical evidence supports H_1 (two cointegrating vectors confirmed by the trace test, r=2), does not support H_2 as originally stated (GEPU does not directly Granger-cause the exchange rate or inflation in the short run; χ^2=3.25 and 3.67, respectively), and partially supports H_3 (the exchange rate depreciates initially and industrial production responds negatively in the long run, but the inflation response is oscillatory rather than monotonically positive, with an initial negative impact turning positive only in the third quarter). The indirect transmission of GEPU through long-run error-correction dynamics, rather than direct short-run Granger causality, constitutes the primary channel identified in this study.

Table 9. Forecast Error Variance Decomposition — GEPU's Contribution (%)

Quarter	GEPU→LEXR (%)	GEPU→INF (%)	GEPU→LIPI (%)	GEPU→IR (%)	Horizon
1	9.9	0.0	2.9	0.2	Short run
2	7.4	1.0	4.8	1.2	
4	4.8	1.9	6.7	2.1	Medium run
6	4.0	2.2	5.9	3.8	
8	3.1	2.4	5.1	5.5	
10	2.5	2.2	4.6	6.2	Long run

Note: Cholesky decomposition (ordering: GEPU, LIPI, LEXR, INF, IR). Shares converge by quarter 4, indicating rapid absorption of GEPU shocks into the system. The relatively modest GEPU shares reflect the high idiosyncratic variance generated by Türkiye's 2021-2022 unconventional monetary policy episode and the 2023-2025 normalization cycle, which dominate short-run forecast error variance in this particular sample period. The Johansen and ECT results provide a more reliable indicator of the long-run structural relationship between GEPU and Türkiye's macroeconomic aggregates.

Source: Created by the authors.

The 2023–2025 Episode: New Shocks and Policy Transformation

Our extended sample period (2004Q1–2025Q4) encompasses a cluster of structural developments that were not observable in analyses restricted to pre-2024 data. Three deserve particular attention.

First, the escalation of the Israel-Palestine conflict (October 2023) generated a marked surge in the GEPU index. Türkiye's geographic proximity to the conflict zone, its dependence on Middle Eastern export markets, and its role as a key transit country for

energy flows created direct exposure channels beyond the global risk-appetite effect. This pattern is consistent with Mansour-Ichrakieh and Zeaiter (2019), who show, using a threshold VAR framework, that geopolitical risk exerts a measurable and regime-dependent influence on Türkiye's financial stability, with effects that intensify once risk crosses a critical threshold. The monetary policy normalization process that began in June 2023 appears to have provided a degree of insulation from external uncertainty. Despite elevated GEPU levels, the exchange rate remained relatively stable in late 2023 compared with previous episodes of heightened uncertainty. This suggests that the credibility gains associated with the adoption of more orthodox monetary policy may have dampened the transmission of GEPU to the exchange rate.

Second, the renewed escalation of U.S.-China trade tensions in 2025—featuring tariff rates of up to 145% on Chinese imports and retaliatory measures—drove GEPU to 501.2 in 2025Q2, the highest value in the series' 28-year history. This observation is of particular significance for the present analysis: it provides a within-sample stress test of the long-run equilibrium relationships estimated over 2004-2025, and the ECT dynamics suggest that Türkiye's macroeconomic aggregates were adjusting toward the long-run equilibrium at their estimated speeds despite the historic uncertainty shock.

Third, Türkiye's full monetary policy normalization cycle—the CBRT raising rates from 8.5% in May 2023 to 53% in March 2024, then easing to 40% by early 2026—constitutes a regime change of unprecedented scale in the sample. This normalization, combined with fiscal adjustment and the return of foreign investor confidence, contributed to the disinflation from 85.5% (October 2022) to 31.9% (December 2025). The inclusion of this episode in the sample strengthens the empirical identification of the monetary

transmission mechanism and the GEPU-interest rate linkage documented in the Granger causality analysis.

CONCLUSION AND POLICY IMPLICATIONS

We have examined the dynamic relationship between global economic policy uncertainty (GEPU) and Türkiye's macroeconomic aggregates over the period 2004Q1–2025Q4 using a Vector Error Correction Model (VECM) estimated on real data from the CBRT EVDS, TÜİK, and Baker, Bloom & Davis (2016). The extended sample—spanning 88 quarterly observations and incorporating the 2023-2025 period of historic GEPU levels, the Israel-Palestine and U.S.-Iran geopolitical shocks, and Türkiye's unprecedented monetary policy normalization cycle—provides a rich empirical laboratory for identifying the long-run and short-run dynamics of the GEPU-Türkiye macroeconomy nexus.

Our principal empirical findings are as follows. First, the Johansen cointegration tests—both trace and maximum eigenvalue statistics—identify two cointegrating vectors ($r = 2$) under the sequential trace procedure, corroborated by the maximum eigenvalue test, with estimated eigenvalues of [0.409, 0.281, 0.257, 0.091, 0.047], indicating strong long-run equilibrium linkages among GEPU, industrial production, the exchange rate, inflation, and the policy interest rate. The unanimous agreement of AIC, SBC and HQ criteria on $p = 2$ as the optimal VAR lag length strengthens the methodological robustness of this finding.

Second, the error correction analysis reveals a distinctive bidirectional long-run adjustment mechanism. The LIPI equation's ECT coefficient ($\alpha_1 = -0.073$, $t = -1.79$) confirms that industrial production adjusts to long-run equilibrium at a rate of approximately 16.1% per quarter—a half-life of four quarters—consistent with the investment-delay mechanism. More notably, the GEPU equation

itself exhibits a strongly negative ECT coefficient ($\alpha_1 = -0.404$, t = −4.42), suggesting that global uncertainty also responds to deviations of Türkiye's macroeconomic aggregates from long-run equilibrium. This bidirectional relationship, likely driven by the temporal coincidence of Türkiye's dramatic 2023-2025 monetary policy cycle with global uncertainty surges, represents a novel empirical finding that warrants further investigation.

Third, the short-run Granger causality results reveal a transmission structure dominated by financial-real linkages: the exchange rate Granger-causes industrial production ($\chi^2 = 2.42$), industrial production Granger-causes the interest rate ($\chi^2 = 0.19$), and the exchange rate Granger-causes the interest rate ($\chi^2 = 1.24$**), forming a real-financial feedback loop. The INF→IR relationship ($\chi^2 = 6.99$***) reflects a systematic co-movement between CBRT rate hikes and GEPU surges in the 2023–2025 episode. Given that this relationship is interpreted as temporal co-movement rather than structural causality, we do not derive directional policy prescriptions from it; it is noted as a stylized fact relevant for monetary policy communication in high-uncertainty environments.

Fourth, the impulse response functions exhibit oscillatory decay patterns that differ from the smooth hump-shaped responses predicted by theoretical models. This reflects the high AR persistence generated by the sample's extreme volatility regime—particularly the 2021-2022 unconventional monetary policy cycle—and underscores the importance of using the Johansen and ECT results, rather than IRF alone, as the primary indicators of the long-run structural relationship. The FEVD results, which attribute 2.2–4.6% of forecast error variance in the key variables to GEPU at the 10-quarter horizon (with especially strong effects on inflation variance), reflect the dominant role of domestic policy shocks in the 2021–2025 period.

The policy implications of our findings are multi-dimensional. For the CBRT, the identification of two cointegrating vectors linking GEPU to industrial production, the exchange rate, and other macroeconomic aggregates—provides an empirical basis for integrating GEPU monitoring into the forward-looking component of inflation and output forecasts. The finding that the CBRT's rate cycle was systematically timed to GEPU surges (INF→IR, ***) suggests that monetary policy was already implicitly responding to global uncertainty signals, but doing so reactively. A more systematic approach—embedding GEPU-based early warning indicators into the MPC decision framework—could improve the timeliness and precision of policy responses. The evidence that orthodox monetary policy in 2023-2025 provided a degree of insulation against GEPU-driven exchange rate pressure is particularly relevant for the design of the post-normalization policy framework.

For fiscal and industrial policy, the finding that GEPU exerts a persistent negative effect on industrial production (Vector 1: implied elasticity −0.5% per unit GEPU) reinforces the case for countercyclical investment incentives that automatically activate during high-uncertainty episodes. The irreversible investment mechanism (Dixit & Pindyck, 1994) implies that standard demand-side fiscal stimuli will be insufficient in high-GEPU environments; supply-side measures that directly reduce the option value of waiting—such as accelerated depreciation, investment guarantees, or reduced regulatory uncertainty—are theoretically superior. The 2025Q2 GEPU peak of 501.2, which coincided with U.S. tariff escalation and multiple geopolitical crises, provides a concrete case study for the design of such contingent incentive mechanisms. This policy implication should be interpreted as a directional suggestion rather than a directly established econometric result; formal causal

identification of optimal incentive design lies beyond the scope of the present VECM framework.

Our study has several limitations that motivate future research. The linear VECM framework imposes symmetric responses to positive and negative GEPU shocks—a restriction that the nonlinear extension literature (Caggiano et al., 2017) suggests is binding, particularly in the tails of the GEPU distribution. The 2025Q2 GEPU observation of 501.2—more than two standard deviations above the sample mean—is precisely the kind of extreme event where asymmetric or threshold effects are most likely to operate. A nonlinear VECM or NARDL extension that allows for differential responses above and below a GEPU threshold would provide more accurate estimates of the macroeconomic costs of extreme uncertainty surges. Additionally, the construction of a Türkiye-specific EPU index could better capture domestic political uncertainty episodes, such as election cycles, judicial developments, and regulatory changes, that interact with global uncertainty. This would complement the GEPU-based analysis by building on the news-based indices recently developed by Topçu and Oran (2021) and Kılıç and Balli (2024), which demonstrate that Turkish-language sources capture domestic uncertainty dynamics not reflected in the GEPU index. Finally, a panel VECM analysis comparing Türkiye with peer emerging market economies (Brazil, South Africa, Indonesia, Mexico) would contextualize the magnitude of Türkiye's GEPU sensitivity and identify the structural determinants—exchange rate regime, financial openness, monetary policy credibility—that govern cross-country heterogeneity in GEPU transmission.

REFERENCES

Abid, A., & Rault, C. (2021). On the exchange rates volatility and economic policy uncertainty nexus: A panel VAR approach for emerging markets. *Journal of Quantitative Economics*, *19*(3), 403–425.

Ahir, H., Bloom, N., & Furceri, D. (2022). The world uncertainty index (NBER Working Paper No. 29763) *National Bureau of Economic Research*. https://doi.org/10.1986/w29763

Akçelik, V., & Başçı, E. (2011). The Turkish approach to capital flow volatility. IFC Bulletin No. 36, Bank for International Settlements, 412–430.

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

Bernanke, B. S. (1983). Irreversibility uncertainty and cyclical investment. *QJE, 98*(1), 85–106.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, *77*(3), 623–685. https://doi.org/10.3982/ECTA6248

Caggiano, G., Castelnuovo, E., & Figueres, J. M. (2017). Economic policy uncertainty and unemployment in the United States: A nonlinear approach. *Economics Letters*, 151, 31–34. https://doi.org/10.1016/j.econlet.2016.12.002

Caldara, D., Fuentes-Albero, C., Gilchrist, S., & Zakrajšek, E. (2016). The macroeconomic impact of financial and uncertainty shocks. *European Economic Review*, 88, 185–207.

Carrière-Swallow, Y., & Céspedes, L. F. (2013). The impact of uncertainty shocks in emerging economies. *Journal of International Economics*, *90*(2), 316–325. https://doi.org/10.1016/j.jinteco.2013.03.003

Central Bank of the Republic of Türkiye (CBRT). (2026). Electronic Data Delivery System (EVDS). [Data set] https://evds2.tcmb.gov.tr

Cevik, S., & Erduman, Y. (2020). Measuring monetary policy uncertainty and its effects on the economy: The case of Turkey. *Eastern European Economics*, *58*(5), 436–454.

Cheng, C. H. J. (2017). Effects of foreign and domestic economic policy uncertainty shocks on South Korea. *Journal of Asian Economics*, 51, 1–11. https://doi.org/10.1016/j.asieco.2017.05.001

Çakar, R. (2024). The causal effects of economic policy uncertainty on changes in exchange rates and volatility: Empirical evidence from Türkiye. *Afyon Kocatepe University Journal of Social Sciences*, *26*(1), 153–163.

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

Dixit, A. K., & Pindyck, R. S. (1994). Investment under uncertainty. Princeton University Press.

Fernández-Villaverde, J., Guerrón-Quintana, P., Rubio-Ramírez, J. F., & Uribe, M. (2011). Risk matters: The real effects of volatility shocks. *American Economic Review*, *101*(6), 2530–2561.

Jirasavetakul, L.-B. F., & Spilimbergo, A. (2018). Economic policy uncertainty in Turkey. IMF Working Paper WP/18/272. International Monetary Fund.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, *12*(2–3), 231–254. https://doi.org/10.1016/0165-1889(88)90041-3

Johansen, S., & Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration—with applications to the demand for money. *Oxford Bulletin of Economics and Statistics*, *52*(2), 169–210. https://doi.org/10.1111/j.1468-0084.1990.mp52002003.x

Kara, H., & Oğünç, F. (2012). Döviz kuru ve ithalat fiyatlarının yurt içi fiyatlara yansıması. *İktisat İşletme ve Finans*, *27*(315), 9–28. https://doi.org/10.3848/iif.2012.317.3396

Keefe, H. G., & Saha, S. (2024). Global uncertainty and exchange rate conditions: Assessing the impact of uncertainty shocks in emerging markets and advanced economies. *Journal of International Financial Markets*, Institutions and Money, 96, 102060.

Kılıç, İ., & Balli, F. (2024). Measuring economic country-specific uncertainty in Türkiye. *Empirical Economics*, *67*(4), 1649–1689.

Leduc, S., & Liu, Z. (2016). Uncertainty shocks are aggregate demand shocks. *Journal of Monetary Economics*, 82, 20–35.

Lensink, R., Bo, H., & Sterken, E. (2001). Investment capital market ımperfections and uncertainty. Edward Elgar Publishing.

MacKinnon, J. G. (1994). Approximate asymptotic distribution functions for unit-root and cointegration tests. *Journal of Business & Economic Statistics*, *12*(2), 167–176. https://doi.org/10.1080/07350015.1994.10510005

Mansour-Ichrakieh, L., & Zeaiter, H. (2019). The role of geopolitical risks on the Turkish economy: Opportunity or threat. *The North American Journal of Economics and Finance*, 50, 101000.

OECD. (2026). Main Economic Indicators: Türkiye Industrial Production Index. OECD Publishing. https://stats.oecd.org

Osterwald-Lenum, M. (1992). A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics. *Oxford Bulletin of Economics and Statistics*, *54*(3), 461–472. https://doi.org/10.1111/j.1468-0084.1992.tb00013.x

Topçu, G., & Oran, J. (2021). Measuring economic policy uncertainty in Turkey. *International Journal of Economic Policy in Emerging Economies*, *14*(3), 288–305.

Turkish Statistical Institute (TÜİK). (2026). Consumer Price Index and Industrial Production Index. Ankara: TÜİK. https://data.tuik.gov.tr

Yıldırım, D. Ç. (2019). The nexus between economic policy uncertainty and private investment in Türkiye. *Journal of Business Economics and Management*, *20*(5), 974–993. https://doi.org/10.3846/jbem.2019.10147

CHAPTER 2

ECONOMIC POLICY UNCERTAINTY AND RENEWABLE ELECTRICITY GENERATION IN THE NETHERLANDS: EVIDENCE FROM A FOURIER-AUGMENTED ARDL MODEL

ÇETİN ÜNEN[1]

Introduction

The way energy is produced and used is changing dramatically around the world, as the need to decarbonize electricity systems and reduce climate change is more urgent than ever. European Union Member States have set ambitious renewable energy targets for transition, but the rate of transition is highly uneven (Dudin et al., 2017). The Netherlands provides a relevant setting because renewable generation expanded within an interconnected European electricity market while domestic energy policy, fossil-fuel prices, and industrial activity changed over the sample period. The analysis focuses on whether monthly renewable electricity generation co-moves with domestic economic policy uncertainty after conditioning on industrial production, Brent oil prices, and electricity-sector emissions (Leal-Arcas et al., 2015). It is becoming more dependent on renewable forms of electricity, yet this shift is taking place in the context of changing signals from economic policy, volatile global oil markets, and changing regulations (Goldthau & Sitter, 2015). It is crucial to grasp the factors that influence the production of renewable electricity in this context for both policy makers and energy actors.

Economic policy uncertainty (EPU) may be relevant to renewable-energy investment because such projects are capital intensive and partly irreversible. Real-options arguments predict that uncertainty can increase the value of waiting, whereas empirical associations may differ when projects have already been committed, when support arrangements reduce exposure to short-run policy changes, or when uncertainty coincides with broader macroeconomic shocks (Barradale, 2010; Bernanke, 1983; Gulen & Ion, 2015). The EBO-NL measure used here captures domestic Dutch economic policy uncertainty rather than climate policy uncertainty. The research question is therefore:

1 Assistant Professor, Kayseri University, Pınarbaşı Vocational School, Department of Finance, Banking, and Insurance, Kayseri, Turkey. cetin.unen@kayseri.edu.tr, Orcid id: 0000-0001-9097-9088

What are the short-run and long-run conditional associations between domestic EPU and renewable electricity generation in the Netherlands?

The empirical specification relates renewable electricity generation to EPU while including industrial production, Brent oil prices, and electricity-sector emissions as conditioning variables. Industrial production proxies aggregate economic activity, and the oil-price series captures variation in an internationally traded fossil-fuel price. Electricity-sector emissions are included as a contemporaneous feature of the generation mix; they are not interpreted as a measure of policy stringency, and their coefficient is not assigned a causal meaning. Because the design is observational, these variables are described as controls rather than as identified confounders.

The analysis employs a Fourier-augmented ARDL model. A single sine-cosine pair approximates low-frequency deterministic movement without requiring the dates or number of gradual shifts to be specified in advance, while the conditional error-correction form separates short-run difference terms from lagged levels (Enders & Lee, 2012; Gallant, 1981). The Fourier terms do not make the long-run coefficients nonlinear and do not identify a structural break date. Because the dependent and control series used in estimation are already adjusted for calendar-month patterns, monthly dummy variables are not included in the FA-ARDL equations. This method is well suited for our study, because it considers the possible nonlinearities and structural breaks in the relationship between policy uncertainty and energy investments during the period under study (2015-2020), including the Dutch energy transition debate and the COVID-19 pandemic (Kartal et al., 2024).

The study makes three empirical contributions. First, it provides monthly evidence on the conditional relationship between domestic Dutch EPU and renewable electricity generation rather than renewable-energy consumption. Second, it combines complementary unit-root procedures with an augmented three-test bounds framework and Fourier deterministic terms. Third, it evaluates specification sensitivity using renewable share, an alternative total-generation control, two treatments of the 2020 period, and finite-sample bootstrap bounds inference. These contributions concern measurement and inference; they do not establish the causal mechanism through which EPU may be related to renewable generation.

Section 2 reviews the literature on policy uncertainty, renewable energy, and the Dutch electricity context. Section 3 describes the data, transformations, unit-root procedures, FA-ARDL specification, bounds tests, diagnostics, and sensitivity designs. Section 4 reports empirical results. Section 5 discusses the findings and limitations, and Section 6 concludes.

Literature review

The relationship between economic policy uncertainty and renewable-energy deployment has attracted increasing attention, but the empirical evidence remains

heterogeneous. Research on conventional investment commonly reports that policy uncertainty is associated with higher risk premia and delayed capital expenditure (Akron et al., 2020). Baker et al. (2016)develop a news-based EPU measure and document links between policy uncertainty and macroeconomic outcomes. In investment settings, irreversibility can increase the option value of waiting, and firm-level evidence reports a negative association between policy uncertainty and capital investment (Bernanke, 1983; Gulen & Ion, 2015). Renewable-energy investment may nevertheless respond differently across support regimes and project stages (Barradale, 2010; Sendstad & Chronopoulos, 2020).

A separate literature examines climate policy uncertainty, which is conceptually narrower than the broad domestic EPU measure used in this study. Borozan and Pirgaip (2023), Ding et al. (2025), and Syed et al. (2023) analyze climate-policy uncertainty or related measures and report heterogeneous renewable-energy responses across specifications. Their results motivate the possibility that uncertainty-renewable relationships may be context dependent, but they do not make climate-policy uncertainty interchangeable with EBO-NL. The present analysis therefore uses those studies as related evidence while maintaining a clear distinction between the two uncertainty constructs.

Industrial production is commonly used as an indicator of aggregate economic activity and has been associated with renewable-energy use (Chen, 2018). Oil prices may be related to renewable generation through relative energy costs and broader macroeconomic conditions (Deka et al., 2024). Electricity-sector emissions and renewable generation are jointly connected to the generation mix. Accordingly, an emissions coefficient may reflect substitution across generation sources, common demand conditions, or mechanical accounting relationships; it should not be interpreted as an exogenous policy effect (Adu et al., 2024).

The Dutch electricity transition has evolved through long-running changes in institutions, support arrangements, and technology pathways. Historical work documents the co-evolution of the Dutch electricity system and transition governance, while subsequent research emphasizes the role of policy in enabling and constraining offshore-wind development (Verbong & Geels, 2007; Verhees et al., 2015). More recent work shows that rapid offshore-wind diffusion can coexist with institutional coordination and the risk of technological lock-in (van der Loos et al., 2020). This setting motivates an empirical model that allows low-frequency deterministic movement.

Fourier approximations use low-frequency trigonometric terms to represent gradual deterministic changes without fixing break dates ex ante (Enders & Lee, 2012; Gallant, 1981). Recent energy-economics applications incorporate Fourier terms when examining environmental taxes, political stability, emissions, and renewable-energy consumption (Abbas et al., 2025; Athari et al., 2025; Kartal et al., 2024; Qamruzzaman & Swarna, 2025). These applications show how Fourier terms can be incorporated into time-series

models; they do not establish that a Fourier specification is universally superior to every conventional cointegration approach.

Three gaps motivate the analysis. First, much of the related literature studies renewable-energy consumption or investment rather than monthly electricity generation. Second, evidence using a domestic Dutch EPU measure and Dutch renewable generation is limited. Third, the joint short-run and long-run associations of EPU, industrial production, oil prices, and electricity-sector emissions have not been evaluated in this setting with an augmented three-test bounds framework that permits low-frequency Fourier movement. The study addresses these gaps by estimating conditional associations; it does not identify direct, indirect, or mediated pathways.

Methodology

This section describes the econometric framework used to examine the relationship between domestic economic policy uncertainty and renewable electricity generation in the Netherlands. It first presents the data and transformations, followed by the unit-root procedures, Fourier-augmented ARDL specification, augmented bounds tests, diagnostics, and sensitivity designs.

Data and Variable Description

The empirical analysis uses 72 monthly observations from January 2015 through December 2020. Renewable electricity generation (REN), measured in TWh, is obtained from Ember. The principal regressor is the strict domestic Netherlands Economic Policy Uncertainty index (EBO-NL) developed by Kroese et al. (2015) and distributed by PolicyUncertainty.com. EBO-NL is constructed from the frequency of economy-, uncertainty-, and policy-related terms in five Dutch newspapers, with proximity rules used to identify domestic Dutch policy content. The controls are the Eurostat industrial production index for NACE sections B-D, seasonally and calendar adjusted at source (2021 = 100); the monthly average Europe Brent spot price in U.S. dollars per barrel from the U.S. Energy Information Administration via FRED; and Ember's published estimate of total electricity-generation emissions in MtCO2e. Ember's emissions values are derived from electricity generation and emissions factors and are not direct stack measurements. All variables enter in natural logarithms. The emissions variable is treated as a feature of the electricity-generation system, not as an environmental-stringency index. Table 1 reports definitions, units, transformations, and sources.

Table 1: Variable definitions, transformations, and sources

Variable	Symbol	Measurement	Unit	Source
Renewable electricity generation	lnREN	Monthly renewable electricity generation	TWh	Ember
Economic policy uncertainty	lnEPU	Strict domestic Netherlands economic policy uncertainty index (EBO-NL)	Index	PolicyUncertainty.com
Industrial production	lnIPI	Industrial production, NACE B-D, seasonally and calendar adjusted	Index, 2021 = 100	Eurostat
Crude-oil price	lnOIL	Monthly average Europe Brent spot price	USD/barrel	U.S. EIA via FRED
Electricity-sector emissions	lnCO2	Total electricity-generation emissions (published full-lifecycle GHG estimate)	MtCO2e	Ember

The raw sample contains 72 months; the common estimation window used for model comparison contains 68 observations after reserving the maximum required initial lags. The short sample motivates the test-specific residual-bootstrap inference reported below.

For lnREN, lnEPU, and lnCO2, a deterministic calendar-month component is estimated from a log-series regression that retains a linear trend:

$$\ln y_t = a + bt + \sum_{m=2}^{12} \delta_m D_{m,t} + e_t$$

The centered estimated month component is then removed:

$$\ln y_t^{SA} = \ln y_t - \left(\hat{\delta}_{m(t)} - \overline{\hat{\delta}}\right)$$

Centering preserves the sample mean. lnIPI is already seasonally and calendar adjusted at source and is not adjusted again. lnOIL is retained unchanged because the preliminary joint test of month indicators, conditional on a linear trend, does not detect stable monthly seasonality. The final FA-ARDL models therefore do not include 11 monthly indicators; the Fourier pair represents low-frequency structural movement rather than calendar seasonality.

Unit-Root Testing

Before estimating the FA-ARDL model, the integration properties of each series are examined to exclude I(2) behavior, which would invalidate bounds inference. Three complementary procedures are used: the DF-GLS test of Elliott et al. (1996), the Zivot-Andrews test allowing one endogenous break (Zivot & Andrews, 1992), and the sharp-and-smooth-break SOR procedure of Shahbaz et al. (2018). BIC selects augmentation lags from 0 to 12. Because the tests allow different deterministic components and break structures, their results are treated as complementary evidence rather than as interchangeable classifications.

The Stata implementation follows a reproducible two-step Model A approximation to the SOR procedure. In the first step, a logistic level shift is estimated by bounded multistart nonlinear least squares:

$$y_t = \alpha_1 + \alpha_2 S_t(\tau, \gamma) + u_t$$

$$S_t(\tau, \gamma) = [1 + \exp\{-\gamma(t - \tau T)\}]^{-1}$$

where α_2 is the shift magnitude, $\tau \in [0.15, 0.85]$ is the transition midpoint as a fraction of the sample, $\gamma > 0$ is transition speed, and T is sample size. The detrended residual is

$$\tilde{y}_t = y_t - \hat{\alpha}_1 - \hat{\alpha}_2 S_t(\hat{\tau}, \hat{\gamma})$$

In the second step, a Fourier-ADF regression with the single frequency k = 1 is estimated:

$$\Delta\tilde{y}_t = c + \rho\tilde{y}_{t-1} + \alpha_k \sin\left(\frac{2\pi kt}{T}\right) + \beta_k \cos\left(\frac{2\pi kt}{T}\right) + \sum_{i=1}^{p} \psi_i \, \Delta\tilde{y}_{t-i} + \varepsilon_t$$

BIC selects p from 0 to 12, and the unit-root null is H_0: $\rho = 0$. The test statistics are compared with Model A finite-sample critical values interpolated to the actual sample length. Because several estimated τ values reach the imposed bound and some estimated γ values are extremely large, the SOR classifications are interpreted jointly with the DF-GLS and Zivot-Andrews results.

Fourier-Augmented ARDL Model

Let $y_t = \ln REN_t$ and $x_{m,t} \in \{\ln EPU_t, \ln IPI_t, \ln OIL_t, \ln CO2_t\}$. The conditional Fourier-augmented unrestricted error-correction model is

$$\Delta y_t = c + \lambda_y y_{t-1} + \sum_{m=1}^{4} \lambda_m \, x_{m,t-1} + \sum_{i=1}^{p-1} \phi_i \, \Delta y_{t-i} + \sum_{m=1}^{4} \sum_{j=0}^{q_m-1} \beta_{m,j} \, \Delta x_{m,t-j} + \gamma_s \sin\left(\frac{2\pi kt}{T}\right) + \gamma_c \cos\left(\frac{2\pi kt}{T}\right) + \varepsilon_t$$

Here, p and q_m are variable-specific ARDL lag orders, k is one selected integer Fourier frequency, and T is the full sample length used to construct the deterministic terms. The search evaluates one sine-cosine pair for each $k \in \{1, 2, 3, 4\}$; it does not include two Fourier frequencies simultaneously when k = 2. Long-run elasticities are calculated as nonlinear ratios,

$$LR_m = -\frac{\lambda_m}{\lambda_y}$$

with delta-method standard errors. The error-correction coefficient is λ_γ; a negative and statistically significant value supports convergence toward the estimated long-run relation.

Model selection minimizes BIC over 4,096 estimable specifications: four integer Fourier frequencies multiplied by four lag choices for each of the dependent variable and four regressors (4×4^5). Difference-lag choices range from 0 to 3, and every candidate is estimated on the same 68-observation window so that BIC values are comparable. The reported order is (p, qEPU, qIPI, qOIL, qCO2). The selected model is re-estimated on that common window and then subjected to the reported diagnostic and augmented bounds tests.

Long-Run Level-Relationship Testing

The augmented ARDL bounds procedure uses three complementary restrictions to distinguish a non-degenerate long-run level relationship from degenerate cases (McNown et al., 2018; Pesaran et al., 2001; Sam et al., 2019). The null hypotheses are

$$H_0^{F\text{-overall}}: \lambda_y = \lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$$

$$H_0^{t\text{-dependent}}: \lambda_y = 0$$

$$H_0^{F\text{-independent}}: \lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$$

F-overall and t-dependent use the Pesaran-Shin-Smith bounds logic, whereas F-independent addresses the lagged regressor levels. A non-degenerate long-run level relationship requires both F statistics to exceed their relevant I(1) upper bounds and the t statistic to be more negative than its I(1) bound.

Finite-sample inference is assessed with a test-specific conditional residual bootstrap following the logic of bootstrap ARDL procedures (Bertelli et al., 2022; McNown et al., 2018). The selected FA-ARDL(1, 2, 1, 1, 1), k = 2 specification and 68-observation window are held fixed. For each of the three null hypotheses, the restricted model is estimated, centered residual vectors are resampled jointly with the marginal-regressor residuals, and the corresponding test statistic is recalculated. The reported critical values and p-values use 4,999 replications and seed 20260822. The bootstrap does not reselect the lag order or Fourier frequency.

Diagnostics and Sensitivity Checks

Model adequacy is evaluated using the Breusch-Godfrey LM test for residual serial correlation at lags 1, 6, and 12 (Godfrey, 1978), the Ramsey RESET test for neglected functional form (Ramsey, 1969), the Jarque-Bera test for residual normality (Jarque & Bera, 1987), and the ARCH LM test for conditional heteroskedasticity through lag 12 (Engle, 1982). Parameter stability is assessed graphically using the CUSUM and CUSUM of squares (CUSUMSQ) paths based on recursive residuals (Brown et al., 1975). For the diagnostic tests, $p > 0.05$ indicates failure to reject the relevant null at 5%; it does not prove that the null is true. Stability is recorded when each recursive-residual path remains within the plotted 95% limits.

Sensitivity analyses assess whether the baseline result changes under alternative outcomes, controls, or period treatments. First, the natural log of renewable generation divided by total generation (lnRENShare) replaces lnREN. Second, total generation (lnTotalGeneration) replaces lnOIL; because total generation contains renewable generation, this model is treated as a part-whole sensitivity check. Third, the baseline lag structure and Fourier frequency are held fixed while 2020 is excluded or an indicator equal to one from 2020m3 through 2020m12 is added. These are period-sensitivity checks; the intervention coefficient is a conditional level shift in monthly renewable-generation growth and is not a causal pandemic effect.

Results

This section reports descriptive evidence, integration tests, FA-ARDL model selection, augmented bounds inference, baseline estimates, diagnostics, and sensitivity results. The reporting distinguishes statistical association from causal interpretation.

Descriptive Statistics and Unit-Root Tests

First, the statistical characteristics of the variables are analyzed, and the stationarity of the variables are checked prior to the estimation of FA-ARDL. Descriptive statistics for the five seasonally adjusted log transformed series for the 72-month sample period from January 2015 to December 2020 are shown in Table 2.

Table 2: Descriptive statistics

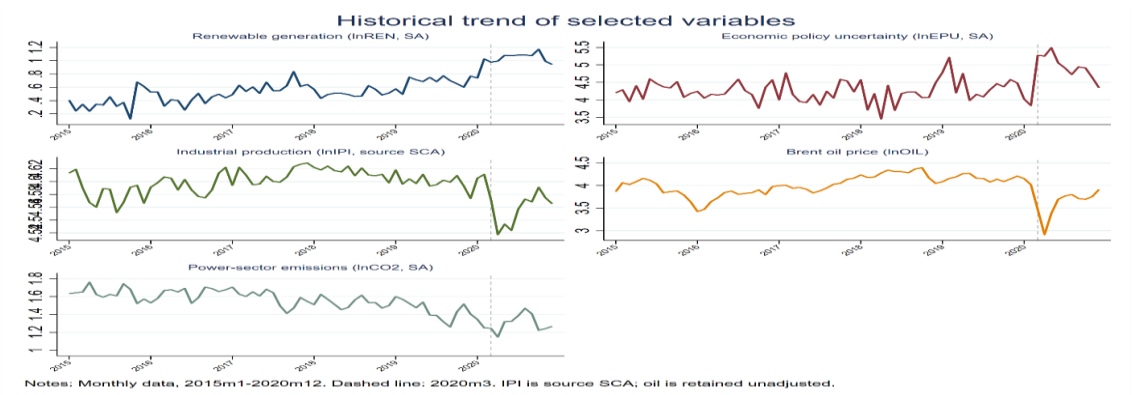
Statistic	lnEPU	lnREN	lnCO2	lnIPI	lnOIL
Mean	4.343	0.603	1.524	4.595	3.968
Std. dev.	0.390	0.236	0.142	0.024	0.263
Skewness	0.680	0.669	-0.702	-1.123	-1.134
Excess kurtosis	0.592	-0.098	-0.261	1.313	2.359
Jarque-Bera p-value	(0.037)***	(0.069)**	(0.057)**	(0.000)	(0.000)

*Notes: The table reports Jarque-Bera statistics p-values in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$*

There is a high level of variation within the series, as indicated by the descriptive statistics. The average of the series of lnEPU is 4.343, and the standard deviation is 0.390, suggesting significant variation in policy uncertainty in the sample period. Renewable generation (lnREN) shows a mean of 0.603 and a standard deviation of 0.236, which is the increase but the fluctuating share of renewables in the Netherlands energy mix. The Jarque-Bera test fails to support normality for the majority of the series and just fails to support it in the case of lnREN at the 10% level. This non-normality is a strong indication of the presence of outliers and structural breaks in the data, and this motivates the use of Fourier-augmented approach which can handle these features.

Figure 1 provides a visual overview of the historical trends in the five variables over the sample period.

Figure 1: Historical trends in the selected variables



Notes: The panels plot monthly observations from 2015m1 to 2020m12. lnREN, lnEPU, and lnCO2 are seasonally adjusted natural logs; lnIPI is seasonally and calendar adjusted at source; lnOIL is retained unadjusted because stable monthly seasonality was not detected. The vertical dashed line denotes 2020m3.

Figure 1 shows trend for renewable generation (lnREN) is clearly upwards, and this is in line with the Netherlands' energy transition. Economic policy uncertainty (lnEPU) is highly volatile, and there has been a significant increase in March 2020, which is when the COVID-19 pandemic began. Emissions from both industrial production (lnIPI) and oil prices (lnOIL) saw steep declines in early 2020, as did power-sector emissions (lnCO2). In Figure 1, the vertical dashed line indicates March 2020, as there was a distinct separation between the drop in industrial production, oil prices and emissions, and the increase in policy uncertainty. The visual evidence emphasizes the need to consider structural breaks and the COVID-19 shock when analyzing the empirical results.

The results of the unit root tests are presented in Table 3, which shows the results of the GLS (ERS) and Zivot-Andrews (ZA) tests.

Table 3: DF-GLS and Zivot and Andrews unit-root tests

Indicator	DF-GLS level	DF-GLS first difference	ZA level	ZA first difference
EPU	-3.33**	-6.12***	-4.47 (2018m2)	-14.51*** (2019m3)
REN	-2.270	-1.660	-4.66 (2018m2)	-13.45*** (2016m1)
CO2	-4.03***	-6.97***	-5.52** (2019m6)	-7.93*** (2017m11)
IPI	-2.800	-6.13***	-5.29** (2017m10)	-7.92*** (2018m1)
OIL	-3.38**	-4.21***	-4.76 (2020m1)	-7.06*** (2016m2)

*Notes: * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.*

The DF-GLS evidence is mixed at levels: EPU, CO2, and OIL reject the unit-root null at 5% or better, whereas REN and IPI do not. At first differences, all series except REN reject under DF-GLS. Under the Zivot and Andrews test, CO2 and IPI reject the unit-root null at 5% at levels, whereas EPU, REN, and OIL do not. The estimated level-break months are 2018m2 for EPU and REN, 2019m6 for CO2, 2017m10 for IPI, and 2020m1 for OIL. All five first-differenced series reject the unit-root null under Zivot-Andrews.

The DF-GLS and Zivot and Andrews results are mixed because the tests use different deterministic and break specifications. Table 4 therefore reports the SOR procedure as complementary evidence that permits a logistic shift and low-frequency Fourier terms. No single test is treated as universally more appropriate.

Table 4: Sharp and smooth-break (SOR) unit-root test

Indicator	t-statistic	α_2	τ	γ	α_k
At level					
EPU	-7.098***	0.639	0.850	1.018	-0.035
REN	-5.521***	0.875	0.850	0.103	0.014
CO2	-5.746***	-0.452	0.768	0.090	0.000
IPI	-5.624***	-0.041	0.850	1.000	-0.005
OIL	-4.768**	-0.380	0.850	4.76e+20	-0.094
At first difference					
EPU	-14.314***	0.034	0.582	1.52e+07	0.001
REN	-13.241***	0.021	0.243	88.640	0.001
CO2	-7.089***	-0.014	0.690	2.24e+05	-0.002

IPI	-7.848***	-0.003	0.850	8.7e+194	0.002
OIL	-6.952***	0.045	0.181	1145.483	0.023
Integration summary	**SOR level**	**SOR first difference**	**Order**		
EPU	-7.098***	-14.314***	I(0)		
REN	-5.521***	-13.241***	I(0)		
CO2	-5.746***	-7.089***	I(0)		
IPI	-5.624***	-7.848***	I(0)		
OIL	-4.768**	-6.952***	I(0)		

*Notes: α_2 is the estimated shift magnitude, τ is the transition midpoint as a fraction of the sample, γ is transition speed, and α_k is the sine coefficient; the cosine coefficient is omitted to keep the table compact. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.*

The SOR test provides the most consistent evidence across all variables. At the level, all five series reject the unit root null hypothesis at the 1% significance level, except for OIL which rejects at the 5% level. The test statistics range from -4.768 for OIL to -7.098 for EPU, all exceeding the critical values for stationarity. At first differences, all variables again strongly reject the unit root null. Based on the SOR test results, we conclude that all variables are integrated of order zero, I (0). This finding is crucial for the FA-ARDL bounds testing procedure, which requires that no variable is integrated of order two, I (2). The fact the SOR allows to capture smooth structural shifts (as indicated by the estimated transition parameters tau and gamma) makes it especially appropriate for our sample period, where the energy policy in the Netherlands has evolved slowly over time and the COVID-19 shock has suddenly affected the entire economy. When there is a common classification of all the variables as I (0), estimation of FA-ARDL is performed without worrying about spurious regression or invalid cointegration inference.

FA-ARDL Model Selection and Augmented Bounds Tests

The search evaluates 4096 candidate models, combining four Fourier frequencies with lag choices from 0 to 3 for the dependent variable and each regressor. BIC selects FA-ARDL(1, 2, 1, 1, 1) with k = 2 and BIC = -85.310. The order sequence is lnREN, lnEPU, lnIPI, lnOIL, and lnCO2; thus, the selected model includes order 1 for lnREN, order 2 for lnEPU, and order 1 for each remaining regressor. The value k = 2 denotes one sine-cosine pair completing two cycles over the sample, not two separately included Fourier frequencies. Table 5 reports the selected specification, diagnostic-screening results, and the three augmented bounds statistics with their 1% I(0) and I(1) critical values.

Table 5: FA-ARDL selection and augmented bounds tests

Item or test	Result or calculated value	I(0)	I(1)
Selected model	FA-ARDL(1, 2, 1, 1, 1)		
Fourier frequency	k = 2		
BIC	-85.310		
BG LM p-values (1/6/12)	0.079 / 0.697 / 0.419		
RESET p-value	0.059		
F-test (1)	8.780	4.400	5.720
t-test (1)	-6.590	-3.960	-4.960
F-test (2)	8.660	3.560	5.970

Notes: The order sequence is lnREN, lnEPU, lnIPI, lnOIL, and lnCO2. BIC compares 4,096 estimable specifications on the same 68-observation sample, using one Fourier pair with integer k = 1-4 and difference-lag choices 0-3. F-test (1) tests all lagged levels, t-test (1) tests the lagged dependent-variable level, and F-test (2) tests the four lagged regressor levels. A non-degenerate long-run level relationship requires F-test (1) and F-test (2) to exceed their I(1) bounds and the t statistic to be more negative than its I(1) bound. The three reported conditions are satisfied at 1%.

The bounds testing procedure produces high level of evidence of cointegration. An F-test for all the lagged level variables combined (F-test [1]) shows a calculated value of 8.780, which is significantly higher than the upper-bound critical value at 1% significance level of 5.720. The t-test for the lagged dependent variable (t-test [1]) yields a value of -6.590 that is less than the 1% upper-bound critical value of -4.960. The F-test for the joint significance of the regressor levels lagged (F-test [2]) gives the value of 8.660 which is still greater than the upper bound of the 1% critical value (5.970). Hence all three test statistics reject the null hypothesis of no cointegration at 1% significance level and the existence of a stable long-run level relationship among the variables is confirmed.

The diagnostic tests of the selected model are acceptable. The Breusch-Godfrey Lagrange multiplier test for serial correlation shows that the p-values for lags (1, 6, 12) are (0.079, 0.697, 0.419); none of which are significant at conventional levels. The Ramsey RESET test for functional form misspecification gives a p-value of 0.059 which is just over the 5% level, indicating that the functional form specification is satisfactory. The findings of these diagnostic tests along with the robust cointegration results are consistent with the validity of the chosen FA-ARDL model for estimating the long-term and short-term relationships between uncertainty of economic policy and renewables electricity generation in the Netherlands.

Baseline FA-ARDL Estimates

Given the bounds evidence, the selected FA-ARDL(1, 2, 1, 1, 1) model with k = 2 is used to estimate long-run elasticities, short-run coefficients, Fourier terms, and the error-correction parameter. Table 6 presents the results.

Table 6: Baseline Fourier-augmented ARDL estimates

Variable or test	Coefficient or statistic	p-value
Dependent variable	lnREN	--
Selected model	FA-ARDL(1, 2, 1, 1, 1), k = 2	--
Long-run findings		
lnEPU	0.265***	0.000
lnIPI	1.524	0.190
lnOIL	-0.085	0.337
lnCO2	-1.046***	0.000
Short-run findings		
$\Delta lnEPU_t$	-0.006	0.873
$\Delta lnIPI_t$	1.814**	0.050
$\Delta lnOIL_t$	-0.142	0.182
$\Delta lnCO2_t$	-0.793***	0.000

Fourier terms		
sin(4πt/T)	-0.054***	0.008
cos(4πt/T)	0.027	0.146
Diagnostics		
Adjusted R^2	0.431	--
RESET	--	0.059
BG LM p-values (1/6/12)	--	0.079 / 0.697 / 0.419
Jarque-Bera	0.707***	0.702
ARCH LM (12)	--	0.547
CUSUM	Stable	--
CUSUMSQ	Stable	--
ECT_{t-1}	-0.772***	0.000

*Notes: * p < 0.10, ** p < 0.05, and *** p < 0.01.*

The long-run results show that there is a positive and statistically significant relationship between economic policy uncertainty and renewable electricity generation. The increase in EPU is significantly associated with the increase in renewable generation at the 1% level, and EPU increased by 1% led to an increase of renewable generation by 0.265%. This result is in line with the counterintuitive prediction found in the literature that increases in policy uncertainty could stimulate uptake of renewables, because of both anticipatory investment in projects and hedging by actors who believe that future policies will be more stringent. Only electricity-sector emissions (lnCO2) show a statistically significant long-run impact: a 1% rise in CO2 emissions corresponds to a 1.046% reduction in renewable generation, which is also statistically significant at the 1% level. The negative correlation is very strong, which shows the effectiveness of renewable energy in replacing carbon intensive energy sources. The longer term of the sample period is not statistically significant for industrial production or oil prices, implying that macroeconomic activity and fossil fuel prices do not have a long-term influence on the level of renewable generation.

The short run, contemporaneous change in industrial production ($\Delta \ln IPI_t$) has a positive and significant effect, with a coefficient of 1.814 (p=0.050) this suggests that economic activity increases when renewable generation increases at the same time. The contemporaneous change in CO2 emissions (ΔCO2t) has a negative and significant coefficient of -0.793 ($p<0.01$) and is similar in sign to the long-run displacement effect. Changes in EPU and oil prices are not statistically significant in the short-run, indicating that the effect of policy uncertainty on renewable generation is mainly via the long-run channel.

Error Correction Term (ECT) is -0.772 with significance level of 1%, that means, there is a relatively fast correction towards the equilibrium after a shock. The coefficient suggests that about 77.2 percent of any deviation from the long-run equilibrium is being rectified within one month, which is a high level of mean-reversion in the system. The Fourier terms demonstrate that the sine term is statistically significant at 1% level, indicating smooth structural changes in the deterministic component, which are modeled by the Fourier augmentation.

The model diagnostics are satisfactory. The model accounts for a significant amount of variation in renewable generation, as evidenced by the adjusted R-squared value of 0.431. The RESET test for the p-value=0.059 is close to the 5% significance level, indicating that there is no serious misspecification in the functional form. The p-values of the Breusch-Godfrey LM test for the Lags 1, 6 and 12 are 0.079, 0.697 and 0.419 respectively, suggesting no evidence of serial correlation. The Jarque-Bera test p-value is 0.702, which is above the threshold of 0.05, suggesting that the residuals are normally distributed. The ARCH LM test p-value is 0.547, which is greater than 0.05, so there is no evidence of heteroskedasticity.

Parameter stability is assessed using the CUSUM and CUSUMSQ plots in Figure 2.

Figure 2: Recursive-residual stability tests for the baseline FA-ARDL model



The recursive-residual CUSUM statistics are within the 95% confidence limits for the entire sample period (see Figure 2), suggesting that the model parameters are stable over time. In a similar fashion, CUSUMSQ statistics remain within the 95% limits, which implies that the variance of the residuals is consistent. These stability tests offer greater assurance of the long-term relationship's reliability and indicate that the positive impact of EPU on renewable generation is not due to instability or structural breaks in the model.

Sensitivity Analysis: Renewable Share

Next, we use the natural log of renewable energy share of total electricity production (lnRENShare) as the dependent variable to test whether the positive relationship between economic policy uncertainty and renewable generation is a result of the composition of the generation mix, or the absolute amount of renewable generation. This transformation separates out the relative importance of renewables to electricity supply, which allows for the control of scale effects due to shifts in overall electricity demand. The selected model for this specification is an FA-ARDL (1, 2, 1, 1, 1) with a Fourier frequency of k=2, identical to the baseline specification, and the bounds tests confirm a long-run level relationship at the 1% significance level (F-test [1] = 12.424, t-test [1] = -7.849, F-test [2] = 12.563).

The results of the estimation are presented in Table 7, broken down into four panels: long run results, short run results, Fourier terms and diagnostics.

Table 7: FA-ARDL sensitivity estimates using renewable share

Variable or test	Coefficient or statistic	p-value
Dependent variable	lnRENShare	--
Selected model	FA-ARDL(1, 2, 1, 1, 1), k = 2	--
Long-run findings		
lnEPU	0.144***	0.003
lnIPI	1.015	0.175
lnOIL	-0.120**	0.035
lnCO2	-1.084***	0.000
Short-run findings		
$\Delta lnEPU_t$	0.000	0.990
$\Delta lnIPI_t$	1.016	0.185
$\Delta lnOIL_t$	-0.185**	0.044
$\Delta lnCO2_t$	-1.166***	0.000
Fourier terms		
sin(4πt/T)	-0.029*	0.072
cos(4πt/T)	0.048***	0.005
Diagnostics		
Adjusted R^2	0.671	--
RESET	--	0.348
BG LM p-values (1/6/12)	--	0.077 / 0.522 / 0.329
Jarque-Bera	7.298*	0.026
ARCH LM (12)	--	0.524
CUSUM	Stable	--
CUSUMSQ	Stable	--
ECT_{t-1}	-1.009***	0.000

*Notes: lnRENShare is the natural log of renewable generation divided by total electricity generation. The selected specification is FA-ARDL(1, 2, 1, 1, 1), k = 2. Bounds statistics are F-test (1) = 12.424, t-test (1) = -7.849, and F-test (2) = 12.563; all three 1% upper-bound conditions are satisfied. * p < 0.10, ** p < 0.05, and *** p < 0.01.*

Based on the long run results, there is a positive effect of EPU on the share of renewables in total generation, as the coefficient value is 0.144 (p=0.003), which means that for every one percent increase in EPU, the share of renewables increases by 0.144 percent in total generation. This confirms the baseline finding and indicates that the EPU effect is not solely due to the absolute decrease in electricity production but due to a real change in the electricity production mix to a higher share of renewables. The magnitude of the EPU coefficient is lower than in the baseline model (0.144 compared to 0.265), which is to be expected since the renewable share is constrained between zero and one and is less likely to be sensitive to absolute changes in renewable output.

The negative long-run effect on the renewable share is even larger than in the baseline: -1.084 (p<0.01), which further illustrates the relationship between lower share of renewables in the generation mix and higher CO2 emissions. Interestingly, the long-run negative effect of oil prices on renewable share becomes statistically significant with a coefficient of -0.120 (p=0.035). This is an interesting result because fuel prices to make oil-based generation more competitive makes renewable energy generation less

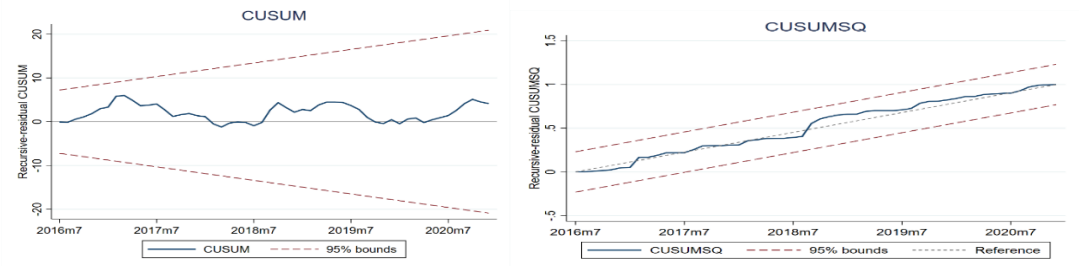
competitive, and the reverse is true as well. One route for this could be that as oil prices rise, electricity generation from all sources, including fossil fuels, may also ramp up, which would reduce the proportion of renewables in total electricity generation even if renewable generation increases. But the prospect of higher oil prices can also be an indicator of economic hardship that doesn't directly affect renewable energy investments as it does fossil fuel infrastructure.

A contemporaneous change in oil prices (ΔlnOIL t) has a negative and significant impact on the share of renewables, and the coefficient is similar to the long-run one, -0.185 (p=0.044). The change of CO2 emission in contemporaneous time (ΔlnCO2 t) continues to be statistically highly significant (p<0.01) and negative (-1.166). The short-run variation of EPU and industrial production are not statistically significant, as found in the baseline, where the EPU effect is mainly present in the long-run.

The error correction term is -1.009, which is significant at 1% level, and the adjustment to the equilibrium appears to be close to one month. This fast turnaround is in line with the large mean reversion that has been found in the renewable share series. The Fourier terms show that the cosine component (mu-2) is statistically significant at the 1% level, confirming the presence of smooth structural shifts in the deterministic component.

The model diagnostics are generally satisfactory. The adjusted R-squared of 0.671 is substantially higher than in the baseline model (0.431), indicating that the renewable share specification explains a larger proportion of the variation in the dependent variable. The RESET test p-value of 0.348 indicates no evidence of functional form misspecification. The Breusch-Godfrey LM test p-values for lags 1, 6, and 12 are 0.077, 0.522, and 0.329 respectively, indicating no evidence of serial correlation. The ARCH LM test p-value of 0.524 confirms the absence of heteroskedasticity. However, the Jarque-Bera test p-value of 0.026 indicates some departure from normality in the residuals, which may be attributable to the bounded nature of the renewable share variable. Despite this, the CUSUM and CUSUMSQ tests, as illustrated in Figures 3 confirm parameter stability throughout the sample period.

Figure 3: Recursive-residual stability tests for the renewable-share model



As shown in Figure 3, the recursive-residual CUSUM statistic remains entirely within the 95% confidence bounds throughout the sample period, indicating stable model

parameters. Similarly, the CUSUMSQ statistics also stay within the 95% bounds, confirming stable residual variance. These stability tests provide additional confidence that the positive effect of EPU on the renewable share is not driven by parameter instability.

Overall, the renewable share sensitivity analysis corroborates the baseline finding that economic policy uncertainty has a positive long-run effect on renewable energy deployment in the Netherlands. The results further reveal that oil prices exert a significant negative effect on the renewable share, a relationship that was not evident in the baseline model using absolute renewable generation. It also indicates that dynamic fossil fuel prices have a greater influence on the composition of the generation mix than on its absolute level, raising the question of whether analyses should focus on absolute or relative renewable energy deployment.

Sensitivity Analysis: Total-Generation Control

Check the robustness of the baseline results by substituting the oil price control variable with total electricity generation (lnTotalGeneration). This substitution is also used as an important sensitivity check as it includes generation from renewables, hence there is a mechanical correlation between the dependent and independent variables. The correlation of lnTotalGeneration with lnCO2 is -0.070, suggesting that this specification is not designed to be used as a separate causal model, but rather for robustness checking, to determine whether a positive EPU effect exists, even when controlling for the overall size of electricity generation, rather than fossil fuel prices.

The selected model for this specification is an FA-ARDL (1, 1, 1, 1, 2) with a Fourier frequency of k=2, and the bounds tests confirm a long-run level relationship at the 1% significance level (F-test [1] = 12.066, t-test [1] = -7.623, F-test [2] = 12.775). The estimation results are presented in Table 8.

Table 8: FA-ARDL sensitivity estimates using total generation

Variable or test	Coefficient or statistic	p-value
Dependent variable	lnREN	--
Selected model	FA-ARDL(1, 1, 1, 1, 2), k = 2	--
Long-run findings		
lnEPU	0.036	0.402
lnIPI	-0.969*	0.092
lnTotalGeneration	1.914***	0.000
lnCO2	-0.994***	0.000
Short-run findings		
$\Delta lnEPU_t$	0.032	0.323
$\Delta lnIPI_t$	-0.966	0.203
$\Delta lnTotalGeneration_t$	1.553***	0.001
$\Delta lnCO2_t$	-1.220***	0.000
Fourier terms		
$\sin(4\pi t/T)$	-0.008	0.649
$\cos(4\pi t/T)$	0.052***	0.002
Diagnostics		
Adjusted R^2	0.610	--

RESET	--	0.080
BG LM p-values (1/6/12)	--	0.497 / 0.882 / 0.884
Jarque-Bera	1.475	0.478
ARCH LM (12)	--	0.506
CUSUM	Stable	--
CUSUMSQ	Stable	--
ECT_{t-1}	-1.005***	0.000

*Notes: lnTotalGeneration replaces lnOIL. Total generation contains renewable generation, which creates a part-whole mechanical relationship with the dependent variable; this specification is therefore a sensitivity check rather than a separate causal model. The selected model is FA-ARDL(1, 1, 1, 1, 2), k = 2. Bounds statistics are F-test (1) = 12.066, t-test (1) = -7.623, and F-test (2) = 12.775; * p < 0.10, ** p < 0.05, and *** p < 0.01.*

The long run returns show a dramatic shift: the EPU coefficient is no longer statistically significant (coefficient = 0.036, p = 0.402). This finding suggests that when the overall scale of electricity generation is controlled for, the positive effect of policy uncertainty on renewable generation disappears. Careful interpretation of this result is needed. The model accounts for variation in generation, including renewables, which means that the effect of EPU is not necessarily the cause of the variation (this may be captured by the size of the dependent variable itself). That is, the positive EPU effect seen in the baseline model can happen either because of a change in overall electricity production or because of a direct change in the energy mix, from non-renewables to renewables. This interpretation is confirmed by the renewable share sensitivity analysis, which found a smaller but not negligible EPU coefficient, indicating that EPU has an impact on the absolute value, as well as the relative share of renewables, but the impact on the share is less pronounced.

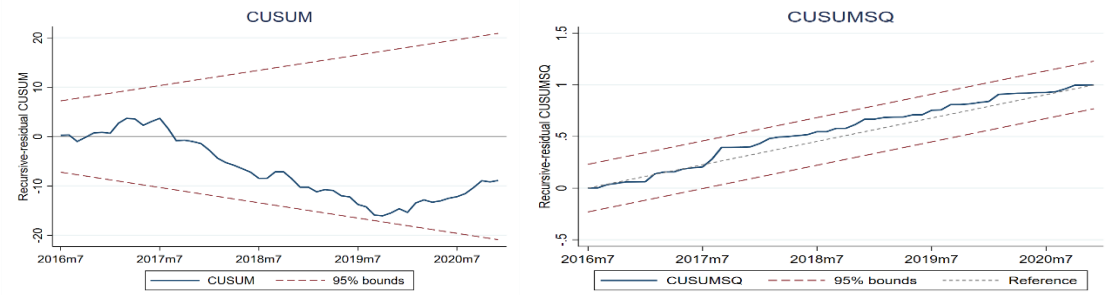
Total generation has a strong positive long run effect on renewable generation with coefficient 1.914 ($p<0.01$) with a one percent increase in generation being nearly associated with a two percent increase in renewable generation. This finding is consistent with the inclusion of renewable generation in total generation and represents the elasticity of renewable generation to total electricity demand. CO2 emissions continued to negatively impact as expected with a coefficient of -0.994, also statistically significant at $p < 0.01$. The long-run effect of industrial production is negative and marginally significant at the p = 0.092 level (coefficient = -0.969). This could be due to the countercyclical nature of investments in renewables compared to overall investment.

The contemporaneous change in total generation (ΔlnTotalGeneration$t$) is positive and significant (coefficient = 1.553, p = 0.001) while the contemporaneous change in CO2 emissions (ΔlnCO2t) is negative and highly significant (coefficient = -1.220, $p < 0.01$). The change in EPU and industrial production in the short run is not statistically significant, confirming the baseline result that the EPU effect primarily occurs through the long run.

The error correction term is -1.005, significant at 1% indicating that the adjustment to equilibrium is essentially over in one month. The Fourier terms indicate that the cosine term (mu-2) is statistically significant at the 1% level giving evidence of smooth structural

shifts. The model diagnostics are satisfactory: adjusted R-squared value of 0.610 is greater than the baseline model, the p-value of the RESET test is 0.080, which implies that there is no serious functional form misspecification, and the Breusch-Godfrey LM test p-values of lags 1, 6, and 12 are 0.497, 0.882, and 0.884 respectively, indicating no evidence of serial correlation. The Jarque-Bera test p-value of 0.478 confirms normality of the residuals, and the ARCH LM test p-value of 0.506 confirms that there is no evidence of heteroskedasticity.

Figure 4: Recursive-residual stability tests for the total-generation-control model



The CUSUM statistics for the recursive residual is within the 95% bounds for the entire sample period, as seen in Figure 4, which suggests that the model parameters are stable during the sample period. Also, CUSUMSQ remains within the 95% limits, indicating the stability of the residual variance. These stability tests are useful to give extra confidence that the model specification is correct and that the estimated relations are not caused by instability of the parameters.

The total generation control sensitivity analysis shows that sensitivity to the positive EPU effect is due to the use of total generation as a control variable. The impact of policy uncertainty on renewables generation is negligible when the overall scale of electricity generation is considered. The results indicate that the EPU effect that is found in the baseline model could be acting through other mechanisms, such as a shift in the composition of electricity production, and not the direct effect of a shift in electricity production toward renewables. It should be noted, however, that this specification includes a mechanical relationship between the dependent variable and the control variable: renewable generation as part of total generation. Thus, the sensitivity analysis results should be treated with care and do not necessarily contradict trends from the baseline analysis. Rather, they emphasize the need to explore the effects of various policy uncertainty specifications when studying the policy uncertainty–renewable energy deployment link.

Period Sensitivity

Two complementary sensitivity analyses to check the stability of the baseline analysis in the face of the unprecedented economic and policy shock due to the COVID-

19 pandemic. The first thing is to remove the entire COVID-19 period because only 56 observations remain if that is done, from January 2015 to December 2019. Second, we retain the full sample but introduce a dummy variable for COVID-19 to indicate the presence of an intervention in the period March 2020 to December 2020 and otherwise zero, which allows for the control of the conditional effect of the change in monthly renewable generation growth during the pandemic. Both specifications use the baseline FA-ARDL (1, 2, 1, 1, 1) model, where the Fourier frequency of k=2 means that any shift in the coefficient estimates is due to the period treatment, and not to a change in lag selection.

Both sensitivity analyses are presented in Table 9 with four panels: Long Run Results, Short Run Results, Fourier Terms and Cointegration and Diagnostics.

Table 9: Period sensitivity of the baseline FA-ARDL specification

Variable or test	**2020 excluded**	**COVID intervention**
Dependent variable	lnREN	lnREN
Sample	2015m1-2019m12	2015m1-2020m12
Observations	56	68
Model held fixed	FA-ARDL(1, 2, 1, 1, 1), k = 2	FA-ARDL(1, 2, 1, 1, 1), k = 2
Long-run findings		
lnEPU	0.118* (0.083)	0.181** (0.011)
lnIPI	2.831*** (0.003)	1.988** (0.045)
lnOIL	-0.084 (0.252)	-0.014 (0.866)
lnCO2	-0.868*** (0.000)	-0.857*** (0.000)
Short-run findings		
$\Delta lnEPU_t$	-0.056 (0.222)	-0.029 (0.460)
$\Delta lnIPI_t$	2.274** (0.035)	2.206** (0.017)
$\Delta lnOIL_t$	-0.133 (0.411)	-0.094 (0.373)
$\Delta lnCO2_t$	-0.730*** (0.000)	-0.737*** (0.000)
COVID-period indicator		0.188** (0.040)
Fourier terms		
sin(4πt/T)	-0.027 (0.140)	-0.059*** (0.003)
cos(4πt/T)	0.059*** (0.006)	0.006 (0.772)
Bounds and diagnostics		
Bounds F1/t1/F2	10.048 / -6.977 / 7.001	10.200 / -6.988 / 7.552
Long-run level relationship	Supported	Supported
Adjusted R^2	0.511	0.464
RESET p-value	0.126	0.091
BG LM p-values (1/6/12)	0.923 / 0.506 / 0.300	0.356 / 0.902 / 0.383
Jarque-Bera statistic (p)	0.225 (0.894)	0.745 (0.689)
ARCH LM (12) p-value	0.456	0.337
CUSUM	Stable	Stable
CUSUMSQ	Stable	Stable
ECT_{t-1}	-0.997*** (0.000)	-0.898*** (0.000)

*Notes: Coefficient cells report estimates with p-values in parentheses. Both columns hold the baseline oil-control lag structure and k = 2 fixed; the models are not reselected. The 2020-excluded column covers 2015m1-2019m12 and has 56 effective observations after initial lags. The intervention equals one from 2020m3 through 2020m12 and zero otherwise. Because the dependent variable in the UECM is $\Delta lnREN_t$, the intervention coefficient is a conditional shift in monthly renewable-generation growth, not a causal pandemic effect. Bounds entries report F-test (1), t-test (1), and F-test (2). A long-run level relationship requires F-test (1) > 5.720, t-test (1) < -4.960, and F-test (2) > 5.970. * p < 0.10, ** p < 0.05, and *** p < 0.01.*

The COVID-19 period treatment is found to be highly sensitive in the long-run EPU coefficient. If the pandemic period is not included, then the coefficient of the EPU is positive but not statistically significant at the 10% level (coefficient = 0.118, p = 0.083). This decreased statistical significance reflects that the high positive EPU effect in the baseline model may be a result of the COVID-19 pandemic, when policy uncertainty increased significantly, and renewable generation increased continuously. When the COVID-19 intervention dummy is added to the full sample, the coefficient on EPU is positive and statistically significant at the 5% level (0.181, p = 0.011) but is smaller than the baseline value of 0.265. The growth in renewable generation remains conditionally higher during the COVID-19 period, with its dummy being significant and positive (coefficient = 0.188; p = 0.040), even after accounting for the effects of EPU and other factors.

The long-run coefficients for the control variables are generally robust across both specifications. Industrial production (lnIPI) becomes statistically significant in both period sensitivity models, with coefficients of 2.831 (p = 0.003) in the COVID-excluded sample and 1.988 (p = 0.045) in the COVID-intervention model, suggesting that the baseline model may have underestimated the long-run effect of economic activity on renewable generation due to the confounding influence of the pandemic. The negative effect of CO2 emissions remains highly significant in both specifications, with coefficients of -0.868 and -0.857 (both $p < 0.01$), confirming the robust displacement relationship between emissions and renewable generation. Oil prices remain statistically insignificant in both models, consistent with the baseline findings.

The short-run dynamics are also consistent across the two period sensitivity specifications. The contemporaneous change in industrial production (ΔlnIPIt) is positive and significant in both models, with coefficients of 2.274 (p = 0.035) and 2.206 (p = 0.017), while the contemporaneous change in CO2 emissions (ΔlnCO2t) remains negative and highly significant, with coefficients of -0.730 and -0.737 (both $p < 0.01$). The short-run changes in EPU and oil prices are not statistically significant in either specification, reinforcing the baseline finding that the EPU effect operates primarily through the long-run channel.

The error correction terms are -0.997 and -0.898 in the COVID-excluded and COVID-intervention models respectively, both significant at the 1% level, indicating rapid adjustment back to equilibrium following shocks. The Fourier terms show some variation across specifications: the sine component ($\sin(4\pi t/T)$) is significant only in the COVID-intervention model, while the cosine component ($\cos(4\pi t/T)$) is significant only in the COVID-excluded model, suggesting that the pandemic period alters the pattern of smooth structural shifts captured by the Fourier terms.

The cointegration tests confirm a long-run level relationship in both specifications, with all three bounds test statistics exceeding the 1% upper-bound critical values. The model diagnostics are satisfactory across both models: the adjusted R-squared values are

0.511 and 0.464, the RESET test p-values are 0.126 and 0.091, and the Breusch-Godfrey LM test p-values for lags 1, 6, and 12 all exceed 0.05, indicating no evidence of serial correlation. The Jarque-Bera test confirms normality of the residuals, and the ARCH LM test indicates no evidence of heteroskedasticity. Both the CUSUM and CUSUMSQ tests confirm parameter stability in both specifications.

The period sensitivity analysis reveals that the positive long-run effect of EPU on renewable generation is robust to the inclusion of a COVID-19 intervention dummy but becomes marginally significant when the pandemic period is excluded. This pattern suggests that the COVID-19 pandemic amplified the positive relationship between policy uncertainty and renewable generation, possibly because the pandemic-induced surge in EPU coincided with continued investment in renewable capacity that had been committed prior to the crisis. The positive and significant COVID-19 dummy coefficient further supports this interpretation, indicating that renewable generation growth was conditionally higher during the pandemic, potentially due to the operation of pre-existing renewable capacity and the relative resilience of renewable energy investments compared to fossil fuel alternatives during the economic downturn.

Table 10 reports the fixed-specification finite-sample bootstrap assessment. F-overall is 8.777, above its 1% critical value of 5.565 ($p = 0.0008$). The t-dependent statistics are -6.591, more negative than -4.603 ($p = 0.0006$). F-independent is 8.660, above 5.093 ($p = 0.0004$). All three nulls are rejected at 1%, confirming the non-degenerate long-run level relationship with test-specific finite-sample critical values. Across the reported models, the negative lnCO2 coefficient remains significant. The positive long-run lnEPU coefficient is significant in the baseline, renewable-share, and period-sensitivity specifications, but not when total generation replaces oil.

Bootstrap Critical Values for Bounds Tests

To assess finite-sample bounds inference, 4,999 test-specific conditional residual-bootstrap replications are computed for the fixed 68-observation baseline specification using seed 20260822.

Table 10 reports the calculated statistics, test-specific bootstrap critical values at 10%, 5%, and 1%, bootstrap p-values, and decisions.

Table 10: Bootstrap critical values for the FA-ARDL bounds tests

Test	Statistic	CV10	CV5	CV1	Bootstrap p-value	Decision
F-overall	8.777***	3.570	4.209	5.565	0.0008	Reject H_0 at 1%
t-dependent	-6.591***	-3.447	-3.888	-4.603	0.0006	Reject H_0 at 1%
F-independent	8.660***	2.912	3.565	5.093	0.0004	Reject H_0 at 1%

Notes: The selected FA-ARDL(1, 2, 1, 1, 1), k = 2 specification and 68-observation common sample are held fixed. Results use 4,999 test-specific conditional residual-bootstrap replications with seed 20260822. CV10, CV5, and CV1 denote 10%, 5%, and 1% bootstrap critical values. F-overall tests all lagged levels,

*t-dependent tests the lagged dependent-variable level, and F-independent tests the four lagged regressor levels. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.*

The bootstrap results strongly confirm the cointegration findings from the standard bounds testing procedure. The F-overall test statistic of 8.777 exceeds the 1% bootstrap critical value of 5.565, with a bootstrap p-value of 0.0008, indicating that the null hypothesis of no cointegration is rejected at the 1% significance level. Similarly, the t-dependent statistic of -6.591 is more negative than the 1% bootstrap critical value of -4.603, with a bootstrap p-value of 0.0006, confirming the presence of a long-run level relationship. The F-independent statistic of 8.660 exceeds the 1% bootstrap critical value of 5.093, with a bootstrap p-value of 0.0004, providing additional evidence against the null hypothesis.

The finite-sample adjustments using the bootstrap procedure are evident in the lower critical values, as compared to the asymptotic critical values in Table 5. For instance, the 1% bootstrap critical value of the F-overall test is 5.565, while the upper-bound asymptotic critical value is 5.720. Though lower thresholds, all three test statistics greatly outperform their respective bootstrap critical values if they are sensitive to the choice of critical value methodology, they do offer strong evidence of cointegration.

All the p-values obtained by the bootstrap method are less than 0.001, meaning under the null hypothesis that there is no cointegration the probability of observing such extreme test statistics is less than 0.001. This is a robust rejection of the null hypothesis, which confirms the presence of a stable long run relationship between renewable generation and economic policy uncertainty, industrial production, oil prices, and electricity-sector emissions in the Netherlands. The overlapping of the results of the bootstrap methods with the results of the asymptotic methods lends further credibility to the cointegration results and provides further validation to the FA-ARDL model for long and short-run representation of renewable electricity generation.

DISCUSSION

The baseline estimates document a positive long-run conditional association between domestic Dutch EPU and renewable generation. This sign differs from the standard real-options prediction that uncertainty raises the value of waiting for irreversible investment (Bernanke, 1983; Dixit & Pindyck, 1994; Gulen & Ion, 2015). The result therefore should not be presented as confirmation of real-options theory or as proof of an anticipatory-investment or hedging mechanism. It may be consistent with several observational explanations, including the timing of pre-committed renewable capacity, policy-support arrangements, common low-frequency movement, or omitted institutional factors. The present design cannot distinguish among these channels. The positive COVID-period indicator likewise captures a conditional growth shift and does not demonstrate that pandemic uncertainty accelerated renewable investment.

The estimates do not imply that policy uncertainty is desirable or that policy stability is unnecessary. Predictable policy can remain relevant to financing and irreversible investment even when the observed long-run EPU coefficient is positive (Barradale, 2010; Sendstad & Chronopoulos, 2020). The appropriate empirical implication is narrower: within this sample and specification, renewable generation is positively associated with EBO-NL after conditioning on the included variables, but the coefficient becomes insignificant when total generation replaces oil. The negative lnCO2 coefficient is also a conditional association within the electricity-generation system; because emissions and renewable generation are jointly tied to the generation mix, it cannot be used as causal evidence that the estimated model measures the emissions avoided by renewables.

The study has five limitations. First, 72 raw monthly observations and 68 baseline estimation observations restrict precision; bootstrap inference addresses finite-sample critical values but does not create additional information. Second, the sample ends in December 2020 and includes only ten months after the intervention onset, limiting period-specific inference and external validity. Third, EBO-NL is a broad, newspaper-based domestic EPU measure and does not isolate fiscal, monetary, energy, or climate-policy uncertainty. Fourth, electricity-sector emissions are jointly determined with the generation mix, and total generation contains renewable generation; coefficients in models containing these variables should not be interpreted causally. Fifth, the Fourier terms approximate deterministic low-frequency movement but do not identify break dates, nonlinear EPU coefficients, or causal transmission mechanisms. The SOR implementation also produces boundary transition estimates for some series, which is why its results are interpreted jointly with DF-GLS and Zivot-Andrews evidence.

Future work could extend the sample, compare domestic EPU with broader or sector-specific uncertainty measures, and evaluate additional countries under a common data definition. Designs with policy-event timing, project-level investment commitments, or credible external instruments would be required to distinguish causal mechanisms. Longer samples would also permit more flexible nonlinearity and multiple-regime analysis without overfitting the present 72-month series.

CONCLUSION

This study estimates the short-run and long-run conditional relationships between domestic economic policy uncertainty and renewable electricity generation in the Netherlands from 2015m1 to 2020m12. The baseline FA-ARDL model yields a positive long-run EPU elasticity of 0.265 ($p < 0.01$), a negative lnCO2 elasticity of -1.046 ($p < 0.01$), and a negative error-correction coefficient of -0.772 ($p < 0.01$). The short-run EPU coefficient is insignificant. Augmented asymptotic bounds tests and test-specific bootstrap critical values support a non-degenerate long-run level relationship. The positive EPU coefficient remains significant with renewable share and in the COVID-

intervention model, is significant at 10% when 2020 is excluded, and is insignificant when total generation replaces oil. The evidence therefore supports a positive but specification-sensitive long-run association; it does not show that uncertainty causes renewable deployment or that policy stability is unimportant.

The findings are specific to the variables, period, and institutional setting examined. Longer samples, alternative uncertainty measures, and research designs capable of separating project commitments from contemporaneous macroeconomic shocks are needed before causal or general policy conclusions can be drawn.

Further studies are recommended to broaden the analysis to several countries having different institutional settings and policy contexts and to understand what conditions enable the policy uncertainty to either drive or impede the deployment of renewable energy. Decomposing the EPU index into its fiscal, monetary, and regulatory components would enable a finer-grained understanding of the EPU effect on renewable energy results and more business cycle data would provide better information on whether the EPU effect is a general phenomenon or dependent on large exogenous shocks. The Netherlands' experience points out that well-designed policies and long-term energy contracts can shield renewable energy generation from short-term policy volatility, offering lessons to policymymakers seeking to increasingly drive an energy transition.

REFERENCES

Abbas, Q., Imran, M., & Sattar, A. (2025). From linear to circular: the impact of economic policies and technological innovations on greenhouse gas emissions in the Netherlands. *Carbon Balance and Management*, *20*(1), 11. https://doi.org/10.1186/s13021-025-00297-1

Adu, D., Jianguo, D., Asomani, S. N., & Abbey, A. (2024). Energy generation and carbon dioxide emission—The role of renewable energy for green development. *Energy Reports*, *12*, 1420–1430. https://doi.org/10.1016/j.egyr.2024.07.013

Akron, S., Demir, E., Díez-Esteban, J. M., & García-Gómez, C. D. (2020). Economic policy uncertainty and corporate investment: Evidence from the U.S. hospitality industry. *Tourism Management*, *77*, 104019. https://doi.org/10.1016/j.tourman.2019.104019

Athari, S. A., Addia, K., Kirikkaleli, D., Al Geitany, S. H., Al Fadhel, L., & Saliba, C. (2025). Do Environmental Taxes and Green Electricity Matter for Environmental Quality? Fresh Evidence in France Based on Fourier Methods. *Energies*, *18*(19), 5046. https://doi.org/10.3390/en18195046

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

Barradale, M. J. (2010). Impact of public policy uncertainty on renewable energy investment: Wind power and the production tax credit. *Energy Policy*, *38*(12), 7698–7709. https://doi.org/10.1016/j.enpol.2010.08.021

Bernanke, B. S. (1983). Irreversibility, Uncertainty, and Cyclical Investment. *The Quarterly Journal of Economics*, *98*(1), 85. https://doi.org/10.2307/1885568

Bertelli, S., Vacca, G., & Zoia, M. (2022). Bootstrap cointegration tests in ARDL models. *Economic Modelling*, *116*, 105987. https://doi.org/10.1016/j.econmod.2022.105987

Borozan, D., & Pirgaip, B. (2023). Does uncertainty in climate policy play a role in shaping renewable energy use? Evidence from the USA. *Environment, Development and Sustainability*, *27*(1), 2259–2284. https://doi.org/10.1007/s10668-023-03965-9

Brown, R. L., Durbin, J., & Evans, J. M. (1975). Techniques for Testing the Constancy of Regression Relationships Over Time. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *37*(2), 149–163. https://doi.org/10.1111/j.2517-6161.1975.tb01532.x

Chen, Y. (2018). Factors influencing renewable energy consumption in China: An empirical analysis based on provincial panel data. *Journal of Cleaner Production*, *174*, 605–615. https://doi.org/10.1016/j.jclepro.2017.11.011

Deka, A., Özdeşer, H., & Seraj, M. (2024). The impact of oil prices, financial development and economic growth on renewable energy use. *International Journal of Energy Sector Management*, *18*(2), 351–368. https://doi.org/10.1108/IJESM-09-2022-0008

Ding, Q., Zhu, X., & Zhang, S. (2025). Asymmetric effects of climate policy uncertainty on the renewable energy consumption. *Environment, Development and Sustainability*. https://doi.org/10.1007/s10668-025-06144-0

Dixit, R. K., & Pindyck, R. S. (1994). *Investment under Uncertainty*. Princeton University Press. https://doi.org/10.1515/9781400830176

Dudin, M. N., Frolova, E. E., Sidorenko, V. N., Pogrebinskaya, E. A., & Nikishina, I. V. (2017). Energy policy of the European Union: Challenges and possible development paths. *International Journal of Energy Economics and Policy*, *7*(3), 294–299.

Elliott, G., Rothenberg, T. J., & Stock, J. H. (1996). Efficient Tests for an Autoregressive Unit Root. *Econometrica*, *64*(4), 813. https://doi.org/10.2307/2171846

Enders, W., & Lee, J. (2012). A Unit Root Test Using a Fourier Series to Approximate Smooth Breaks*. *Oxford Bulletin of Economics and Statistics*, *74*(4), 574–599. https://doi.org/10.1111/j.1468-0084.2011.00662.x

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, *50*(4), 987. https://doi.org/10.2307/1912773

Gallant, A. R. (1981). On the bias in flexible functional forms and an essentially unbiased form. *Journal of Econometrics*, *15*(2), 211–245. https://doi.org/10.1016/0304-4076(81)90115-9

Godfrey, L. G. (1978). Testing Against General Autoregressive and Moving Average Error Models when the Regressors Include Lagged Dependent Variables. *Econometrica*, *46*(6), 1293. https://doi.org/10.2307/1913829

Goldthau, A., & Sitter, N. (2015). Soft power with a hard edge: EU policy tools and energy security. *Review of International Political Economy*, *22*(5), 941–965. https://doi.org/10.1080/09692290.2015.1008547

Gulen, H., & Ion, M. (2015). Policy Uncertainty and Corporate Investment. *Review of Financial Studies*, hhv050. https://doi.org/10.1093/rfs/hhv050

Jarque, C. M., & Bera, A. K. (1987). A Test for Normality of Observations and Regression Residuals. *International Statistical Review / Revue Internationale de Statistique*, *55*(2), 163. https://doi.org/10.2307/1403192

Kartal, M. T., Kirikkaleli, D., & Kılıç Depren, S. (2024). Investigating political stability effect on the environment in the Netherlands by Fourier-based approaches. *International Journal of*

Sustainable Development & World Ecology, *31*(2), 125–136. https://doi.org/10.1080/13504509.2023.2256716

Kroese, L., Kok, S., & Parlevliet, J. (2015). Beleidsonzekerheid in Nederland. *Jaargang*, *100*(4715), 464–467. https://esb.nu/wp-content/uploads/2022/11/464-467_KROESE.pdf

Leal-Arcas, R., Ríos, J. A., & Grasso, C. (2015). The European Union and its energy security challenges. *The Journal of World Energy Law & Business*, *8*(4), 291–336. https://doi.org/10.1093/jwelb/jwv020

McNown, R., Sam, C. Y., & Goh, S. K. (2018). Bootstrapping the autoregressive distributed lag test for cointegration. *Applied Economics*, *50*(13), 1509–1521. https://doi.org/10.1080/00036846.2017.1366643

Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, *16*(3), 289–326. https://doi.org/10.1002/jae.616

Qamruzzaman, M., & Swarna, S. R. (2025). Innovation Led Renewable Energy Consumption in OECD Nations: Evidence from Fourier Functions. *International Journal of Energy Economics and Policy*, *15*(5), 427–449. https://doi.org/10.32479/ijeep.19798

Ramsey, J. B. (1969). Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *31*(2), 350–371. https://doi.org/10.1111/j.2517-6161.1969.tb00796.x

Sam, C. Y., McNown, R., & Goh, S. K. (2019). An augmented autoregressive distributed lag bounds test for cointegration. *Economic Modelling*, *80*, 130–141. https://doi.org/10.1016/j.econmod.2018.11.001

Sendstad, L. H., & Chronopoulos, M. (2020). Sequential investment in renewable energy technologies under policy uncertainty. *Energy Policy*, *137*, 111152. https://doi.org/10.1016/j.enpol.2019.111152

Shahbaz, M., Omay, T., & Roubaud, D. (2018). Sharp and smooth breaks in unit root testing of renewable energy consumption: The way forward. *Journal of Energy and Development*, *44*(1), 5–39.

Syed, Q. R., Apergis, N., & Goh, S. K. (2023). The dynamic relationship between climate policy uncertainty and renewable energy in the US: Applying the novel Fourier augmented autoregressive distributed lags approach. *Energy*, *275*, 127383. https://doi.org/10.1016/j.energy.2023.127383

van der Loos, H. Z. A., Negro, S. O., & Hekkert, M. P. (2020). Low-carbon lock-in? Exploring transformative innovation policy and offshore wind energy pathways in the Netherlands. *Energy Research & Social Science*, *69*, 101640. https://doi.org/10.1016/j.erss.2020.101640

Verbong, G., & Geels, F. (2007). The ongoing energy transition: Lessons from a socio-technical, multi-level analysis of the Dutch electricity system (1960–2004). *Energy Policy*, *35*(2), 1025–1037. https://doi.org/10.1016/j.enpol.2006.02.010

Verhees, B., Raven, R., Kern, F., & Smith, A. (2015). The role of policy in shielding, nurturing and enabling offshore wind in The Netherlands (1973–2013). *Renewable and Sustainable Energy Reviews*, *47*, 816–829. https://doi.org/10.1016/j.rser.2015.02.036

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

DATA SOURCES

Ember. Monthly Electricity Data, Netherlands renewable electricity generation and total electricity-generation emissions. https://ember-energy.org/data/monthly-electricity-data/ (accessed 01 August 2026). Generation download: https://files.ember-energy.org/public-downloads/generation/outputs/release_generation_monthly_lower.csv. Emissions methodology: https://files.ember-energy.org/public-downloads/ember_electricity_data_methodology.pdf.

Eurostat. Production in industry, monthly, Netherlands, NACE Rev. 2 sections B-D, seasonally and calendar adjusted, 2021 = 100 [dataset code: sts_inpr_m]. https://ec.europa.eu/eurostat/databrowser/view/sts_inpr_m/default/table (accessed 01 August 2026). Exact API selection: https://ec.europa.eu/eurostat/api/dissemination/statistics/1.0/data/sts_inpr_m?freq=M&unit=I21&s_adj=SCA&nace_r2=B-D&geo=NL&lang=en.

Kok, S., Kroese, L., & Parlevliet, J. Netherlands monthly EPU indices, strict domestic EBO-NL series. https://www.policyuncertainty.com/netherlands_monthly.html (accessed 01 August 2026). Data file: https://www.policyuncertainty.com/media/Netherlands_Policy_Uncertainty_Data.xlsx.

U.S. Energy Information Administration. Crude Oil Prices: Brent - Europe [MCOILBRENTEU], monthly average, retrieved via FRED, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/MCOILBRENTEU (accessed 01 August 2026).

CHAPTER 3

THE EFFECT OF GOVERNMENTAL FACTORS ON ENVIRONMENT: A SYSTEM GMM APPROACH

İBRAHİM TEKİNER[1]

1-Introduction

As environmental pollution progresses to dangerous levels, the intensity of studies on factors and of methods that will slow the pollution down is also increasing. First barrier of this type of studies is to decide which measures are proper to be used as reprecentative of environmental quality. One of the most reliable representative of that quailty is Carbon Emission Efficiency (CEE) which also called in some contexts Green Total Factor Productivity. Efficiency conceptually or semantically expresses the same thing as productivity. However, the two are formulated differently in practice. Productivity is the ratio of the total output of a system to the inputs used to produce those outputs. Formally, it is written as Output/Input. Productivity refers more to creativity that is obtaining more and more output from a given input. Therefore, it has no limit. The productivity of a system can be measured alone; it measures the contribution of human effort, which is not planned for production,

[1] Dr, Zonguldak Bülent Ecevit Üniversitesi, İktisat, Orcid: 0000-0002-1185-5974

and it does not require comparison. Efficiency, on the other hand, is the highest level of output that can be achieved with a given level of input, so it is a Pareto concept. In other words, there is a point where no more output can be obtained by changing the input distribution. Formally, it can be written as (actual output) / (planned output). The most important feature that formally distinguishes the concept of efficiency from productivity is the property of absoluteness. While the productivity of an economic unit can be calculated absolutely on its own, efficiency requires a comparable froniter. Therefore, it is measured with methods called frontier approaches. Thus, it is possible to speak of the efficiencies of many companies in comparison with each other. CEE is the measurement of efficiency in the context of the environment. It is of critical importance for sustainable economic growth and was first expressed by Kaya and Yokobori (1993) as GDP/CO_2 (GDP produced as a result of the consumption of a unit of CO_2 (Pan et al. 2019:329). This single-factor CEE is simple, easy to calculate, and directly reflects the impact of CO_2 (Tol et al. 2009; Cheng et al. 2018a:331). However, since it does not show the effects of inputs, this single-factor indicator fails to fully reflect the nature of CEE, and makes it difficult to understand the deeper connection between the economy and environmental degradation (Zhu et al. 2021:2). Although single-factor indices are easy to understand, they are inadequate when considering the complex functions between social and environmental factors and the multidimensional nature of CEE. To overcome this problem, a total factor CEE index has been calculated. This index can be linked to economic activities such as capital, labor, energy, etc. (Xie et al. 2021:2).

An enhancement in CEE indicates that economic growth and reduced carbon emissions take place at the same time (Zheng et al. 2022:3). This is due to the fact that higher CEE levels demonstrate the progress of technologies aimed at cutting carbon emissions (Pan

et al. 2019:327). After the data envelopment analysis began to be used in efficiency calculations, CEE has also been calculated with this method and its derivatives. The most important distinguishing feature of environmental efficiency measurement made with linear programming (Data Envelopment Analysis) is the inclusion of an undesirable output variable. Halkos and Petrau (2019) have mentioned four different options for dealing with undesirable outputs: the first is to neglect the undesirable output in the production process; the second is to treat it as a normal input (Reinhard et al., 1999); the third is to take its inverse and use it as a normal output (Lovell et al. 1995; Seiford and Zhu 2002); and the fourth is to make some necessary transformations (Fare and Groskopf 1999). This last approach has leaded to the development of various methods depending on the production possibility set and the characteristics of the undesirable output (e.g., weak disposability). The use of Data Envelopment Analysis (DEA) and its variants in calculating CEE has become increasingly common.

The first DEA model is the boundary model, named CCR after its developers Charnes, Cooper, and Rhodes (1978). The CCR model is based on Farrel's work 1957 (Shahverdi and Ebrahimnejad 2014:383). Banker, Charnes, and Cooper (1984) (BCC) developed new linear programming solution by adding the property of "variable returns to scale," and Tone (2001) developed the Slacks-based Measure (SBM) model. These simple DEA models do not consider undesirable outputs. Seiford and Zhu (2002) state that treating undesirable outputs as inputs would distort the real production process and lead to a loss of the relationship between inputs and outputs. The radial DDF proposed by Chambers et al. (1996) has been considered a way to deal with undesirable outputs. The contribution of DDF to the DEA literature is that it allows for simultaneous increases in desired outputs and decreases in undesirable outputs. However, since this model is applied to all

inputs and outputs in the same way, it can lead to errors in efficiency scores. To obtain more accurate results, Zhou et al. (2012) introduced the non-radial DDF model (Zhu et al. 2021; Chen et al. 2018; Zhou et al. 2019).

Therefore, total greenhouse gas emissions are considered as undesirable output, CEE is seen as a typical input-output relationship, and it is calculated using the DDF-ML (Malmquist-Luenberger) method. The CEE calculated with this model provides relative efficiency over time and across decision-making units (DMUs) (Wang et al. 2022). Subsequently, the relationship between this calculated CEE and regulations is analyzed using the system GMM method, along with control variables. The data analyzed are from 34 countries for the period 2000-2016. Our study explains the relationship between governmental policies and CEE. CEE series that calculated by Directional Distance Function (DDF), Non-Radial DDF and Malmquist Luenberger indices in six different ways were used to represent the environmental efficiency. Then, by applying the system GMM method to sample of 34 countries, the effects of governmental factors on these six CEE series were examined together with control variables. The subsequent sections of the study proceed as follows: Section 2 introduces and summarizes important studies related to the DEA literature in a table; Section 3 presents theoretical framework of utilized models; Section 4 introduces the econometric analysis applied and discusses the analysis findings; and Section 5 summarizes the main findings of the study and outlines the policy implications derived from these findings.

2-Literature Rewiev

As a relative efficiency analysis, the calculation of the CEE index requires the construction of a production frontier. The most commonly used frontier methods are DEA and SFA (stochastic

frontier analysis). However, SFA has some limitations in application because it requires a specific function and its assumptions are very rigid. On the other hand, DEA is frequently used in environmental performance research because it does not require model construction or parameter estimation (Cheng et al. 2018a:331). Since the emergence of the CCR model, numerous DEA models have been developed. For example, Jahanshahloo et al. (2005) proposed a non-radial model that includes undesirable inputs and outputs, and Fare and Groskopf (2010) combined Tone's (2001) SBM model, which has an additive structure, with the DDF function, which also has an additive structure, to develop the SBM-DDF model. In this model, they showed that a DMU can be efficient in terms of SBM only if it is also efficient in terms of DDF. Another study contributing to the DDF model is by Fukuyama and Weber (2009). Cheng et al. (2018a) proposed a new integrated approach combining the environmentally extended input-output model and the DEA model. Cheng et al. (2018a) developed the extended NDDF model. Therefore, there are many approaches in DEA analysis such as BCC, SBM, DDF, and NDDF.

DEA models and their derivatives work with cross-sectional data, comparing the situations of DMUs at a specific point in time. Therefore, they do not allow for comparisons across different time periods. Models that allow for both intertemporal and cross-DMU efficiency comparisons have been developed, such as SBM-ML. The ML index also permits the consideration of undesirable outputs (Li and Chen 2021:16). Another indices that jointly considers both the time and DMU dimensions is the NDDF-ML index. One of the important advantage of this index is its flexibility, as it allows for differences in input and output slacks across units (either Cross sectionally or over time). Therefore, the ML index, which allows for simultaneous comparisons across both DMUs and time periods, is frequently used in the literature. The results of studies presented in

the literature on CEE and its influencing factors, calculated using DEA, SBM, DDF, and other linear programming methods, are shown in Table 1.

As seen in the table, a large number of significant analyses have been conducted using a combination of various methods. A significant portion of recent studies have focused on the regions of China. Ma et al. (2022) measured green efficiency for 279 cities in China and then applied the Spatial Durbin model (SDM). As a result of this study, it was determined that the spillover contribution of environmental regulations is non-linear: while the relationship between formal environmental regulations and CEE is inverted U-shaped, the effect of informal environmental regulations is normal U-shaped. Pan et al. (2019) calculated CEE based on DDF and then analyzed the impact of environmental policies implemented in China on CEE. They found that environmental policies, while varying regionally, increased CEE. Tang and Qin (2022), with the efficiency data they obtained through a three-stage DEA analysis they conducted for 30 regions of China, and the results of the SFA analysis they conducted, showed that inter-local government competition significantly constrained green economic growth.

Zhu et al. (2021:4) calculated CEE in China's energy-intensive industries for the period 2005-2017 using a three-stage DEA model, and then examined the factors affecting CEE through temporal-spatial distribution, spatial autocorrelation, and the Tobit model. The first stage of the three-stage DEA model is based on the BCC model. The slack obtained from this stage is used to adjust the input and output data. Afterward, the adjusted data is used to reapply the initial BCC model to calculate the final CEE. They found an inverse U-shaped relationship between carbon emissions and economic growth. Another study applying the three-stage DEA model, by Chen et al. (2021:4), showed that the northeastern region of China has a higher CEE than the eastern region.

Li et al. (2018:4) consider the increasing CEE as an indicator that production technologies and systems are becoming more sustainable. Cheng et al. (2018b) emphasize that technological progress plays a significant role in the increase of CEE. Zhao et al. (2022) applied the Spatial Durbin model to the efficiency statistics obtained using the Epsilon-Based Measure (EBM) and demonstrated that government regulation of the transportation sector contributes to green growth. Li and Chen (2021:2) investigated the Spatial Distribution and temporal development patterns of green efficiency; for this purpose, they first created a new model by combining the SBM model with undesirable outputs and ML, referring to this new method as the SBM-ML model. After calculating the efficiency of each period using this method, the CEE of other periods is calculated based on a reference period.

Studies on the relationship between CEE and environmental policies can be divided into three categories: those that argue there is a positive relationship (Chen et al. 2018), those that argue there is a negative relationship, and those that argue the relationship is non-linear (e.g., Ma et al. 2022). Chen et al. (2018) categorized regulations into three groups and examined their effects on GTFV. According to their findings, the impact of regulations in all three categories on environmental quality was positive. Table 1 presents a summary of key studies in the literature. This information includes the methods used to calculate CEE, the analysis of CEE, and the results obtained.

3-Theory

In general terms, there are three popular methods for measuring efficiency: which are index measurements, linear programming, and econometric models. Index approaches measure the output-input ratio, such as growth accounting. Linear programming, which is also used in DEA, determines a production

frontier and assesses the efficiency of DMUs based on this frontier, which is why it is also called a frontier approach. In econometric models, data sets are analyzed using statistical methods, and efficiency is estimated. The SFA model is an example of econometric frontier approaches. DDF and NDDF models used in this study are DEA models that employ linear programming (Streimikis and Saraji 2022:821). DEA is based on the concept of relative efficiency and convex analysis performed using linear programming. DEA has proven effective in situations involving multiple inputs and outputs. However, the problem of undesirable outputs is not addressed in traditional models (Liu et al. 2016). The first DEA method is Charnes, Cooper and Rhodes (CCR) model (Charnes et al. 1978). The CCR model, involves maximizing a simple output/input ratio using linear programming under certain constraints (Ding et al. 2019; Kutlar and Bakırcı 2018; Alperen 2008). CCR is a non-parametric mathematical programming model. Using this model, the efficiency score of many decision makers can be calculated. Accordingly, the distance of each decision-making unit to the production frontier is important (Xie et al. 2021:5). The main advantages of DEA models: they maximize many outputs simultaneously, whereas other methods can maximize only one output. In addition, since the model is non-parametric, it does not need a specific production function. This makes DEA more flexible and applicable. (Streimikis and Saraji 2022:821).

The BCC model which is the second DEA approach, on the other hand, defines production set with variable returns to scale (Zhu et al. 2021; Kutlar and Bakırcı 2018). Both CCR and BCC models are radial models that ignore slacks. Subsequently, non-radial models that include slacks were developed to overcome these limitations. Tone (2001), building upon BCC, developed the SBM (Slack Based Model) where slacks are weighted. The SBM model is modeled as a weighting of slacks in the BCC model (Kutlar and

Bakırcı, 2018:165). However, the efficiency scores of DMUs are generally equal to 1 in both DEA and SBM models, making it difficult to distinguish and compare efficient DMUs. Andersen and Petersen (1993) introduced the super-efficiency DEA model, which allows DMU values to be greater than 1. Tone (2001) subsequently introduced the super-SBM model by combining the advantages of the SBM and super-DEA models. The SBM model has been modified to include undesirable outputs in the production process. The super-SBM model addresses undesirable outputs and compares the efficiency of each DMU. This model is used to measure indexes such as TFP, ecological efficiency, and energy efficiency (Zheng et al. 2022:5). In our study we uses directional distance function and Malmquist- Luenberger Productivity index.

Directional Distance Function

Briefly, the environmental technology discussed in this study assumes free disposability of good outputs, but weak disposability of the joint production of good and bad outputs. The production technology is labeled "environmental" owing to the existence of undesirable outputs (Cheng et al. 2018a).

The radial DDF approach adjusts all inputs and outputs by the same proportion until the production frontier is attained, which can lead to biased CEE estimates (Wang et al. 2022; Cheng et al., 2019:332). As a result, non-radial DEA models are able to more accurately detect the slacks related to both desirable and undesirable outputs. The NDDF model is formulated as:

$$\vec{D}(K, L, E, Y, A, C; g) = sup\{w^T\beta: ((K, L, E, Y, A, C) + diag(\beta)g)\epsilon T\} \quad (1)$$

In equation (1) $w^T = (w_K, w_L, w_E, w_Y, w_A, w_C)$ is the standardized weight vector. $g = (-g_K, -g_L, -g_E, -g_Y, g_A, -g_C)$ is directional vector and $\beta = (\beta_K, \beta_L, \beta_E, \beta_Y, \beta_C) \geq 0$ is the scale

vector to be maximized. DDF has ability to improve desired outputs while at the same time abating undesirable outputs. To derive NDDF, the subsequent DEA model needs to be solved:

$$\vec{D}(K, L, E, Y, C; g) = max w_K \beta_K + w_L \beta_L + w_E \beta_E + w_Y \beta_Y + w_A \beta_A + w_C \beta_C \qquad (2)$$

$$s.t. \begin{cases} \sum_{j=1}^{N} \lambda_j K_j \leq K - \beta_K g_K \\ \sum_{j=1}^{N} \lambda_j L_j \leq L - \beta_L g_L \\ \sum_{j=1}^{N} \lambda_j E_j \leq E - \beta_E g_E \\ \sum_{j=1}^{N} \lambda_j Y_j \geq Y - \beta_Y g_Y \\ \sum_{j=1}^{N} \lambda_j A_j \geq A - \beta_A g_A \\ \sum_{j=1}^{N} \lambda_j C_j = C - \beta_C g_C \\ \beta_Y \geq 0 \\ 0 \leq \beta_K, \beta_L, \beta_E, \beta_Y, \beta_C < 1 \\ \lambda_j \geq 0, j = 1,2, \dots, N \\ \sum_{j=1}^{n} \lambda_j = 1 \end{cases}$$

The directional vector quantifies how far a unit lies from the frontier. Consequently, the DDF captures inefficiency and is bounded between zero and one. By contrast, the NDDF aggregates every slack and can therefore exceed one, reflecting a greater degree of inefficiency for the DMU. In the present study the Malmquist-Luenberger (ML) efficiency index was also computed together with the DDF and NDDF.

Malmquist-Luenberger Productivity Index

The conventional approach to efficiency measurement, which focuses only on the output-input ratio, is seen as providing a biased efficiency measure by many researchers due to its disregard of undesirable outputs. Chung et al. (1997) proposed an efficiency index based on the radial DEA analysis. Referred to as the MLPI, this index simultaneously encourages a reduction in undesirable outputs and an expansion of desirable ones. Taking s and t as two

successive periods, and letting the input vector be denoted by $x = (K, L, E)$, the desirable-output vector by $y = (Y, A)$, and the undesirable-output vector by c, with the directional vector set to $g(x, y, c) = (0, y, -c)$, the output-oriented MLPI is then expressed as (Wang et al., 2022:110):

$$ML_t^{t+1} = \left\{\frac{1+D_j^t(K^t,L^tE^t,Y^t,A^t,C^t:g^t)}{1+D_j^t(K^{t+1},L^{t+1}E^{t+1},Y^{t+1},A^{t+1},C^{t+1}:g^{t+1})} x \frac{1+D_j^{t+1}(K^t,L^tE^t,Y^t,A^t,C^t:g^t)}{1+D_j^{t+1}(K^{t+1},L^{t+1}E^{t+1},Y^{t+1},A^{t+1},C^{t+1}:g^{t+1})}\right\}^{1/2} \quad (8)$$

$$\text{MLEFFCH} = \frac{1+D_j^t(K^t,L^tE^t,Y^t,A^t,C^t:g^t)}{1+D_j^t(K^{t+1},L^{t+1}E^{t+1},Y^{t+1},A^{t+1},C^{t+1}:g^{t+1})} \quad (3)$$

$$MLTECH = \left\{\frac{1+D_j^{t+1}(K^t,L^tE^t,Y^t,A^t,C^t:g^t)}{1+D_j^t(K^t,L^tE^t,Y^t,A^t,C^t:g^t)} x \frac{1+D_j^{t+1}(K^{t+1},L^{t+1}E^{t+1},Y^{t+1},A^{t+1},C^{t+1}:g^{t+1})}{1+D_j^t(K^{t+1},L^{t+1}E^{t+1},Y^{t+1},A^{t+1},C^{t+1}:g^{t+1})}\right\}^{1/2} \quad (4)$$

The ML index can be broken down into two elements: technical efficiency and technological change. It accounts for undesirable outputs as well as a temporal aspect. Evaluating efficiency over successive periods endows the index with a dynamic character (Li and Chen 2021: 4). In the present analysis we computed and examined three variants of the ML index—radial ML, global ML, and non-radial global ML. The global ML efficiency measure, introduced by Oh (2010), satisfies circularity and resolves the possible infeasibility that can arise when linear programming is used to evaluate successive-period DDFs (Chen et al. 2018).

4-Data Set and Econometric Analysis

Generalized Methods of Moments

To address the primary research question, this study utilizes panel data for 34 countries spanning the period 2000–2016, dictated by data availability. All variable definitions are detailed in Table 1. Methodologically, we first construct six distinct Carbon Emission Efficiency (CEE) series: three based on the Directional Distance Function (DDF) and three on the Malmquist-Luenberger (ML) index, incorporating global, radial, and non-radial specifications. Subsequently, these CEE series are analyzed using the dynamic System GMM estimator. Developed by Arellano and Bover (1995) and Blundell and Bond (1998) as an extension of the Difference GMM framework (Holtz-Eakin et al., 1988; Arellano and Bond, 1991), System GMM estimates a system of two equations (level and transformed). Both estimators are designed for panels with small T (time dimension) and large N (cross-sectional dimension). The Arellano-Bond estimator first transforms all variables and then takes their differences before applying the GMM method, hence named difference GMM. Arellano and Bover (1995) and Blundell and Bond (1998) developed a new estimator by assuming that the first difference of the instruments is uncorrelated with the fixed effects. This leads to more instruments and increases efficiency. The Arellano-Bond estimator relies on a system of two equations: the original equation and the transformed equation. This method is called system GMM (Roodman 2009:86). The model that is set up is as below:

$$lnY_{it} = \phi lnY_{it-1} + \gamma Z'_{it} + \beta X'_{it} + d_t + \varepsilon_{it} \quad (5)$$

Where Z'_{it} stands for kontrol variables, X'_{it} for explanatory variables and d_t for fixed effect.

Findings and Discussion

Table 2 provides descriptions of the variables used and the sources from which they were obtained. The CEE measures were calculated using data envelopment analysis, as detailed above, with

total greenhouse gas emissions as an undesirable output. The first three calculations represent DDF; when calculating DDF, capital per employee (Cap_emp), employment (emp), and energy use (enuse) were considered as inputs, GDP and the number of articles (artic) as outputs, and total greenhouse gas (Green_H) emissions as an undesirable output.

Table 1: Variables

Variable	Description	Source
pat	Resident plus non-resident patent applications.	World
Pop	Population in millions	Penn W
GDP	Real GDP at constant 2017 national prices (in mil. 2017US$)	PWT
emp	Employment	PWT
Green_H	Total greenhouse gas emissions (kt of CO_2 equivalent)	WB
Cap_emp	Capital per worker: Calculated by taking capital stock and employment data from PWT.	PWT
k_GDP	GDP per capita	PWT
Com_re_wa	Combustable renewables and waste: Combustible renewables and waste comprise solid biomass, liquid biomass, biogas, industrial waste, and municipal waste, measured as a percentage of total energy use.	WB
Corr	Control of Corruption captures perceptions of the extent to which public power is exercised for private gain. Ranging from approximately -2.5 to 2.5.	WB
Enuse	Energy use (kg of oil equivalent) per $1,000 GDP (constant 2017 PPP)	WB
Gini	The Gini index.	WB
Gov_Eff	Government Effectiveness captures perceptions of the quality of public services and the degree of its independence from political pressures. Estimate gives the country's score on the aggregate indicator, in units of a standard normal distribution, i.e. ranging from approximately -2.5 to 2.5.	WB
ICT	Information and Communicaiton Technology goods exports (% of total goods exports)	
Ind_v	Industry (including construction), value added (% of GDP)	WB
Tour	International tourism, expenditures (% of total imports)	WB
Hi_tec	The proportion of medium and high-tech industry value added in total value added of manufacturing	WB
Ex_hi	Merchandise exports to high-income economies (% of total merchandise exports)	WB
Ex_lo_mi	Merchandise exports to low- and middle-income economies in East Asia & Pacific (% of total merchandise exports)	WB

Pop_den	People per sq. km of land area	WB
R_e_c	Renewable energy consumption (% of total final energy consumption)	WB
law	Rule of Law: Ranging from approximately -2.5 to 2.5.	WB
artic	Scientific and technical journal articles.	WB
second	Secondary general pupils.	WB
tax	Tax revenue (% of GDP)	WB
trade	Trade is the sum of exports and imports of goods and services measured as a share of gross domestic product. % of GDP	WB
HHI	Herfindahl–Hirschman index	World Trade S
Env_tax	Environmentally related tax revenue	OECD
Dv1	Global non-radial DDF	-
Dv2	Global radial DDF	-
Dv3	Non-global Radial DDF	-
tfp1	Global non-radial TFP	-
tfp2	Global radial TFP	-
tfp3	Non-global radial TFP	-

Table 2 presents the 17-year average carbon emissions efficiency values for the 34 countries that subjected in the study. Since Dv1 is calculated using a non-radial DDF, it can take values greater than 1

Table 2: Average CEE Values

country	Dv1	Dv2	Dv3	tfpch1	tfpch2	tfpch3
Algeria	27.74124	0.749488	0.725671	2.425405882	0.941176471	0.941212
Australia	1.641606	0.084724	0.026035	0.077335294	0.958447059	0.972129
Bahrain	17.02186	0.500394	0.487318	-0.207717647	0.938741176	0.935882
Belgium	2.212729	0.104776	0.052759	0.030423529	0.948052941	0.948112
Canada	1.550253	0.060871	0.0245	0.085623529	0.961552941	0.9664
Chile	9.226476	0.566218	0.534518	0.507688235	0.941241176	0.941335
Colombia	16.76401	0.568876	0.5105	1.778188235	0.945223529	0.954524
Finland	2.451618	0.053406	0	0.034817647	0.960764706	0.722176
France	1.964753	0.137582	0.090376	0.035717647	0.948723529	0.948865
Germany	0.688553	0.026112	0	0.069464706	0.948452941	0.959524
Greece	3.391947	0.229765	0.179888	0.109482353	0.962852941	0.974888
Iran	10.81377	0.486035	0.445641	1.629423529	0.971941176	0.980859
Ireland	2.528765	0.104418	0.046388	0.102705882	0.965029412	0.979371

Israel	2.726659	0.034406	0	0.000305882	0.944205882	0.873489
Italy	1.2612	0.089641	0.059465	0.111135294	0.965582353	0.975776
Japan	0.332088	0.0108	0	0.054158824	0.946352941	0.953971
Kuwait	10.8051	0.276047	0.214259	-0.286205882	0.934611765	0.934159
Lebanon	9.8366	0.574194	0.543841	0.360323529	0.941523529	0.942353
Lithuania	8.637647	0.214694	0	0.508764706	0.941317647	0.942282
Mexico	5.524324	0.284459	0.224724	0.394347059	0.954858824	0.955376
Morocco	20.25089	0.792853	0.763482	0.719652941	0.941323529	0.941559
Netherlands	1.331806	0.037759	0.000588	0.062829412	0.949605882	0.958029
Norway	1.675306	0.017347	0	0.107158824	0.947658824	0.954676
Qatar	16.29879	0.157524	0.117006	1.643188235	0.922917647	0.658635
ROK	2.511947	0.196376	0.091265	0.234041176	0.972735294	0.987488
S.Arabia	11.1557	0.342747	0.294824	0.383864706	0.927594118	0.925582
Slovakia	5.643818	0.463741	0.423406	0.215611765	0.948788235	0.951824
Spain	1.7692	0.115512	0.059465	0.078364706	0.960947059	0.977871
Sweden	0.836324	0.0299	0	0.031441176	0.948311765	0.951829
Switzerland	0.519941	0.009953	0	0.029917647	0.944141176	0.959241
Tunisia	16.37142	0.304106	0	0.9052	0.941482353	0.941588
Turkey	7.831618	0.080441	0.018665	0.669258824	0.966482353	0.973435
UK	0.544865	0.027059	0	0.044252941	0.948729412	0.954718
USA	0.402918	0.020541	0	0.037564706	0.951335294	0

In the table 3 below the descritive and abstract informations can be seen. Coloumns contains variables shor name, number of obsorvations, means, standart deviations, minimum and maximum values respectively.

Table 3: Descriptive Statistics

Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
lnpop	578	2.891.878	1.362.531	-0.5234584	5.777.702
lnpat	492	8.360.368	2.147.872	3.912.023	1.331.393
lngdp	578	1.324.473	1.350.402	1.043.082	1.676.507
lnk_gdp	578	1.035.285	0.679489	8.312.688	1.170.146
lncom_re_wa	488	0.6811006	185.247	-5.758.186	3.181.966
lncorr	402	-0.081424	1.009.231	-5.117.703	0.897765

ln_enuse	533	7.979.632	0.988651	4.489.177	997.211
lngini	354	3.534.012	0.183989	3.206.803	407.244
lngov_eff	441	-0.013879	1.051.698	-5.794.152	2.539.948
lnICT	568	1.226.049	4.238.443	-1.591.271	2.164.754
lnind_V	551	3.281.246	0.352007	2.267.932	4.314.981
lntour	454	2.524.834	4.123.013	-0.2124116	2.372.726
lnhi_tec	578	3.500.643	0.587057	0.2397029	4.575.464
lnex_hi	578	4.120.004	0.789877	-0.3425423	456.909
lnex_lo_mi	578	1.479.466	1.705.593	-2.407.225	8.627.081
lnr_e_c	544	1.852.876	2.157.655	-460.517	8.712.359
lnlaw	428	0.4429615	2.884.659	-4.860.177	1.388.963
lnartic	578	9.210.205	1.933.653	3.785.552	1.296.799
lnsecond	521	1.362.708	2.227	3.867.995	1.702.263
lntax	442	2.862.812	0.525303	0.8387057	4.314.621
lntrade	578	4.466.008	11.964	2.973.466	1.082.157
lnHHI	569	-2.437.144	0.651757	-3.401.813	-0.34597
lnDv1	532	1.319.019	1.305.148	-4.110.474	3.988.021
lntfp1	422	-1.713.643	1.641.534	-6.812.445	256.324
lntfp2	544	0.008375	0.037635	-0.2604561	0.329807
lnDv2	515	-2.094.151	1.537.127	-8.517.193	-0.20579
lnDv3	320	-1.623.519	1.200.553	-7.600.903	-0.24987
lntfp3	511	0.0139999	0.053448	-0.3438998	0.400654
lnenv_tax	448	0.6706575	0.446803	-2.302.585	1.386.294

When the total factor productivity series are also examined, it is observed that, with a few exceptions such as Iran, developed countries have higher CEE series, as expected. This strengthens our confidence in the quality of our calculated series. The series have been prepared in the way explained, for econometric analysis. The results of the analysis using the system GMM estimator are presented in Tables 4-8.

Table 4 shows the results of the system GMM analysis examining the impact of environmental taxes (env_tax) on CEE (Y). The table includes both short-term and long-term (value in the last

column) effects, along with control variables. The second part of the table shows the results of diagnostic tests. Here, l.lnY is the endogenous instrument, while others are exogenous instruments. The Dv series (Dv1, Dv2, and Dv3) representing CEE indicate the distance vector, meaning the distance from the efficiency frontier. The tfp series (tfp1, tfp2, and tfp3) represent green total factor productivity (GTFP). Therefore, the results related to these series representing CEE should be the opposite of each other. In other words, factors that decrease the Dv series are expected to increase tfp. In other words, high Dv means low CEE, and high tfp means high CEE. Therefore, it is expected that the signs of Dv and tfp will be opposite in the established models. The lagged effect of Y, l.lnY, is significant and positive, except for the last model.

The impact of environmental taxes on CEE is significant in all models and, as we have detailed above, is mostly positive, aligning with our expectations. When Dv1 and Dv3 are dependent variables, a negative short-term impact is observed, meaning it increases CEE; however, the long-term impact is positive and highly significant. When tfp is the dependent variable, the impact of environmental taxes on CEE becomes more pronounced; both in the short and long term, environmental taxes are observed to increase environmental tfp, or CEE. This demonstrates that environmental taxes have a strong positive impact on CEE and environmental quality. It strongly highlights the importance of the role of governments in combating environmental pollution. Additionally, the lagged effect of CEE, l.lnY, is significant and positive in all models except the last one. This reveals the dynamic nature of the model. According to this, the past values of the dependent variable affect the current state, in other words, CEE is influenced by its own past. Because of factors such as technological infrastructure, regulatory policies, and institutions, when CEE increases, this

increase tends to be permanent. Diagnostic tests conducted on our model also reveal that the model is highly valid.

Table 4: The effect of environmental taxes on CEE

Y	l.Y	lnenv_tax	lnpop	lnhi_tec	lnex_hi	lnind_V	lngini	lnr_e_c	lntrade	uz. Dön
lnDv1	1.135	-0.0528	-0.0519	0.1719	-0.1576	-0.164	-0.2427	-0.0128	-0.1544	0.3917
	0.000***	0.009***	0.002***	0.000***	0.03**	0.054*	0.023**	0.072*	0.000***	0.001***
lnDv2	0.9709	0.0501	-0.0147	-0.0769	-0.1055	0.2497	0.2778	-0.0074	-0.022	1.726
	0.000***	0.234	0.239	0.441	0.237	0.01**	0.14	0.717	0.668	0.703
lnDv3	1.309	-0.075	0.005	-0.047	0.929	-0.483	-0.821	0.041	-1.758	0.244
	0.000***	0.869	0.948	0.845	0.398	0.166	0.749	0.563	0.765	0.859
lntfp1	0.123	0.467	0.202	-0.713	0.346	1.481	2.696	0.06	0.586	0.552
	0.001***	0.000***	0.000***	0.000***	0.101	0.000***	0.000***	0.342	0.000***	0.000***
lntfp2	0.05	0.005	0.002	0.001	-0.004	0.01	-0.013	-0.002	-0.001	0.005
	0.000***	0.002***	0.017**	0.667	113	0.024**	0.016**	0.01**	0.609	0.001***
lntfp3	-0.039	0.008	0.004	0.002	-0.006	0.022	-0.021	-0.003	-0.001	0.008
	0.041**	0.002***	0.051*	0.68	0.273	0.000***	0.03**	0.017	0.883	0.000***

Y	NOI	NOG	AR(1)	AR(2)	Sargan	Hansen
lnDv1	25	27	0.088	0.477	0.002	0.536
lnDv2	25	27	0.133	0.246	0.000***	0.411
lnDv3	25	15	0.193	0.376	0.1	0.938
lntfp1	24	26	0.01	0.684	0.017	0.353
lntfp2	24	27	0.044	0.515	0.19	0.57
lntfp3	24	26	0.041	0.341	0.002	0.461

Given that a two-step system GMM estimation is employed, the Hansen test statistic is more crucial. Nevertheless, the results of both Sargan, Hansen, and Arellano-Bond- AR(1) and AR(2)-autocorrelation tests are presented here. The purpose of the Arellano-Bond test is to determine whether the error term is autocorrelated. In all models, the number of instruments is smaller than the number of groups. This strengthens the validity of the model. The results of

autocorrelation tests are also largely positive. Additionally, a comparison of the number of instruments (NOI) and the number of groups (NOG) is provided.

The null hypothesis of the AR(1) test is "H_0: There is first-order autocorrelation". When this hypothesis is tested against the alternative hypothesis "H_1: There is no first-order autocorrelation", it is rejected in some models but accepted in others. However, when we consider the AR(2) test, which is more crucial in this context, all models indicate that there is no second-order autocorrelation. The null hypothesis of the AR(2) test is "H_0: There is no second-order autocorrelation", and it is tested against the alternative hypothesis "H_1: There is second-order autocorrelation". In all models, the null hypothesis is accepted, indicating that there is no autocorrelation in the model.

The null hypothesis of the Hansen test is "H_0: the model is valid". We need to test this hypothesis against the alternative hypothesis which suggests that the model is invalid. As shown in the second part of Table 4, the alternative hypothesis is strongly rejected in all models, and the null hypothesis, which states that the model is valid, is accepted. This indicates that the established models are quite robust, meaning that the instruments are uncorrelated with the error term and the excluded variables are correctly excluded. Table 5 shows the results regarding the impact of normal taxes on CEE.

Table 5: Impact of regular taxes on CEE

Y	l.Y	lntax	lnpop	lnhi_tec	lnex_hi	lnind_V	lngini	lnr_e_c	lntrade	uz. Dön.
lnDv1	1.156	-0.252	-0.057	0.161	-0.15	-0.367	0.193	-0.01	0.0123	1.612
	0.000**	0.01**	0.003**	0.008**	0.026*	0.000**	0.051	0.428	0.000**	0.069*
lnDv2	1.03	0.67	0.042	0.053	0.061	0.539	0.276	0.036	-0.053	-21.857
	0.000**	0.009**	0.263	0.696	0.552	0.000**	0.018**	0.196	0.371	0.576
lnDv3	1.225	0.3	0.007	0.169	0.45	-0.414	0.512	-0.011	0.016	-1.334

	0.000***	0.42	0.875	0.174	0.34	0.023	0.577	0.734	0.94	0.488
lntfp1	0.2	301	0.18	-0.693	0.182	1.536	1.305	-0.005	0.482	0.376
	0.000***	0.424	0.017**	0.000***	0.345	0.000***	0.000*	0.939	0.000***	0.412
lntfp2	0.047	0.022	0.003	0.007	-0.008	0.031	-0.016	-0.002	-0.005	0.023
	0.000***	0.000***	0.005**	0.019**	0.083*	0.000***	0.031**	0.007**	0.019**	0.000***
lntfp3	-0.008	0.032	0.003	0.008	-0.009	0.045	-0.018	-0.007	-0.007	0.032
	0.571	0.03**	0.072*	0.175	0.028**	0.001***	0.033**	0.197	0.038**	0.019**

Y	NOI	NOG	AR(1)	AR(2)	Sargan	Hansen
lnDv1	25	26	0.087	0.49	0.000***	0.474
lnDv2	25	26	0.0142	0.21	0.000***	0.459
lnDv3	25	16	0.202	0.495	0.426	0.963
lntfp1	24	26	0.011	0.926	0.056	0.217
lntfp2	24	26	0.045	0.522	0.361	0.353
lntfp3	24	25	0.039	0.364	0.005	0.364

As seen in the table, although the effect of normal taxes on CEE is not as strong as that of environmental taxes, it is significant and positive in both the short and long run in models represented by tfp. However, in models represented by dv, while it is somewhat ambiguous, the negative effect, especially in the long run, dominates. The impact of normal taxes on environmental quality, while not direct, occurs indirectly through controlling consumption and economic expansion through high taxes. In all models, the number of instruments exceeds the number of groups. According to the Arellano-Bond test results, there is no second-order autocorrelation in the model, and according to the Hansen test results, all models are valid.

Table 6: The impact of rule of law on CEE

Y	l.Y	lnlaw	lnpop	lnhi_tec	lnex_hi	lnind_V	lngini	lnr_e_c	lntrade	uz. Dön

lnDv1	1.215	0.019	-0.014	0.245	-0.023	-0.073	0.057	0.014	-0.127	-0.087
	0.000***	0.128	0.489	0.000***	0.851	0.414	0.82	0.253	0.03**	0.162
lnDv2	0.936	-0.126	-0.037	-0.03	-0.529	0.032	-0.328	-0.008	-0.051	-2.007
	0.000***	0.000***	0.609	0.898	0.4	0.84	0.821	0.935	0.855	0.541
lnDv3	0.854	-0.638	-0.073	0.172	0.48	4.045	0.779	0.034	-0.826	-4.374
	0.24	0.758	0.827	0.962	0.946	0.517	0.983	0.943	0.781	0.883
lntfp1	0.208	-0.067	0.185	-0.61	0.393	0.983	1.765	0.036	0.707	-0.085
	0.000***	0.447	0.002***	0.003***	0.287	0.000***	0.001***	0.521	0.001	0.428
lntfp2	0.058	0.002	-0.004	-0.002	0.004	0.006	-0.027	-0.004	-0.005	0.003
	0.000***	0.000***	0.472	0.644	0.129	0.076**	0.000***	0.000***	0.002***	0.000***
lntfp3	-0.005	0.004	0.001	-0.004	0.001	0.013	-0.034	-0.005	-0.006	0.005
	0.753	0.001***	0.42	0.531	0.792	0.206	0.004***	0.001	0.052	0.000***

Y	NOI	NOG	AR(1)	AR(2)	Sargan	Hansen
lnDv1	25	24	0.102	0.401	0	0.468
lnDv2	25	24	0.164	0.288	0	0.366
lnDv3	25	11	0.242	0.475	0.552	1
lntfp1	23	23	0.025	0.355	0.002	0.212
lntfp2	24	24	0.078	0.416	0.092	0.499
lntfp3	24	23	0.06	0.347	0	0.576

According to the results in Table 6, the legal system has a net positive effect on CEE. In models with DV, the coefficients of the law variable are negative in both the short and long run, meaning that as the level of the legal system develops, the distance from the frontier decreases. In models with tfp, however, there is a positive effect in both the short and long run. Additionally, although the Arellano-Bond and Hansen tests show that the models pass the tests for second-order autocorrelation and validity, the number of instruments exceeds the number of groups. Similar results are

observed in Table 7 for the government effectiveness series; as government effectiveness increases, CEE and environmental quality improve.

Table 7: Impact of effectiveness of government on CEE

Y	l.Y	lnGov_eff	lnpop	lnhi_tec	lnex_hi	lnind_V	lngini	lnr_e_c	lntrade	uz. Dön
lnDv1	0.407	-0.843	-0.553	-0.608	-0.966	1.52	0.278	-0.2	-0.098	-1.432
	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.552	0.007***	0.061**	0.000***
lnDv2	0.883	-0.17	-0.051	-0.089	-0.19	-0.288	0.037	-0.02	-0.024	-1.537
	0.000***	0.000***	0.008***	0.387	0.045**	0.129	0.863	0.631	0.483	0.104
lnDv3	2.718	-0.948	1.62	3.131	8.862	-5.582	23.854	-0.556	4.78	0.552
	0.073*	0.147	0.287	0.209	0.248	0.467	0.271	0.368	0.336	0.048**
lntfp1	0.188	-0.122	0.161	-0.636	0.357	1.145	1.52	-0.014	0.65	-0.151
	0.000***	0.445	0.004***	0.003***	0.283	0.000***	0.002***	0.805	0.001***	0.428
lntfp2	0.045	0.005	-0.003	-0.004	-0.008	0.004	-0.025	-0.004	-0.003	0.005
	0.000***	0.000***	0.65	0.306	0.071*	0.34	0.001***	0.001	0.122	0.000***
lntfp3	-0.013	0.007	0.002	-0.007	-0.005	0.011	-0.035	-0.004	-0.001	0.007
	0.332	0.005***	0.116	0.333	0.244	0.169	0.013	0.006***	0.627	0.002***

Y	NOI	NOG	AR(1)	AR(2)	Sargan	Hansen
lnDv1	25	24	0.475	0.694	0	0.506
lnDv2	25	24	0.146	0.261	0	0.271
lnDv3	25	11	0.281	0.45	0.4	1
lntfp1	24	23	0.02	0.428	0	0.119
lntfp2	24	24	0.08	0.411	0.134	0.585
lntfp3	24	23	0.062	0.347	0.001	0.542

Table 8 demonstrates that the model established to test the impact of corruption on CEE is both valid and free from autocorrelation. Furthermore, only in one model, the number of

instruments exceeds the number of groups, indicating that the model is valid. The coefficients are significant in both the short and long run. In models using DV, the corruption variable exhibits a positive coefficient in both the short and long run, implying that as corruption increases, environmental quality deteriorates. Conversely, in models using tfp, the coefficient of the corruption variable is negative in only one model. However, in models where coefficients are positive, the coefficients are remarkably low which means corruption has little effect on green TFP.

Table 8: Impact of corruption on CEE

Y	l.Y	lncorr	lnpop	lnhi_tec	lnex_hi	lnind_V	lngini	lnr_e_c	lntrade	uz. Dön.
lnDv1	0.225	0.378	-1.248	-0.871	-2.586	0.004	-0.695	-0.498	-1.237	0.486
	0.034**	0.031**	0.000***	0.000***	0.001***	0.989	0.328	0.000***	0.009***	0.034**
lnDv2	-0.671	1.761	-0.369	-15.031	-3.224	0.75	-15.082	-2.297	1.219	1.053
	0.014**	0.017**	0.709	0.005***	0.002***	0.706	0.001***	0.000***	0.000***	0.000***
lnDv3	2.845	1.389	0.35	0.183	2.505	-9.922	6.086	-0.477	2.439	-0.752
	0.114	0.232	0.651	0.953	0.191	0.449	0.11	0.516	0.378	0.001***
lntfp1	0.307	-1.366	-3.482	4.965	-4.251	-5.716	3.18	-0.314	-3.987	-1.972
	0.001***	0.026**	0.002***	0.007***	0.005***	0.003***	0.234	0.211	0.003***	0.01**
lntfp2	0.011	0.002	0.015	-0.02	0.014	0.032	-0.028	-0.002	0.018	0.003
	0.034**	0.005***	0.079*	0.076*	0.338	0.006***	0.022**	0.111	0.092*	0.002***
lntfp3	-0.031	0.01	-0.023	0.012	-0.054	-0.011	-0.024	-0.01	-0.033	0.01
	0.056*	0.002***	0.007***	0.312	0.001***	0.373	0.31	0.000***	0.000***	0.000***

Y	NOI	NOG	AR(1)	AR(2)	Sargan	Hansen
lnDv1	22	23	0.271	0.995	0.001	0.401
lnDv2	22	23	0.763	0.77	0.432	0.472
lnDv3	22	11	0.5	0.37	0.705	1
lntfp1	21	22	0.02	0.553	0.247	0.328

lntfp2	22	23	0.079	0.232	0.073	0.296
lntfp3	22	22	0.062	0.337	0	0.442

Results of system GMM analysis demonstrated above clearly reveal that governmental factors which also contain regulatory characteristic have significant and strong, positive on environmental goodness. Especially environmental taxes positively influences environment both in short and long term in all models. This leads the result that government may actively involve in fight againist environmental pollution.

5-Conclutions and Recommendations

The primary motivation behind this study is to investigate whether regulatory policies implemented by public authorities which represented here by environmental taxes, regular taxes, rule of law, effectiveness of government and corruption, have strong impact on CEE and, consequently, environmental quality. This question has not been addressed in the literature sufficiently. Boyce (1994) was the first to examine the link between environmental pollution and social structure based on the different harms or benefits that different segments of society experience from pollution. However, the study discusses a result that arises from necessity rather than deliberate choices made by opposing classes. This is an outcome that benefits the rich class who both pollute the environment and have the means to protect themselves from environmental pollution. Our initial hypothesis is that regulatory policies with public sanctions are effective about the environmental goodness. The calculations and analyses conducted strongly support this hypothesis. In the first stage of the study, six CEE data sets were calculated using DDF and ML indexes, which are various versions of DEA. Subsequently, these CEE series were analyzed using the system GMM estimator.

The necessity of achieving a sustainable balance between economic prosperity and environmental pollution has become an increasingly accepted fact among researchers. CEE is a good indicator of this balance. Various indices are used to measure CEE.

However, among the methods used to calculate CEE, DEA-based methods are the most widely accepted by researchers. In our study, we used DDF and ML, which are some of the latest versions of these methods. For each index, radial, non-radial, and global versions were calculated. The hypothesis we examined for these six different CEE series was tested using the system GMM method. The reason for preferring this dynamic method is that it solves problems that arises from the fact that the CEE series is influenced by its past due to permanent factors such as technological infrastructure, and system GMM solves problems such as endogeneity, unobserved heterogeneity, and omitted variable bias, as detailed above.

The summary of the results of analysis presented in the tables above: As expected, regulatory factors have a clear and strong impact on CEE or environmental quality. Environmental taxes, normal taxes, improvements in the legal system, and increased government effectiveness positively impact CEE and, consequently, environmental quality. Corruption, on the other hand, has a clear negative impact on CEE. According to the analysis, environmental taxes increase CEE in both the long and short term. Normal taxes also increase CEE, but this effect is weaker. While increased government effectiveness increases CEE, increased corruption has the opposite effect. Therefore, it can be said that implementing measures to increase production levels without damaging environmental quality requires government intervention. Governments should make the necessary regulations for this purpose.

References

Alperen, Ö. (2018). Veri zarflama analizi ve bir uygulama (Master's thesis). Yıldız Technical University. ProQuest https://acikbilim.yok.gov.tr/bitstream/handle/20.500.12812/390141/yokAcikBilim_313781.pdf?sequence=-1

Andersen, P., & Petersen, N. C. (1993). A procedure for ranking efficient units in data envelopment analysis. Management science, 39(10), 1261-1264. https://doi.org/10.1287/mnsc.39.10.1261

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. The review of economic studies, 58(2), 277-297. https://doi.org/10.2307/2297968

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. Journal of econometrics, 68(1), 29-51. https://doi.org/10.1016/0304-4076(94)01642-D

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. Management science, 30(9), 1078-1092. https://doi.org/10.1287/mnsc.30.9.1078

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. Journal of econometrics, 87(1), 115-143. https://doi.org/10.1016/S0304-4076(98)00009-8

Boyce, J. K. (1994). Inequality as a cause of environmental degradation. Ecological economics, 11(3), 169-178. https://doi.org/10.1016/0921-8009(94)90198-8

Chambers, R. G., Chung, Y., & Färe, R. (1996). Benefit and distance functions. Journal of economic theory, 70(2), 407-419. https://doi.org/10.1006/jeth.1996.0096

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. European journal of operational research, 2(6), 429-444. https://doi.org/10.1016/0377-2217(78)90138-8

Chen, Y., Miao, J., & Zhu, Z. (2021). Measuring green total factor productivity of China's agricultural sector: A three-stage SBM-DEA model with non-point source pollution and CO_2 emissions. Journal of Cleaner Production, 318, 128543. https://doi.org/10.1016/j.jclepro.2021.128543

Chen, C., Lan, Q., Gao, M., & Sun, Y. (2018). Green total factor productivity growth and its determinants in China's industrial economy. Sustainability, 10(4), 1052. https://doi.org/10.3390/su10041052

Cheng, Y., Lv, K., Wang, J., & Xu, H. (2019). Energy efficiency, carbon dioxide emissions efficiency, and related abatement costs in regional China: a synthesis of input–output analysis and DEA. Energy Efficiency, 12, 863-877. https://doi.org/10.1007/s12053-018-9695-8

Cheng, Z., Li, L., Liu, J., & Zhang, H. (2018a). Total-factor carbon emissions efficiency of China's provincial industrial sector and its dynamic evolution. Renewable and Sustainable Energy Reviews, 94, 330-339. https://doi.org/10.1016/j.rser.2018.06.015

Cheng, Z., Li, L., & Liu, J. (2018b). Industrial structure, technical progress and carbon intensity in China's provinces. Renewable and Sustainable Energy Reviews, 81, 2935-2946. http://dx.doi.org/10.1016/j.rser.2017.06.103

Chung, Y. H., Färe, R., & Grosskopf, S. (1997). Productivity and undesirable outputs: a directional distance function approach. journal of Environmental Management, 51(3), 229-240. https://doi.org/10.1006/jema.1997.0146

Ding, L., Yang, Y., Wang, W., & Calin, A. C. (2019). Regional carbon emissions efficiency and its dynamic evolution in China: A novel cross efficiency-malmquist productivity index. Journal of Cleaner

Production, 241, 118260. https://doi.org/10.1016/j.jclepro.2019.118260

Färe, R., Grosskopf, S., & Roos, P. (1999). Price Indexes for Nonmarketed Goods. Data Envelopment Analysis in the Service Sector, 121-132.

Färe, R., & Grosskopf, S. (2010). Directional distance functions and slacks-based measures of efficiency. European journal of operational research, 200(1), 320-322. https://doi.org/10.1016/j.ejor.2009.01.031

Fukuyama, H., & Weber, W. L. (2009). A directional slacks-based measure of technical inefficiency. Socio-Economic Planning Sciences, 43(4), 274-287. https://doi.org/10.1016/j.seps.2008.12.001

Halkos, G., & Petrou, K. N. (2019). Treating undesirable outputs in DEA: A critical review. Economic Analysis and Policy, 62, 97-104. https://doi.org/10.1016/j.eap.2019.01.005

Holtz-Eakin, D., Newey, W., & Rosen, H. S. (1988). Estimating vector autoregressions with panel data. Econometrica: Journal of the econometric society, 1371-1395. https://doi.org/10.2307/1913103

Jahanshahloo, G. R., Lotfi, F. H., Shoja, N., Tohidi, G., & Razavyan, S. (2005). Undesirable inputs and outputs in DEA models. Applied mathematics and computation, 169(2), 917-925. https://doi.org/10.1016/j.amc.2004.09.069

Kaya, Y. & Yokobori, K. (1993). Environment, Energy and Economy:Strategies for Sustainability. Delhi:Bookwell Publications, pp. 165-177.

Kutlar, A. & Bakırcı, F. (2018). Veri Zarflama Analizi: Teori ve uygulama (1th ed.). Orion Kitapevi.

Li, H., Zhang, J., Osei, E., & Yu, M. (2018). Sustainable development of China's industrial economy: An empirical study of the period 2001–2011. Sustainability, 10(3), 764. https://doi.org/10.3390/su10030764

Li, Y., & Chen, Y. (2021). Development of an SBM-ML model for the measurement of green total factor productivity: The case of pearl river delta urban agglomeration. Renewable and Sustainable Energy Reviews, 145, 111131.

Liu, Y., Zhao, G., & Zhao, Y. (2016). An analysis of Chinese provincial carbon dioxide emissions efficiencies based on energy consumption structure. Energy Policy, 96, 524-533. http://dx.doi.org/10.1016/j.enpol.2016.06.028

Lovell, C. K., Pastor, J. T., & Turner, J. A. (1995). Measuring macroeconomic performance in the OECD: A comparison of European and non-European countries. European journal of operational research, 87(3), 507-518. https://doi.org/10.1016/0377-2217(95)00226-X

Ma, Y., Lin, T., & Xiao, Q. (2022). The Relationship between Environmental Regulation, Green-Technology Innovation and Green Total-Factor Productivity—Evidence from 279 Cities in China. International Journal of Environmental Research and Publich Health, 19(23), 1-22. https://doi.org/10.3390/ijerph192316290

Oh, D. H. (2010). A global Malmquist-Luenberger productivity index. Journal of productivity analysis, 34, 183-197. DOI 10.1007/s11123-010-0178-y

Pan, X., Pan, X., Li, C., Song, J., & Zhang, J. (2019). Effects of China's environmental policy on carbon emissions efficiency. International Journal of Climate Change Strategies and Management, 11(3), 326-340. https://doi.org/10.1108/IJCCSM-12-2017-0206

Reinhard, S., Lovell, C. K., & Thijssen, G. (1999). Econometric estimation of technical and environmental efficiency: an application to Dutch dairy farms. american Journal of agricultural economics, 81(1), 44-60. https://doi.org/10.2307/1244449

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. The stata journal, 9(1), 86-136. https://doi.org/10.1177/1536867X0900900106

Seiford, L. M., & Zhu, J. (2002). Modeling undesirable factors in efficiency evaluation. European journal of operational research, 142(1), 16-20. https://doi.org/10.1016/S0377-2217(01)00293-4

Shahverdi, R., & Ebrahimnejad, A. (2014). DEA and Malmquist productivity indices for measuring group performance in two periods. International Journal of Industrial and Systems Engineering, 16(3), 382-395. https://doi.org/10.1504/IJISE.2014.060135

Streimikis, J., & Saraji, M. K. (2022). Green productivity and undesirable outputs in agriculture: A systematic review of DEA approach and policy recommendations. Economic research-Ekonomska istraživanja, 35(1), 819-853. https://doi.org/10.1080/1331677X.2021.1942947

Tang, J., & Qin, F. (2022). Analyzing the impact of local government competition on green total factor productivity from the factor market distortion perspective: based on the three stage DEA model. Environment, Development and Sustainability, 1-29. https://doi.org/10.1007/s10668-021-02032-5

Tol, R. S., Pacala, S. W., & Socolow, R. H. (2009). Understanding long-term energy use and carbon dioxide emissions in the USA. Journal of Policy Modeling, 31(3), 425-445. doi:10.1016/j.jpolmod.2008.12.002

Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. European journal of operational research, 130(3), 498-509. https://doi.org/10.1016/S0377-2217(99)00407-5

Wang, D., Du, K., & Zhang, N. (2022). Measuring technical efficiency and total factor productivity change with undesirable outputs in Stata. The Stata Journal, 22(1), 103-124. https://doi.org/10.1177/1536867X221083886

Xie, Z., Wu, R., & Wang, S. (2021). How technological progress affects the carbon emissions efficiency? Evidence from national panel quantile regression. Journal of Cleaner Production, 307, 127133. https://doi.org/10.1016/j.jclepro.2021.127133

Zhao, P., Zeng, L., Li, P., Lu, H., Hu, H., Li, C., ... & Qi, Y. (2022). China's transportation sector carbon dioxide emissions efficiency and its influencing factors based on the EBM DEA model with undesirable outputs and spatial Durbin model. Energy, 238, 121934. https://doi.org/10.1016/j.energy.2021.121934

Zheng, R., Cheng, Y., Liu, H., Chen, W., Chen, X., & Wang, Y. (2022). The spatiotemporal distribution and drivers of urban carbon emissions efficiency: the role of technological innovation. International Journal of Environmental Research and Public Health, 19(15), 9111. https://doi.org/10.3390/ijerph19159111

Zhou, Z., Wu, H., Ding, T., & Xia, Q. (2019). Nonradial directional distance function for measuring the environmental efficiency of the Chinese iron and steel industry. Tropical Conservation Science, 12, 1940082919853755. https://doi.org/10.1177/1940082919853755

Zhou, P., Ang, B. W., & Wang, H. (2012). Energy and CO_2 emissions performance in electricity generation: a non-radial directional distance function approach. European journal of operational research, 221(3), 625-635. http://dx.doi.org/10.1016/j.ejor.2012.04.022

Zhu, R., Zhao, R., Sun, J., Xiao, L., Jiao, S., Chuai, X., ... & Yang, Q. (2021). Temporospatial pattern of carbon emissions efficiency of China's energy-intensive industries and its policy implications. Journal of Cleaner Production, 286, 125507. https://doi.org/10.1016/j.jclepro.2020.125507

CHAPTER 4

The Nexus of Government Debt and Income Inequality: A Panel Data Analysis of Five European Economies

ARZU ALVAN[1]
BARTU YURDAY[2]

Introduction

Income inequality has emerged as a critical socioeconomic challenge in the 21st century, with far-reaching implications for economic growth, social cohesion, and political stability (Piketty, 2014; Stiglitz, 2012). While the determinants of income inequality are multifaceted, the role of fiscal policy, particularly government debt, has gained increasing attention in recent years (Salti, 2015; Azzimonti et al., 2014). This study aims to elucidate the complex relationship between income inequality and general government debt, focusing on five diverse European economies: Turkey, Netherlands, Switzerland, Poland, and Spain.

[1] Assoc. Prof. Dr., Cyprus Science University, Faculty of Economics, Administrative and Social Sciences, [Department to be confirmed], ORCID: 0000-0003-2183-0209
[2] [Title to be confirmed], [Institution to be confirmed], [Department to be confirmed], ORCID: [not provided in source]

The selection of these countries is deliberate, representing a spectrum of economic development, fiscal policies, and social structures within the European context. Turkey, as an emerging economy, offers insights into the dynamics of rapidly evolving markets. The Netherlands and Switzerland, as advanced economies with distinct social models, provide contrasting perspectives. Poland, a post-transition economy, and Spain, a southern European nation with recent fiscal challenges, further enrich the comparative analysis.

The research employs a sophisticated panel data analysis approach, utilizing yearly data from 2010 to 2021. This timeframe encompasses the aftermath of the global financial crisis and the subsequent period of economic recovery and transformation, allowing for a nuanced examination of the evolving relationship between inequality and government debt. Data are sourced from authoritative databases including the Organisation for Economic Co-operation and Development (OECD), the World Bank, and the International Monetary Fund (IMF), ensuring reliability and comparability across nations.

The methodological framework incorporates advanced econometric techniques, including panel unit root tests to assess stationarity, panel cointegration tests to examine long-run relationships, and Fully Modified Ordinary Least Squares (FMOLS) estimation to quantify the magnitude and direction of these relationships. This rigorous approach allows for a comprehensive analysis that accounts for both cross-sectional heterogeneity and temporal dynamics.

Building upon the theoretical foundations laid by scholars such as Ostry et al. (2014) and Kumhof et al. (2015), who have explored the macroeconomic implications of inequality, this study contributes to the existing literature in several key ways. First, it provides a focused analysis on the nexus between government debt

and income inequality in a diverse set of European economies, addressing a gap in comparative studies within this region. Second, by employing panel data techniques, the study offers robust insights into both short-term fluctuations and long-term structural relationships. Third, the research has significant policy implications, potentially informing fiscal strategies aimed at mitigating inequality while managing public debt sustainably.

Conceptual and Theoretical Framework

Income inequality has long been a central topic in economics, as it directly impacts economic growth, social stability, and the well-being of individuals. Over the years, researchers have explored the causes and consequences of income inequality, as well as its relationship with various economic and social factors, including government debt. This section reviews the relevant literature and key theoretical frameworks that underpin the analysis, while identifying gaps that this study aims to address.

Income Inequality and Economic Growth

One of the most widely studied aspects of income inequality is its relationship with economic growth. Brueckner and Lederman (2018), in their study titled "Inequality and Economic Growth: The Role of Initial Income and Prices," published in the Journal of Economic Growth, provide a comprehensive analysis of this relationship. Using a panel dataset of 119 countries from 1960 to 2000, the authors find that income inequality negatively impacts economic growth, both directly through reduced aggregate demand, and indirectly through its adverse impact on investment.

Similarly, Alesina, Di Tella, and MacCulloch (2004), in their work "Inequality and Economic Growth: The Role of Institutions," published in the Journal of Economic Growth, emphasize the importance of institutional quality in mediating the relationship between inequality and growth. Their study argues that high levels

of income inequality can hinder long-term economic growth, particularly in countries with weak institutions, as inequality often leads to political instability, which in turn discourages investment and innovation.

More recently, Quispe-Torreblanca et al. (2021) conducted a comprehensive review of the empirical literature on the economic consequences of income inequality. Their study, titled "A Review of the Empirical Literature on the Economic Consequences of Income Inequality," examines the impact of income inequality on a wide range of economic outcomes, including growth, labor market dynamics, health, social capital, and political stability. They conclude that the majority of studies find a negative relationship between income inequality and these outcomes, reinforcing the argument that reducing inequality is essential for achieving broader economic and social goals.

Government Policies and Income Distribution

Government policies play a crucial role in shaping income distribution and addressing inequality. Burkhauser, Larrimore, and Simon (2013), in their study "Inequality and Government Policy," provide a detailed review of the literature on the relationship between government policies and income inequality. The authors find that government interventions, such as progressive taxation, social welfare programs, and labor market regulations, can significantly reduce income inequality. However, they also caution that poorly designed policies may exacerbate inequality or create inefficiencies that hinder economic growth.

Ostry, Berg, and Tsangarides (2014), in their influential paper "Redistribution, Inequality, and Growth," published by the International Monetary Fund, explore the trade-offs between redistribution and economic growth. Their findings challenge the conventional view that redistribution necessarily hampers growth.

Instead, they argue that well-targeted redistributive policies can reduce inequality without adversely affecting growth, and in some cases, may even enhance it.

Government Debt and Income Inequality

While the relationship between income inequality and government policies has been extensively studied, the specific link between income inequality and government debt remains relatively underexplored. Government debt can influence income inequality through various channels, including fiscal policy, public spending, and taxation. For instance, high levels of government debt may constrain a country's ability to invest in social programs and infrastructure, thereby exacerbating inequality. Conversely, targeted borrowing to fund redistributive policies or public investments can help reduce inequality.

A study by Woo et al. (2013), titled "Distributional Consequences of Fiscal Consolidation," published in the IMF Economic Review, examines the impact of fiscal consolidation on income inequality. The authors find that austerity measures, often implemented to reduce government debt, tend to disproportionately affect low-income households, thereby increasing inequality. This suggests that the relationship between government debt and income inequality is complex and context-dependent, requiring careful consideration of the distributional effects of fiscal policies.

Furceri, Loungani, and Ostry (2017), in their study titled "The Aggregate and Distributional Effects of Fiscal Consolidation," analyze the short- and long-term effects of fiscal consolidation on income inequality across a panel of countries. They find that fiscal consolidation often leads to higher inequality, particularly in the absence of compensatory measures such as targeted social spending or progressive taxation. These findings underscore the importance of

designing fiscal policies that minimize the adverse distributional effects of debt reduction strategies.

Despite the growing body of research on income inequality and its economic implications, the relationship between income inequality and government debt remains an underexplored area. Most existing studies focus on the impact of fiscal policies on inequality or the role of inequality in shaping economic growth, with limited attention to the bidirectional relationship between inequality and government debt. Moreover, there is a lack of comparative studies examining this relationship across countries with similar economic indicators. The current study aims to address these gaps by analyzing the relationship between general government debt and income inequality in a diverse set of European economies.

Theoretical Frameworks

The study of income inequality and its relationship with various economic and social variables has been guided by several theoretical frameworks. These frameworks provide valuable insights into the mechanisms driving inequality and its implications for economic development, government policies, and institutional structures.

The Kuznets Curve Hypothesis

One of the earliest and most influential theoretical frameworks for understanding income inequality is the Kuznets curve hypothesis, proposed by Simon Kuznets (1955). The Kuznets curve posits an inverted-U relationship between income inequality and economic development: income inequality initially rises during the early stages of economic development as industrialization and urbanization create disparities between rural and urban incomes, but eventually declines due to structural changes such as expanded education, social mobility, and redistributive policies. While widely used, the Kuznets curve has faced criticism for its inconsistent

empirical support and limited applicability to contemporary economies, particularly in the context of globalization and technological change (Piketty, 2014).

The Institutional Approach

The institutional approach offers an alternative perspective on the relationship between income inequality and economic development, emphasizing the role of institutions in shaping economic outcomes. This framework argues that institutions such as property rights, the rule of law, and the strength of the state play a critical role in determining the distribution of income within a society. Strong institutions can promote equitable growth by ensuring fair access to resources and fostering social cohesion, while weak or extractive institutions can exacerbate income inequality by enabling rent-seeking behavior and the concentration of wealth (Acemoglu & Robinson, 2012). Alesina, Di Tella, and MacCulloch (2004) provide a comprehensive overview of this approach, highlighting the importance of institutional quality in mediating the effects of economic growth on inequality.

The Dual Economy Theory

The dual economy theory, originally developed to explain structural differences between sectors in developing economies, has also been applied to the study of income inequality (Lewis, 1954). This framework posits that many economies are characterized by a dual structure consisting of a modern, high-productivity sector and a traditional, low-productivity sector, creating significant income disparities between the two. Duflo and Banerjee (2011) argue that the persistence of dual economies is a major driver of inequality in both developing and developed countries, highlighting the importance of policies aimed at reducing structural barriers to economic mobility.

The Behavioral Approach

The behavioral approach to income inequality draws on insights from psychology and sociology to understand how individual and group behaviors influence income distribution. This framework emphasizes the role of cultural norms, social attitudes, and individual decision-making in shaping income inequality (Thaler & Sunstein, 2008). Research has shown that high levels of inequality can erode social trust, increase feelings of relative deprivation, and contribute to social unrest (Wilkinson & Pickett, 2010). These findings suggest that addressing income inequality requires not only economic and institutional reforms but also efforts to change societal attitudes and behaviors.

Integrating Theoretical Frameworks

While each of the theoretical frameworks discussed above offers valuable insights into the dynamics of income inequality, they are not mutually exclusive. A comprehensive understanding of income inequality requires an integrative approach that combines elements from multiple frameworks. For example, the Kuznets curve hypothesis can be enriched by incorporating insights from the institutional approach, which highlights the role of governance and policy in shaping the relationship between inequality and development. Similarly, the dual economy theory can be complemented by the behavioral approach, which provides a deeper understanding of the social and cultural factors that sustain economic duality. The current study adopts this integrative perspective, drawing on these frameworks to analyze the relationship between income inequality and government debt across diverse European economies.

Research Methodology

This study employs a panel data analysis to investigate the relationship between income inequality and general government

debt, incorporating a range of control variables to account for potential confounding factors. The primary variable of interest is the GINI index, a widely used measure of income inequality ranging from 0 (perfect equality) to 1 (complete inequality).

Model Specification

The relationship between income inequality and general government debt is modeled using the following functional form:

GI_i = f(GGD_i, GDPP_i, UNMR_i, ER_i, IR_i, PD_i, HE_i)

Where: GI_i is the GINI index (income inequality) for country i; GGD_i is general government debt as a percentage of GDP; GDPP_i is GDP per capita; UNMR_i is the unemployment rate; ER_i is the employment rate; IR_i is the inflation rate; PD_i is population density; and HE_i is health expenditure as a percentage of GDP.

To ensure robust and reliable estimates, the variables were transformed into their natural logarithmic (log-linear) forms. This transformation reduces the sharpness of fluctuations in time series data, mitigates heteroscedasticity, and allows for the interpretation of coefficients as long-run elasticities (Shahbaz, 2010). The log-linear model is expressed as follows:

LNGIit = a0 + b1(LNGGDit) + b2(LNGDPPit) + b3(LNUNMRit) + b4(LNIRit) + b5(LNERit) + b6(LNPDit) + b7(LNHEit) + eit

Where a0 is the intercept term; t is the time period; i represents cross-sectional units (countries); eit is the error disturbance term; LN denotes the natural logarithm operator; and b1 to b7 are coefficients to be estimated. This log-linear specification allows for the interpretation of the coefficients as elasticities,

providing insights into the percentage change in income inequality resulting from a 1% change in each explanatory variable.

Variables and Their Definitions

The model includes several variables alongside general government debt to control for other factors that may influence income inequality. General Government Debt (GGD) measures the total debt of the government as a percentage of GDP, serving as the primary independent variable. GDP Per Capita (GDPP) is included to account for the level of economic development. The Unemployment Rate (UNMR) measures the percentage of the labor force that is unemployed and actively seeking work. The Inflation Rate (IR) measures the percentage change in the price of a basket of goods and services. The Employment Rate (ER) measures the percentage of the working-age population that is employed. Population Density (PD) captures the potential effects of urbanization and population distribution on income inequality. Health Expenditure (HE) measures the percentage of GDP spent on healthcare, with higher health spending often associated with improved social outcomes and reduced inequality.

Data Collection and Sources

The study utilizes annual panel data for the period 2010-2021, covering five European economies: Turkey, the Netherlands, Switzerland, Poland, and Spain. The data were sourced from three reputable international organizations. First, the Organisation for Economic Co-operation and Development (OECD) provides comprehensive data on income inequality, government debt, and other macroeconomic indicators for its member countries. Second, the World Bank offers a wide range of economic and social indicators, including GDP per capita, unemployment rates, and population density. Third, the International Monetary Fund (IMF) provides data on fiscal variables such as government debt and health

expenditure, as well as inflation rates. These sources were selected for their reliability, consistency, and global coverage, ensuring the robustness of the dataset used in the analysis (OECD, n.d.; World Bank, n.d.; International Monetary Fund, n.d.).

Estimation Technique

The study employs panel data analysis, which combines cross-sectional and time-series data to account for both country-specific and temporal variations. Panel data analysis offers several advantages, including the ability to control for unobserved heterogeneity through fixed or random effects, improved efficiency through the increased number of observations, and the capacity for dynamic analysis of changes over time.

To estimate long-term coefficients of the cointegrated variables, the Fully Modified Ordinary Least Squares (FMOLS) and Dynamic Ordinary Least Squares (DOLS) methods are employed. These methods address issues of serial correlation and endogeneity, providing robust estimates of the long-term relationships (Pedroni, 2004). Additionally, multiple panel unit root tests are employed including the Levin, Lin, and Chu (LLC) test (Levin et al., 2002), the Im, Pesaran, and Shin (IPS) test (Im et al., 2003), the ADF-Fisher test (Maddala & Wu, 1999), and the PP-Fisher Chi-square test (Choi, 2001).

The decision to transform the variables into their natural logarithmic forms is based on several methodological considerations. The log-linear specification reduces the sharpness of fluctuations in time series data, allows for the interpretation of coefficients as elasticities, and helps address issues of non-linearity and heteroscedasticity common in macroeconomic data (Shahbaz, 2010). With respect to sampling, this study employs a purposive, non-random sample of five countries selected based on their diverse characteristics within the European context. This limits the

generalizability of the findings to other regions or countries with substantially different economic and institutional profiles.

Findings

Panel Unit Root Tests

Testing for stationarity is a critical step in panel data analysis to ensure the validity of subsequent econometric estimations. Non-stationary data can lead to spurious regression results, making it essential to confirm the stationarity of the variables before proceeding with further analysis. In this study, multiple panel unit root tests were employed, including the Levin, Lin, and Chu (LLC) test, the Im, Pesaran, and Shin (IPS) test, the ADF-Fisher test, and the PP-Fisher Chi-square test. These tests share the null hypothesis that the series contains a unit root, while the alternative hypothesis suggests stationarity.

The results of the panel unit root tests at levels (I(0)) are presented in Table 1, and at first differences (I(1)) in Table 2. The findings indicate that most variables are non-stationary at their levels but become stationary after first differencing, suggesting that the variables are integrated of order one — a necessary condition for conducting cointegration analysis.

Table 1. Results of Panel Unit Root Tests (Level, I(0))

Variable	**LLC (Intercept)**	**LLC (Trend)**	**IPS (Intercept)**	**IPS (Trend)**	**ADF-Fisher**	**PP-Fisher**
LGI	-926.751 (0.0000) **	-916.766 (0.0000) **	-604.972 (0.0000) **	-838.256 (0.5776)	37.7691 (0.0000) **	7.41341 (0.6859)
LGDPP	-0.7580 (0.7758)	-1.57771 (0.0573) *	2.10629 (0.9824)	0.70875 (0.7608)	5.43727 (0.8601)	49.1375 (0.0000) **
LUNMR	-2.50465 (0.0061) **	-4.4765 (0.0000) **	-0.54619 (0.2925)	-1.9089 (0.0281) **	11.7054 (0.3053)	7.32041 (0.6949)

Variable	LLC (Intercept)	LLC (Trend)	IPS (Intercept)	IPS (Trend)	ADF-Fisher	PP-Fisher
LIR	-1.47963 (0.0695) *	-2.40726 (0.0080) **	-0.66427 (0.2533)	0.18881 (0.5749)	12.8154 (0.2342)	11.2309 (0.3398)

*Note: * and ** indicate significance at the 5% and 1% levels, respectively. Values in parentheses are p-values.*

Table 2. Results of Panel Unit Root Tests (First Difference, I(1))

Variable	LLC (Intercept)	LLC (Trend)	IPS (Intercept)	IPS (Trend)	ADF-Fisher	PP-Fisher
ΔLGI	-774.968 (0.0000) **	6989.38 (1.0000)	-586.339 (0.0000) **	-532.658 (0.0000) **	49.1355 (0.0000) **	73.2616 (0.0000) **
ΔLGDPP	0.11004 (0.5438)	0.65320 (0.7432)	-0.67507 (0.2498)	-0.28358 (0.3884)	6.54710 (0.7674)	25.1648 (0.0050) **
ΔLUNMR	-2.80631 (0.0025) **	-0.2478 (0.4021)	-1.80869 (0.0352) *	0.27934 (0.6100)	19.4989 (0.0344) **	32.6435 (0.0003) **

*Note: * and ** indicate significance at the 5% and 1% levels, respectively. Values in parentheses are p-values. The results confirm that all variables are stationary at the first difference, satisfying the precondition for cointegration analysis.*

Panel Cointegration Test

After confirming the stationarity of the variables, the next step involves testing for cointegration to determine whether a long-term relationship exists among the variables. The Pedroni Residual Cointegration Test (Pedroni, 2004) was employed for this purpose. This method allows for heterogeneity across panel members and provides two types of tests: the intradimensional approach (Panel tests) and the interdimensional approach (Group tests). The null hypothesis assumes no cointegration, while the alternative

hypothesis suggests the presence of cointegration. The results of the Pedroni cointegration test are presented in Table 3.

Table 3. Results of Pedroni Cointegration Test

Test Statistic	Without Trend	With Trend
Panel v-Statistic	0.9816	0.9950
Panel rho-Statistic	0.9988	0.9990
Panel PP-Statistic	0.0005**	0.0000**
Panel ADF-Statistic	0.1855	0.0000**
Group rho-Statistic	0.9993	0.9999
Group PP-Statistic	0.0000**	0.0000**
Group ADF-Statistic	0.0000**	0.0000**

*Note: The null hypothesis is that the variables are not cointegrated. ** indicates significance at the 1% level.*

The results indicate that most test statistics reject the null hypothesis of no cointegration at the 1% significance level, confirming the existence of a long-term relationship among the variables.

Panel Long-Run Elasticity

To estimate the long-term coefficients of the cointegrated variables, the Fully Modified Ordinary Least Squares (FMOLS) and Dynamic Ordinary Least Squares (DOLS) methods were employed. These methods address issues of serial correlation and endogeneity, providing robust estimates of the long-term relationships (Pedroni, 2004). The results of the panel cointegrated regression are presented in Table 4.

Table 4. Results of Panel Cointegrated Regression (FMOLS)

Variable	Coefficient	p-value
LGDPP	-0.302653	0.7638
LUNMR	0.829687	0.4120
LIR	1.146587	0.2589
LGGD	4.097944**	0.0002
LER	1.652946	0.1068
LPD	3.810124**	0.0005
LHE	-3.942875**	0.0003

*Note: ** indicates significance at the 1% level.*

The FMOLS results reveal several key findings. With respect to General Government Debt (LGGD), a 1% increase in general government debt is associated with a 4.09% increase in income inequality in the long run. This finding aligns with studies such as Woo et al. (2013), which highlight the regressive effects of fiscal consolidation on income distribution. With respect to Population Density (LPD), a 1% increase in population density is associated with a 3.81% rise in income inequality, consistent with the findings of Kuznets (1955) who noted that urbanization and population concentration often exacerbate income disparities. With respect to Health Expenditure (LHE), a 1% increase in health expenditure is associated with a 3.94% reduction in income inequality, supporting the argument that social spending can mitigate inequality by improving access to essential services (Ostry et al., 2014). The remaining variables — GDP per capita, unemployment rate, inflation rate, and employment rate — do not exhibit statistically significant associations with income inequality at conventional significance levels in this model.

Discussion

The findings provide robust evidence of a long-term predictive relationship between income inequality and general government debt. The statistically significant positive association between government debt and inequality suggests that fiscal policies need to be carefully designed to avoid exacerbating income disparities. This aligns with the findings of Furceri et al. (2017), who argue that austerity measures often disproportionately affect low-income households, and with Woo et al. (2013), who document the regressive distributional consequences of fiscal consolidation.

The significant mitigating role of health expenditure in reducing inequality highlights the importance of social spending in

promoting equity. This is consistent with the work of Wilkinson and Pickett (2010), who emphasize the role of public investment in reducing social disparities, and with Ostry et al. (2014), who argue that well-targeted redistributive policies can reduce inequality without adversely affecting growth. The finding that a 1% increase in health expenditure is associated with a 3.94% reduction in inequality underscores the importance of maintaining social spending even during periods of fiscal adjustment.

Population density also emerges as a statistically significant predictor of income inequality, with a 1% increase associated with a 3.81% rise in inequality. This finding is consistent with the urban economics literature noting that population concentration can exacerbate income disparities through labor market segmentation and differential access to public services. The dual economy framework of Lewis (1954) offers theoretical grounding for this result, as urbanization processes tend to widen the gap between high-productivity and low-productivity workers.

It is important to note that while the FMOLS results indicate statistically significant associations, these should be interpreted as predictive relationships rather than causal mechanisms. The explanatory power of the model captures only a portion of the variation in income inequality, and several other potential factors may also play important roles. These include the quality of institutions, the structure of the tax system, the efficiency of public spending, workload and labor market dynamics, organizational support structures, leadership quality in public governance, and the broader work climate in each country. These factors were not directly captured in the current model and represent an important avenue for future research.

Practical Recommendations

Based on the findings of this study, several practical recommendations emerge for policymakers seeking to address the dual challenges of managing public debt and mitigating income disparities. These recommendations are proportionate to the explanatory power of the study's findings and should be considered alongside broader policy considerations.

First, fiscal consolidation efforts should be carefully designed to avoid exacerbating existing income gaps. Policymakers should consider targeted measures to protect vulnerable groups during periods of fiscal adjustment, potentially through maintaining or even increasing social expenditures, particularly in healthcare. The significant negative association between health expenditure and inequality suggests that cuts to health spending as part of austerity measures could have particularly adverse distributional consequences.

Second, the positive association between government debt and inequality suggests that the composition and purpose of government borrowing matter for income distribution outcomes. Targeted borrowing to fund redistributive policies, public health investments, or productivity-enhancing infrastructure may have different distributional consequences than debt accumulated during economic downturns or through inefficient public spending. Policymakers should consider the distributional implications of debt financing alongside traditional fiscal sustainability concerns.

Third, given the role of population density and urbanization in amplifying income disparities, regional development policies that promote balanced growth across urban and rural areas may help mitigate inequality. Investments in regional infrastructure, education, and service delivery in less densely populated areas could help reduce urban-rural income gaps. However, these recommendations should not be applied in isolation from broader institutional, cultural, and economic context-specific considerations.

Limitations and Future Research

This study is subject to several limitations that should be considered when interpreting its findings. First, while the panel data approach allows for cross-country comparisons, it may not fully capture the nuanced institutional and cultural factors that influence income distribution in each country. Country-specific institutional arrangements, tax structures, and social protection systems could all mediate the relationship between government debt and inequality in ways that aggregate panel data analysis cannot fully account for.

Second, the time frame of the study (2010-2021) encompasses a period of significant economic volatility, including the aftermath of the global financial crisis and the onset of the COVID-19 pandemic. These extraordinary events may have influenced the relationship between government debt and inequality in ways that might not be representative of more stable economic periods.

Third, while the model controls for several macroeconomic variables, there may be other unobserved factors influencing the relationship between government debt and income inequality. Institutional quality, tax system structure, and public spending efficiency are all factors that could mediate this relationship but were not directly captured in the model. The cross-sectional nature of the panel data and reliance on self-reported macroeconomic indicators also limit the study's ability to establish causal mechanisms. The study's use of a non-random, purposive sample of five countries further limits generalizability to other contexts.

Future research could address these limitations in several ways. Case studies focusing on individual countries could provide more detailed insights into the mechanisms through which government debt affects income inequality, taking into account country-specific institutional arrangements and policy responses.

Extending the time frame of the analysis to include periods of both economic stability and crisis could offer a more comprehensive understanding of how this relationship evolves under different economic conditions. Longitudinal designs with comprehensive causal models and the incorporation of qualitative methods could also enrich understanding of the mechanisms at work. Additionally, future studies could explore potential non-linear effects of government debt on income inequality, as the relationship may change at different levels of debt. Incorporating measures of wealth inequality alongside income inequality could provide a more holistic view of the distributional impacts of fiscal policies. Finally, extending the analysis to different regions and country contexts — including middle-income economies and countries outside Europe — would help establish the generalizability of these findings.

Conclusion

This study has provided a comprehensive analysis of the relationship between income inequality and general government debt in five diverse European economies: Turkey, the Netherlands, Switzerland, Poland, and Spain. Employing advanced panel data techniques and incorporating a range of macroeconomic control variables, the study sheds light on the complex interplay between fiscal policy and income distribution in the post-financial crisis era.

The findings reveal a statistically significant positive predictive relationship between general government debt and income inequality, with a 1% increase in government debt associated with a 4.09% rise in inequality in the long run. The study also highlights the mitigating role of health expenditure, with a 1% increase in health spending associated with a 3.94% reduction in inequality. These results underscore the importance of considering distributional effects when formulating fiscal policies, particularly during periods of economic volatility or fiscal consolidation.

These findings contribute to the growing literature on the economic consequences of income inequality by elucidating the role of government debt in shaping income distribution. As nations continue to grapple with the economic fallout of global crises and seek to build more resilient and equitable economies, understanding the complex relationship between fiscal policy and income inequality remains a critical area for both research and policy. The findings call for a careful and equity-conscious approach to fiscal management that considers the distributional consequences of public debt and social expenditure decisions.

References

Acemoglu, D., & Robinson, J. A. (2012). Why nations fail: The origins of power, prosperity, and poverty. Crown Business.

Alesina, A., Di Tella, R., & MacCulloch, R. (2004). Inequality and economic growth: The role of institutions. Journal of Economic Growth, 9(2), 145-178.

Brueckner, J. K., & Lederman, D. (2018). Inequality and economic growth: The role of initial income and prices. Journal of Economic Growth, 23(1), 1-30.

Burkhauser, R. V., Larrimore, J., & Simon, K. (2013). Inequality and government policy. Journal of Economic Literature, 51(1), 1-34.

Choi, I. (2001). Unit root tests for panel data. Journal of International Money and Finance, 20(2), 249-272.

Duflo, E., & Banerjee, A. (2011). Poor economics: A radical rethinking of the way to fight global poverty. PublicAffairs.

Furceri, D., Loungani, P., & Ostry, J. D. (2017). The aggregate and distributional effects of fiscal consolidation. World Development, 98, 352-368.

Im, K. S., Pesaran, M. H., & Shin, Y. (2003). Testing for unit roots in heterogeneous panels. Journal of Econometrics, 115(1), 53-74.

International Monetary Fund. (n.d.). IMF data. Retrieved from https://www.imf.org/en/Data

Kuznets, S. (1955). Economic growth and income inequality. The American Economic Review, 45(1), 1-28.

Levin, A., Lin, C. F., & Chu, C. S. J. (2002). Unit root tests in panel data: Asymptotic and finite-sample properties. Journal of Econometrics, 108(1), 1-24.

Lewis, W. A. (1954). Economic development with unlimited supplies of labour. The Manchester School, 22(2), 139-191.

Maddala, G. S., & Wu, S. (1999). A comparative study of unit root tests with panel data and a new simple test. Oxford Bulletin of Economics and Statistics, 61(S1), 631-652.

OECD. (n.d.). OECD data. Retrieved from https://data.oecd.org/

Ostry, J. D., Berg, A., & Tsangarides, C. G. (2014). Redistribution, inequality, and growth. IMF Staff Discussion Note.

Pedroni, P. (2004). Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. Econometric Theory, 20(3), 597-625.

Piketty, T. (2014). Capital in the twenty-first century. Harvard University Press.

Quispe-Torreblanca, R., Brown, M., Boyce, W. J., Wood, A., & De Neve, J. (2021). A review of the empirical literature on the economic consequences of income inequality. Journal of Economic Surveys, 35(4), 1001-1030.

Shahbaz, M. (2010). Income inequality-economic growth nexus: The role of financial development. American Economic Review, 100(2), 1-15.

Thaler, R. H., & Sunstein, C. R. (2008). Nudge: Improving decisions about health, wealth, and happiness. Yale University Press.

Wilkinson, R., & Pickett, K. (2010). The spirit level: Why greater equality makes societies stronger. Bloomsbury Press.

Woo, J., Bova, E., Kinda, T., & Zhang, Y. S. (2013). Distributional consequences of fiscal consolidation. IMF Economic Review, 61(2), 271-309.

World Bank. (n.d.). World Bank data. Retrieved from https://data.worldbank.org/

CHAPTER 5

THE IMPACT OF DEFENSE EXPENDITURE, CLIMATE CHANGE AND ECONOMIC CRISES

MUSTAFA TORUSDAĞ [1]

Introduction

The global economy may be seriously threatened by climate change. The economic impacts of climate change are already being felt around the world but the extent of future potential fragility depends upon the size and structure of economies, the durability of institutions and physical infrastructure and the capacity to mitigate and adapt to climate change. It also depends on the particular shocks each country is exposed to. The Global Financial Crisis (GFC); Covid-19 pandemic and the recent war in Ukraine (with some countries reverting to use of non-renewable energy sources based on new renewables) are restarting the debate on the compatibility between economic development and environmental protection (Jalles, 2024). One of the most important environmental and socio-economic challenges facing humanity is mitigation of possible effects of climate change. There is a wide scientific consensus that temperatures should not exceed 2ºC by the year 2010; if they do,

[1] Van Yüzüncü Yıl University, Faculty of Economics and Administrative Sciences, Department of Economics, Türkiye, (0000-0002-8839-0562), mustafatorusdag@yyu.edu.tr.

very dangerous effects could occur on ecosystems and human societies. This also implies that the concentrations of greenhouse gasses (GHGs) in the atmosphere have to be kept below 400 parts per million (ppm), while the concentration was 375 ppm in 2005 (IPCC 2007a). In fact, achieving these levels of concentration will require prompt and radical reductions in emissions, together with unprecedented systemic changes in existing behaviors, social organization and technologies. In this context, it is urgent to develop, implement and disseminate low carbon technologies in the short, medium and long term, especially in the energy and transport sectors (Del Río and Labandeira, 2009).

Climate change is a systemic risk: 'new, large, global, long-term, permanent and uncertain' (Stern, 2007: 25). Unlike the financial and economic crisis, climate change is a medium- and long-term threat that will grow cumulatively every year if not addressed. This can be termed a 'transformative crisis' (Standing, 2009). Nevertheless, they have at least two things in common: Both represent the global crisis and both represent the colossal failure of market mechanisms. "Climate change poses a unique challenge for the economy: it is the biggest and widest-ranging market failure ever seen" (Stern, 2007: 1; Gough, 2010).

The aim of this study is to examine how rising global defense spending, coupled with economic crises exacerbated by the growing climate crisis, is disrupting national public budgets and, indirectly, national development goals. It also hinders national environmental sustainability objectives. Therefore, this complex interplay of three necessitates a holistic analysis. The significance of this study lies in its ability to integrate the interplay of these relationships—security needs, environmental risks, and the need for economic stability—within a common framework.

2. Theoretical Framework

Financial crises are confusing the long- and short-term environmental effects. For some, short-term reductions in emissions notwithstanding, economic crises are seldom good for the environment in the years of crisis. The main argument is that periods of stagnation—including those involving investments in low-carbon technologies—negatively affect emission reduction efforts because they make access to capital difficult and act as a deterrent to investment (Del Río ve Labandeira 2009). Governments and the private sector are moving away from climate policies, with the focus on recovery and tweaking their own budgets. 'Yeşil' bir şirkete ve ekonomiye geçiş yerine krizden çıkmak, krizler nedeniyle, o dönemdeki hedef olmuştur; çevresel projelerin ertelenmesine neden olur. And, the larger carbon lock-in is during economic crisis. That's more likely. Also, cheap energy prices amid a crisis reduce the economic viability of cleaner technologies (Wooders and Runnalls 2008; Egenhofer, and Alessi, 2013). Governments generally avoid imposing additional costs and regulations on businesses when the economy is fragile and risks exist. Economic crises can provide a good opportunity to develop and invest in low carbon technologies that can provide a way out of stagnation (Greenpeace 2008). In this view, the long life spans of most energy infrastructure and technologies mean that crises should not be missed as opportunities to replace carbon-intensive technologies with cleaner alternatives. As Papandreou (2015) suggests, crises may create the possibility for new institutional trajectories, by inducing changes in existing norms and institutions through the forces they generate or by rebalancing conflicting political and economic interests. "Crises" questions institutions, forms of governance and the theories that legitimize them from a new critical perspective. Crises should reinforce the need for designing adequate climate policies that lead to cost effective emission reductions in an intertemporal perspective given the increasing competition over limited resources and short term

priorities for the use of these resources. Crises can be used to ramp up our efforts to achieve low-carbon economic growth. Crises provide the opportunity and the impetus to improve the efficiency of energy and material use, to move toward more sustainable production and to develop new green businesses and industries. Crises can also provide an incentive for much needed structural reforms, if there are opportunities for both economic and environmental benefits. Crises are an opportunity to reform or eliminate policies that may be expensive, inefficient, and environmentally harmful. The policy reforms will also boost incentives for innovation by removing market distortions (Jalles, 2024).

The primary aim of defense services is to safeguard the country's sovereignty. The size of defense expenditures is determined by many factors, such as the geopolitical position of countries, changes in population, wars, civil wars, external threats (problems related to foreign policy), changes in defense expenditures of neighboring or rival countries (arms race), and internal threats (Collier and Hoeffler, 2002: 4; Krugman, 2005: 522). The level of defense expenditures may change to the extent that these dynamics and threats threaten national security. The primary purpose of defense spending is to provide and maintain national security, but it is also a fact that the economic consequences of such spending are significant.

The relationship between defense spending and economic crises is not only explained by the size of public expenditures, but also through their indirect effects on macroeconomic balances. In particular, increases in defense spending are shown to increase budget deficits, expand public debt stocks, strengthen inflationary pressures and put pressure on the balance of payments, which are among the main mechanisms that increase economic fragility (Dunne & Tian, 2015). The Keynesian approach argues that the rise in public expenditure may help total demand and production in the short run but in the economies with limited production capacity or if

the expenditure is financed through borrowing or monetary expansion, demand-oriented inflation may take place (Keynes, 1936; Barro, 1990). In particular, the impact of defense expenditures on inflation may be more evident in developing countries. High inflation can lead to an increase in interest rates, a decrease in investment spending, and a disruption of macroeconomic stability, thus increasing the likelihood of economic crisis (Friedman, 1968; Sargent & Wallace, 1981). In contrast, an increase in military spending in countries that are highly dependent on imported intermediate goods and high technology in the defense industry can widen the current account deficit and put pressure on the balance of payments through increasing the imports of arms and defense equipment (Dunne, Nikolaidou, & Smith, 2002). The widening of the current account deficit increases the demand for foreign financing, but it also increases the volatility of the exchange rate; this situation can trigger crisis dynamics through capital outflows, reserve losses and financial instability (Krugman, 1979; Obstfeld, 1996). High level of borrowing to finance defense spending also creates uncertainty in the market about the sustainability of public debt. Deterioration of financial sustainability increases the likelihood of economic crisis by leading to consequences such as increasing risk premiums, increasing public financing costs and weakening investor confidence (Reinhart & Rogoff, 2010; Aizenman & Marion, 2011). In this regard, the link between defense expenditure and economic crises is established through macroeconomic transmission channels such as economic balance, public debt, inflation, current account balance and exchange rate, rather than direct causality.

Defense expenditures are linked to inflation, and their effects are debated. The idea that military spending causes inflation is based on the economic nature of military spending. This is founded on the premise that defense spending is unproductive as is other spending, and that defense expenditures have a greater inflationary effect than

other government expenditures (Starr et al., 1984: 107; Deger and Sen, 1995: 279). On the other hand, it is also discussed whether defense expenditures create a new purchasing power. The problem with defense spending is that demand can be met without surplus. The increase in defense expenditures may not affect the increase in demand in other public expenditures, although it supports the supply (Fallows, 1981: 7).

Defense spending creates more demand rather than supply, which causes inflation. It can be stated that defense expenditures increase demand without increasing supply and therefore do not contribute to current or future living standards (Looney, 1989: 150). Namely, if increased defense spending encounters limited capacity or inelastic supply in the narrow-based defense industry, this may increase inflation. Because the increase in defense expenditures will increase the demand for labor and capital for companies. Since the source of qualified labor and capital will not change in the short term, increased defense expenditures will cause an increase in wages and prices. In the long run, the supply of certain factors (e.g. educated engineers) may increase prices in the long run as they cannot respond to the sharply increasing demand (Starr et al., 1984: 108). Therefore, to the extent that the defense industry and its associated skilled labor market are characterized by monopolistic and/or monopsony pricing practices, increases in defense spending may bring about price increases that spread from these markets to the rest of the economy. Thus, defense expenditures may be the cause of a disproportionate cost-driven inflation (Nourzad, 1987:58). For example, an increase in aircraft orders related to the defense industry will lead to an increase in employment in the aviation industry, which may affect the wage level in that sector. Over time, these wage increases may spread to other sectors, which will lead manufacturing companies to increase prices (due to the increase in production costs). Price increases will tend to increase pressure for further wage increases (Stekler, 1979: 21). On the demand side, increased defense

expenditures contribute to the increase in nominal demand growth, and if this increase is not balanced by tax increases or restrictive monetary activities, it may negatively affect inflation by contributing to the acceleration of nominal demand growth. However, if there is an output gap, that is, if the economy is not at full employment equilibrium, the increase in defense expenditures (as in other public expenditures) may not necessarily produce an inflationary effect (Schultze, 1981: 3; Nourzad, 1987: 58).

The impact of increases in defense spending on the balance of payments is another way that inflation can be affected. For example, imported military goods and services may lead to an increase in balance of payments deficits. The impact of a balance of payments deficit on inflation depends on the size of the deficit and how it is financed. A large deficit causes the national currency to depreciate, which can make imports more expensive and exports more competitive. The increasing cost of imports and the increase in demand for export goods may put pressure on prices and cause a slight increase in the inflation rate (Starr et al., 1984: 109). If the increase in defense expenditures cannot be covered by taxes, borrowing will be required. Budget deficits will increase due to both the increase in defense expenditures and the increase in interest expenses, the cost of borrowing. Because policymakers are reluctant to raise taxes or reduce other spending to finance additional spending, their resort to budget deficits and public debt tends to create greater inflationary pressure (Looney, 1989: 150).

Increasing economic integration through trade and financial flows between countries increases the likelihood of a global business cycle. Two standard methods are applied to determine the peaks and troughs of both national and global business cycles: statistical methods and judgmental methods. Both methods, when applied to the global business cycle, show 4 troughs in the last 50 years, occurring in 1975, 1982, 1991 and 2008. These troughs are considered periods of global economic recession. However, the

Great Depression, which resulted from the intensifying recession of the early 1930s, is considered the most severe economic downturn ever seen in advanced economies such as the United States, the United Kingdom, Germany, and France. On the other hand, the sharp economic declines experienced in developing economies in 1997-1998, from which they are still trying to recover, and the subsequent declines experienced in developed economies in 2001, are not considered global recessions, as they showed growth in developed economies and strong performance in emerging markets. Although there is no official definition of economic recession, the term generally refers to troughs or periods of decline in economic activities (Obani and Gupta, 2016).

Most modern economies are heavily dependent on fossil fuels, which account for 81% of the primary energy mix (International Energy Agency - IEA, 2011). The burning of fossil fuels is the single largest source of carbon dioxide (CO2) emissions and accounts for approximately 60% of all greenhouse gas emissions (Parker, Blodgett, and Yacobucci, 2011; World Resources Institute, 2009). In 2004, CO2 accounted for 77% of all anthropogenic greenhouse gas emissions (Metz, et al., 2007). If economic activities are not decoupled from CO2 emissions, economic growth and energy intensity remain the largest sources of greenhouse gas emissions worldwide (Andreoni & Galmarini, 2012; Cutlip & Fath, 2012). This makes the possible limited availability of fossil fuels an important factor in mitigating human-induced climate change (Höök & Tang, 2013). Likewise, the decline in economic activity that characterizes a recession has the potential to reduce greenhouse gas emissions. This trend is clearly seen in the total greenhouse gas emissions figures between 1990 and 2010; these figures show a sharp decline in emissions immediately following the 2008 recession. This trend is also reflected in the global CO2 emission figures resulting from fossil fuels and cement production between 2008 (8,748,000 Mtons) and 2009 (8,626,351 Mtons) (Obani and Gupta, 2016).

According to the first category, economic downturns lower carbon emissions by reducing aggregate demand, industrial production, and transportation. According to Peters et al. (2012), a drop in the use of fossil fuels during the 2008–2009 financial crisis led to a large fall in worldwide CO2 emissions. In a similar vein, Le Quéré et al. (2020) calculated that during the height of COVID-19 lockdowns, daily CO2 emissions globally dropped by about 17%. Reduced demand for transportation is the reason for this decline. Environmental gains are often transient, according to a second group in the literature. Global emissions rose by around 6% as industrial production recovered in the immediate aftermath of the financial crisis, according to a research by Peter et al. (2012).

Literature Review

Researchers are paying increasing attention to the link between economic crises and carbon dioxide emissions, particularly in the wake of the COVID-19 pandemic and the 2008 Global Financial Crisis. Current research in the literature falls into four categories: the influence of energy consumption, environmental, and financial policies; rebound effects during the recovery phase; and emission reductions resulting from the crisis. New studies covering the 2024–2026 period have emerged in the literature—addressing both economic and climate dimensions—driven by NATO's decisions to increase defense spending, the Russia-Ukraine War, and geopolitical risks. There is virtually no research on the relationship extending from financial crises to climate change. The closest area of research concerns the impact of recessions on pollutant emissions. Giedraitis et al. (2010) and Stavytskyy et al. (2016) support the claim that CO2 emission patterns during the economic downturns of the 1870s and 1930s were similar. Siddiqi (2000), examining the Asian Financial Crisis, noted that it yielded some positive outcomes for the global environment. Research addressing the relationship between defense spending and public debt/budget balance has identified a dynamic

relationship between defense spending and public debt. Dudzevičiūtė et al. (2021) analyzed NATO-EU countries using panel data analysis. In an article by Calduch Cervera (2012), the relationship between economic crisis and defense spending in European and NATO countries was analyzed using a comparative method; significant changes in defense budgets and the implementation of fiscal constraints were observed following the 2008 crisis. Bardakas et al. (2023), in a study using non-linear time series analysis to examine the relationship between defense spending and public debt in Greece, indicated that defense spending [affects] fiscal vulnerability through public debt... According to Caruso and Di Domizio (2015) concluded that US military spending increases European public debt. Cabrillac et al. (2023) concluded that climate change negatively affects debt sustainability in developing countries, particularly in vulnerable ones. Weidinger and Ranger (2025) argues that climate adaptation investments can strengthen debt sustainability and reduce financing costs in the long run. Barrage (2024) posits that climate change adversely affects debt dynamics by disrupting public finances. Agarwala et al. (2021) argue that climate risks pose a systemic risk to fiscal sustainability. According to Dudzevičiūtė et al. (2021) a bidirectional causal relationship exists between defense spending and public debt.

In their study using panel data analysis, Dong et al. (2025) state that military spending on a global scale increases greenhouse gas emissions and complicates the achievement of climate targets, that is, concluded that expenditures jeopardize 1.5°C and 2°C climate targets by increasing CO2 emission intensity. Akusta (2025), using the ARDL method, argues that renewable energy mitigates the environmental impact of defense spending. Cabrillac et al. (2023) found a bidirectional causal relationship between public debt and climate. Balázs Markó (2024) states that military expenditures increase CO_2 emissions, reduce green patents, and crowd out renewable energy investments.

Conclusion

The last century, which began with the Industrial Revolution, also witnessed major advancements in the field of technology. The 20th century—a period that encompassed two world wars—saw numerous conflicts and wars of varying intensity and scale, many of which continue to this day. These global wars and conflicts, along with military and political crises, the fight against terrorism, arms races between nations, and technological developments and modernization efforts in weaponry systems, have driven an increase in national defense expenditures. The third fundamental pillar of the ecological crisis consists of oil crises. As a primary energy source, oil acts both as a trigger for crises and as a critical production input affected by them. Holding strategic, political, and economic significance for nations, oil is viewed as a key driver of multifaceted ecological crises. Furthermore, the consumption of oil contributes significantly to climate change and global warming through the emission of carbon dioxide.

The global economy has recently witnessed a global financial crisis, followed by a period of stagnation—manifesting as a debt crisis in Europe—alongside ecological crises brought about by climate change. The simultaneous emergence of these issues has necessitated a multidimensional solution encompassing economic, social, and environmental aspects. Consequently, the "green economy" has been proposed as a prudent remedy for both economic crisis and ecological degradation.

As previously noted in the discussion on global defense spending, while there was a downward trend in such expenditures during the post-Cold War era, they are currently on the rise. The primary reason for this shift is undoubtedly the growing global demand for security. Escalating acts of terrorism, the chaos and uncertainty prevailing in countries torn by civil war, refugee crises, the arms race driven by nations seeking to gain an advantage over one another, and various

political and economic crises have all contributed to this heightened need for security. Military activities are highly energy-intensive, and since most of this energy is derived from fossil fuels, they result in significant greenhouse gas (GHG) emissions. However, the contribution to climate change is not limited to direct emissions from fuel-consuming combat aircraft, warships, and armored vehicles; it also encompasses a vast array of indirect carbon emissions stemming from the movement of equipment, complex and massive supply chains—which include the energy-intensive production of materials such as steel and aluminum—and the impacts of war itself.

Military spending encompasses investments made for defense and security purposes. Another aspect that affects environmental degradation—yet receives less attention in the literature—is defense and military expenditure. The impact of military spending on environmental degradation occurs through various channels. The first channel arises from the desire of nations to pursue further industrialization to achieve their development goals. The extraordinary pace of industrialization over the last few decades has significantly increased the pressure on natural resources—a subject of ongoing debate among environmental scientists and policymakers. The demands brought about by rapid industrialization have led to the overexploitation and depletion of natural resources. This, in turn, has increased the tendency of nations to invest in military operations aimed at securing control over these natural resources. Such military operations, however, pose potential risks to the environment, as the overexploitation of natural resources can disrupt the natural balance and cause damage to ecosystems. The second channel arises from considering the effects of military spending on environmental degradation through the Wheel of Destruction Theory. The Wheel of Destruction Theory suggests that the depletion of natural resources and the increase in harmful effects on the environment are caused by economic growth and industrialization. In this context, an increase in military spending

could also lead to overconsumption of natural resources and increased environmental degradation.

The rise in consumption is considered the most significant factor in the emergence of the global ecological crisis. The depletion of the world's stock of natural wealth - often in an irreversible manner - has detrimental effects on the well-being of future generations and creates challenges for the future. Furthermore, excessive consumption—driven by economic growth that accelerates the accumulation of physical, financial, and human capital—leads to capital imbalances, which in turn negatively impact the ecosystem.

During the 2008 global financial crisis, a number of other crises—including food, oil, and water crises—occurred simultaneously; these are collectively referred to as 'multiple crises'. As rapid industrialization led to an increase in fossil fuel consumption and a sharp rise in carbon dioxide emissions, issues such as global warming and climate change emerged, making environmental crises increasingly apparent.

Theoretical evaluation and policy implications should be associated with the reviewed literature, and defense expenditures and climate expenditures to combat climate change should be planned in a coordinated manner. Budget resilience against possible simultaneous climate crisis and defense expenditure shocks should be calculated, and the sustainability and environmental costs of expenditures should also be taken into account.

Source

Aizenman, J., & Marion, N. (2011). Using inflation to erode the US public debt. Journal of Macroeconomics, 33 (4), 524–541.

Agarwala M, Burke M, Klusak P, Mohaddes K, Volz U, Zenghelis D. (2021). Climate Change And Fiscal Sustaınability: Rısks And Opportunities. National Institute Economic Review. 258, 28-46. Doi:10.1017/Nie.2021.37

Akusta, E. (2025). Can renewable energy sources alleviate the pressure of military expenditures on the environment? Empirical evidence from Turkiye. Environmental Research and Technology, 8 (2), 410–424. https://doi.org/10.35208/ert.1502424

Andreoni, V., & Galmarini, S. (2012). Decoupling economic growth from carbon dioxide emissions: A decomposition analysis of Italian energy consumption. Energy, 44 (1), 682–691. https://doi.org/10.1016/j.energy.2012.05.024

Bardakas, I., Doulos, D., & Zombanakis, G. A. (2023). Defence expenditure and public debt in Greece: A non-linear relationship. Security and Defence Quarterly, 41 (1), 1–14. https://doi.org/10.35467/sdq/155960

Barro, R. J. (1990). Government Spending in a Simple Model of Endogenous Growth. Journal of Political Economy, 98 (5), S103–S125.

Cabrillac, B., Fabre, C., & Jacolin, L. (2023). Debt sustainability and climate change (Policy Bradford, J. H., & Stoner, A. M. (2017). The treadmill of destruction in comparative perspective: a panel study of military spending and carbon emissions, 1960-2014. Journal of World-Systems Research, 23 (2), 298-325.

Brief No. B249). Fondation pour les Études et Recherches sur le Développement International (FERDI). https://ferdi.fr/publications/soutenabilite-de-la-dette-et-changement-climatique

Calduch Cervera, R. (2012). Impact of the economic crisis on defence policies: A comparative study. Journal of the Spanish Institute for Strategic Studies, 0. https://revista.ieee.es/en/article/view/403

Caruso, R., & Di Domizio, M. (2015). The impact of US military spending on public debt in Europe (1992–2013): A note. Peace Economics, Peace Science and Public Policy, 21 (4), 459–466. https://doi.org/10.1515/peps-2015-0040

Collier, P., & Hoeffler, A. (2002). Military expenditure: Threats, aid, and arms races (Policy Research Working Paper No. 2927). World Bank. https://doi.org/10.1596/1813-9450-2927

Cutlip, L., & Fath, B. D. (2012). Relationship between carbon emissions and economic development: Case study of six countries. Environment, Development and Sustainability, 14 (3), 433–453. https://doi.org/10.1007/s10668-011-9328-2

Çolak, O., Özuyar, S. E. G., & Bölükbaşı, Ö. F. (2022). Asymmetric effects of the defense burden on environmental degradation: Evidence from NATO countries. Sustainability, 15 (1), 573.

Deger, S., & Sen, S. (1995). Military expenditure and developing countries. In K. Hartley & T. Sandler (Eds.), Handbook of defense economics, Vol. 1, pp. 275–307.

Del Río, P., & Labandeira, X. (2009). Climate change at times of economic crisis. Estudios Económicos 05-09). FEDEA. https://documentos.fedea.net/pubs/ee/2009/05-09.pdf

Dong, W., Ran, Q., Liu, F. *et al.* (2025). Rising military spending jeopardizes climate targets. *Nat Commun* 16, 4766. https://doi.org/10.1038/s41467-025-59877-x

Dudzevičiūtė, G., Česnuitytė, V., & Prakapienė, D. (2021). Defence expenditure–government debt nexus in the context of sustainability in selected small European Union countries. Sustainability, 13 (12), Article 6669. https://doi.org/10.3390/su13126669

Dunne, J. P., & Tian, N. (2015). Military Expenditure, Economic Growth and Heterogeneity. Defence and Peace Economics, 26 (1), 15–31. https://doi.org/10.1080/10242694.2013.848575

Dunne, J. P., Nikolaidou, E., & Smith, R. P. (2002). Military Spending, Investment and Economic Growth in Small Industrialising Economies. South African Journal of Economics, 70 (5), 789–808.

Egenhofer, C., & Alessi, M. (2013). EU policy on climate change mitigation since Copenhagen and the economic crisis. CEPS Working Document, 380.

Fallows, J. (1981, August). Defense, taxes and the budget. The Atlantic, 7–8.

Friedman, M. (1968). The Role of Monetary Policy. American Economic Review, 58 (1), 1–17.

Giedraitis, V. R., Girdenas, S., & Rovas, A. (2010). Feeling the heat: Financial crises and their impact on global climate change. Perspectives of Innovations, Economics and Business, 4 (1), 7–10. https://doi.org/10.15208/pieb.2010.1

Gough, I. (2010). Economic crisis, climate change and the future of welfare states. Twenty-First Century Society, 5 (1), 51–64.

Greenpeace International. (2008). Energy [r]evolution: A sustainable EU 27 energy outlook. https://archivo-es.greenpeace.org/espana/Global/espana/report/cambio_climatico/eu-energy-r-evolution-report.pdf

Höök, M., & Tang, X. (2013). Depletion of fossil fuels and anthropogenic climate change—A review. Energy Policy, 52, 797–809.

Intergovernmental Panel on Climate Change. (2007a). Climate change 2007: Synthesis report. Summary for policymakers. IPCC.

International Energy Agency. (2011). World energy outlook 2011. https://www.iea.org/reports/world-energy-outlook-2011

Jalles, J. T. (2024). Financial crises and climate change. Comparative Economic Studies, 66 (1), 166–190. https://doi.org/10.1057/s41294-023-00209-7

Keynes, J. M. (1936). The General Theory of Employment, Interest and Money. London: Macmillan.

Krugman, P. (1979). A Model of Balance-of-Payments Crises. Journal of Money, Credit and Banking, 11 (3), 311–325.

Krugman, P. (2005). Is fiscal policy poised for a comeback? Oxford Review of Economic Policy, 21 (4), 515–523. https://doi.org/10.1093/oxrep/gri029

Le Quéré, C., Jackson, R. B., Jones, M. W., Smith, A. J. P., Abernethy, S., Andrew, R. M., De-Gol, A. J., Willis, D. R., Shan, Y., Canadell, J. G., Friedlingstein, P., Creutzig, F., & Peters, G. P. (2020). Temporary reduction in daily global CO2 emissions during the COVID-19 forced confinement. Nature Climate Change, 10 (7), 647–653. https://doi.org/10.1038/s41558-020-0797-x

Looney, R. E. (1989). Impact of arms production on income distribution and growth in the Third World. Economic Development and Cultural Change, 38 (1), 145–153.

Markó, B. (2024). The green peace dividend: The effects of militarization on emissions and the green transition (Version 1) [Preprint]. arXiv. https://arxiv.org/abs/2408.16419v1

Metz, B., Davidson, O. R., Bosch, P. R., Dave, R., & Meyer, L. A. (Eds.). (2007). Climate change 2007: Mitigation of climate change. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press. https://archive.ipcc.ch/publications_and_data/ar4/wg3/en/contents.html

Nourzad, F. (1987). A reexamination of the effect of rapid military spending on inflation. Quarterly Journal of Business and Economics, 26 (3), 57–66.

Obani, P. C., & Gupta, J. (2016). The impact of economic recession on climate change: eight trends. Climate and Development, 8 (3), 211–223., https://doi.org/10.1080/17565529.2015.1034226

Obstfeld, M. (1996). Models of Currency Crises with Self-Fulfilling Features. European Economic Review, 40 (3–5), 1037–1047.

Papandreou, A. (2015). The Great Recession and the transition to a low-carbon economy (Working Paper No. 88). Financialisation, Economy, Society & Sustainable Development (FESSUD) Project.

Parker, L., Blodgett, J., & Yacobucci, B. D. (2011). U.S. global climate change policy: Evolving views on cost, competitiveness, and comprehensiveness (CRS Report No. RL30024). Congressional Research Service. https://digital.library.unt.edu/ark:/67531/metadc103150/

Peters, G. P., Marland, G., Le Quéré, C., Boden, T., Canadell, J. G., & Raupach, M. R. (2012). Rapid growth in CO2 emissions after the 2008–2009 global financial crisis. Nature Climate Change, 2 (1), 2–4. https://doi.org/10.1038/nclimate1332

Reinhart, C. M., & Rogoff, K. S. (2010). Growth in a Time of Debt. American Economic Review: Papers & Proceedings, 100 (2), 573–578.

Sargent, T. J., & Wallace, N. (1981). Some Unpleasant Monetarist Arithmetic. Federal Reserve Bank of Minneapolis Quarterly Review, 5 (3), 1–17.

Schultze, C. L. (1981). Economic effects of the defense budget. The Brookings Bulletin, 18 (2), 1–5.

Shear, B. W. (2014). Making the green economy: politics, desire, and economic possibility. Journal of Political Ecology, 21 (1).

Siddiqi, T. A. (2000). The Asian financial crisis - Is it good for the global environment? Global Environmental Change, 10 (1), 1–7. https://doi.org/10.1016/S0959-3780(00)00003-0

Smith, R. P. (1980). Military Expenditure and Investment in OECD Countries, 1954–1973. Journal of Comparative Economics, 4 (1), 19–32.

Standing, G. (2009). Work after globalization: Building occupational citizenship. Edward Elgar Publishing.

Starr, H., Hoole, F. W., Hart, J. A., & Freeman, J. R. (1984). The relationship between defense spending and inflation. Journal of Conflict Resolution, 28 (1), 103–122.

Stavytskyy, A., Giedraitis, V., Sakalauskas, D., & Huettinger, M. (2016). Economic crises and emission of pollutants: A historical review of select economies amid two economic recessions. Ekonomika, 95 (1), 7–21.

Stekler, H. O. (1979). The relationship between inflation and defense expenditures (IDA Paper P-1457). Institute for Defense Analyses.

Stern, N. (2007). The economics of climate change: The Stern review. Cambridge University Press.

Tienhaara, Kyla (2010). A Tale of Two Crises: What the Global Financial Crisis Means for the Global Environment Crisis. Environmental Policy and Governance, 20, 197-208.

UNEP (2011), Towards a Green Economy: Pathways to Sustainable Development and Poverty Eradiction (France).

UNEP (2009a). Global Green New Deal: An Update for the G20 Pittsburgh Summit, UNEP Green Economy Initiative. http://wedocs.unep.org/handle/20.500.11822/7736

Weidinger, M., & Ranger, N. (2025). The fiscal case for adaptation and improved sustainability analysis. In Compendium of practice from a global community of ministries of finance and leading organizations. Coalition of Finance Ministers for Climate Action. https://greenandresilienteconomics.org/compendium-of-practice/contributions/the-fiscal-case-for-adaptation-and-improved-sustainability-analysis/

Wooders, P., & Runnalls, D. (2008). The financial crisis and our response to climate change. International Institute for Sustainable Development. https://www.iisd.org/articles/policy-analysis/financial-crisis-and-our-response-climate-change

World Resources Institute. (2009). World greenhouse gas emissions in 2005 [Chart]. https://www.wri.org/data/world-greenhouse-gas-emissions-2005

CHAPTER 6

CURRENCY COMPOSITION AND MEASUREMENT SENSITIVITY IN TURKEY'S CREDIT GAP

BİLGE KAĞAN ÖZDEMİR[1]

Introduction

Over the past three decades Turkey has gone through five episodes of financial stress, namely the 1994 currency crisis, the twin banking and currency crisis of 2000-2001, the spillover of the global crisis in 2008-2009, the currency and corporate debt crisis of 2018, and the policy-induced stress of 2021-2023. Each had a different origin, and credit played a different part in each. In 1994 the decisive factor was short-term public debt rather than private credit. By 2001 exchange rate risk had built up in the banks' own open positions. The rapid credit growth before 2008 was a deepening that started from a very low base. In 2018 it was the foreign currency debt of corporates that sat at the centre of the crisis.

This variety is a problem for the credit-to-GDP gap, which Basel III takes as the reference for the countercyclical capital buffer (CCyB). The gap measures how far total credit deviates from its

[1]Prof. Dr., Anadolu University, Faculty of Economics and Administrative Sciences, Department of Economics, ORCID: 0000-0002-8716-9305

long-run trend. It says nothing about the currency in which credit is extended, who receives it or what it is used for. In an economy where credit has played such different roles in crisis dynamics, a single aggregate indicator is likely either to miss vulnerabilities or to send false alarms too often. Whether the gap, or indicators that take account of the currency composition of credit and of external fragility, can anticipate these episodes in real time is in the end a question about prediction. This chapter does not try to answer it directly. It starts one step earlier, with the measurement itself.

Measuring the credit-to-GDP gap looks simple enough, but reading it in economic terms is harder. The credit stock is divided by annual nominal income and an estimated trend is subtracted from the resulting ratio. A positive value means the ratio is above a reference derived from its own history. That does not by itself mean excessive debt, just as a negative value does not mean that balance sheets are safe. Change the coverage of credit, the valuation method, the income denominator or the way the trend is computed, and the level of the indicator can shift. Sometimes even its sign does.

Economies where borrowing in foreign currency is common face a further problem of interpretation. When the local currency depreciates, the local currency value of existing foreign currency debt goes up even if not a single new loan is extended. Rapid growth in nominal income can work the other way, pulling the credit-to-GDP ratio down while the nominal credit stock keeps rising.

This chapter takes up three questions for the Turkish economy. The first is how separating credit by currency changes the reading of the aggregate credit ratio and its gap. The second is in which periods a foreign currency credit ratio scaled by the real effective exchange rate (REER) departs from the ratio calculated at the current exchange rate. The third is how sensitive the size and sign of the gaps are to the filter parameter, the start date and the choice of normalisation. All three are measurement questions and can be

answered with observed data. Whether any of these indicators predicts future systemic banking crises better than the others is a separate research question, which we leave to a later extension of this study.

The sample runs from 2003Q1 to 2026Q2. One advantage of the dataset is that the split between Turkish lira and foreign currency exposures can be followed over a long period from a single banking source. The price paid for this is that coverage cannot be restricted to domestic non-financial private borrowers. The total gap computed here is therefore not a replication of the credit gap published by the BIS, nor of the official CCyB indicator used by national authorities. This is more than a footnote to the data section, since it sets the limits within which the results can be applied.

There are three main findings. Peak dates for foreign currency credit ratios scaled by the current exchange rate differ from those of ratios scaled by the REER. A negative aggregate gap does not mean that every component gap is negative, because the components can offset each other. And the choice of filter, particularly towards the end of the sample, makes a difference large enough to change the economic story. Taken together, the results argue for reporting the ratios and the alternative gaps side by side instead of treating any single gap as a direct measure of risk. The chapter does not offer a new history of crises, and it makes no claim that some untested early warning model would do better. What it does is to set out the mathematical conditions under which measurements drawn from the same data can be compared, and to show which claims the results can support. A simple debt-service framework derived from Minsky's financial instability hypothesis shows where the aggregate credit indicator and borrower vulnerability part ways. The empirical sections then report the properties of the observed ratios and gaps without assuming that this distance has been closed.

Literature and the Measurement Problem

Large cross-country panels show that the credit-to-GDP gap signals banking crises several years in advance (Borio and Lowe, 2002; Drehmann, Borio and Tsatsaronis, 2011; Drehmann and Juselius, 2014). This evidence is what led the Basel Committee to adopt the gap as the common reference indicator for the CCyB (BCBS, 2010). Even so, the framework treats the gap as a starting point. A mechanical threshold rule is not considered adequate for every country and period, and authorities are expected to base their assessment on a wider set of information (BCBS, 2010).

Drehmann and Tsatsaronis (2014) make a related point when they separate the monitoring of a build-up in financial vulnerability from turning the indicator into a mechanical decision rule at any given moment. The same line is followed here. A positive gap is not automatically coded as a call for policy tightening, and a negative gap is not coded as a signal to release the buffer. A step of that kind would require the decision objective, the other risk indicators and the costs of wrong decisions to be defined separately.

BIS credit gap series are built on total credit to the private non-financial sector, and in that definition the borrower's sector and residence matter, as do non-bank and cross-border sources of funding (BIS, 2026). An indicator built from bank loan data published by the Banking Regulation and Supervision Agency (BRSA) overlaps with this concept without being identical to it. Any discrepancy found in the Turkish data may therefore come from the currency split, or from differences in which borrowers and lenders are covered. We have not spliced different datasets into a single continuous series. The ambiguity is stated openly instead.

For emerging market economies (EMEs), three questions remain open. The first is how well the indicator performs in individual EMEs. The BIS calibration rests largely on the experience

of advanced economies. Where financial deepening is still under way, a slow-moving trend lags behind the structural rise in the credit-to-GDP ratio and may label healthy deepening as excess credit (Drehmann and Tsatsaronis, 2014).

Foreign currency debt and borrower balance sheets

The second open question is related with the currency composition of credit. In economies constrained by Eıchengreen and Hausman's original sin (Eichengreen and Hausmann, 1999), a firm that earns in Turkish lira but borrows in foreign currency is structurally more fragile than a firm owing the same amount in lira. Foreign currency debt matters for financial stability for reasons that go beyond its share in total debt. When a debt is denominated in a foreign currency, exchange rate movements can change the borrower's capacity to repay. Models of balance sheet effects build this channel into their assessment of exchange rate policy and financial conditions (Céspedes, Chang and Velasco, 2004).

Two firms with the same amount of foreign currency debt can face quite different risks, depending on their foreign currency revenues, liquid assets, derivative positions and debt maturities. Nor are the banks' foreign currency positions the same as those of the firms they lend to. Exchange rate risk that sits with the borrowing firm can come back to the bank as credit risk, and how much of it does so depends heavily on collateral, capital and loss-sharing arrangements.

For Turkey, Çapacıoğlu and Kara (2024) examine how a regulation aimed at firm-level currency mismatches changed the composition of borrowing. The 2018 regulation they study ties the foreign currency borrowing of firms below a certain foreign currency loan threshold to their past foreign currency revenues. It does not follow that every foreign currency borrower is fully hedged, or that exchange rate risk was eliminated as of some particular date.

When macro-level credit series are interpreted, the group targeted by the regulation has to be kept apart from the aggregate coverage of the data.

Broad cross-country comparisons such as Gourinchas and Obstfeld (2012) do not break the credit variable down by currency. Aldasoro, Borio and Drehmann (2018) do test foreign currency debt as a possible early warning indicator in a panel of countries, but they find that it performs less well than the traditional indicators and keep only the cross-border credit measure for the rest of their analysis. Their focus is on predictive content across countries, not on how the measured foreign currency credit gap of a single country depends on valuation and filter choices. Bringing currency information into early warning work is therefore not new in itself. Our contribution is narrower. We compare how current and REER-scaled ratios, both calculated within the scope of consolidated Turkish banking data, behave under the same filter design.

The third open question is how the trend is extracted. The Hodrick-Prescott (HP) filter balances closeness to the observed series against the smoothness of the trend. (Edge and Meisenzahl, 2011). Hamilton (2018) argues that the filter can generate spurious dynamics and has particular problems at the end of the sample, and he proposes a regression-based alternative. Drehmann and Yetman (2020) compare several univariate filters by how well they predict crises and find that Hamilton gaps estimated country by country perform poorly in real time and fail to beat the BIS gap. A method that looks attractive in theory, or in a statistical sense, is therefore not guaranteed to forecast better in every application.

This chapter does not run a "filter contest". The main calculation uses $\lambda = 400{,}000$, the common reference value in the literature, and the two lower values are there mainly to show how much the results depend on the degree of smoothing. None of the three values was chosen for its crisis prediction performance.

Comparing them is a way of checking how stable the measure is, not of finding the "right" setting.

One-sided calculation is not the same as using real-time information. That a filter uses no observations beyond period *t* does not imply that the data for period *t* had been published by that date, or that they were not revised afterwards.

Debt Service and Aggregation

Financial fragility has to do both with the stock of debt and with the cash available to service it. Minsky's distinction between hedge, speculative and Ponzi finance, which is central to his financial instability hypothesis (Minsky, 1992), captures exactly this difference. Equation (1) gives a one-period version of the idea. It is a partial expression and should not be read as a complete model of the firm and its balance sheet.

Let $D_i > 0$ be the debt of firm i at the beginning of the period, Π_i the operating cash flow available for debt service, r_i the effective interest rate and a_i the share of the debt that falls due as principal in the same period. Defining cash flow per unit of debt as $\kappa_i = \Pi_i / D_i$, we obtain three zones.

$$\kappa_i \geq r_i + a_i \quad \text{(hedge finance)}$$

$$r_i \leq \kappa_i < r_i + a_i \quad \text{(speculative finance)}$$

$$\kappa_i < r_i \ \text{(Ponzi finance)} \qquad (1)$$

Under this classification, the classification is made by the ratio of cash flows that cover interest payment and prinicipal. Hedge finance is the case when cash flow covers both interest and principal. Speculative finance is the case when it covers only the interest. Ponzi finance is the case when it does not cover even the interest. The thresholds are definitional. To place a firm in one of the zones empirically, one needs data on its debt service and its cash flow. Selling liquid assets or bringing in new equity can keep a firm from

default despite a cash shortfall, so the classification is not the same thing as default.

Suppose all firms face a common interest rate r, and let $F^D(\kappa)$ be the debt-weighted cumulative distribution of κ. The share of debt in Ponzi positions is then $F^D(r)$. Holding the distribution fixed, how sensitive this share is to a small rise in the interest rate depends on the density $f^D(r)$. This does not mean that a higher aggregate credit-to-GDP ratio necessarily raises $F^D(r)$. New credit may also go to borrowers with strong cash flows, and nominal income and the sectoral distribution can change at the same time.

Nor does a leftward shift of the distribution necessarily raise the density at every interest rate threshold. One cannot infer from the credit gap that "the concentration of vulnerable firms around the threshold has increased". The link between the aggregate indicator and the micro-level distribution is a plausible channel in theory, but it is not tested directly with this dataset.

The exchange rate threshold and its conditions

Let D^L and D^f be the firm's debt in domestic and in foreign currency, r^L and r^f the corresponding interest rates, e the domestic currency price of foreign currency, and Π^L and Π^f the operating cash flows in domestic and foreign currency. The operating cash surplus after interest, S, can then be written as in equation (2).

$$S(e) = \Pi^L + e\Pi^f - r^L D^L - e r^f D^f = A - eN. \quad (2)$$

Here $A = \Pi^L - r^L D^L$ and $N = r^f D^f - \Pi^f$. $N > 0$ describes the case in which foreign currency operating flows cannot cover the foreign currency interest bill. If $A > 0$, interest-paying capacity runs out at $e^P = A/N$. When $A \geq e_0 N$ holds at the initial exchange rate e_0, the firm is able to cover its interest at the start. If e rises above e^P while everything else stays the same, operating cash flow falls below

interest payments. The ratio δ^P in equation (3) gives the proportional depreciation needed to reach that point.

$$e^P = A/N, \qquad \delta^P = e^P/e_0 - 1$$

$$A > 0,\ N > 0,\ A \geq e_0 N \quad (3)$$

This threshold marks a change in the financing position that is triggered by an exchange rate shock. We call it a "currency-contingent Ponzi" case, by which we mean a situation where, under the stated assumptions, operating cash flow is not enough to pay interest. Because the exchange rate can affect export volumes, pricing, the cost of imported inputs and interest rates, a change in the rate also moves the threshold itself. A more exact calculation would have to include hedging derivatives and the cash flows from foreign currency assets as well.

When $N \leq 0$, a higher exchange rate does not lower the cash surplus in this simplified, interest-only equation. That does not make the firm fully hedged against all kinds of exchange rate risk, since principal repayments, translation risk, margin calls and the timing of revenues are all outside the equation. In the same way, $N > 0$ does not by itself mean that the firm is close to default. What actually happens depends on its liquidity and its access to funding.

From the firm to the macro indicator

The case for separating credit by currency is more modest. Domestic and foreign currency debts are affected by different valuation and debt-service mechanisms, and following the two components separately shows which kind of credit is driving the aggregate ratio. This helps to decide where firm-level risk analysis should look more closely. It does not tell us how well borrowers are hedged, and it does not measure the causal effect of shocks.

The same distinction sets the limits of how the empirical results can be phrased. Statements such as “the foreign currency ratio has increased”, “the component gap has diverged from the total” or “the result is sensitive to the filter” describe things that are measured here. Claims such as “currency mismatch has intensified”, “the crisis was anticipated” or “the risk has been eliminated” would need additional data and a different research design.

Data and Coverage

Credit stocks variable were taken from the “Total Loans” table of the BRSA Monthly Bulletin. Quarterly end-of-period stocks were built from the March, June, September and December figures of each year. The domestic and foreign currency columns are denominated both in millions of TL. The main calculation uses the sum of these two columns. The largest difference between this sum and the “Total” column of the source is TL 1 million, which is within rounding.

GDP is the quarterly series at current prices produced by TurkStat and retrieved through the Electronic Data Delivery System (EVDS) of the Central Bank of the Republic of Türkiye (CBRT).. Each credit stock is divided by the sum of GDP in the current and the previous three quarters, so that a stock and a flow are compared on a common annual basis. To compute this denominator for 2003Q1, the GDP series starts a few quarters before the credit sample.

The REER is the CBRT's CPI-based real effective exchange rate index, taken for the last month of each quarter. A rise in the index means a real appreciation of the Turkish lira. The base of the index is 2025 = 100.

Table 1 Main data sources

Variable	Source and unit	Treatment
TL and FX loans	BRSA Monthly Bulletin, TL million	End-of-quarter stock
Nominal GDP	TurkStat / EVDS, TL thousand	Converted to TL million, sum of the last four quarters
CPI-based REER	CBRT / EVDS, 2025 = 100	Last month of the quarter

Note: The credit sample covers 2003Q1-2026Q2 (94 observations). Source codes and the order of calculation are given in Appendix A.

The credit series covers bank lending in a broad sense, excluding interbank loans. Public sector borrowers, non-bank financial institutions and non-resident borrowers can not be removed with one consistent method over the whole sample. The reporting perimeter of the banks also does not exactly match some CBRT credit tables that are based on the residence of the borrower. For these reasons the results are not labelled as "domestic non-financial private sector" credit.

More detailed data on loans to financial institutions and non-residents exist for recent years, but they are not enough to rebuild the earlier years with the same coverage. Using the broad series before 2022 and a narrower one afterwards would create an artificial break in the data.

Participation banks enter the data in 2005, and this has to be taken into account too. Since the sample starts in 2003, an alternative calculation that starts in 2005 can serve as a check on how this change in coverage affects the filter. The check does not mean that every reporting change has been dealt with. What it measures is how much the known difference in the starting point changes the results.

Number of observations and data vintage

The dataset covers 94 consecutive quarters, starting in 2003Q1. The first 40 observations, from 2003Q1 to 2012Q4, serve

as a burn-in period for the initial trend estimates. The main analysis and its interpretation concern the period from 2013Q1 onwards.

Because the data come from the latest available vintage, historical revisions to GDP and the REER may have changed past figures. A look at the major revision documents of TurkStat and the REER update announcements of the Central Bank of the Republic of Türkiye (CBRT) shows that one series code does not necessarily stand for a dataset that stays consistent over time. The retrieval date is therefore part of what makes the results reproducible. Giving the observation period alone is not enough.

Calculating the Ratios and Gaps

Let L_t denote the stock of Turkish lira credit, F_t the reported Turkish lira equivalent of foreign currency loans, and Y_t quarterly nominal GDP. The annual denominator is defined as $G_t = Y_t + Y_{t-1} + Y_{t-2} + Y_{t-3}$. All ratios are expressed in per cent.

$$R_t^L = 100L_t/G_t; \quad R_t^f = 100F_t/G_t$$

$$R_t^T = 100(L_t + F_t)/G_t = R_t^L + R_t^f \quad (4)$$

The total ratio at the current exchange rate is the benchmark for the comparison, and it is consistent with the TL credit stock reported by banks. Because of its broad coverage, however, we do not call it the official CCyB reference series. Showing the current foreign currency ratio separately makes it possible to see directly which way the foreign currency component of the total is moving.

Scaling by the REER

Converting foreign currency loans into TL at a single nominal exchange rate over the whole period is problematic in a long sample with high inflation. The numerator is then held at a fixed nominal conversion while nominal GDP keeps growing, so a fall in such a ratio may reflect a mismatch in nominal scale that has nothing

to do with the amount of credit. Dividing foreign currency credit by GDP expressed in foreign currency does not solve the problem either. With the same current exchange rate on both sides, one simply gets the current TL ratio back.

In this chapter the alternative series is obtained by scaling the current TL ratio of foreign currency loans by the REER. Let Q_t denote the REER index and q_0 a positive normalisation constant.

$$R_t^R = R_t^f \, Q_t/q_0; \quad q_0 = (1/T) \sum_t Q_t. \quad (5)$$

In a simple model with a single foreign currency and consistent price indices, Q_t is proportional to the ratio of the domestic price level to the foreign price level converted at e_t. The product e_tQ_t then links changes in the nominal exchange rate to relative prices. In practice, though, the CBRT's REER is a trade-weighted multilateral index, and there is no reason to assume that it carries the same weights as the banks' foreign currency loan basket. F_t, for its part, contains the reported TL value of loans in more than one currency.

R^R_t is therefore a comparison series scaled by the REER, and not, strictly speaking, “the real amount of credit valued at a constant real exchange rate”. The transformation is useful for studying how current valuation relates to relative price movements. Without observing the currencies and amounts of loans one by one, it cannot be claimed to isolate a pure exchange rate effect. An exact constant-exchange-rate calculation would need stocks by currency together with the matching exchange rates.

It would be just as wrong to treat all exchange rate valuation as economically irrelevant. If the borrower earns in domestic currency, a rise in the current exchange rate can erode real repayment capacity. That is why the scaled ratio is presented alongside the current ratio, not in place of it. The scaled ratio offers an alternative valuation perspective, and the current ratio shows the reported debt

burden as it stands. Which series is more useful depends on the kind of risk being monitored and the kind of decision at hand.

Normalisation and additivity

The simple sample mean of the REER is 130.004. Using this constant does not make the means of the scaled and current foreign currency ratios equal. For that, the covariance between the credit ratio and the REER would have to be zero. The actual means are 15.10 per cent for the current ratio and 14.16 per cent for the scaled one. The difference is not a computational error, it follows from the way the normalization is defined.

If the two ratios are to have exactly the same sample mean, the normalization constant has to be weighted by the credit ratio. The value q_m below does this, and in our sample it equals 121.991.

$$q_m = (\sum_t R_t^f Q_t)/(\sum_t R_t^f). \quad (6)$$

The REER mean over 2003-2007 is 152.449. Each of these constants rescales the same Q_t series by a positive number. Under the same linear filter, the sign and turning points of the scaled gap stay the same and only its magnitude changes. A signal that relies on a fixed threshold, such as "exceeding 2 percentage points", can change, though. For a threshold rule, the choice of normalisation is far from innocuous.

The current total is not equal to the sum of the TL ratio and the scaled FX ratio. The difference should be given an explicit name and reported separately.

$$V_t = R_t^T - R_t^L - R_t^R = R_t^f - R_t^R = R_t^f(1 - Q_t/q_0). \quad (7)$$

We refer to V_t as the "REER scaling difference". It is an accounting difference arising from the transformation and should not be read as a full causal valuation effect. The level of V_t and the deviation of this series from its trend are two different things. The

first is the distance between the two ratios. The second tells us where that distance stands relative to its own trend.

The one-sided HP calculation

For each ratio the HP trend is re-estimated using only the observations up to period t. The objective function minimises a weighted sum of squared deviations from the observations and squared second differences of the trend. The baseline value is λ = 400,000. No future observations of the ratio enter the calculation.

$$\min\{\sum_{s=1}^{t}(R_s - \tau_s)^2 + \lambda\sum_{s=3}^{t}(\tau_s - 2\tau_{s-1} + \tau_{s-2})^2\}$$

$$g_t = R_t - \hat{\tau}_t|_t \quad (8)$$

At each date only the deviation at the last point is kept, and the procedure is repeated on the expanded sample in the following quarter. All components start on the same date and use the same λ. No gap is reported for the first two observations. The filter is not restarted in 2013 in order to interpret the post-2013 period, so the earlier observations stay in the trend calculation.

Since the HP transformation is linear, for the same sample and parameter the current total gap equals the sum of the TL gap and the current FX gap. Once the scaled FX ratio is used, a third difference component is needed.

$$g_t^T = g_t^L + g_t^f; \;\; g_t^T = g_t^L + g_t^R + g_t^V. \quad (9)$$

These identities follow from a property of linear transformations and say nothing about economic causation. The same equality cannot be expected when gaps obtained from different start dates are added up, or when different filters are applied to the components. In our calculations the largest numerical error in the current total identity is below 10^{-12} percentage points, which confirms that there is no adding-up discrepancy apart from rounding.

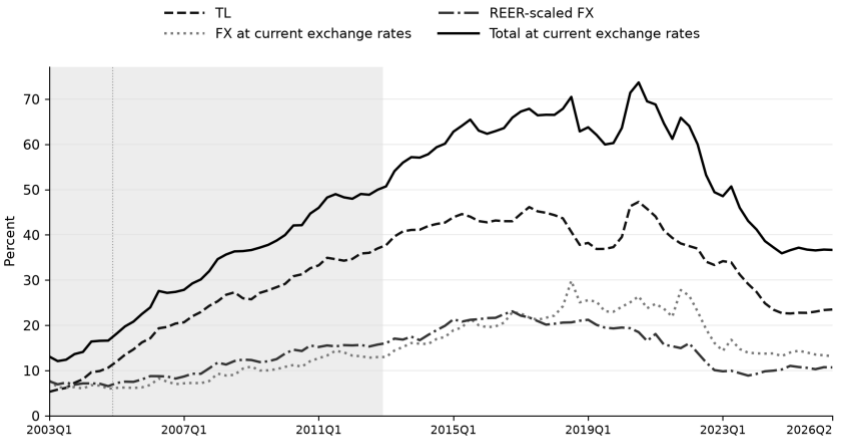
Three kinds of check were carried out. In the first, λ is set to 25,000, 125,000 and 400,000 in turn. In the second, the series start in 2005Q1, which means they are calculated from the period in which participation banks are included. The third compares three REER normalisations, namely the full-sample mean, the 2003-2007 mean and the mean-equalising constant. In the filter and start-date checks all other calculation choices are kept as they are.

Start dates are compared over the common period 2015Q1-2026Q2. In this way the series beginning in 2005 also has 40 earlier quarters in its trend calculation. The common period contains 46 observations. Sign agreement with the main indicator is the share of quarters in which the two gaps fall in the same class, the classes being whether the gap is above zero or not. A gap of exactly zero is counted as non-positive. Sign agreement should not be read as success in capturing crises.

Empirical Findings

Figure 1 shows the four ratios together. The TL ratio and the FX ratio at the current exchange rate add up to the total ratio at the current exchange rate. The REER-scaled FX ratio is shown as a separate comparison and is not part of this sum. The shaded area marks the period before 2013, which is used to initialise the trend for the gap calculations.

Figure 1 Credit Ratios by Currency



Source: author's calculations.

Credit stocks are divided by the sum of nominal GDP in the current and the previous three quarters. The data are BRSA bank credit with broad coverage, excluding interbank loans, and public sector and other borrower groups could not be separated for the whole sample. The REER-scaled FX ratio is the FX ratio at the current exchange rate multiplied by $Q_t/130.004$. The shaded area covers 2003-2012, and the dotted vertical line marks the change in data coverage in 2005. Sources: BRSA, CBRT EVDS and author's calculations.

As it can be seen in Figure 1, the FX ratio at the current exchange rate reaches its sample peak in 2018Q3, at 29.83%. The REER-scaled ratio peaks earlier, at 23.09% in 2016Q4. Changing only the way FX credit is valued thus moves the peak seven quarters earlier. The time profile is clearly sensitive to the conversion method, but this does not mean that some definite part of the 2018 increase is due to the exchange rate as such. The TL ratio and the total ratio both peak in 2020Q3, at 47.29% and 73.70% respectively. That the total and the FX component peak at different dates suggests that the

components do not move within one common financial cycle. “Peak” here simply means the highest value observed in the current data and sample.

In 2026Q2 the TL ratio is 23.50%, the FX ratio at the current exchange rate is 13.18% and the total ratio is 36.68%. The REER-scaled FX ratio stands at 10.69%. The difference between the current total and the sum of the TL ratio and the scaled FX ratio is 2.48 percentage points.

Table 2 Credit Ratios in Selected Quarters

Quarter	**TL**	**FX current**	**FX REER-scaled**	**Total current**
2007Q4	24.31	7.68	10.51	31.99
2012Q4	37.00	12.98	15.74	49.98
2016Q4	43.02	22.90	23.09	65.92
2017Q4	44.87	21.68	20.15	66.55
2018Q3	40.67	29.83	20.65	70.50
2020Q3	47.29	26.41	18.48	73.70
2021Q3	39.32	21.90	15.37	61.22
2024Q2	24.87	13.75	9.83	38.62
2026Q2	23.50	13.18	10.69	36.68

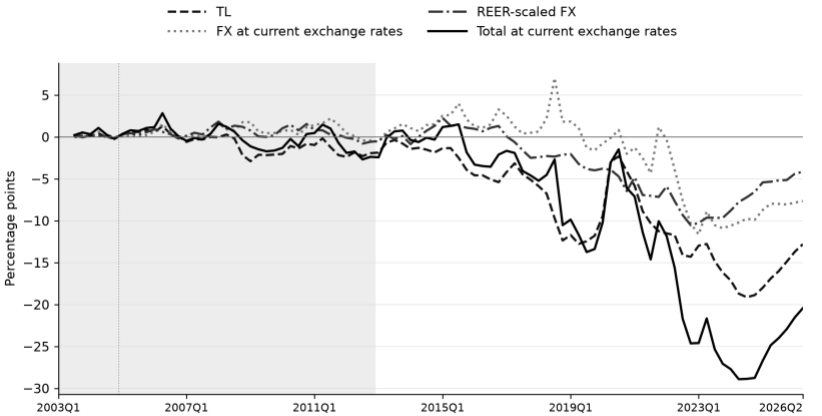
Unit: per cent of GDP. Source: author's calculations.

The quarters in Table 2 are not an event test for the whole period. They give a comparable picture of different stages of the long sample and of the peaks in the ratios. The contrast between 2016Q4 and 2018Q3 shows clearly that the current and the scaled ratios answer different questions. Because the contributions of credit volumes, cross exchange rates, classification changes and income between these dates have not been decomposed seperately, the outcome is not attributed to any single factor.

Figure 2 applies the same one-sided HP procedure to every series, and each gap is the difference between the observed ratio and its estimated trend, in percentage points. A ratio at its sample peak

need not have a positive gap, since the trend may be even higher at that point. When the total ratio peaks in 2020Q3, for instance, the total gap is -1.48 percentage points.

Figure 2 One-Sided Credit Gaps by Currency



Source: author's calculations.

λ = 400,000, common starting point 2003Q1. Each quarter the end-point deviation is recalculated with observations from the current data vintage up to that quarter. The shaded area is the 40-quarter initialisation period. This is not a real-time test based on historical vintages. Credit coverage and the REER transformation are the same as in Figure 1. Sources: BRSA, CBRT EVDS and author's calculations.

In 2016Q4 the TL gap is -5.39, the FX gap at the current exchange rate is +3.29 and the total gap is -2.10 percentage points, while the REER-scaled FX gap is +1.32. A negative total gap can thus coexist with a positive component gap. The aggregate is not wrong. Netting simply folds different movements in the components into one number.

In 2018Q3 the FX gap at the current exchange rate climbs to +7.03 percentage points, but the REER-scaled FX gap is -2.34. With a TL gap of -9.71, the total gap comes to -2.68. Both aggregation and the choice of transformation therefore matter for interpretation. The current-rate FX indicator reflects the high reported value of foreign currency debt, and the scaled indicator shows where the REER-transformed ratio stands against its own history. Neither should be called the "correct" version of the other.

In 2026Q2 the total gap under the baseline filter is -20.45 percentage points, of which the TL component accounts for -12.80 and the FX component at the current rate for -7.64. With a scaled FX gap of -4.15, the gap of the scaling difference in equation (9) is -3.50 percentage points. That the corresponding level difference is +2.48 on the same date is not a contradiction, because a variable can be positive in level and still below its own trend.

Table 3 Credit Gaps in Domestic and Foreign Currencies

Quarter	TL	FX current	FX REER-scaled	Total current
2007Q4	0.01	0.37	1.13	0.38
2012Q4	-1.94	-0.43	-0.57	-2.37
2016Q4	-5.39	3.29	1.32	-2.10
2017Q4	-5.10	0.51	-2.49	-4.59
2018Q3	-9.71	7.03	-2.34	-2.68
2020Q3	-2.31	0.83	-4.71	-1.48
2021Q3	-10.29	-4.32	-7.02	-14.61
2024Q2	-18.69	-10.21	-7.76	-28.90
2026Q2	-12.80	-7.64	-4.15	-20.45

Unit: percentage points. λ = 400,000, starting point 2003Q1. Source: author's calculations.

Splitting credit by currency is useful for description. Date ranges picked after looking at the results are no substitute for a validation sample.

Between 2020Q3 and 2026Q2 the total credit-to-GDP ratio falls from 73.70% to 36.68%. Over the same period total nominal

credit grows to about 7.54 times its starting level, while trailing four-quarter nominal GDP grows to about 15.15 times. The ratio falls because the denominator grows faster then the numerator.

The relationship can be written exactly as a logarithmic accounting identity. With total credit defined as $C_t = L_t + F_t$, the log change in the ratio equals the log change in credit minus the log change in annual nominal income.

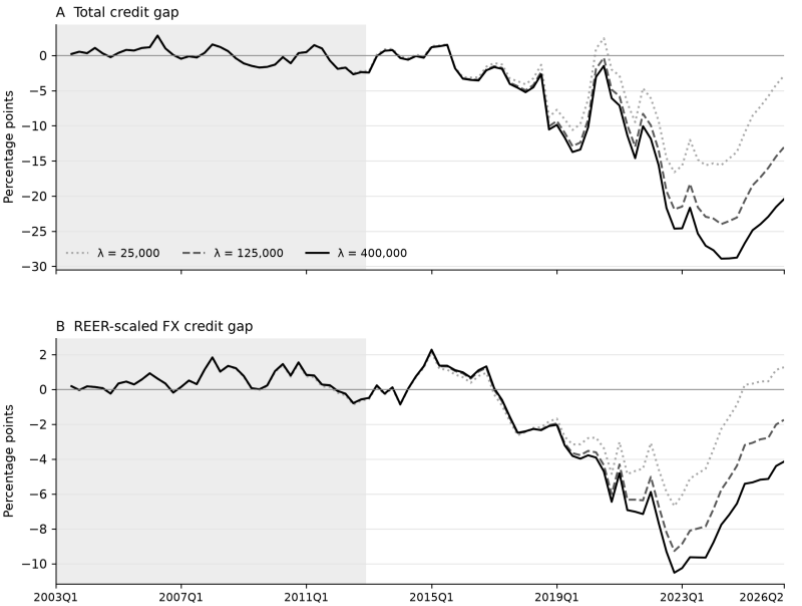
$$\Delta \ln(C/G) = \Delta \ln C - \Delta \ln G. \quad (10)$$

Between the two dates the log increase is 2.020 for total credit and 2.718 for nominal GDP, a difference of -0.698. This accounts for the fall in the ratio in terms of its numerator and its denominator. It does not say how much of the GDP growth reflects prices and how much reflects real activity, which would need a consistent split into real GDP and the deflator.

Results from Sensitivity Test

Figure 3 compares the gaps produced by different values of λ. Here the baseline ratio definitions left unchanged. A lower λ lets the trend follow the observed series more quickly. How large the effect is in practice is an empirical matter, and the figures below show it.

Figure 3 Sensitivity to the filter parameter



Source: author's calculations.

All series start in 2003Q1, and the normalization and data vintage are the same in every calculation. The alternative λ values were not chosen for predictive performance, they are used only to show measurement sensitivity. The period between 2003 and 2012 is accepted as the initialization period.

In 2026Q2 the total gap is -2.90 percentage points with λ = 25,000, -13.06 with λ = 125,000 and -20.45 with λ = 400,000. The sign stays negative throughout, but the size changes a great deal. These figures should not be read as a confidence interval around one estimate. They are different indicators produced by different definitions of the trend.

The divergence is larger for the REER-scaled FX gap. On the same date it is +1.27 percentage points with λ = 25,000, -1.75 with λ = 125,000 and -4.15 with λ = 400,000. Because here even the sign changes, the statement that "FX credit is below trend" cannot be defended without saying which filter is meant. This is one of the main findings that limit the policy reading of the study.

Over the common period 2015Q1-2026Q2, the sign classifications under λ = 25,000 and under the baseline filter agree in 95.7% of quarters for the total gap and in 84.8% for the REER-scaled FX gap. High overall agreement does not make the difference in magnitudes at the end of the sample unimportant. Equally, an agreement of 84.8% does not mean that the remaining quarters are irrelevant for policy. When the divergence occurs, and how far the gap is from a decision threshold, both need to be assessed separately.

Table 4 Sensitivity to the Filter and Starting Point

Start	**λ**	**Total gap**	**Scaled FX gap**	**Total sign agreement %**	**FX sign agreement %**
2003Q1	25,000	-2.90	1.27	95.65	84.78
2003Q1	125,000	-13.06	-1.75	100.00	97.83
2003Q1	400,000	-20.45	-4.15	100.00	100.00
2005Q1	400,000	-19.19	-3.83	100.00	97.83

Gaps are in percentage points for 2026Q2. Sign agreement is measured against the calculation that starts in 2003 with λ = 400,000, over the 46 observations of the common period 2015Q1-2026Q2. It is not a measure of crisis prediction performance.

With a 2005 start and λ = 400,000, the 2026Q2 total gap is -19.19 and the REER-scaled FX gap -3.83 percentage points, that is 1.26 and 0.32 points away from the baseline. Over the 46 common quarters, sign agreement is 100% for the total gap and 97.8% for

scaled FX credit. For this sample and this alternative start date the results move less than they do when the filter parameter changes.

In the normalization exercise the REER-scaled FX ratio in 2026Q2 is 10.69% with the full-sample mean, 9.12% with the 2003-2007 mean and 11.40% with the mean-equalising constant. The corresponding baseline-filter gaps are -4.15, -3.54 and -4.42 percentage points. Since all of these changes come from scaling by a positive constant, signs and peak dates stay the same.

Interpreting the Results and Moving Towards Early Warning

The obtained empirical results show that the credit gap is not an economic quantity that exists independently of data coverage and method. Once the current total ratio, the TL and FX components, the alternative FX conversion and the filter sensitivity are read side by side, it becomes much easier to tell which interpretation comes from the data and which comes from a methodological choice.

Splitting credit by currency generally helps in two situations. The first is when the total gap nets out components that are moving in opposite directions. The second is when the FX ratio at current exchange rates and the REER-scaled ratio trace out different paths over time. In the first case a closer look at the sectoral or borrower level is called for, whereas the second means that valuation and income channels have to be tracked separately. Such findings tell us where the analysis should focus. They do not produce policy decisions automatically.

Stripping out current exchange rate effects is not an improvement for every purpose. If the aim is to follow how an exchange rate shock feeds into borrowers' capacity to repay, the debt burden measured at current rates may well be indispensable. If the aim is instead to track the build-up in the volume of credit, stock and flow data expressed in foreign currency are probably the better fit.

Rather than folding both objectives into one transformed series, it is better to say plainly which question each indicator actually answers.

By the same token, a negative total gap does not show that the banking system needs no additional loss-absorbing capacity. Capital adequacy, asset quality, FX liquidity, borrower mismatches and funding conditions are separate dimensions of risk. Nothing in the calculations presented in this chapter points to a particular CCyB rate, or to a regulatory intervention that ought to be made on a particular date.

At the next stage of the research, first decision has to be how the event to be predicted is defined. A systemic banking crisis, pressure in the foreign exchange market and broad financial stress are not the same dependent variable. That an indicator picks up exchange rate volatility says little about whether it would do equally well for banking crises. The event definition should therefore be fixed without looking at how the credit indicators perform, and different kinds of events should not be pooled into a single success rate.

The second condition is data availability. . The end-of-quarter credit stock, GDP for the same quarter and the matching REER observation are not published on the same day.

The forecast horizon also constrains the possible end of the sample. If the question is whether an event occurs within 4-12 quarters after dat *t*, a complete negative label requires observations up to t+12.

In a short Turkish sample with limited coverage, the number of independent events is far smaller than the number of quarters. Nine overlapping quarters ahead of the same event are not nine independent crises. Penalty terms, the Firth correction or resampling cannot create new information about events that were never

observed. Comparing many variables and models on the same handful of events can therefore easily end in overfitting.

A workable follow-up design would compare a small number of pre-specified indicators against a simple benchmark. Fixed percentage-point thresholds should not be carried over automatically to subcomponents or to rescaled series, which are measured in different units. Threshold selection should be confined to the training period, and the frequency of false alarms, the number of independent events captured, the lead time and the uncertainty around all of these should be reported together. Since this design was written after the historical sample had already been seen, it cannot be presented as a pre-registration for past data. It can, however, serve as a pre-specified validation plan for new observations.

Conclusion

Interpreting the credit gap in Turkey requires looking at the currency composition of credit and at the calculation choices together. With the same banking data, the FX credit ratio at the current exchange rate peaks in 2018Q3, whereas the REER-scaled ratio peaks in 2016Q4. The valuation method alone thus changes the time profile of FX credit, although the analysis does not attribute any particular part of this difference to the exchange rate as such. Since the components of the aggregate gap can also move in opposite directions, the aggregate gap on its own does not give the full picture.

None of this makes one indicator superior to the others. Since trade weights differ from the currency composition of credit, REER conversion is only an approximate comparison. Gross credit stocks do not reflect borrowers' income, hedging positions or debt maturities either. Therefore, the link between the theoretical debt-service threshold and the macro credit gap is indirect and needs further information.

The sensitivity analysis shows that some of the results depend heavily on methodological choices. Under a lower λ the REER-scaled FX gap turns positive in 2026Q2. And a negative value from one filter is not a sound basis for a general story of declining risk. Normalization keeps the sign but can still decide whether a threshold is crossed. The ratio, the gap and the threshold are best seen as three separate steps of measurement.

References

Aldasoro, I., Borio, C. and Drehmann, M. (2018). Early warning indicators of banking crises: Expanding the family. BIS Quarterly Review, March, 29-45.

Bank for International Settlements [BIS]. (2026). Credit-to-GDP gaps: Overview and methodology. BIS Data Portal. Accessed 8 September 2026. https://data.bis.org/topics/CREDIT_GAPS

Banking Regulation and Supervision Agency [BRSA]. (2026a). Türk Bankacılık Sektörü Aylık Bülten: Krediler [Monthly Bulletin of the Turkish Banking Sector: Loans]. Observations for March 2003-June 2026. Data retrieved 14 September 2026. https://www.bddk.org.tr/BultenAylik/tr/Home/Index

Banking Regulation and Supervision Agency [BRSA]. (2026b). Aylık Bülten metaverisi [Monthly Bulletin metadata]. Version accessed 7 October 2026. https://www.bddk.org.tr/BultenDosyalari/Home/Index/Aylik-MetaVeri

Basel Committee on Banking Supervision [BCBS]. (2010). Guidance for national authorities operating the countercyclical capital buffer. Bank for International Settlements, December. https://www.bis.org/publ/bcbs187.htm

Borio, C. E., & Lowe, P. W. (2002). Asset prices, financial and monetary stability: exploring the nexus.

Central Bank of the Republic of Türkiye [CBRT]. (2026a). Electronic Data Delivery System (EVDS): CPI-based real effective exchange rate, TP.RK.T1.Y.

Central Bank of the Republic of Türkiye [CBRT]. (2026b). Real effective exchange rate statistics: Methodology and update announcements.

Céspedes, L. F., Chang, R., & Velasco, A. (2004). Balance sheets and exchange rate policy. American Economic Review, 94(4), 1183-1193.

Çapacıoğlu, T. and Kara, H. (2024). Taming corporate sector currency mismatches: Reflections from a quasi-natural (macroprudential) experiment. CBRT Working Paper No. 24/06. h

Drehmann, M., Borio, C. and Tsatsaronis, K. (2011). Anchoring countercyclical capital buffers: The role of credit aggregates. International Journal of Central Banking, 7(4), 189-240.

Drehmann, M. and Juselius, M. (2014). Evaluating early warning indicators of banking crises: Satisfying policy requirements. International Journal of Forecasting, 30(3), July-September 2014

Drehmann, M. and Tsatsaronis, K. (2014). The credit-to-GDP gap and countercyclical capital buffers: Questions and answers. BIS Quarterly Review, March, 55-73.

Drehmann, M. and Yetman, J. (2020). Which credit gap is better at predicting financial crises? A comparison of univariate filters. BIS Working Papers No. 878

Edge, R. M. and Meisenzahl, R. R. (2011). The unreliability of credit-to-GDP ratio gaps in real time: Implications for countercyclical capital buffers. International Journal of Central Banking, 7(4), 261-298.

Eichengreen, B. and Hausmann, R. (1999). Exchange rates and financial fragility. In New Challenges for Monetary Policy (pp. 329-368). Federal Reserve Bank of Kansas City.

Gourinchas, P. O., & Obstfeld, M. (2012). Stories of the twentieth century for the twenty-first. American Economic Journal: Macroeconomics, 4(1), 226-265.

Hamilton, J. D. (2018). Why you should never use the Hodrick-Prescott filter. The Review of Economics and Statistics, 100(5), 831-843.

Minsky, H. P. (1992). The financial instability hypothesis. Levy Economics Institute Working Paper No. 74.

Turkish Statistical Institute [TurkStat]. (2025). Ulusal hesaplar ana revizyonu 2025 [Main revision of national accounts, 2025].

Turkish Statistical Institute [TurkStat]. (2026). Gross domestic product at current prices, quarterly. Series TP.GSYIH20.BY.B1GQ via CBRT EVDS.

Appendix A. Source and output variables

Table A1 Variable Glossary

Quantity	Source or output definition
L_t and F_t	BRSA total loans table, TL and FX columns (TL million)
Y_t	EVDS TP.GSYIH20.BY.B1GQ / 1,000 (TL million)
Q_t	EVDS TP.RK.T1.Y, last month of the quarter
G_t	Sum of Y_t and the three preceding quarters
R^L, R^f, R^T	100 × relevant credit stock / G_t
R^R	$R^f \times Q_t$ / 130.0037234
g^L, g^f, g^R, g^T	Recursive HP end-point deviation of each ratio
V and g^V	Difference between the current and scaled FX ratios, and between their gaps

The data file may also contain auxiliary USD/TRY and corporate foreign currency debt series, but the core ratios in this chapter do not rely on them. Because FX credit is already observed in its TL equivalent, it has not been converted back into US dollars only to be multiplied again by the same exchange rate. The fact that these auxiliary series are available does not mean that their economic coverage matches that of the main credit table.

CHAPTER 7

E-COMMERCE AND REGIONAL INCOME CONVERGENCE IN TURKIYE

Burcu KILINÇ SAVRUL[1]
Mesut SAVRUL[2]

Introduction

Whether e-commerce narrows regional income gaps depends on how lower trading costs interact with uneven productive capacity. Digital technologies reduce search and verification costs, creating opportunities to connect buyers and sellers across locations (Goldfarb & Tucker, 2019). Online trade is also less sensitive to geographical distance than conventional trade, as evidence from eBay demonstrates (Lendle et al., 2016). These changes could improve access to markets in disadvantaged regions. However,

[1] Prof.Dr.Çanakkale Onsekiz Mart Üniversitesi, İ.İ.B.F. İktisat Bölümü, kilincburcu@hotmail.com, Çanakkale, Türkiye. 0000-0001-7114-0084
[2] Doç.Dr.Çanakkale Onsekiz Mart Üniversitesi, İ.İ.B.F. Ekonometri Bölümü, msavrul@comu.edu.tr, Çanakkale, Türkiye. 0000-0003-1513-1874.

evidence that broadband infrastructure supports aggregate economic growth addresses a different question from the geographical distribution of its gains (Czernich et al., 2011).

Regional development theory provides reasons for uneven responses. Economies of scale and the interaction between demand and production can concentrate activity despite changes in transport costs (Krugman, 1991). Digital adoption may consequently reward places possessing complementary capabilities. Advanced internet investment was associated with substantial wage growth mainly in already advantaged US counties (Forman et al., 2012), while Norwegian evidence shows that broadband adoption improves skilled workers' labour-market outcomes but worsens those of unskilled workers (Akerman et al., 2015). Although these studies examine broader digital infrastructure, they caution against assuming that expanded online access necessarily favours poorer locations.

Evidence specific to e-commerce further separates consumer benefits from producer-income gains. Fan et al. (2018) find that online trade can alleviate spatial consumption inequality in China, with relatively larger welfare gains in smaller and more remote cities. A randomised rural expansion programme nevertheless produced little evidence of income gains for producers and workers, with benefits concentrated in lower living costs for some households (Couture et al., 2021). Liu et al. (2024) similarly distinguish reductions in consumption inequality from changes in income inequality. Regional output convergence therefore captures only one dimension of e-commerce's distributional consequences.

Türkiye offers a relevant setting for examining this distinction. Domestic e-commerce transaction volume rose from 2.7% of GDP in 2019 to 6.5% in 2024, although the path included a decline after 2023 (Ministry of Trade, 2026). This expansion

occurred within an economy whose provincial accounts reveal substantial differences in output per person (Turkish Statistical Institute [TURKSTAT], 2025). Earlier Turkish research also found that declining overall regional inequality could coexist with increasingly dominant spatial dependence (Gezici & Hewings, 2007). National expansion and regional catch-up consequently require separate empirical assessment.

The unresolved question concerns how predetermined provincial capacity relates to income trajectories during this common expansion, including differences by initial income and associations extending across provincial boundaries. Convergence analysis asks whether initially poorer economies catch up with richer ones (Barro & Sala-i-Martin, 1992), but a negative convergence coefficient cannot attribute that process to e-commerce. Spatial convergence research likewise shows why neighbouring outcomes cannot be treated as independent (Rey & Montouri, 1999). Establishing an exposure-related indirect association requires computed spatial responses, rather than inferring transmission from income clustering alone (LeSage & Pace, 2009).

This study examines these relationships using a balanced panel of 81 provinces over 2019–2024. The outcome is provincial GDP per capita relative to its national counterpart, expressed in logarithms at current prices. Predetermined readiness combines 2016–2018 information-and-communication enterprise shares, log population density, literacy and light-commercial vehicles per 1,000 residents. Interacting readiness with national expansion generates synthetic differential exposure. The analysis evaluates three hypotheses: an exposure association with relative income, with its direction left open; a more favourable association in initially poorer provinces; and a non-zero indirect association across provinces. The

outcome concerns relative nominal output rather than real household welfare.

Province and year fixed effects (FE) provide the benchmark; the spatial Durbin model accommodates outcome feedback and neighbouring covariates (Elhorst, 2010). Complete interactions distinguish readiness-related heterogeneity from initial-income differences. Controlled FE tests the overall exposure relationship. Spatial derivatives distinguish own-province associations from network responses generated by poorer and richer sources. The 2025 adaptation comparison assesses measurement relevance; earlier-year profiles and sensitivity checks assess credibility. Predetermined exposure may still correlate with omitted differential shocks, so estimates remain conditional associations (Goldsmith-Pinkham et al., 2020).

The findings separate provincial catch-up from exposure associations. Income dispersion declined during 2019–2024, while clustering persisted. The positive exposure coefficient becomes imprecise with covariate trends. Neither the FE poorer–richer slope difference nor the spatial source-total contrast establishes a favourable differential; source totals concern network responses, not gains received by those groups. The overall indirect association is not established after Holm adjustment, while joint pre-period diagnostics constrain attribution. Explicit income-group comparisons and computed spatial associations contribute a provincial output perspective.

Section 2 develops the literature and hypotheses. Section 3 explains the data and methodology, Section 4 presents and discusses the empirical results, and Section 5 concludes.

Literature Review and Hypotheses Development

Regional Development and the Economics of E-Commerce

Regional income convergence depends on whether changing economic opportunities weaken or reinforce existing spatial advantages. Neoclassical growth theory links catch-up to initial income and structural characteristics (Barro & Sala-i-Martin, 1992), but faster growth in poorer economies need not eliminate persistent distributional stratification (Quah, 1996). Heterogeneous transitions can also generate convergence clubs rather than a common destination (Phillips & Sul, 2007). Increasing returns and the interaction between market size and transport costs offer a competing explanation for sustained concentration (Krugman, 1991). These perspectives imply that technological expansion alone cannot establish regional equalisation.

National e-commerce expansion can alter regional economic opportunities by changing the costs of reaching customers and sourcing goods. Goldfarb and Tucker (2019) explain how digital technologies reduce search, replication, transportation, tracking and verification costs. Evidence that internet development stimulates international trade supports the market-access channel, although it does not imply that geography becomes irrelevant (Freund & Weinhold, 2004). Lendle et al. (2016) find a substantially smaller distance effect in eBay transactions than in conventional international trade. Together, these studies establish a basis for wider market participation, while leaving open where additional production and income accrue.

The geographic distribution of consumer benefits illustrates this distinction. Online purchasing depends partly on access to local retail alternatives (Forman et al., 2009). Fan et al. (2018) show that e-commerce can reduce spatial consumption inequality, with larger modelled welfare gains in smaller and more remote Chinese cities.

However, lower prices and greater product variety need not increase income where consumers reside. In a rural Chinese experiment, Couture et al. (2021) find limited producer and worker income gains; benefits arise principally through lower living costs for a subset of households. Provincial GDP per person consequently captures a different margin from consumption welfare.

Provincial Capabilities, Agglomeration and Spatial Transmission

Who benefits from wider market access depends on existing capabilities. At the firm level, absorptive capacity depends on prior knowledge that enables organisations to use external information (Cohen & Levinthal, 1990). Competition also redistributes activity: Goldmanis et al. (2010) show that e-commerce reallocates market shares towards lower-cost producers. These mechanisms suggest that commercial expansion may benefit capable suppliers while exposing other local businesses to stronger competition. A readiness measure consequently needs to represent complementary capabilities rather than equating participation with internet access alone.

Digital infrastructure evidence reinforces this conditional interpretation. Forman et al. (2012) associate advanced internet investment with wage growth primarily in US counties already possessing favourable economic and educational characteristics. Akerman et al. (2015) find that broadband complements skilled workers but can worsen outcomes for unskilled workers in Norway. Conversely, Hjort and Poulsen (2019) document employment gains following fast internet arrival in Africa, including gains among less-educated workers. The contrast indicates that skills and productive opportunities shape distributional responses; connectivity alone does not imply a common regional outcome.

Regional evidence reflects this ambiguity. Yao and Qiu (2024) report that China's e-commerce demonstration-city policy narrows regional economic disparities, with heterogeneous responses across city characteristics. Conversely, S. Wu et al. (2022) find that internet development can widen Chinese intercity disparities through differences in penetration and city size. Y. Wu et al. (2025) document persistent geographic differences and heterogeneous convergence patterns in e-commerce development itself. These studies concern different interventions and outcomes: a targeted programme may supply complementary support, while convergence in commercial participation need not produce income convergence. Their findings warrant evaluating distributional responses rather than assuming an equalising effect.

Spatial relationships provide a further channel of transmission. Rey and Montouri (1999) show that ignoring spatially correlated disturbances can misspecify regional convergence regressions. In spatial models, indirect responses provide a basis for assessing spillovers that individual coefficients cannot supply (Elhorst, 2010; LeSage & Pace, 2009). Such responses may reflect complementary demand, knowledge diffusion or competitive displacement; aggregate provincial data do not separately identify these mechanisms.

Digital Infrastructure and Regional Income Convergence in Turkiye

Turkish evidence connects these issues directly to digital infrastructure. Celbis and de Crombrugghe (2018) find that internet infrastructure increases convergence speed across 26 level-2 regions of the Nomenclature of Territorial Units for Statistics (NUTS) during 1999–2011. Their result supports an equalising possibility, but concerns infrastructure rather than e-commerce readiness. At the provincial level, Ursavaş and Mendez (2023) identify six

convergence clubs and spatially structured income dynamics during 2007–2019. More recently, Ursavaş et al. (2026) document slow convergence alongside persistent spatial inequality and neighbour-conditioned income mobility during 2004–2023. These findings suggest that catch-up can occur within a durable geographic hierarchy.

The unresolved question is whether exposure to national e-commerce expansion has a more favourable relative income association in initially poorer provinces, after accounting for readiness and spatial relationships. The present study addresses this question across 81 provinces during 2019–2024. Its outcome is nominal GDP per person relative to the national value, so the analysis concerns provincial output positions rather than household purchasing power or directly measured consumer welfare. Combining readiness with initial income and decomposing spatial associations allows the analysis to distinguish differential exposure from general provincial catch-up. Predetermined measurement preserves temporal ordering without establishing exogeneity.

Hypotheses Development

Lower search and transaction costs can expand market access, while stronger competition can redistribute production away from less competitive suppliers (Goldfarb & Tucker, 2019; Goldmanis et al., 2010). These opposing mechanisms leave the income association empirically open. H1 states that differential e-commerce exposure is associated with provincial relative income, conditional on initial income and the included covariates. The controlled fixed-effects benchmark provides its primary assessment; spatial models examine the separate conditional associations and network comparisons.

Wider market access may favour poorer provinces with adequate capabilities. Yao and Qiu's (2024) regional-disparity

evidence supports this possibility, while Fan et al. (2018) examine consumer welfare; poverty also differs from remoteness. H2 states that the exposure association is more favourable in initially poorer provinces than in initially richer provinces. FE compares own-province slopes at fixed baseline gap cutoffs. The planned spatial source-total contrast complements this comparison through network responses generated by poorer and richer origins, without directly measuring gains received by those groups.

Market integration and competitive displacement may also transmit income associations across provincial boundaries. H3 states that neighbouring e-commerce exposure has a non-zero indirect association with provincial relative income. Following spatial impact decomposition, the relevant claim is a non-zero overall average indirect exposure association; spatial clustering alone cannot establish this relationship (Elhorst, 2010; LeSage & Pace, 2009). All three hypotheses concern conditional associations and do not presume that national expansion or provincial readiness is exogenous.

Data and Methodology

The analysis distinguishes descriptive provincial catch-up from the conditional exposure associations underlying H1–H3. The controlled FE benchmark assesses the exposure relationship; spatial models examine own-province and source-network associations. Both retain initial-income heterogeneity and predetermined provincial controls.

Data Sources and Sample Construction

The main estimation sample is a balanced panel of 81 Turkish provinces over 2019–2024, comprising 486 province-year observations. Provincial GDP per capita and its national counterpart are measured in current Turkish lira. Demographic, enterprise, education and transport series originate from the regional statistics

workbooks of the Turkish Statistical Institute (TURKSTAT). These sources provide the population, economic structure and capacity measures used in the analysis (TURKSTAT, n.d., 2025). The archived panel covers 2007–2024, with earlier observations supporting baseline construction and differential-trend diagnostics. Province names were normalised and matched to the supplied tur_admin1 layer's administrative codes. Its attributes provide fixed areas and centre coordinates (v01; recorded valid_on date: 1 January 2022; WGS84).

National e-commerce series and the separate 81-province adaptation index are taken from Türkiye'de E-Ticaretin Görünümü Raporu 2025, published by the Ministry of Trade (2026). The national series covers 2019–2025; observations through 2024 enter estimation, whereas the 2025 observation supplies descriptive context. The provincial adaptation index refers to 2025 and is reserved for measurement validation. It is not used to explain income outcomes ending in 2024. Raw source observations are preserved without interpolation, imputation or winsorisation, and national exposure measures remain missing before 2019. The main specifications use a common complete sample; sample exclusions are identified separately in the sensitivity analyses.

Variable Construction and Readiness Validation

The primary outcome, r_{it}, is the natural logarithm of provincial GDP per capita relative to the national value, stored as ln_rel_income:

$$r_{it} = \ln\left(Y_{it}/Y_t^{\mathrm{TR}}\right) \quad (1)$$

Here, Y_{it} denotes provincial GDP per capita and Y_t^{TR} its national counterpart. A positive value indicates an above-national output position. Common monetary rescaling cancels in the ratio, but province-specific price differences remain. The outcome therefore

measures relative nominal output per person rather than real household income or purchasing power (TURKSTAT, 2025). Supporting outcomes are annual changes in log relative income multiplied by 100 (rel_growth_pct), the positive log income shortfall $\max(-r_{it}, 0)$ (income_shortfall), and the absolute log gap $|r_{it}|$ (income_gap_abs). The absolute gap combines shortfalls and above-national excesses, so its interpretation differs from that of the positive shortfall.

Predetermined readiness R (ready_pre) combines 2016–2018 averages of information-and-communication enterprise shares, used as a proxy for information and communication technology (ICT), log population density, literacy among residents aged six and older, and light-commercial vehicles per 1,000 residents. Enterprise shares and literacy are percentages. Components are standardised across 81 provinces using sample standard deviations; their equal-weight mean is standardised again to mean zero and sample standard deviation one. These proxies capture enterprise structure, agglomeration, basic education and transport capacity, alongside general development; literacy does not measure digital skills, nor do vehicle stocks measure delivery networks. Transparent aggregation and component checks follow composite-indicator guidance (Organisation for Economic Co-operation and Development et al., 2008). Negative scores mean below-average capacity, not negative activity.

The initial income gap, g_i (gap_c_pre), is minus the 2016–2018 average of log relative income, centred at the unweighted mean across the 81 provinces. Larger values identify initially poorer provinces. National expansion, S_t (shock_gdp), equals the domestic e-commerce share of GDP, q_t (nat_ec_gdp_pct), minus its 2019 value of 2.7%, and is measured in percentage points (Ministry of Trade, 2026). The exposure block is constructed as:

$$E_{it} = R_i S_t, \quad G_{it} = g_i S_t, \quad H_{it} = R_i g_i S_t = E_{it} g_i \qquad (2)$$

These terms correspond to ec_exposure, shock_gap and ec_exp_gap, respectively. E_{it} is synthetic differential exposure, rather than observed provincial online sales. G_{it} allows initial income to interact with the national path independently of readiness, while H_{it} permits the exposure association to vary with initial income.

The predetermined controls are 2016–2018 averages of the age-dependency ratio, adult tertiary attainment and the manufacturing-enterprise share. Adult tertiary attainment (tertiary15_pct) is the percentage of all residents aged 15 and older whose highest qualification is an associate/bachelor's degree, master's degree or doctorate; unknown attainment remains in the denominator. Manufacturing is measured as a percentage of registered enterprises. Each baseline control is interacted with t − 2019, producing trend_age, trend_tertiary15 and trend_manufact. The centre-population share (center_share_pct) measures residents in provincial and district centres, an administrative settlement classification. Road density (road_map_density) measures state and provincial road kilometres per 1,000 km² of fixed boundary area. Both variables enter validation or sensitivity exercises; boundary area is distinguished from official land-area measures.

Measurement relevance uses the standardised 2025 adaptation index (z_adapt): provincial sales-to-purchases ratios, e-commerce/GDP, transaction volume per e-commerce enterprise and e-commerce enterprise shares (Ministry of Trade, 2026). Regressions include readiness alone, then standardised baseline relative income and centre-population share, and finally tertiary attainment and road density. Standardisation uses 81 provinces; Pearson and Spearman correlations complement inference using the HC3 heteroskedasticity-consistent covariance estimator

(MacKinnon & White, 1985). This later commercial criterion neither separates capacity from general development nor validates 2019–2024 exposure or exogeneity. ICT-only and leave-one-component-out measures remain sensitivity checks.

Descriptive Convergence and Spatial Structure

Descriptive convergence is assessed separately from exposure. Sigma is the population-weighted standard deviation of log relative income. Theil T uses population shares and GDP per capita normalised by the explicitly calculated population-weighted provincial mean, avoiding assumptions about its equivalence to the published national denominator. Measures cover 2019–2024, with 2013 and 2018 benchmarks. Beta regressions for 2013–2018 and 2019–2024 relate the five-year log relative-income change divided by five to starting log relative income; controlled versions add starting-year age dependency, tertiary attainment and manufacturing share. A negative starting-income coefficient indicates catch-up, without attributing it to e-commerce (Barro & Sala-i-Martin, 1992).

The fixed queen matrix W links polygons sharing a boundary or vertex, has a zero diagonal and is row-standardised. Geographic proximity approximates connectivity, not observed trade links. Its 81 × 81 structure has 199 undirected links, 4.914 neighbours on average, a one-to-nine range and no islands. The alternative connects the four nearest boundary-layer centre points by great-circle distance, symmetrises and row-standardises, yielding 188 links. Neither measures commercial networks; row-standardised weights can be asymmetric despite symmetric links.

Global Moran's I describes log relative-income levels and annual changes (Moran, 1950). Permutation inference reassigns values across fixed locations 9,999 times, compares absolute departures from $-1/(81-1)$, and applies a plus-one correction (Phipson & Smyth, 2010). It tests spatial exchangeability; the native

intercept-only asymptotic test is reported separately. Income clustering is distinct from H3's estimated indirect exposure association (Rey & Montouri, 1999).

Exposure Specification and Identification

The controlled non-spatial benchmark relates relative-income levels to the complete exposure block:

$$r_{it} = \alpha_i + \tau_t + \beta_E E_{it} + \beta_G G_{it} + \beta_H H_{it} + (t - 2019)\boldsymbol{c}_i'\boldsymbol{\gamma} + \varepsilon_{it} \quad (3)$$

Province effects, α_i, absorb time-invariant readiness, initial income, their product and other fixed provincial characteristics. Year effects, τ_t, absorb the common national expansion term and other shocks shared across provinces. The vector $\boldsymbol{c}_i$ contains the three predetermined controls. Its interaction with calendar time permits different secular income paths across provinces. Retaining E_{it}, G_{it} and H_{it} together preserves the time-varying lower-order terms of the readiness–gap–expansion interaction. A homogeneous-readiness specification omits H_{it}, followed by the full interaction specification, the controlled benchmark and a model replacing common year effects with NUTS-1 region-by-year effects.

The conditional exposure slope is $\beta_E + \beta_H g_i$. H1's primary assessment jointly tests the E and H coefficients in the controlled FE benchmark; the homogeneous and other FE specifications provide comparisons. H2 is assessed through the interaction and a formal contrast between slopes at the 75th and 25th percentiles of g_i. This cutoff contrast equals $\beta_H(g_{75} - g_{25})$. Because g_i is larger in poorer provinces, a positive contrast indicates a more favourable association there. Differences between the significance levels of separate slopes cannot establish a significant difference between them.

Identification comes from provinces with different fixed readiness and initial income following a common national e-commerce path. Predetermination establishes temporal ordering, but does not guarantee that readiness is unrelated to omitted differential income trends. Exposure-design research emphasises the substantive assumptions needed when baseline characteristics interact with common aggregate changes (Goldsmith-Pinkham et al., 2020). Here, inflation, pandemic disruptions and other national developments may coincide with the e-commerce trajectory, while affecting provinces differently. The estimates are therefore interpreted as conditional associations. The main equation excludes lagged income, avoiding the conventional short-panel fixed-effects bias associated with a lagged dependent variable (Nickell, 1981); descriptive beta regressions supply the separate catch-up assessment.

Spatial Estimation and Conditional Associations

The primary spatial specification is a spatial Durbin model (SDM), which accommodates an endogenous spatial outcome lag and spatially lagged covariates (Elhorst, 2010):

$$r_t = \rho W r_t + X_t\beta + WX_t\theta + \alpha + \tau_t 1 + \varepsilon_t \quad (4)$$

The vector r_t contains provincial log relative incomes, and X_t contains E, G, H and the three baseline covariate trends. Province and year effects are retained. The SDM lags each substantive covariate, constructing W(RgS) from provincial products rather than separately lagged components. Spatial autoregressive (SAR) and spatial error (SEM) models provide likelihood comparisons; the spatially lagged explanatory-variable model (SLX) includes X and WX without the endogenous outcome lag. Quasi-maximum likelihood (QML) estimation uses xsmle with fe type(both) and vce(cluster province_id); the SDM's explicit Durbin list contains the substantive covariates, excluding year and regional-year indicators.

No leeyu transformation is enabled. SLX uses the within estimator (Belotti et al., 2017).

Specification tests examine ρ = 0, the joint WX restrictions θ = 0 associated with the SAR reduction, and the common-factor restrictions θ = −ρβ associated with the SEM representation. AIC and BIC are compared only among likelihood models estimated on identical observations. The SDM remains the primary spatial specification rather than being selected according to significance. Numerical convergence and the position of ρ within the matrix-specific invertibility interval are checked before interpreting computed associations. With six years, finite-sample accuracy remains uncertain; numerical convergence and admissible spatial parameters do not establish unbiased estimation (Lee & Yu, 2010).

Spatial coefficients alone do not describe the response to local exposure because outcome feedback transmits changes through W (LeSage & Pace, 2009). Let D_g denote the diagonal matrix of observed baseline gaps. A conditional change in E changes H = Eg simultaneously, while G is held fixed. The corresponding derivative matrix is:

$$J_E = (I - \rho W)^{-1}[\beta_E I + \beta_H D_g + \theta_E W + \theta_H W D_g] \qquad (5)$$

Element (i, j) of J_E gives the response in destination i to unit exposure originating in source j. Direct associations average the diagonal, totals average source column sums, and indirect associations equal total minus direct, retaining observed gaps and spatial feedback. E combines standardised readiness with national expansion in percentage points; its unit represents neither a percentage increase in provincial sales nor a common national increase in S.

The planned spatial comparison uses gaps at or above the 75th percentile and at or below the 25th percentile, with 21 sources

each. Direct contrasts concern own-province responses; source totals concern responses generated across destinations, not gains received by poorer provinces. Observed gaps and feedback distinguish these averages from FE cutoff contrasts. H3 tests the overall indirect association; Figure 3 summarises source quartiles. Joint delta-method uncertainty covers local and spatial E/H coefficients and ρ; gradients are checked against central finite differences.

Inference, Validity Diagnostics and Sensitivity

Province-clustered covariance permits within-province serial dependence but assumes independence across provinces (Cameron & Miller, 2015). FE/SLX use cluster-based t or F references; QML coefficients and spatial associations use normal or chi-square references. Validation, beta convergence and long differences use HC3 covariance (MacKinnon & White, 1985). Intervals are 95%. Holm adjustment covers the poorer-minus-richer source-total contrast and H3's overall indirect association (Holm, 1979), with raw and adjusted two-sided p-values. Favourable source comparisons require positive contrasts. H1 is assessed separately in the controlled FE benchmark; adjustment does not cover a joint family of all three hypotheses.

Spatial heteroskedasticity and autocorrelation consistent (HAC) covariance comparisons for FE and SLX sum scores over six years within each province before applying a Bartlett distance kernel at 200 and 400 km (Conley, 1999). The prespecified cutoffs compare local and broader distance dependence, without representing estimated transmission thresholds; covariance eigenvalues are checked and invalid intervals withheld. Null-imposed, studentised wild-cluster tests of E and H enumerate all 4,096 Rademacher patterns across 12 NUTS-1 regions. Enumeration removes simulation error without ensuring exact inference with twelve clusters (MacKinnon et al., 2023). A single national trajectory also

lacks the independent sectoral shocks required for corresponding shift-share inference (Adão et al., 2019).

Unrestricted 2013–2024 profiles jointly interact readiness, initial gaps and their product with year, normalised to 2019. Six pre-2019 coefficients per block and all eighteen together are tested; profiles repeat with 2013–2015 baselines. These fixed-effects diagnostics differ from the controlled benchmark. Neither component non-rejection nor joint rejection identifies an e-commerce effect. Design rank, scaled condition numbers, leverage, serial residual correlation and Moran coefficients supplement assessment. Residual summaries concern full-dummy OLS, not SDM innovations; Moran coefficients receive no uncalibrated significance test.

Sensitivity checks use 2013–2015 baselines or the centred 2018 gap, ICT-only and leave-one-component-out readiness, alternative shifters, readiness, gap and product trends, and regional-year effects. Pandemic controls interact baseline manufacturing and tourism with 2020 and 2021 indicators. Shifters include log retail-transaction growth relative to 2019 and log dollar-volume growth. Connectivity uses symmetrised four-nearest-neighbour weights; samples exclude eleven earthquake provinces throughout 2019–2024 or Istanbul, Ankara and Izmir. Matrices are rebuilt and row-normalised, with nearest-neighbour fallback for islands. Exclusions retain primary gap cutoffs; alternative baselines use their own. Tables A5 and A6 specify the comparisons.

Supporting analyses use annual log relative-income changes multiplied by 100, absolute gaps and positive shortfall among provinces with below-zero 2016–2018 mean log relative income (66 provinces; 396 observations). The 2019–2024 long-difference model uses changes in E, G and H with baseline controls. Year exclusions assess dependence on individual years; fixed 2018 population

weights provide a further comparison. Table 6 summarises validity and sensitivity, with full results in Appendix Tables A1–A8. Sensitivity estimates retain their own units and unadjusted inference.

Empirical Results

The results separate provincial catch-up from the income association with e-commerce exposure. During 2019–2024, income dispersion declined, initially poorer provinces grew faster in relative income, and readiness was associated with later adaptation. Yet the controlled models give no stable evidence of a favourable exposure differential for poorer provinces, and neither primary SDM claim is established after Holm adjustment. Hypothesis assessment therefore rests on the prespecified exposure tests, poorer–richer contrasts and indirect associations, rather than descriptive convergence alone.

Descriptive Patterns and Readiness Validation

Table 1 summarises 486 province-year observations and assesses readiness's measurement relevance. Mean log relative income is −0.312, an unweighted provincial average. Readiness ranges from −2.118 to 3.830 standard deviations. Its standard deviation is one across 81 distinct provinces and slightly lower when fixed scores repeat over six years.

Table 1. Descriptive Statistics and Readiness Validation

Panel A1 Main Panel

Variable	N	Mean	SD	Minimum	Maximum
Log relative income	486	−0.312	0.339	−1.202	0.566
Annual log relative-income change × 100	486	0.306	6.978	−32.205	27.165
Positive income shortfall	486	0.346	0.285	0.000	1.202
Absolute income gap	486	0.380	0.260	0.004	1.202
National e-commerce share of GDP (%)	486	5.050	1.392	2.700	6.800
Predetermined readiness R	486	0.000	0.995	−2.118	3.830

Variable	N	Mean	SD	Minimum	Maximum
Centred initial income gap g	486	0.000	0.350	−0.824	0.883
Exposure E = S × R	486	0.000	2.716	−8.683	15.702
Gap exposure G = S × g	486	0.000	0.956	−3.377	3.620
Interaction H = S × R × g	486	−0.535	1.333	−12.872	2.198
Age dependency × time	486	124.606	88.682	0.000	388.767
Tertiary attainment × time	486	36.163	26.248	0.000	130.897
Manufacturing share × time	486	24.243	18.011	0.000	82.260

Panel A2 Distinct Provincial Baseline Values

Variable	N	Mean	SD	Minimum	Maximum
Predetermined readiness R	81	0.000	1.000	−2.118	3.830
Centred initial income gap g	81	0.000	0.352	−0.824	0.883
Baseline age dependency ratio	81	49.842	8.158	38.230	77.753
Baseline adult tertiary attainment (%)	81	14.465	2.921	8.569	26.179
Baseline manufacturing enterprise share (%)	81	9.697	2.337	4.848	16.452
Population in 2018	81	1,012,394	1,821,603	82,274	15,067,724

Panel B Validation Regressions For 2025

Regressor or statistic	(1) Bivariate	(2) Income and centre share	(3) Full validation controls
Readiness R	0.7071*** (0.1008)	0.5422*** (0.1553)	0.5777*** (0.1502)
Baseline log relative income	—	0.0386 (0.0922)	0.0159 (0.1377)
Baseline centre population share	—	0.2614*** (0.0912)	0.2331** (0.0976)
Baseline adult tertiary attainment	—	—	0.0293 (0.1247)
Baseline road density	—	—	−0.1441* (0.0784)
Observations	81	81	81
Adjusted R squared	0.494	0.530	0.539
Overall HC3 Wald F	49.23	22.86	15.01

Regressor or statistic	(1) Bivariate	(2) Income and centre share	(3) Full validation controls
Model p-value	<0.0001	<0.0001	<0.0001

Panel C Alternative Readiness Measures

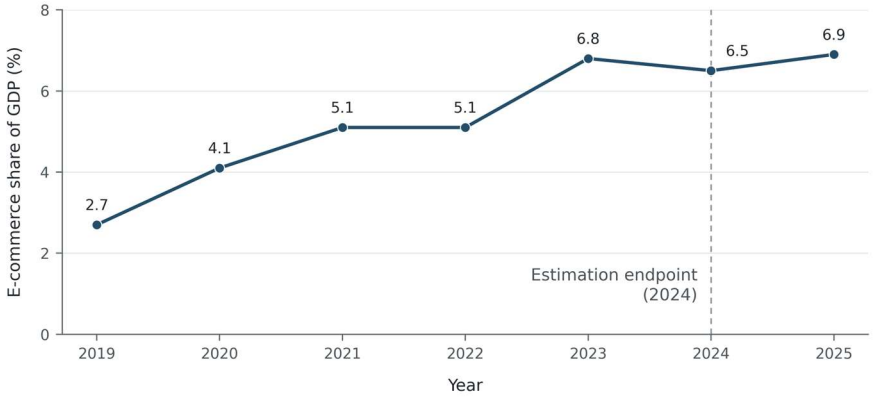
Readiness measure	Pearson r	p-value	Spearman rho	p-value	N
Composite readiness	0.707	<0.0001	0.638	<0.0001	81
ICT component only	0.319	0.0037	0.195	0.081	81
Excluding ICT	0.606	<0.0001	0.600	<0.0001	81
Excluding density	0.605	<0.0001	0.576	<0.0001	81
Excluding literacy	0.712	<0.0001	0.627	<0.0001	81
Excluding the vehicle component	0.700	<0.0001	0.600	<0.0001	81

Note. The panel covers 2019–2024; baselines use 2016–2018 averages unless indicated. Panel B uses the standardised 2025 adaptation index; all validation variables are standardised across 81 provinces. HC3 standard errors are in parentheses; intercepts are omitted. Correlation p-values are unadjusted. Throughout, * $p < 0.10$, ** $p < 0.05$ and *** $p < 0.01$.

Panels B and C assess readiness against the separate 2025 adaptation index. Its commercial components and shared provincial characteristics support a criterion comparison, not a validation of 2019–2024 provincial exposure or its exogeneity.

The bivariate standardised readiness coefficient is 0.7071; with full controls it remains 0.5777 (95% interval 0.278 to 0.877). A one-standard-deviation readiness difference is thus associated with approximately 0.58 standard deviations in later adaptation. The composite Pearson correlation is 0.707, compared with 0.319 for ICT alone, whose rank correlation is not statistically significant at the 5% level. The composite has a larger observed correlation than ICT alone, but no formal between-measure test is reported. Figure 1 presents the national e-commerce trajectory and the endpoint of the income estimation sample.

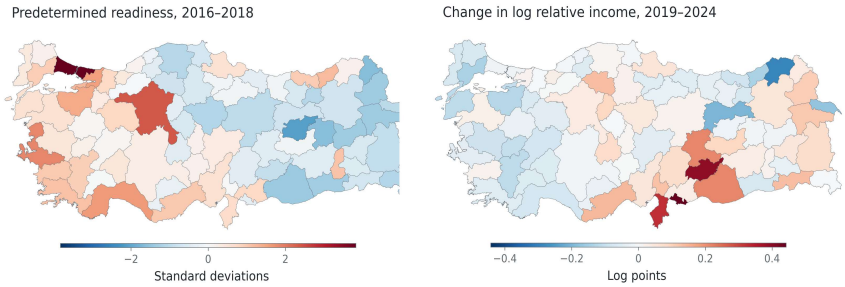
Figure 1. National E-Commerce Share of GDP in Turkiye, 2019–2025



Note. The dashed line marks the 2024 estimation endpoint; 2025 is descriptive only. Source: Ministry of Trade (2026).

The e-commerce share of GDP rose from 2.7% in 2019 to 6.5% in 2024, after reaching 6.8% in 2023 (Ministry of Trade, 2026). This expansion supplies the common national trajectory used to construct differential provincial exposure. Figure 2 compares baseline provincial readiness with changes in log relative income during 2019–2024.

Figure 2. Provincial E-Commerce Readiness and Changes in Relative Income in Turkiye



Note. Left: standardised 2016–2018 readiness. Right: the 2019–2024 change in the log ratio of provincial to national nominal GDP per capita. Separate scales are

centred at zero, in standard deviations (left) and log points (right). Sources: study dataset; tur_admin1 boundaries (v01; valid_on: 1 January 2022; WGS84).

Readiness is generally higher in western and prominent metropolitan provinces, while several eastern and southeastern provinces show relative-income gains. These patterns suggest that readiness and catch-up need not share a geographic ordering.

Income convergence and spatial clustering

Table 2 documents an uneven reduction in income dispersion alongside persistent spatial clustering. Population-weighted sigma falls from 0.4589 in 2019 to 0.4187 in 2024, while Theil T declines from 0.0900 to 0.0806. Both measures temporarily increase in 2021, so the endpoint improvement does not describe a monotonic convergence process. The two five-year regression windows assess whether lower starting income is associated with faster subsequent growth, following the descriptive beta-convergence approach of Barro and Sala-i-Martin (1992).

Table 2. Income Dispersion Convergence and Spatial Clustering

Panel A Annual Distribution and Spatial Dependence

Year	Sigma	Theil T	Income Moran I	Permutation p	Growth Moran I	Permutation p
2013	0.4529	0.0925	0.6742	0.0001	0.0405	0.442
2018	0.4685	0.0939	0.6934	0.0001	0.0565	0.337
2019	0.4589	0.0900	0.6657	0.0001	0.2310	0.0007
2020	0.4380	0.0843	0.6462	0.0001	0.1126	0.056
2021	0.4694	0.0936	0.6549	0.0001	0.4019	0.0001
2022	0.4605	0.0898	0.7151	0.0001	0.2232	0.0010
2023	0.4471	0.0890	0.7264	0.0001	0.0963	0.121
2024	0.4187	0.0806	0.6967	0.0001	0.3424	0.0001

Panel B Five-Year Growth Regressions

Regressor or statistic	2013–2018 Unconditional	2013–2018 Conditional	2019–2024 Unconditional	2019–2024 Conditional
Initial log relative income	0.0021 (0.0071)	0.0008 (0.0146)	−0.0379*** (0.0087)	−0.0591*** (0.0174)
Initial age dependency	—	−0.0011*** (0.0003)	—	0.0002 (0.0005)
Initial tertiary attainment (%)	—	−0.0028* (0.0015)	—	0.0031*** (0.0011)
Initial manufacturing share (%)	—	−0.0024** (0.0011)	—	0.0018* (0.0010)
Constant	0.0028 (0.0021)	0.1101*** (0.0319)	−0.0111*** (0.0029)	−0.0920*** (0.0312)
N	81	81	81	81
Adjusted R squared	−0.011	0.184	0.319	0.390

Note. N = 81 per year. Sigma is the population-weighted standard deviation of log relative income; Theil T also uses population weights. Moran's I uses log relative-income levels and annual changes with row-normalised queen weights. Two-sided permutation tests use 9,999 draws around −0.0125 (minimum p = 0.0001); native income-level asymptotic p-values are below 0.0001. Beta regressions use five-year log relative-income changes divided by five; conditional models add the covariates shown. HC3 standard errors are in parentheses.

The post-2019 regressions indicate catch-up: the initial-income coefficient is −0.0379 without controls and −0.0591 with the three starting-year covariates. The corresponding 2013–2018 estimates are close to zero and imprecise. Evidence within the later window is therefore clearer, although significance in one window and insignificance in another do not themselves establish a statistically significant difference between the coefficients. Combined with the dispersion measures, the estimates support descriptive convergence over 2019–2024 without assigning it to e-commerce or implying a structural convergence speed.

Income levels remain strongly clustered, with annual Moran coefficients between 0.6462 and 0.7264 during the main window. Every income-level permutation test reaches the reported simulation floor of 0.0001. Growth clustering varies more, with strong dependence in 2021 and 2024 and weaker evidence in 2020 and 2023. These patterns resemble the coexistence of declining overall inequality and persistent spatial dependence found for earlier Turkish data by Gezici and Hewings (2007). They also accord with Rey and Montouri's (1999) argument that regional convergence should be examined with attention to spatial interdependence. Spatial similarity nevertheless supplies no direct evidence for H3, which concerns the estimated indirect exposure association.

Differential exposure and initial-income heterogeneity

Having documented provincial catch-up, Table 3 assesses whether relative-income trajectories are associated with readiness-weighted exposure. Specifications retain the gap–expansion term, then add the interaction, covariate trends and NUTS-1-by-year effects. H2 uses explicit own-province slope differences at fixed gap cutoffs, rather than comparisons of separate significance levels.

Table 3. Non-Spatial Panel Associations

Regressor or statistic	(1) E and G	(2) E G and H	(3) Covariate trends	(4) Region by year
Exposure E = S × R	0.0088** (0.0038)	0.0073* (0.0043)	0.0043 (0.0046)	0.0040 (0.0039)
Gap exposure G = S × g	0.0440*** (0.0146)	0.0431*** (0.0144)	0.0341 (0.0224)	0.0684*** (0.0237)
Interaction H = S × R × g	—	−0.0064 (0.0051)	−0.0037 (0.0054)	−0.0143* (0.0080)
Age dependency × time	—	—	0.000668 (0.000498)	−0.001016 (0.000788)
Tertiary attainment × time	—	—	0.001132 (0.000936)	0.000138 (0.001124)
Manufacturing share × time	—	—	0.001396 (0.001333)	0.000057 (0.001199)

Regressor or statistic	(1) E and G	(2) E G and H	(3) Covariate trends	(4) Region by year
Province fixed effects	Yes	Yes	Yes	Yes
Common year effects	Yes	Yes	Yes	Replaced
NUTS-1 by year effects	No	No	No	Yes
Observations / provinces	486 / 81	486 / 81	486 / 81	486 / 81
Within R squared	0.129	0.132	0.143	0.413
Core block Wald F	4.575	4.011	3.066	4.088
Core block df / residual df	2 / 80	3 / 80	6 / 80	6 / 80
Core block p-value	0.013	0.010	0.0094	0.0013
E and H joint p-value	0.024	0.012	0.263	0.065
H p-value	—	0.212	0.496	0.078

Conditional Exposure Slopes

Conditional association	Column 2	Column 3	Column 4
Slope at g = Q25	0.0086** (0.0039)	0.0051 (0.0041)	0.0069* (0.0038)
Slope at g = Q75	0.0059 (0.0050)	0.0035 (0.0053)	0.0008 (0.0047)
Q75 slope minus Q25 slope	−0.0027 (0.0022)	−0.0016 (0.0023)	−0.0061* (0.0034)

Note. Outcome: log relative income. Province-clustered standard errors are in parentheses. Core Wald tests cover substantive regressors only; column 1's E/H test covers E alone. Slopes use Q25 = −0.201087 and Q75 = 0.222703; their contrast differs from Table 5's source-group averages. Column 4's full-model F is unavailable because of covariance-rank limitations.

The positive exposure coefficient in the homogeneous specification, 0.0088, weakens once heterogeneity and covariate trends are accommodated. In the controlled benchmark, the exposure coefficient is 0.0043 (p = 0.346), and the interaction coefficient is −0.0037 (p = 0.496). The joint exposure–interaction test has p = 0.263. Significance of the wider six-regressor block therefore cannot be interpreted as support for H1, because that block includes the gap–expansion term and covariate trends. The data provide some evidence of an exposure association in the least controlled

specification, but H1 receives no stable support from the prespecified controlled benchmark.

The controlled exposure slopes are 0.0051 at the richer cutoff and 0.0035 at the poorer cutoff. Their difference is −0.0016, with a 95% confidence interval of −0.0061 to 0.0030. The point estimate thus runs against H2's predicted favourable poorer-province differential, while the interval permits either sign. Regional year effects produce a more negative difference of −0.0061, but its p-value of 0.078 does not establish the difference at 5%. The appropriate conclusion is that the proposed favourable differential is unsupported, rather than that poorer provinces necessarily experience an adverse association.

Spatial specifications and model comparisons

Persistent income clustering motivates extending the FE benchmark to geographic interdependence. Table 4 compares the planned SDM with SLX, SAR and SEM. Restriction tests assess model compatibility; Table 5 reports conditional spatial associations rather than a full SDM test of H1.

Table 4. Spatial models and Specification Tests

Regressor or statistic	SLX	SAR	SDM	SEM
Exposure E = S × R	0.0016 (0.0037)	0.0050 (0.0040)	0.0023 (0.0033)	0.0045 (0.0039)
Gap exposure G = S × g	0.0333** (0.0145)	0.0381* (0.0217)	0.0421*** (0.0140)	0.0538** (0.0255)
Interaction H = S × R × g	−0.0061 (0.0045)	−0.0051 (0.0050)	−0.0069 (0.0044)	−0.0077 (0.0049)
Age dependency × time	0.000119 (0.000593)	0.000180 (0.000461)	0.000039 (0.000614)	0.000093 (0.000563)
Tertiary attainment × time	0.000083 (0.001187)	0.000688 (0.000842)	0.000118 (0.001115)	0.000674 (0.000957)
Manufacturing share × time	−0.000244 (0.001058)	0.000782 (0.001347)	−0.000727 (0.001024)	−0.000585 (0.001253)

Regressor or statistic	SLX	SAR	SDM	SEM
W × Exposure E = S × R	−0.0172* (0.0087)	—	−0.0155* (0.0081)	—
W × Gap exposure G = S × g	−0.0844*** (0.0209)	—	−0.0801*** (0.0214)	—
W × Interaction H = S × R × g	0.0092 (0.0169)	—	0.0088 (0.0163)	—
W × Age dependency × time	0.003757*** (0.001085)	—	0.003046*** (0.001114)	—
W × Tertiary attainment × time	0.005644** (0.002673)	—	0.005346** (0.002541)	—
W × Manufacturing share × time	0.007620*** (0.002295)	—	0.006733*** (0.002288)	—
Outcome spatial parameter rho	—	0.3570*** (0.0624)	0.3340*** (0.0573)	—
Error spatial parameter lambda	—	—	—	0.4361*** (0.0634)
Log pseudolikelihood	—	703.797	739.292	708.329
AIC	—	−1391.59	−1450.58	−1400.66
BIC	—	−1358.11	−1391.98	−1367.17
Within R squared for SLX	0.264	—	—	—

Spatial Specification Restrictions

Model	Restriction	Reference statistic	Value	p-value
SLX	All six WX coefficients = 0	F(6, 80)	7.260	<0.0001
SAR	rho = 0	Chi squared (1)	32.683	<0.0001
SDM	rho = 0	Chi squared (1)	33.975	<0.0001
SDM	SDM to SAR: all theta = 0	Chi squared (6)	31.483	<0.0001
SDM	SDM to SEM: theta = −rho beta	Chi squared (6)	22.821	0.0009

Note. All models use 486 observations, 81 provinces, province and year effects, and row-normalised queen weights. Standard errors are province-clustered. AIC/BIC compare SAR, SDM and SEM QML fits using coefficient-vector counts (BIC: N = 486). Stars test coefficients; dashes denote unavailable or inapplicable entries.

The SDM spatial outcome parameter is positive at 0.3340 (standard error 0.0573). Joint tests reject both the SAR reduction and the SEM common-factor restrictions, with $p < 0.0001$ and $p = 0.0009$, respectively. The SDM also has the lowest AIC and BIC among the three comparable QML fits. These results support retaining spatial covariate terms within the maintained model class. They are consistent with the case for using an unrestricted SDM to assess indirect associations discussed by Elhorst (2010).

The substantive exposure evidence remains less decisive than the spatial specification evidence. The SDM local exposure coefficient is 0.0023 and its interaction coefficient is −0.0069, neither statistically significant at the 5% level. The neighbouring exposure coefficient is negative, whereas the neighbouring interaction coefficient is positive and imprecise. Neither these coefficients nor the significant outcome parameter can independently establish H3. Outcome feedback combines local and neighbouring coefficients, and heterogeneity requires the interaction to change with exposure. Consequently, the hypothesis assessment relies on the conditional derivative calculations in Table 5 rather than interpreting spatial regression coefficients as marginal associations (LeSage & Pace, 2009).

Conditional direct and indirect exposure associations

Table 5 retains feedback and the exposure–gap interaction for unit local exposure changes. Each extreme group contains 21 sources. The source-total contrast and overall indirect association form the planned Holm-adjusted family (Holm, 1979). Source totals measure responses generated across destinations, not gains received within source groups.

Table 5. Conditional spatial exposure associations

Panel A Source group associations

Source group	Association	SDM	SAR	SLX
All provinces	Direct	0.0011 (0.0033)	0.0052 (0.0042)	0.0016 (0.0037)
	Indirect	−0.0210* (0.0109)	0.0026 (0.0021)	−0.0173** (0.0086)
	Total	−0.0199* (0.0115)	0.0078 (0.0062)	−0.0157* (0.0082)
Poorer source group	Direct	−0.0017 (0.0044)	0.0028 (0.0056)	−0.0011 (0.0048)
	Indirect	−0.0154 (0.0175)	0.0013 (0.0026)	−0.0119 (0.0134)
	Total	−0.0171 (0.0180)	0.0042 (0.0082)	−0.0130 (0.0126)
Richer source group	Direct	0.0037 (0.0030)	0.0074** (0.0037)	0.0042 (0.0033)
	Indirect	−0.0252*** (0.0085)	0.0038** (0.0018)	−0.0216*** (0.0065)
	Total	−0.0215** (0.0092)	0.0112** (0.0053)	−0.0174*** (0.0061)
Poorer minus richer	Direct	−0.0054 (0.0036)	−0.0046 (0.0045)	−0.0052 (0.0039)
	Indirect	0.0097 (0.0181)	−0.0025 (0.0021)	0.0097 (0.0135)
	Total	0.0044 (0.0183)	−0.0071 (0.0065)	0.0044 (0.0124)

Panel B Prespecified Primary SDM Claims

Claim	Estimate	95% interval	Raw p	Holm p
Overall indirect association	−0.0210	[−0.0424, 0.0004]	0.054	0.109
Poorer minus richer total association	0.0044	[−0.0315, 0.0402]	0.811	0.811

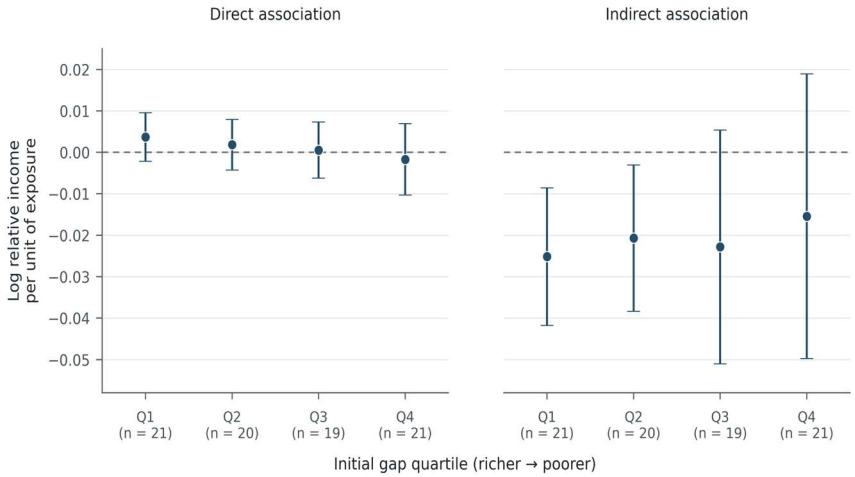
Note. Standard errors in parentheses use province-clustered covariance. Direct associations average derivative diagonals; totals average source column sums; indirect associations equal total minus direct. Source counts are 81 overall and 21 per extreme group. Group totals measure responses generated across all destinations by those sources. SDM/SAR use joint delta-method normal inference; SLX uses t inference with 80 degrees of freedom. Panel A stars are unadjusted;

Panel B reports the two planned Holm-adjusted SDM tests. Units: log relative income per unit of synthetic local exposure.

The SDM average direct association is 0.0011. Its indirect association is −0.0210, with a 95% interval of −0.0424 to 0.0004; the raw p-value of 0.054 becomes 0.109 after Holm adjustment. H3 is therefore not established under the planned rule. These model-conditional intervals use province-clustered covariance; the FE/SLX spatial-HAC comparisons do not validate SDM inference. Units remain log relative income per unit of synthetic local exposure, rather than a percentage increase in provincial sales.

The poorer-minus-richer source-total contrast is 0.0044, with a 95% interval of −0.0315 to 0.0402 and adjusted p = 0.811. It compares network responses generated by poorer and richer origins, not gains received within those groups. The own-province direct contrast is −0.0054, with a 95% interval spanning zero; neither comparison establishes a favourable differential. A significant indirect estimate for richer sources cannot establish a between-group difference. SAR gives a positive overall indirect estimate, whereas SLX gives a negative estimate with unadjusted p = 0.047, indicating model dependence. Figure 3 summarises direct and indirect associations across source-province gap quartiles.

Figure 3. Conditional Exposure Associations Across Initial Gap Quartiles



Note. Points show primary SDM conditional associations; bars show pointwise 95% joint delta-method intervals without multiplicity adjustment. Quartiles group source provinces by the centred initial gap, from richer (Q1) to poorer (Q4). Extreme groups match Table 5.

Own-province direct estimates decline from richer to poorer source quartiles, with all intervals including zero. Indirect source-network estimates are negative; pointwise intervals exclude zero in Q1 and Q2 but widen in Q3 and Q4. These intervals are not simultaneously adjusted and neither establish four separate spillover claims nor show which income groups receive larger gains.

Validity diagnostics and sensitivity analyses

Table 6 summarises design validity, inference sensitivity and specification sensitivity; Appendix Tables A1–A8 provide full results. Interpretation retains the distinction between the FE cutoff contrast and SDM source-group contrast.

Table 6. Summary of Validity Diagnostics and Sensitivity Analyses

Panel A. Validity Diagnostics

Check	Model / comparison	Reported result	p-value
Joint pre-period tests	2016–2018 / 2013–2015 baselines	F = 2.751 / 2.609 df = 18, 80	0.0010 / 0.0018
Within-design diagnostics	Reference / all three trends	Rank: 6/6 / 9/9 Condition: 4.710 / 18.249	—
Residual dependence	Full-dummy OLS	Lag 1 correlation: 0.2976 Moran I: 0.0100–0.3442	—

Panel B. Inference Sensitivity

Check	Model / comparison	Reported result	p-value
FE spatial HAC	E / H; 200 and 400 km	E = 0.0043 H = −0.0037	E: 0.367–0.378 H: 0.381–0.447
Regional wild-cluster	FE; E / H	E = 0.0043 H = −0.0037	0.535 / 0.685
SLX overall indirect	Province cluster / HAC 200 km / HAC 400 km	−0.0173 under all three	0.047 / 0.073 / 0.105

Panel C. Specification Sensitivity

Check	Model / comparison	Reported result	p-value
ICT-only poorer–richer contrast	FE cutoff / SDM source group	−0.0048 / −0.0329	0.015 / 0.0021
Other specification contrasts	15 FE / 14 SDM checks	No contrast significant at 5%	Minimum: 0.078 / 0.054
Supporting outcomes and long difference	FE; E and H	Neither E nor H significant at 5% in any of the four analyses	—
Leave-one-year-out	FE; H	Negative and imprecise in all six fits	—

Note. Paired entries follow the comparison-column order. HAC p-value ranges reflect 200 and 400 km cutoffs; wild-cluster tests use 12 NUTS-1 regions. Other specification checks exclude reference and ICT-only fits. Sensitivity p-values are

unadjusted; residual statistics are descriptive. FE/SLX covariance checks do not validate SDM inference. Tables A1–A8 provide full results.

Full within rank does not establish stable trajectories. Joint pre-period restrictions reject under both baselines; separate readiness and readiness–gap blocks do not reject at 5%. Table A8's unrestricted profiles differ from the controlled benchmark and constrain attribution without testing an e-commerce effect. Predetermination cannot exclude differential shocks (Goldsmith-Pinkham et al., 2020); OLS residual summaries motivate inference comparisons.

FE exposure and interaction estimates remain imprecise under both spatial-HAC cutoffs and regional wild-cluster inference. SLX's overall indirect estimate loses 5% significance under spatial HAC without changing its magnitude. This isolates covariance sensitivity (Conley, 1999), rather than demonstrating a zero association or supplying HAC inference for QML SDM. Tables A3 and A4 report the detailed comparisons.

ICT-only FE cutoff and SDM source-total contrasts are significantly negative (Tables A5 and A6), favouring neither poorer provinces nor poorer origins, respectively. Other contrasts remain imprecise. Shifter units prevent raw-magnitude comparisons. Regional-year effects yield a significant local SDM coefficient, an imprecise source-total contrast and smaller ρ; local significance does not establish a favourable network comparison.

Across the supporting outcomes and long-difference analysis, neither E nor H is significant at the 5% level; year exclusions leave H negative and imprecise (Table A7). Some gap–expansion estimates are significant, but this term contains no readiness and cannot establish H1. Outcome units and favourable directions also differ. Overall, neither observed catch-up nor these

sensitivity results establishes the exposure hypotheses; rejected pre-period tests and covariance sensitivity constrain interpretation.

Discussion of Empirical Results and Comparison with the Literature

The exposure estimates differ from Celbis and de Crombrugghe's (2018) Turkish infrastructure analysis and Yao and Qiu's (2024) Chinese policy evaluation in measurement, spatial scale and identification. Their findings therefore cannot be directly contradicted by the present imprecision. Digital cost reductions widen market access (Goldfarb & Tucker, 2019), while Fan et al. (2018) find larger consumer-welfare gains in smaller and remote cities. Cheaper consumption and greater variety need not raise nominal provincial GDP per person, so remain outside this output-based assessment.

Evidence from directly evaluated e-commerce expansion also cautions against equating connectivity with producer-income gains. Couture et al. (2021) find limited income gains for rural producers and workers in a randomised Chinese programme, with benefits concentrated in lower living costs for some households. The present estimates are compatible with the broader proposition that income responses are contingent, although their observational design cannot reproduce that causal conclusion. Liu et al. (2024) further distinguish income and consumption inequality, finding that China's income-to-consumption inequality ratio rises with e-commerce development. Together, these studies explain why a test based solely on provincial nominal output addresses one part of the distributional question and should not be treated as a comprehensive welfare assessment.

The measurement and sensitivity evidence is compatible with provincial capabilities shaping participation, without identifying their income returns. Forman et al. (2012) associate advanced

internet investment with wage growth mainly in already advantaged US counties. Here, the broad composite has a larger observed correlation with later adaptation than ICT alone, but the difference is not formally tested. Negative ICT-only FE cutoff and SDM source-total contrasts concern different comparisons. Neither contrast identifies a mechanism or supports the hypothesised favourable differential.

Declining dispersion alongside persistent clustering accords with Gezici and Hewings (2007), the convergence clubs of Ursavaş and Mendez (2023), and the slow catch-up and spatial hierarchy reported by Ursavaş et al. (2026). These findings concern income dynamics rather than the present exposure estimand. Rey and Montouri (1999) motivate accounting for spatial dependence, while Elhorst (2010) motivates computed indirect associations. Here, indirect estimates vary by model and inference choice. A significant outcome lag describes maintained dependence without identifying beneficial e-commerce transmission.

The controlled FE benchmark does not support H1. H2's favourable own-province differential is unsupported by the FE contrast and imprecise spatial direct comparison; source totals also fail to establish a favourable network response from poorer origins. H3 is not supported after Holm adjustment. These assessments concern uncertainty, not uniformly zero associations. Descriptive catch-up is clearer than the exposure evidence. Joint pre-period rejection, a short national trajectory and model-conditional spatial inference bound interpretation.

Conclusion

Provincial income convergence in Türkiye during 2019–2024 does not, by itself, establish that e-commerce expansion favoured poorer provinces. Across the balanced panel of 81

provinces, relative nominal GDP per person became less dispersed and initially poorer provinces experienced catch-up, while pronounced spatial clustering persisted. The study examined whether these changes were associated with differential exposure to national e-commerce expansion, using predetermined provincial readiness and initial income gaps. Its findings distinguish observed convergence from the more uncertain relationship between e-commerce exposure and regional economic performance.

H1 has limited support in the least controlled FE specification but is not supported by the controlled joint test. H2's favourable own-province differential is unsupported by the FE contrast and imprecise spatial direct comparison. The positive source-total contrast is also imprecise and measures network responses generated by poorer origins, not gains received within that group. These findings constrain evidence without establishing uniformly zero relationships.

The primary SDM's indirect estimate is negative and imprecise and does not establish H3 after Holm adjustment. Its intervals remain conditional on the model and province-clustered covariance; FE/SLX comparisons do not validate them. Alternative models give different signs, and covariance comparisons weaken some SLX significance. The results sustain no general claim of beneficial or harmful spillovers. A significant outcome coefficient describes maintained dependence rather than the indirect exposure association.

Taken together, the findings qualify the expectation developed in the study that expanded market access, combined with provincial capacity, would favour income catch-up in disadvantaged locations. The analysis does not establish that expected pattern. Readiness is associated with later adaptation, but this criterion comparison does not establish that capacity generated relative

income gains during the sample. The distinction between opportunities for e-commerce participation and observed provincial output therefore remains central to interpretation. The aggregate estimates do not identify the mechanisms through which those opportunities might affect regional performance.

The study integrates provincial catch-up, readiness-related exposure and geographic dependence. Complete interactions separate readiness-weighted exposure from expansion interacted with initial income. FE and spatial direct contrasts assess own-province differences; source totals concern responses generated across destinations. Computed indirect associations make H3 assessable beyond individual coefficients. The adaptation comparison supports measurement relevance, without validating provincial volumes or establishing exogenous exposure.

The policy implications concern the evaluation of regional outcomes. National e-commerce growth should not be treated as sufficient evidence of provincial income equalisation. Evaluation needs to examine relative provincial output, formally compare initially poorer and richer provinces, and account for associations extending beyond provincial borders. The readiness composite can characterise differences in provincial capacity, but the exposure estimates do not identify the income returns to its individual components. The available evidence consequently provides a basis for assessing regional distributional objectives while leaving the effectiveness of particular interventions unresolved.

Six years and a single national trajectory limit identification. Province and year effects absorb fixed differences and common shocks, not omitted provincial changes coinciding with expansion. Combined pre-period restrictions reject although separate readiness and readiness–gap blocks do not reject at 5%. These diagnostics constrain attribution without identifying an e-commerce effect.

Predetermination establishes timing rather than exogeneity; pandemic disruptions, inflation and broader provincial changes further complicate attribution.

Measurement imposes additional limits. The outcome captures relative output per person at current prices, so provincial price differences prevent interpreting convergence as a demonstrated improvement in real income or household welfare. Exposure combines a broad readiness index with national e-commerce intensity rather than observed provincial online transactions. Aggregate provincial data also leave the mechanisms connecting market access, production and consumption unresolved. These constraints narrow the study's conclusions to the measured associations.

Future research should extend the panel using consistent provincial series and observed local e-commerce activity. Compatible provincial chain-volume GDP, population and national output data would permit assessment of real relative output. Defensible sources of exogenous variation would be needed to advance causal interpretation, while explicit poorer–richer contrasts and spatial impacts should remain central to evaluation. These extensions would address the present identification and measurement limits and clarify whether e-commerce participation contributes to provincial catch-up.

References

Adao, R., Kolesar, M., & Morales, E. (2019). Shift-share designs: Theory and inference. *The Quarterly Journal of Economics, 134*(4), 1949–2010. https://doi.org/10.1093/qje/qjz025

Akerman, A., Gaarder, I., & Mogstad, M. (2015). The skill complementarity of broadband Internet. *The Quarterly Journal of Economics, 130*(4), 1781–1824. https://doi.org/10.1093/qje/qjv028

Barro, R. J., & Sala-i-Martin, X. (1992). Convergence. *Journal of Political Economy, 100*(2), 223–251. https://doi.org/10.1086/261816

Belotti, F., Hughes, G., & Piano Mortari, A. (2017). Spatial panel-data models using Stata. *The Stata Journal, 17*(1), 139–180. https://doi.org/10.1177/1536867X1701700109

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *The Journal of Human Resources, 50*(2), 317–372. https://doi.org/10.3368/jhr.50.2.317

Celbis, M. G., & de Crombrugghe, D. (2018). Internet infrastructure and regional convergence: Evidence from Turkey. *Papers in Regional Science, 97*(2), 387–409. https://doi.org/10.1111/pirs.12244

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly, 35*(1), 128–152. https://doi.org/10.2307/2393553

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics, 92*(1), 1–45. https://doi.org/10.1016/S0304-4076(98)00084-0

Couture, V., Faber, B., Gu, Y., & Liu, L. (2021). Connecting the countryside via e-commerce: Evidence from China. *American Economic Review: Insights, 3*(1), 35–50. https://doi.org/10.1257/aeri.20190382

Czernich, N., Falck, O., Kretschmer, T., & Woessmann, L. (2011). Broadband infrastructure and economic growth. *The Economic Journal, 121*(552), 505–532. https://doi.org/10.1111 /j.1468-0297.2011.02420.x

Elhorst, J. P. (2010). Applied spatial econometrics: Raising the bar. *Spatial Economic Analysis, 5*(1), 9–28. https://doi.org /10.1080/17421770903541772

Fan, J., Tang, L., Zhu, W., & Zou, B. (2018). The Alibaba effect: Spatial consumption inequality and the welfare gains from e-commerce. *Journal of International Economics, 114*, 203–220. https://doi.org/10.1016/j.jinteco.2018.07.002

Forman, C., Ghose, A., & Goldfarb, A. (2009). Competition between local and electronic markets: How the benefit of buying online depends on where you live. *Management Science, 55*(1), 47–57. https://doi.org/10.1287/mnsc.1080.0932

Forman, C., Goldfarb, A., & Greenstein, S. (2012). The Internet and local wages: A puzzle. *American Economic Review, 102*(1), 556–575. https://doi.org/10.1257/aer.102.1.556

Freund, C. L., & Weinhold, D. (2004). The effect of the Internet on international trade. *Journal of International Economics, 62*(1), 171–189. https://doi.org/10.1016/S0022-1996(03)00059-X

Gezici, F., & Hewings, G. J. D. (2007). Spatial analysis of regional inequalities in Turkey. *European Planning Studies, 15*(3), 383–403. https://doi.org/10.1080/09654310601017091

Goldfarb, A., & Tucker, C. (2019). Digital economics. *Journal of Economic Literature, 57*(1), 3–43. https://doi.org/ 10.1257/jel.20171452

Goldmanis, M., Hortaçsu, A., Syverson, C., & Emre, Ö. (2010). E-commerce and the market structure of retail industries. *The Economic Journal, 120*(545), 651–682. https://doi.org/ 10.1111/j.1468-0297.2009.02310.x

Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review, 110*(8), 2586–2624. https://doi.org/10.1257/aer.20181047

Hjort, J., & Poulsen, J. (2019). The arrival of fast Internet and employment in Africa. *American Economic Review, 109*(3), 1032–1079. https://doi.org/10.1257/aer.20161385

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65–70. https://www.jstor.org/stable/4615733

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy, 99*(3), 483–499. https://doi.org/10.1086/261763

Lee, L.-F., & Yu, J. (2010). Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics, 154*(2), 165–185. https://doi.org/10.1016/j.jeconom.2009.08.001

Lendle, A., Olarreaga, M., Schropp, S., & Vézina, P.-L. (2016). There goes gravity: eBay and the death of distance. *The Economic Journal, 126*(591), 406–441. https://doi.org/10.1111/ecoj.12286

LeSage, J., & Pace, R. K. (2009). *Introduction to spatial econometrics*. CRC Press. https://doi.org/10.1201/9781420064254

Liu, N., Qian, Y., Gu, X., & Li, G. (2024). Digital technology, e-commerce, and economic inequality: The case of China. *International Review of Economics & Finance, 91*, 259–271. https://doi.org/10.1016/j.iref.2024.01.016

MacKinnon, J. G., Nielsen, M. Ø., & Webb, M. D. (2023). Cluster-robust inference: A guide to empirical practice. *Journal of Econometrics, 232*(2), 272–299. https://doi.org/10.1016/j.jeconom.2022.04.001

MacKinnon, J. G., & White, H. (1985). Some heteros kedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics, 29*(3), 305–325. https://doi.org/10.1016/0304-4076(85)90158-7

TC. Ticaret Bakanlığı. (2026). *Türkiye'de e-ticaretin görünümü raporu 2025* [Report on the outlook of e-commerce in Türkiye 2025]. https://ticaret.gov.tr/data/6a02f2c7269 de183c0b98 bc4/T%C3%BCrkiye%27de%20E-Ticaretin%20G%C3%B6r%C3 %BCn %C3%BCm%C3%BC%20Raporu%202025.pdf

Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika, 37*(1–2), 17–23. https://doi.org/10.1093 /biomet/37.1-2.17

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica, 49*(6), 1417–1426. https://doi.org/10.2307 /1911408

Organisation for Economic Co-operation and Development, European Union, & European Commission, Joint Research Centre. (2008). *Handbook on constructing composite indicators: Methodology and user guide*. OECD Publishing. https://doi.org /10.1787/9789264043466-en

Phillips, P. C. B., & Sul, D. (2007). Transition modeling and econometric convergence tests. *Econometrica, 75*(6), 1771–1855. https://doi.org/10.1111/j.1468-0262.2007.00811.x

Phipson, B., & Smyth, G. K. (2010). Permutation p-values should never be zero: Calculating exact p-values when permutations are randomly drawn. *Statistical Applications in Genetics and Molecular Biology, 9*(1), Article 39. https://doi.org/10.2202/1544-6115.1585

Quah, D. T. (1996). Twin peaks: Growth and convergence in models of distribution dynamics. *The Economic Journal, 106*(437), 1045–1055. https://doi.org/10.2307/2235377

Rey, S. J., & Montouri, B. D. (1999). US regional income convergence: A spatial econometric perspective. *Regional Studies, 33*(2), 143–156. https://doi.org/10.1080/00343409950122945

TUIK, (2026). *Bölgesel istatistikler* [Regional statistics] [Database]. Retrieved October 11, 2026, from https://biruni.tuik.gov.tr/bolgeselistatistik/

TUIK. (2025, December 11). *Gross domestic product by provinces, 2024* (No. 53930) [Statistical bulletin]. https://veriportali.tuik.gov.tr/en/press/53930

Ursavaş, U., Chen, Y., & Rey, S. J. (2026). Spatial inequality and income dynamics across Turkish provinces. *The Annals of Regional Science, 75*, Article 105. https://doi.org/10.1007/s00168-026-01538-y

Ursavaş, U., & Mendez, C. (2023). Regional income convergence and conditioning factors in Turkey: Revisiting the role of spatial dependence and neighbor effects. *The Annals of Regional Science, 71*(2), 363–389. https://doi.org/10.1007/s00168-022-01168-0

Wu, S., Wang, P., & Sun, B. (2022). Can the Internet narrow regional economic disparities? *Regional Studies, 56*(2), 324–337. https://doi.org/10.1080/00343404.2021.1942444

Wu, Y., Wang, H., Guo, D., Li, Y., Zhang, H., & Shan, T. (2025). Spatio-temporal evolution and convergence patterns of e-commerce development in China: Regional disparities and policy implications. *PLOS ONE, 20*(8), Article e0330066. https://doi.org/10.1371/journal.pone.0330066

Yao, T., & Qiu, Z. (2024). Can e-commerce narrow regional economic disparities? Evidence from national e-commerce demonstration cities policy in China. *GeoJournal, 89*, Article 165. https://doi.org/10.1007/s10708-024-11155-x

Appendix A. Detailed validity diagnostics and sensitivity results

Tables A1–A8 detail the diagnostics and sensitivity analyses summarised in Table 6. Appendix p-values and stars are unadjusted (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$). Alternative outcomes and shifters retain their own units; raw coefficient magnitudes are not directly comparable.

Table A1. Joint Pre-Period Tests with 2019 as the Reference

Baseline years	Restriction	F statistic	df	p-value
2016–2018	Readiness × 2013–2018	1.479	6, 80	0.196
2016–2018	Gap × 2013–2018	2.314	6, 80	0.041
2016–2018	Readiness × gap × 2013–2018	0.355	6, 80	0.905
2016–2018	All three pre-period blocks	2.751	18, 80	0.0010
2013–2015	Readiness × 2013–2018	0.809	6, 80	0.566
2013–2015	Gap × 2013–2018	1.269	6, 80	0.281
2013–2015	Readiness × gap × 2013–2018	0.345	6, 80	0.911
2013–2015	All three pre-period blocks	2.609	18, 80	0.0018

Note. Profiles cover 2013–2024 (972 observations), with 2019 as the reference. Joint F tests use province-clustered covariance and 80 residual degrees of freedom. Table A8 reports annual coefficients.

Table A2. Design and Descriptive Residual Diagnostics

Design	**Within rank / columns**	**Scaled condition number**
Reference design	6 / 6	4.710
Design with all three differential trends	9 / 9	18.249
Residual diagnostic	**Value**	**N**
Pooled residual lag 1 correlation	0.2976	405
Maximum leverage from full-dummy OLS	0.4254	486
Residual Moran I in 2019	0.3384	81
Residual Moran I in 2020	0.2434	81
Residual Moran I in 2021	0.1976	81
Residual Moran I in 2022	0.3442	81
Residual Moran I in 2023	0.1652	81
Residual Moran I in 2024	0.0100	81

Note. Condition numbers use scaled within regressors. Full-dummy OLS residual statistics are descriptive; no residual Moran p-values are reported.

Table A3. Coefficient Inference Under Alternative Covariance Estimators

Model	Inference	Target	Estimate	SE	p-value	95% interval
FE	HAC 200 km	E	0.0043	0.0049	0.378	[−0.0053, 0.0140]
FE	HAC 200 km	H	−0.0037	0.0048	0.447	[−0.0132, 0.0058]
FE	HAC 400 km	E	0.0043	0.0048	0.367	[−0.0051, 0.0137]
FE	HAC 400 km	H	−0.0037	0.0042	0.381	[−0.0119, 0.0046]
FE	Wild NUTS-1	E	0.0043	—	0.535	—
FE	Wild NUTS-1	H	−0.0037	—	0.685	—
SLX	HAC 200 km	E	0.0016	0.0035	0.651	[−0.0053, 0.0085]
SLX	HAC 200 km	H	−0.0061	0.0046	0.191	[−0.0151, 0.0030]
SLX	HAC 200 km	W E	−0.0172	0.0098	0.078	[−0.0363, 0.0019]
SLX	HAC 200 km	W H	0.0092	0.0166	0.581	[−0.0234, 0.0418]
LX	HAC 400 km	E	0.0016	0.0035	0.650	[−0.0053, 0.0085]
SLX	HAC 400 km	H	−0.0061	0.0045	0.178	[−0.0149, 0.0028]
SLX	HAC 400 km	W E	−0.0172	0.0108	0.110	[−0.0383, 0.0039]
SLX	HAC 400 km	W H	0.0092	0.0174	0.598	[−0.0249, 0.0433]

Note. Bartlett spatial HAC aggregates province scores over six years; intervals use normal inference. All four HAC covariance matrices have positive minimum eigenvalues. Null-imposed studentised wild tests enumerate all 4,096 Rademacher patterns across 12 NUTS-1 regions and report p-values only.

Table A4. SLX Conditional Associations Under Alternative Covariance Estimators

Inference	Association	Estimate	SE	p-value	95% interval
Province cluster	Overall indirect	−0.0173	0.0086	0.047	[−0.0344, −0.0002]
Province cluster	Poorer minus richer total	0.0044	0.0124	0.720	[−0.0202, 0.0290]
HAC 200 km	Overall indirect	−0.0173	0.0097	0.073	[−0.0362, 0.0016]
HAC 200 km	Poorer minus richer total	0.0044	0.0114	0.696	[−0.0178, 0.0267]
HAC 400 km	Overall indirect	−0.0173	0.0107	0.105	[−0.0382, 0.0036]
HAC 400 km	Poorer minus richer total	0.0044	0.0115	0.698	[−0.0180, 0.0269]

Note. Bartlett spatial HAC follows Table A3 and uses normal-reference intervals. Covariance changes leave point estimates and the definitions of conditional associations unchanged.

Table A5. Non-Spatial Specification Comparisons

Specification	N	E coefficient	H coefficient	Q75 minus Q25 slope	Contrast p
Reference	486	0.0043 (0.0046)	−0.0037 (0.0054)	−0.0016 (0.0023)	0.496
Readiness gap and product trends	486	−0.0124* (0.0063)	−0.0013 (0.0063)	−0.0005 (0.0027)	0.837
2013–2015 baselines	486	0.0022 (0.0040)	−0.0067 (0.0054)	−0.0027 (0.0022)	0.212
Initial gap from 2018	486	0.0044 (0.0045)	−0.0049 (0.0054)	−0.0022 (0.0024)	0.371
ICT component only	486	−0.0020 (0.0026)	−0.0114** (0.0046)	−0.0048** (0.0020)	0.015
Readiness excluding ICT	486	0.0069 (0.0050)	0.0014 (0.0070)	0.0006 (0.0030)	0.845
Readiness excluding density	486	−0.0004 (0.0048)	−0.0062 (0.0050)	−0.0026 (0.0021)	0.222
Readiness excluding literacy	486	−0.0013 (0.0034)	−0.0077 (0.0050)	−0.0033 (0.0021)	0.124
Readiness excluding the vehicle component	486	0.0077* (0.0045)	−0.0014 (0.0059)	−0.0006 (0.0025)	0.813
Retail transaction shifter	486	0.0186 (0.0118)	−0.0016 (0.0143)	−0.0007 (0.0060)	0.912
USD volume shifter	486	0.0108 (0.0145)	−0.0236 (0.0179)	−0.0100 (0.0076)	0.191
NUTS-1 by year effects	486	0.0040 (0.0039)	−0.0143* (0.0080)	−0.0061* (0.0034)	0.078
Pandemic exposure controls	486	0.0032 (0.0046)	−0.0038 (0.0054)	−0.0016 (0.0023)	0.480
Fixed 2018 population weights	486	0.0094 (0.0067)	0.0020 (0.0045)	0.0009 (0.0019)	0.651
Symmetrised four nearest neighbours	486	0.0043 (0.0046)	−0.0037 (0.0054)	−0.0016 (0.0023)	0.496
Exclude 11 earthquake provinces	420	0.0021 (0.0036)	−0.0035 (0.0047)	−0.0015 (0.0020)	0.462
Exclude three major provinces	468	0.0044 (0.0046)	−0.0045 (0.0097)	−0.0019 (0.0041)	0.641

Note. Outcome: log relative income. Province-clustered standard errors are in parentheses. Models retain E, G, H and baseline covariate trends unless indicated; population weights are fixed analytic weights. Contrasts compare Q75 and Q25 cutoff slopes. Changing only the spatial matrix leaves FE estimates unchanged.

Table A6. Spatial Durbin Specification Comparisons

Specification	N	Local E	Local H	Poorer minus richer total	Contrast p
Reference	486	0.0023 (0.0033)	−0.0069 (0.0044)	0.0044 (0.0183)	0.811
Readiness gap and product trends	486	−0.0116 * (0.0068)	0.0066 (0.0076)	−0.0231 (0.0197)	0.242
2013–2015 baselines	486	−0.0002 (0.0034)	−0.0112** (0.0046)	0.0034 (0.0171)	0.842
Initial gap from 2018	486	0.0022 (0.0033)	−0.0078* (0.0045)	0.0052 (0.0187)	0.780
ICT component only	486	0.0059** (0.0025)	−0.0070** (0.0032)	−0.0329*** (0.0107)	0.0021
Readiness excluding ICT	486	0.0001 (0.0044)	−0.0116* (0.0068)	0.0285* (0.0148)	0.054
Readiness excluding density	486	−0.0001 (0.0039)	−0.0043 (0.0051)	0.0100 (0.0159)	0.531
Readiness excluding literacy	486	0.0012 (0.0026)	−0.0073** (0.0036)	−0.0119 (0.0199)	0.548
Readiness excluding the vehicle component	486	0.0064** (0.0032)	−0.0055 (0.0040)	−0.0065 (0.0189)	0.732
Retail transaction shifter	486	0.0083 (0.0084)	−0.0082 (0.0112)	0.0587 (0.0476)	0.218
USD volume shifter	486	0.0093 (0.0114)	−0.0309** (0.0154)	−0.0178 (0.0687)	0.795
NUTS-1 by year effects	486	0.0067** (0.0033)	−0.0111* (0.0065)	−0.0017 (0.0155)	0.914

Specification	N	Local E	Local H	Poorer minus richer total	Contrast p
Pandemic exposure controls	486	0.0008 (0.0035)	−0.0056 (0.0044)	0.0057 (0.0178)	0.747
Symmetrised four nearest neighbours	486	−0.0011 (0.0036)	−0.0109** (0.0044)	−0.0011 (0.0155)	0.941
Exclude 11 earthquake provinces	420	0.0025 (0.0036)	−0.0044 (0.0045)	−0.0081 (0.0198)	0.682
Exclude three major provinces	468	0.0034 (0.0037)	−0.0101 (0.0080)	0.0328 (0.0220)	0.135

Note. Province-clustered standard errors are in parentheses; source-total contrasts use joint delta-method normal inference. Queen weights apply unless indicated. Contrasts use 42 sources, or 35 and 39 after earthquake and major-province exclusions, respectively. Alternative baselines and the 2018 gap use their own cutoffs; exclusions retain primary cutoffs. Additional controls are held fixed in the exposure derivative. With regional-year effects, $\rho = 0.024$ ($p = 0.726$).

Table A7. Supporting Outcomes, Long Differences and Year Exclusions

Specification or outcome	N	E coefficient	G coefficient	H coefficient	Q75 minus Q25 slope
Annual log relative-income change × 100	486	0.3247 (0.3111)	2.2561** (0.9805)	−0.5574 (0.3531)	−0.2362 (0.1496)
Positive shortfall in provinces initially below the national benchmark	396	−0.0013 (0.0047)	−0.0618** (0.0291)	−0.0174 (0.0164)	−0.0074 (0.0070)
Absolute income gap	486	−0.0056 (0.0046)	−0.0253 (0.0237)	0.0031 (0.0055)	0.0013 (0.0023)

Specification or outcome	N	E coefficient	G coefficient	H coefficient	Q75 minus Q25 slope
2019–2024 long difference	81	−0.0014 (0.0055)	0.0758*** (0.0270)	−0.0087 (0.0067)	−0.0037 (0.0028)
Omit 2019	405	0.0125* (0.0069)	0.0683*** (0.0238)	−0.0090 (0.0076)	−0.0038 (0.0032)
Omit 2020	405	0.0018 (0.0043)	0.0274 (0.0215)	−0.0040 (0.0051)	−0.0017 (0.0021)
Omit 2021	405	0.0034 (0.0045)	0.0433* (0.0225)	−0.0052 (0.0055)	−0.0022 (0.0023)
Omit 2022	405	0.0048 (0.0047)	0.0300 (0.0237)	−0.0029 (0.0055)	−0.0012 (0.0023)
Omit 2023	405	0.0040 (0.0052)	0.0414* (0.0234)	−0.0015 (0.0060)	−0.0006 (0.0025)
Omit 2024	405	0.0027 (0.0041)	0.0193 (0.0237)	−0.0032 (0.0054)	−0.0013 (0.0023)

Note. Shortfall models include 66 provinces with negative 2016–2018 mean log relative income. Higher growth and lower shortfalls or absolute gaps are the favourable directions. Long differences use HC3 standard errors; other rows use province-clustered errors. Contrasts compare Q75 and Q25 cutoff slopes.

Table A8. Unrestricted Annual Interaction Profiles

Year	R 2016–2018	R 2013–2015	g 2016–2018	g 2013–2015	R × g 2016–2018	R × g 2013–2015
2013	0.0431** (0.0164)	0.0262* (0.0155)	0.1022* (0.0524)	0.0300 (0.0512)	0.0073 (0.0258)	0.0148 (0.0260)
2014	0.0354** (0.0145)	0.0211 (0.0144)	0.0585 (0.0418)	−0.0006 (0.0431)	0.0132 (0.0221)	0.0191 (0.0215)
2015	0.0308** (0.0150)	0.0179 (0.0145)	0.0472 (0.0403)	0.0010 (0.0425)	0.0179 (0.0212)	0.0196 (0.0206)

Year	R 2016–2018	R 2013–2015	g 2016–2018	g 2013–2015	R × g 2016–2018	R × g 2013–2015
2016	0.0101 (0.0113)	0.0049 (0.0106)	−0.0079 (0.0259)	−0.0225 (0.0295)	0.0129 (0.0156)	0.0151 (0.0162)
2017	0.0014 (0.0094)	−0.0008 (0.0087)	−0.0199 (0.0245)	−0.0208 (0.0249)	0.0132 (0.0149)	0.0128 (0.0147)
2018	0.0025 (0.0078)	0.0015 (0.0078)	−0.0417 * (0.0216)	−0.0420* (0.0220)	0.0028 (0.0100)	0.0028 (0.0095)
2019	Fixed at 0	Fixed at 0	Fixed at 0	Fixed at 0	Fixed at 0	Fixed at 0
2020	−0.0306** (0.0089)	−0.0301*** (0.0077)	−0.0106 (0.0180)	−0.0094 (0.0168)	−0.0082 (0.0101)	−0.0018 (0.0129)
2021	0.0015 (0.0152)	−0.0022 (0.0145)	−0.0505 (0.0415)	−0.0625 (0.0395)	0.0156 (0.0178)	0.0200 (0.0177)
2022	0.0210 (0.0167)	0.0144 (0.0149)	0.0102 (0.0560)	−0.0130 (0.0468)	0.0158 (0.0201)	0.0175 (0.0192)
2023	0.0150 (0.0163)	0.0079 (0.0138)	0.1018* (0.0568)	0.0812* (0.0476)	−0.0248 (0.0185)	−0.0306 (0.0187)
2024	0.0197 (0.0179)	0.0151 (0.0159)	0.2164** (0.0559)	0.2012*** (0.0525)	−0.0277 (0.0223)	−0.0348 (0.0218)

Note. R denotes readiness and g the centred initial gap; headers identify baseline windows. Models include province and year effects and all three annual interaction profiles. Province-clustered standard errors are in parentheses. The 2019 zeros are normalisations; pointwise stars do not replace Table A1's joint pre-period tests.


BİDGE