Berat Kara, PhD

Assistant Professor
Istanbul Medeniyet University, Türkiye
Faculty of Political Science
E-mail: berat.kara@medeniyet.edu.tr
Orcid: https://orcid.org/0000-0002-6948-2197

WHY DO MEDIUM-TERM EXPENDITURE FORECASTS FAIL? NEW EVIDENCE FROM PANEL DATA

UDC / UDK: 336.1:330.101.541]:519.86
JEL classification / JEL klasifikacija: H61, C23, E62, E27
https://doi.org/10.17818/EMIP/2026/21
Original scientific paper / Izvorni znanstveni rad
Received / Primljeno: March 25, 2026 /25. ožujka 2026.
Accepted / Prihvaćeno: April 20, 2026 / 20. travnja 2026.

Abstract

This study examines the macroeconomic determinants of medium-term expenditure forecast errors in advanced economies from 2005 to 2024, focusing on two- and three-year horizons. Using panel data methods, it analyses the roles of unemployment, economic growth, inflation, the current account balance, public debt, fiscal rules, and major crisis periods. The results show that unemployment is the most consistent determinant of forecast errors across all models and horizons, highlighting the importance of labour market conditions for fiscal planning. Inflation is significant only at longer horizons, indicating that its effects accumulate over time. GDP growth has limited and model-dependent relevance, while public debt and fiscal rules do not show a systematic effect on forecast accuracy. Major crises are associated with substantial increases in forecast errors and account for much of the models' explanatory power.

Keywords: medium-term budgeting, budget forecasting, medium-term forecasting, expenditure forecasting, budget accuracy

1. INTRODUCTION

The budget functions as the official legal document that controls all national operations while it serves as the primary instrument for overseeing government financial activities. The budget process serves as an essential part of policy development because all public policies require financial support which originates from budgetary resources (Filc & Scartascini, 2010). The budgetary decision process establishes public service quality which directly impacts both

economic stability and social welfare results (Schick, 1998; Atiyas & Sayın, 1997). When budget preparation focuses exclusively on one fiscal year the process creates a fundamental clash between annual budget needs and the continuous execution of public policies which extend beyond one year.

The process of annual budgeting results in significant operational deficiencies. Public investments and social reforms generate effects that stretch well beyond a single fiscal year, yet departments disregard both procurement timelines and future operating costs which lead to sustainability risks (Filc & Scartascini, 2010). The need to achieve short-term goals block organizations from executing their long-term strategies while they find it difficult to shift their resource allocation (Vlaicu et al., 2014). The practice of annual budgeting results in myopic and conservative budget management according to Wildavsky and Zapico-Goñi (1993). Because it hides future budget obligations and demands assessment of budget programs through their separate components.

The development of a mid-term approach to planning emerged as a solution to these restrictions which several nations faced. The system of Medium-term budgeting (MTB) establishes a connection between annual budgets and multi-year revenue and expenditure forecasts while using strategic priorities to determine resource distribution for more than one fiscal year. The concept originated from development planning work conducted during the 1950s (Schiavo-Campo, 2009) while its first practical use began in OECD countries during the early 1970s (Allen & Tommasi, 2001). International organizations promoted MTB as an essential practice for fiscal management which developed countries adopted from the 1990s onward (OECD, 2023; Dorotinsky & Watkins, 2013).

The implementation of a medium-term perspective enables organizations to maintain fiscal discipline through budget allocation according to their available resources for upcoming years (Kumar & Ter-Minassian, 2007; Vlaicu et al., 2014). The system enables organizations to reorganize their work while establishing processes for detecting financial difficulties at an early stage (Harris et al., 2013; Holmes & Evans, 2003). The system improves organizational efficiency through its requirement for organizations to evaluate their spending priorities in an organized manner (Boex et al., 2000). The government uses multi-year forecasts to determine service costs and evaluate policy implications (Schroeder, 2007; Jena, 2006). The planning of MTB functions as an uncertain environment management solution which prevents unnecessary expenditures that would negatively impact economic development and social welfare (Muzychenko et al., 2017).

Political stability and accurate macroeconomic forecasts and institutional strength and technical system efficiency constitute the essential requirements needed to achieve these advantages. Developing countries experienced failed MTB reforms because they lack these essential requirements. Schiavo-Campo (2009) shows that international institutions bring reforms to developing countries without understanding their local capabilities which results in expensive but ineffective budget systems. Filc and Scartascini (2010) maintain that MTBs should match

specific needs of each country which require appropriate implementation conditions. The essential components of the system require political commitment, accurate forecasting, restrictions on non-budget funds, implementation of international accounting standards, cost-benefit assessment and operational audits.

The quality of medium-term forecasts (MTF) matters for fiscal stability because it has a significant impact on financial systems which depend on future revenue forecasts. The combination of optimistic growth forecasts with revenue forecasts leads to uncontrolled deficits which endanger debt sustainability, while the combination of optimistic growth forecasts with revenue forecasts leads to controlled deficits which endanger debt sustainability, and the combination of conservative forecasts produces funding shortages for essential public services (Uralovich, 2025). Forecast errors grow larger as the horizon lengthens (Plesko, 1988; Penner, 2001; Frankel, 2011). Understanding why medium-term expenditure forecasts go wrong is therefore essential for sound fiscal planning. Yet the existing studies mainly concentrate on revenue forecasts and institutional design except for macroeconomic factors which affect expenditure forecast errors because these aspects remain unstudied.

The research investigates how macroeconomic variables impact medium-term expenditure forecast (MTEF) errors together with their assessment in thirteen advanced economies during the period from 2005 to 2024. The study aims to determine which macroeconomic factors that include unemployment, GDP, inflation, current account balance, public debt and migration patterns, systemically influence forecasting accuracy while examining the impact of fiscal regulations on forecasting results.

The analysis further investigates whether the effects of these determinants vary across two-year and three-year forecast horizons. The study of advanced economies enables to study macroeconomic effects which occur during times of institutional stability because this period allows them to separate economic changes from institutional deficiencies which complicate research on developing nations.

The study uses Driscoll-Kraay standard errors to handle cross-sectional dependence and autocorrelation while applying sequential block-wise modeling to test how various macroeconomic and institutional factors contribute to explanatory power. In doing so, the paper contributes to the literature by providing the first systematic, multi-country panel evidence on the macroeconomic drivers of medium-term expenditure forecast errors, shifting the focus from revenue forecasts and institutional design to the expenditure side of the budget and the distinct role of labor market dynamics and crisis-period disruptions.

The results show that the unemployment rate serves as the strongest factor that determines economic conditions. Changes to the current account system have a two-year period of impact which ends after that time. The prediction becomes accurate only when the forecasting period extends beyond the current time frame. The research demonstrates that public debt changes and net migration patterns do not create any impact while fiscal regulations fail to enhance forecasting precision.

The 2008 financial crisis and the COVID-19 pandemic together account for approximately twenty-five percent of the total forecasting errors. The results demonstrate that organizations need to improve their labor market modeling methods which should work together with experimental crisis response systems and their need to monitor inflation behavior over extended periods.

2. RELATED STUDIES

The research about MTB has developed through two separate yet interconnected research pathways. The first stream investigates fiscal forecast accuracy and bias by studying how macroeconomic assumptions and forecast horizons affect forecasting errors. Plesko (1988) established through his pioneering research that forecast errors grow larger with extended forecast horizons while budget balance forecasts for lengthier periods show consistent optimistic forecasting errors caused by unpredicted policy changes. Frankel (2011) later validated the finding that longer forecast horizons decrease accuracy rates. Breuer (2014) studied revenue forecasting to prove that analysts make revenue forecasting errors because they believe economic growth will occur at higher rates than actual growth.

The research by Kara (2024) and by Kara and Sarıoğlu (2026) developed new research methods for studying medium-term time periods which showed that inflation and economic growth impact forecasting precision. The research findings demonstrate that macroeconomic assumption quality serves as the main factor which determines forecasting accuracy.

The second stream of research investigates forecast results by examining the foundational political and institutional elements that support MTB. The effectiveness of MTB programs depends on three essential factors which include institutional capacity, budget process integration and permanent political support according to case studies and comparative research from Boex et al. (2000), Houerou and Taliercio (2002) and Holmes and Evans (2003). Developing countries experience failure with externally imposed reforms because their administrative systems lack strength according to Schiavo-Campo (2009) and Filc and Scartascini (2010).

Advanced economies exhibit forecasting patterns that respond to political incentives which determine their forecasting methods. Heinemann (2006) shows that electoral considerations systematically bias fiscal forecasts. The evidence on MTB's impact on spending outcomes remains mixed. The research of Vlaicu et al. (2014) shows that countries with advanced MTB frameworks achieve better fiscal results while others believe that implementation practice holds greater importance than design specifications.

Harris et al. (2013) demonstrate that binding expenditure ceilings show two effects because they improve fiscal discipline and they enhance forecast accuracy. Francesco and Barroso (2015) warn that most systems depend on basic

costing methods which lack essential elements. Raudla et al. (2020) further document that even formally adopted spending frameworks are frequently circumvented under political and economic pressure.

Recent contributions increasingly focus on forecast accuracy within MTB systems. The studies of Kara (2024) and Allen et al. (2017) established that major forecasting errors happen because of macroeconomic fluctuations and institutional deficiencies, while Radu (2023) demonstrates that European systems suffer from ongoing operational inconsistencies and research deficiencies

The existing research needs advanced methods to address its current challenges because Uralovich (2025) demands that researchers apply scenario-based forecasting methods to analyze macroeconomic transmission systems and Kara and Sarıoğlu (2026) demonstrate that macroeconomic factors including inflation and growth determine the accuracy of medium-term forecasts.

The literature demonstrates that forecast accuracy in MTB systems depends on two factors which are macroeconomic uncertainty and institutional quality. The current body of research focuses mainly on revenue forecasts and the forecasting practices of developing countries. There are few empirical studies that examine how macroeconomic factors impact MTEF errors in advanced economies and across various forecasting periods. The existing gap requires common methods of systematic comparison which can determine how economic and institutional elements affect medium-term budgeting frameworks.

The available research about MTB has progressed since previous studies because current empirical literature shows more advanced research findings. The findings show institutional differences which lead to different forecasting errors, but the research lacks a unified theoretical framework to support these results. The study uses multiple theoretical frameworks to organize its empirical research.

The analysis uses political economy to demonstrate that fiscal forecasts have political implications because governments in power use optimistic budget forecasts to gain electoral support while deferring their fiscal problems until after elections (Heinemann, 2006; Frankel, 2011; Beetsma et al., 2009). The common pool problem predicts that individual spending ministries will defend budget requests which result in total expenditures that surpass the established budget limits (Alesina & Perotti, 1996). The theory of fiscal illusion shows that voters and policymakers make consistent errors when they try to estimate how much public spending will actually cost which enables them to predict their spending needs for future periods (Buchanan & Wagner, 1977).

On the macroeconomic side, the economic concept of automatic stabilizers creates a framework which predicts that unemployment and economic growth changes will impact spending forecasting errors because social security payments and anti-recession financial programs vary according to employment market conditions (Girouard & André, 2005; Boukari & Veiga, 2018). Inflation establishes a separate path which budget planning processes face through nominal

rigidities because price-level effects will increase their impact over extended future periods (Allan, 1965; Afonso & Silva, 2012). Institutional theory, meanwhile, offers competing predictions regarding fiscal rules: while rules may serve as commitment devices that enhance credibility (Harris et al., 2013; Kumar & Ter-Minassian, 2007), their effectiveness is contingent on design features, enforcement mechanisms, and the broader governance environment (Schick, 1998; Filc & Scartascini, 2010).

The existence of nonbinding rules produces a deceptive perception of discipline which fails to affect spending practices of people. The existing research demonstrates that institutional frameworks establish their effectiveness only under specific conditions which need to be investigated further (Podvorica, Hoti, & Gashi, 2025; Rebić & Arčabić, 2023). The economics of uncertainty provides principles which governments use to handle unexpected financial emergencies and pandemic outbreaks because these events create uncertainty which disrupts established forecasting systems and makes pre-crisis forecsts useless (Knight, 1921; Calitz, Siebrits, & Stuart, 2013).

By grounding the empirical investigation in this multi-theoretical foundation, the study uses multiple theories to conduct empirical research because this approach allows the study to analyze forecast errors. The study investigates macroeconomic and institutional factors by testing their effects on various conditions which determine their strength. The upcoming sections present an empirical strategy that tests the observable results which these theoretical perspectives predict. The study tests four specific elements which include labor market dynamics, inflation, fiscal rules and crisis-period disruptions.

3. DATASET

Thirteen advanced economies (Australia, Canada, Finland, Germany, Hong Kong SAR, Iceland, Ireland, Italy, Netherlands, New Zealand, Sweden, United Kingdom, and United States) was used in the study over the period 2005 to 2024. Three main criteria was used for sample selection. First, the need for consistent and publicly available MTEF for both two-year and three-year horizons. Advanced economies publish MTEF documents, but different countries have different practices regarding the documents' format, content and historical forecast availability. The thirteen countries left in the sample are those for which a continuous series and methodological consistent forecasting of out-year and outer-year expenditure could be constructed from the primary budget documents (Budget Bills, Medium-Term Fiscal Plans, and Stability/Convergence Programmes).

Second, the sample is restricted to advanced economies only to minimize confounding effects of weak budgetary institutions, political instability, or fiscal data measurement errors, which are well documented in developing country experiences (Schiavo-Campo, 2009; Allen et al., 2017). By limiting economies to

those with more or less equal fiscal transparency and reporting standards, the institution noise is separated from the role of macroeconomic fluctuations.

Third, the sample includes a variety of fiscal traditions, such as Anglo-Saxon and continental European models, as well as differences in how fiscal rules are adopted. This means that there is enough diversity to find the marginal effects of macroeconomic and institutional variables. Even though the sample isn't complete, it is a carefully chosen group that was meant to maximize internal validity. Consequently, the results should be regarded as substantial evidence regarding the macroeconomic factors influencing MTEF errors in institutionally stable contexts, rather than as universally applicable to all advanced economies. Future research could broaden the analysis to encompass a wider array of countries as data accessibility enhances.

The year 2005 serves as the starting point because it marks when MTB became established and it guarantees ongoing data access and the period lets researchers track significant worldwide events including the 2008 financial crisis. The MTEF error for each country serves as the dependent variable. The following formula describes how MTEF errors are calculated:

$$\text{MTEF Error}_{i,t+h} = \frac{\text{Forecast}_{i,t+h} - \text{Actual}_{i,t}}{\text{Forecast}_{i,t+h}} \quad (1)$$

The equation shows that negative values demonstrate actual spending surpasses expected amounts because the initial forecast was too optimistic. The positive values demonstrate that actual spending falls short of expected amounts because the initial forecast was too pessimistic. Table 1 shows explanatory variables to determine which major economic factors affect MTEF errors.

Table 1 Variables, Sources and Acronyms

Variable	Description	Source	Acronym
Out-Year Forecast Error (t+2)	Two-year-ahead budget expenditure forecast error, computed as Equation (1).	Budget Bills, MTPs, SCPs, MTFPs	**Out-year**
Outer-Year Forecast Error (t+3)	Three-year-ahead budget expenditure forecast error, computed as Equation (1).	Budget Bills, MTPs, SCPs, MTFPs	**Outer-Year**
GDP Growth	Annual real GDP growth rate (%)	World Bank - WDI	**GDP**
Unemployment	Annual unemployment rate (% of total labor force)	IMF - WEO	**UNP**
Inflation	Inflation rate, end of period consumer prices (annual % change)	IMF - WEO	**INF**
Current Account Balance	Current account balance (% of GDP)	IMF - WEO	**CAD**
Government Debt	Central government total debt (% of GDP)	IMF - WEO	**DBT**
Net Migration	Net migration (total number of migrants)	World Bank - WDI	**MGR**
Expenditure Rule	1 if the country has an expenditure rule in place, 0 otherwise	IMF - Fiscal Rules Dataset	**EXR**
Expenditure Ceiling	1 if the country has a multi-year expenditure ceiling in place, 0 otherwise	IMF - Fiscal Rules Dataset	**EXC**
Covid19	1 for the years 2020 - 2022	Literature	**CVD**
Global Financial Crisis	1 for the years 2008 - 2009	Literature	**CRS**

Source: Author's own elaboration.

The empirical model establishes fiscal planning through its inclusion of macroeconomic variables that research has confirmed to be essential for this purpose. GDP measures total economic activity which affects spending decisions because economic expansion leads to reliable fiscal results while economic contractions create unpredictable fiscal outcomes. The UNP represents current labor market conditions because any UNP changes will directly impact social spending needs and tax revenue forecasts which result in better forecast results. INF measures price fluctuations which create changes in public spending value throughout different periods. Unexpected INF changes disrupt prediction models which include nominal assumptions as part of their framework. CAD serves as an external condition measurement tool because budget imbalances create fiscal performance and forecast accuracy risks that stem from underlying vulnerabilities. DBT shows two aspects of fiscal viability because rising DBT levels restrict financial options while changing spending decision processes. MGR represents population changes which determine public sector requirements for social programs and workforce development and public works projects.

The fiscal policy frameworks use two dummy variables for their measurement. EXR and EXC serve the purpose of establishing control over macroeconomic discipline and environment, which comes with rule-based budgeting. Two crisis dummies capture the macroeconomic disruptions which are brought by CRS and CVD. The study employs these variables to establish their impact on extreme weather conditions, which cause permanent changes to expenditure forecasting during economic downturns.

The sample selection process establishes a methodological trade-off test which measures both internal and external validity. The research analyzes advanced economies which have stable institutional frameworks and dependable fiscal records to decrease errors while studying how macroeconomic factors influence forecasting errors. The chosen approach establishes restrictions on generalizable results, yet it enhances causal understanding because it minimizes the impact of institutional instability and differences in data quality.

4. METHODOLOGY

The analysis of medium-term fiscal outcomes requires panel data methodology because it enables researchers to study both cross-country variations and time-series data. The fiscal performance of MTEF systems depends on two factors which include macroeconomic changes throughout different years and the specific fiscal institutions and budgeting practices that each country develops. The panel framework enables researchers to manage unobserved differences between subjects which leads to better estimation results than methods that rely only on cross-sectional data or time-series data.

The panel data set below is used to estimate how dependent variable links to explanatory variable for identification of macroeconomic factors that cause MTEF errors:

$$Y_{i,t} = \alpha_i + X'_{i,t}\beta + \gamma D_t + \varepsilon_{i,t} \quad (2)$$

$Y_{i,t}$ is the MTEF error of country i in year t, calculated for two-year and three-year forecast horizons. $X_{i,t}$ represents a vector of explanatory variables consisting of macroeconomic indicators and fiscal rule dummy variables. D_t denotes common time dummies capturing global crisis periods that may affect all countries simultaneously. The term α_i represents country-specific effects that are constant over time and capture unobserved heterogeneity across countries, such as long-standing institutional characteristics, fiscal policy traditions, or structural features of national budgeting systems.

The existence of country fixed effects indicates that pooled OLS estimates will produce biased results when country fixed effects correlate with the regressors. Therefore, appropriate panel techniques are required. This choice is particularly important in the context of fiscal forecasting, where unobserved country-specific characteristics, such as institutional quality, fiscal governance traditions, and administrative capacity, are likely to be correlated with macroeconomic variables. Ignoring these effects would result in omitted variable bias and inconsistent estimates, making fixed effects the most appropriate specification for capturing structural heterogeneity across countries.

The determination of stationarity becomes essential because it protects against generating false regression results. First-generation panel unit root tests may not be reliable in the presence of cross-sectional dependence. Hence, I start with testing for cross-sectional dependence of continuous variables using Pesaran's CD test (2004, 2015).

Table 2 CD Test Results

Variable	CD Test	p-value
Out-year	10.65	0.000
Outer-year	11.77	0.000
GDP	24.35	0.000
UNP	15.40	0.000
INF	25.38	0.000
CAD	0.22	0.828
DBT	14.00	0.000
MGR	4.85	0.000

Source: Author's own calculation using Stata 15.

The CD test shows strong cross-sectional dependence in Out-year, Outer-year, GDP, UNP, INF, DBT, and MGR, but not in CAD. This makes first-generation unit root tests unsuitable and requires second-generation tests.

Accordingly, the CADF (Cross-sectionally Augmented Dickey–Fuller) test developed by Pesaran (2007) is employed.

Table 3 CADF Test Results

Variable	t-bar	p-value	result
Out-year	-2.295	0.027	I(0)
Outer-year	-2.589	0.002	I(0)
GDP	-2.402	0.011	I(0)
UNP	-2.155	0.076	I(0)
INF	-1.902	0.295	I(1)
CAD	-1.569	0.739	I(1)
DBT	-0.800	1.000	I(1)
MGR	-1.457	0.850	I(1)

Source: Author's own calculation using Stata 15.

CADF test results show that Out-year, Outer-year, GDP, and UNP are stationary at levels, while INF, DBT, CAD, and MGR contain a unit root. To ensure stationarity, the first differences of these variables are taken for the 2004–2024 period. To improve comparability across variables measured in different units and facilitate interpretation, min–max normalization is applied to all continuous variables, transforming each into the [0,1] interval:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{3}$$

The process of normalization takes place after conducting unit root tests and stationarity tests because this procedure maintains original distribution patterns of variables while eliminating scale differences which impact their estimation accuracy. The complete set of variables which emerged from the transformations will first undergo descriptive statistical analysis followed by correlation analysis before the process of econometric estimation begins.

Table 4 Descriptive Statistics

Variable	Mean	Std. Dev.	Variance	Skewness	Kurtosis
Out-year	0.72873	0.11463	0.01314	-2.12950	11.45006
Outer-Year	0.76766	0.14303	0.02045	-1.88064	8.89419
GDP	0.34786	0.09989	0.00997	0.72553	11.36109
UNP	0.28671	0.18364	0.03372	1.06799	4.62242
INF	0.49211	0.10498	0.01102	-0.32680	9.21653
CAD	0.57236	0.09205	0.00847	-0.58076	18.26429
DBT	0.42331	0.09536	0.00909	1.08675	11.01856
MGR	0.68291	0.06991	0.00488	-3.25843	43.41781
EXR	0.57307	0.49558	0.24560	-0.29548	1.08730
EXC	0.36538	0.48246	0.23277	0.55910	1.3126
CVD	0.15	0.35776	0.12799	1.96039	4.87313
CRS	0.1	0.30057	0.09034	2.66666	8.11111
Obs.:	260	**Min:**	0	**Max:**	1

Source: Author's own calculation using Stata 15.

Table 4 provides descriptive statistical information about the data. The mean Out-year forecast error is −0.054 and Outer-year −0.081, which shows that governments possess systematic optimism bias because they underestimate their future expenditures. The three-year error standard deviation measures 0.122 which exceeds the two-year error standard deviation of 0.103, thus proving that forecast uncertainty increases as forecasting time spans.

The explanatory variables show extensive distribution across normalized GDP and UNP which indicate various economic conditions. First-differenced variables (INF, CAD, DBT, MGR) demonstrate stationarity because their mean values approach zero. Dummy variable means reflect frequency: EXR appears in 57.3% of observations, EXC in 36.5%. Crisis periods account for 15% (CVD) and 10% (CRS) of the sample.

Table 5 Correlation Matrix

Variable	Out-year	Outer-Year	GDP	UNP	INF	CAD	DBT	MGR	EXR	EXC	CVD	CRS
Out-year	1.000											
Outer-Year	0.8734	1.000										
	0.0000											
GDP	-0.0096	-0.0101	1.000									
	0.8770	0.8707										
UNP	0.1547	0.1953	-0.1745	1.000								
	0.0124	0.0016	0.0048									
INF	-0.1098	-0.0678	0.03911	-0.0519	1.000							
	0.0772	0.2758	0.0000	0.4048								
CAD	-0.0892	-0.0229	0.1455	0.0137	-0.0637	1.000						
	0.1514	0.7134	0.0189	0.8257	0.3059							
DBT	-0.0175	-0.0254	-0.5976	0.2677	-0.0168	0.0382	1.000					
	0.7794	0.6832	0.0000	0.0000	0.7870	0.5393						
MGR	0.0695	0.0561	0.1576	-0.0710	0.1852	-0.0230	-0.0900	1.000				
	0.2639	0.3676	0.0109	0.2537	0.0027	0.7125	0.1481					
EXR	-0.0107	0.0349	0.0442	0.2047	0.0371	0.0237	-0.0761	-0.0203	1.000			
	0.8631	0.5754	0.4780	0.0009	0.5516	0.7033	0.2212	0.7440				
EXC	-0.0541	-0.0545	0.0252	-0.1579	0.0114	0.0616	-0.0639	0.0011	0.0574	1.000		
	0.3853	0.3812	0.6861	0.0108	0.8553	0.3225	0.3050	0.9857	0.3562			
CVD	-0.4377	-0.4175	0.0140	-0.0844	0.3790	0.0290	0.0576	0.0589	0.0359	0.0168	1.000	
	0.0000	0.0000	0.8217	0.1750	0.0000	0.6414	0.3546	0.3445	0.5641	0.7878		
CRS	-0.0804	-0.0733	-0.3647	0.0511	-0.0861	0.0404	0.3407	-0.0202	-0.1270	-0.0399	-0.1400	1.000
	0.1961	0.2392	0.0000	0.4123	0.1665	0.5165	0.0000	0.7455	0.0407	0.5215	0.0239	

Source: Author's own calculation using Stata 15.

Table 5 presents the results of correlation analysis. UNP has a positive and significant connection to errors which shows that two-year errors rise by 0.15 and three-year errors increase by 0.20 when labor market conditions become less favorable. CVD shows a strong negative correlation with pandemic-related fiscal disruptions which leads to a correlation coefficient of −0.44 and −0.42. INF has a

weak negative association with other variables which show no statistically significant results. The GDP and DBT relationship shows a moderate negative correlation which reaches −0.60 and this finding might indicate the presence of multicollinearity. The Driscoll-Kraay standard errors and block-wise model specifications have reduced this issue to a minor extent. The study results will remain unaffected by multicollinearity because other correlations show low values.

The analysis has completed its evaluation of descriptive properties and bivariate relationships. The analysis now proceeds to test the econometric model. The selection of estimators depends on two main aspects of the dataset which include the country sample's nonrandom distribution and the panel units' cross-sectional dependence. The sample consists of thirteen advanced economies which represent specific countries instead of being a random sample from a bigger population. The unobserved country-specific differences which include institutional traits and fiscal customs and economic structures will probably associate with the research variables. The random effects model will produce inconsistent results under these particular conditions. Fixed effects represents the correct theoretical choice because it includes all factors which remain constant for a specific country throughout the study into the intercept.

Panel units display cross-sectional dependence which creates another obstacle for econometric analysis. The Pesaran CD test (Table 2) shows strong cross-sectional dependence across countries, making conventional standard errors unreliable. The models use the Driscoll–Kraay (1998) nonparametric covariance matrix estimator to generate standard errors which maintain their accuracy in handling cross-sectional dependence and heteroskedasticity and serial correlation. The maximum lag length in baseline specifications has its limit set to two years because this represents the standard method for handling serial correlation.

This choice is further justified by the nature of macro-panel data, where countries are simultaneously exposed to common global shocks and interconnected economic dynamics. In such settings, standard errors that correct only for within-unit correlation (such as clustered estimators) may remain biased. The Driscoll–Kraay method delivers complete corrections because it simultaneously handles cross-sectional dependence, heteroskedasticity and serial correlation, which makes it appropriate for the dataset used in this research.

The empirical analysis uses a sequential block-wise modeling strategy which assesses how different groups of variables improve the model's ability to explain outcomes. This approach also serves as a diagnostic tool, allowing the identification of incremental explanatory power and helping to assess the stability of coefficients across specifications, thereby reducing the risk of omitted variable bias. The approach achieves two distinct objectives. The first step of adding variables lets researchers check how coefficient stability changes with different model specifications and how omitted variable bias arises. The second step enables administrators to sustain their control over all operational processes.

$$Y_{i,t} = \alpha_i + \beta_1 GDP_{i,t} + \beta_2 UNP_{i,t} + \beta_3 \Delta INF_{i,t} + \varepsilon_{i,t} \quad (4)$$

where $Y_{i,t}$denotes the MTEF error for country i in year t, α_i represents country-specific fixed effects, and $\varepsilon_{i,t}$is the idiosyncratic error term.

The second specification adds the change in CAD to capture external sector conditions. External imbalances show two problems because they reflect macroeconomic weaknesses and show how a country will react to international economic shocks, which makes it difficult to make accurate fiscal forecasts. The current account improvements show better financial fundamentals of the country, while current account deterioration decreases prediction accuracy and increases uncertainty.

$$Y_{i,t} = \alpha_i + \beta_1 GDP_{i,t} + \beta_2 UNP_{i,t} + \beta_3 \Delta INF_{i,t} + \beta_4 \Delta CAD_{i,t} + \varepsilon_{i,t} \quad (5)$$

The third specification adds the DBT to capture fiscal sustainability considerations. The DBT system operates through its dynamics to affect both fiscal policy behavior and the credibility of forecasted outcomes. The market pressures which rising debt creates together with institutional constraints will result in improved accuracy through their disciplinary effects. The existence of debt sustainability issues establishes policy uncertainty which results in difficulties for forecasting.

$$Y_{i,t} = \alpha_i + \beta_1 GDP_{i,t} + \beta_2 UNP_{i,t} + \beta_3 \Delta INF_{i,t} + \beta_4 \Delta CAD_{i,t} + \beta_5 \Delta DBT_{i,t} + \varepsilon_{i,t} \quad (6)$$

The fourth specification adds MGR to account for demographic dynamics. Migration flows impact public spending because they determine how many people will work and which public services will be needed and which infrastructure facilities will be required. Continuous migration patterns will lead to increased financial demands in three areas which include education and healthcare and social assistance programs.

$$Y_{i,t} = \alpha_i + \beta_1 GDP_{i,t} + \beta_2 UNP_{i,t} + \beta_3 \Delta INF_{i,t} + \beta_4 \Delta CAD_{i,t} + \beta_5 \Delta DBT_{i,t} + \beta_6 \Delta MGR_{i,t} + \varepsilon_{i,t} \quad (7)$$

The fifth specification introduces fiscal policy regulations through EXR and EXC which serve two purposes by limiting spending increases and establishing fiscal planning standards. The rules drive government spending limits while they enhance fiscal management processes to produce better forecasting of future budget needs.

$$Y_{i,t} = \alpha_i + \beta_1 GDP_{i,t} + \beta_2 UNP_{i,t} + \beta_3 \Delta INF_{i,t} + \beta_4 \Delta CAD_{i,t} \beta_5 \Delta DBT_{i,t} + \beta_6 \Delta MGR_{i,t} + \beta_7 EXR_{i,t} + \beta_8 EXC_{i,t} + \varepsilon_{i,t} \quad (8)$$

The sixth specification serves as the most extensive research method because it includes major global crises which have the potential to disrupt both macroeconomic conditions and fiscal forecasts. The economic system experienced sudden and extreme changes when CRS and CVD created extraordinary shocks that affected both fiscal policies and public spending activities. The study uses two crisis dummies to manage the effects of these exceptional time periods.

$$Y_{i,t} = \alpha_i + \beta_1 GDP_{i,t} + \beta_2 UNP_{i,t} + \beta_3 \Delta INF_{i,t} + \beta_4 \Delta CAD_{i,t} + \beta_5 \Delta DBT_{i,t} + \beta_6 \Delta MGR_{i,t} + \beta_7 EXR_{i,t} + \beta_8 EXC_{i,t} + \beta_9 CVD_t + \beta_{10} CRS_t + \varepsilon_{i,t} \quad (9)$$

The same six nested specifications are estimated separately for Out-year and Outer-year. The two forecasting periods that studied will show whether macroeconomic factors affect different forecasting lengths and whether longer forecasts show greater response to economic conditions and institutional factors.

5. RESULTS

The fixed effects panel regression results with Driscoll-Kraay standard errors serve as the primary empirical results for this section. The results should first be explained for Out-year and then for Outer-year. Table 6 displays the coefficient values which the six nested models use to forecast Out-year prediction errors.

Table 6 Results for Out-Year Driscoll–Kraay Standard Errors

		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
GDP	**Coef.**	0.0786	0.106	0.0992	0.0907	0.0874	-0.101
	Std.Err.	0.0926	0.0957	0.0983	0.0862	0.0866	0.0743
UNP	**Coef.**	0.0999***	0.103***	0.104***	0.106***	0.111***	0.0778**
	Std.Err.	0.0279	0.0271	0.0281	0.0279	0.0294	0.0284
INF	**Coef.**	-0.140	-0.158	-0.155	-0.171	-0.167	0.0863
	Std.Err.	0.127	0.128	0.132	0.144	0.145	0.0739
CAD	**Coef.**		-0.142**	-0.140**	-0.138**	-0.134*	-0.0581
	Std.Err.		0.0656	0.0671	0.0651	0.0672	0.0594
DBT	**Coef.**			-0.0101	-0.00677	-0.0164	-0.00850
	Std.Err.			0.0446	0.0423	0.0450	0.0855
MGR	**Coef.**				0.156	0.154	0.164
	Std.Err.				0.174	0.172	0.101
EXR	**Coef.**					-0.00931	-0.00844
	Std.Err.					0.0134	0.0123
EXC	**Coef.**					-0.00434	-0.00658
	Std.Err.					0.0153	0.0139
CVD	**Coef.**						-0.155***
	Std.Err.						0.0192
CRS	**Coef.**						-0.0684***
	Std.Err.						0.0102
cons	**Coef.**	0.742***	0.821***	0.826***	0.726***	0.734***	0.661***
	Std.Err.	0.0575	0.0721	0.0673	0.126	0.126	0.106
R^2		0.0382	0.0507	0.0507	0.0594	0.0613	0.251

(*) indicates MacKinnon (1996) one-sided p-values: (***) significance at 1%, (**) at 5% and (*) at the 10%.

Source: Author's own calculation using Stata 15.

Table 6 shows that UNP serves as the most stable factor determining forecasting errors. The coefficient shows a positive relationship which reaches significance at 1% in Models 1 to 5 while Model 6 shows 5% significance. The estimate stays between 0.10 and 0.10 for early models but decreases to 0.078 after adding crisis periods. The positive coefficient shows that increased UNP causes more positive forecast errors which results in less negative forecast errors. The

presence of negative errors shows that forecasters become more cautious when labor markets experience downturns because social spending and automatic stabilizers start to increase.

The model results show that GDP has no impact on forecast accuracy. The coefficient in Model 6 shows values that range close to zero and change between positive and negative. The research shows that short-term growth variations after controlling for fixed effects make no impact on forecasting accuracy.

The change in INF results in no meaningful effects. The coefficient exists as a negative value throughout Models 1 to 5 while showing a positive value in Model 6 yet maintaining its status as an insignificant result. The post-COVID surge period shows higher instability because advanced economies experienced moderate INF throughout most of the period. The use of crisis dummies explains part of the variation that INF used to show because special events need to be monitored for research purposes.

The reduction in CAD shows both negative results and substantial impact across all tested Models 2 to 5, which range from -0.142 to –0.134. The external position shows strength, which suggests that the economy functions well, thus creating positive spending expectations that later exceed actual needs. The Model 6 results show that the coefficient becomes insignificant when crisis dummies are included, while its value decreases to -0.058, which indicates that crises explain most of the CAD variation that researchers previously studied.

The change in DBT never reaches an important level of significance. The coefficient remains close to zero throughout all models while the standard errors appear substantial especially in Model 6. The accuracy of forecasts remains unaffected by short-term changes in DBT. The level or long-term trajectory of debt may still matter because first-differencing eliminates trend components while measuring year-to-year debt changes which show random fluctuations.

The change in MGR is also insignificant in all specifications. Migrants tend to spend money on education and health services and social benefits but their spending patterns create too many delayed effects which make it impossible to track total spending through current first difference assessments. The point estimate remains stable between 0.15 and 0.16 yet the standard errors produce extensive uncertainty.

The Model 5 model tests fiscal rules through EXR and EXC. The results show that both variables lack importance because their coefficients fall below the value of -0.01. The rules existence does not result in different forecast errors after researchers controlled for macroeconomic factors. The rules have some effect but their value depends on institutional design and enforcement practice which requires intricate rule classification that extends beyond the parameters of this research.

The true extent of changes appears with the introduction of crisis dummies for Model 6. Both CRS and CVD show negative results which reach 1% significance level through their coefficients that show values of -0.068 and -0.155. The pandemic effect became more than double the impact which the 2008 crisis brought. Actual spending during both crises exceeded forecasts which created a

situation of excessive over-optimism. The R^2 value increases from 0.061 in Model 5 to 0.251 in Model 6 because extraordinary shocks operate as the primary element that leads to errors despite all macroeconomic variables being taken into account.

The constant term remains positive and highly significant across all models, reflecting the average forecast error after controlling for covariates. The base value of the measurement decreases slightly when include crisis dummies, which shows that some over-optimism continues to exist during typical periods.

Labour market conditions serve as the strongest economic predictor for the two-year period because external imbalances only remain important until research account for crisis situations. The relationship between fiscal rules and demographic changes does not exist. The most crucial finding shows that extraordinary disruptions together with CRS and CVD account for quarter of forecast error variation while they overshadow all other components.

The analysis now turns to Outer-year to see whether these patterns persist or diverge through the process of accumulating uncertainty. The Outer-year results display their findings through the same sequential specification method which Table 7 presents.

Table 7 Results for Outer-Year Driscoll–Kraay Standard Errors

		Model1	Model2	Model3	Model4	Model5	Model6
GDP	**Coef.**	0.0801	0.0922	0.0156	0.0070	0.0044	-0.229**
	Std.Err.	0.110	0.0992	0.0903	0.0843	0.0846	0.0816
UNP	**Coef.**	0.157***	0.158***	0.167***	0.170***	0.169***	0.128***
	Std.Err.	0.0323	0.0322	0.0361	0.0360	0.0407	0.0367
INF	**Coef.**	0.108	0.116	0.0973*	0.103*	0.101**	0.217**
	Std.Err.	0.177	0.176	0.053	0.062	0.051	0.103
CAD	**Coef.**		-0.0629	-0.0444	-0.0417	-0.0384	0.0560
	Std.Err.		0.160	0.161	0.162	0.163	0.164
DBT	**Coef.**			-0.115	-0.111	-0.116	-0.108
	Std.Err.			0.0879	0.0853	0.0830	0.113
MGR	**Coef.**				0.159	0.158	0.170
	Std.Err.				0.199	0.198	0.101
EXR	**Coef.**					-0.0026	-0.0013
	Std.Err.					0.0207	0.0185
EXC	**Coef.**					-0.0066	-0.0094
	Std.Err.					0.0211	0.0191
CVD	**Coef.**						-0.195***
	Std.Err.						0.0193
CRS	**Coef.**						-0.0816***
	Std.Err.						0.0196
cons	**Coef.**	0.748***	0.783***	0.831***	0.730***	0.735***	0.643***
	Std.Err.	0.0863	0.127	0.112	0.183	0.183	0.156
R^2		0.0441	0.0456	0.0488	0.0545	0.0551	0.246

(*) indicates MacKinnon (1996) one-sided p-values: (***) significance at 1%, (**) at 5% and (*) at the 10%.

Source: Author's own calculation using Stata 15.

Table 7 confirms previous findings while providing additional evidence. UNP stands as the strongest predictor throughout all six model tests because its coefficient values maintain positive statistical significance at the 1% level which results in coefficient values ranging from 0.128 to 0.170 across all model tests. The longer time frame has greater forecasting difficulties because labor market dynamics become less predictable and fiscal effects accumulate over time. The coefficient maintains its value across different specifications to demonstrate that labor market conditions dominate the development of fiscal forecasts for the next year.

The Outer-year models show a significant shift in INF when research add more variables to their analysis. The first two specifications show no significant results which become marginally significant at 10% for Models 3 through 4 whereas Models 5 through 6 show results that reach significant status at 5%. The positive sign (0.097–0.217) indicates higher INF is associated with more positive forecast errors. Unanticipated INF creates spending challenges over three years because it affects nominal expenditure through indexation and contracts and wages. Research founds that fiscal rules and crisis dummies control the situation which shows that the effects of fiscal regulations need special examination.

The relationship between GDP and errors remains weak and unstable according to most specifications which duplicate Out-year results. The coefficient in Models 1 to 5 remains insignificant because it changes both its direction and its size. The value in Model 6 reaches a negative value which becomes statistically significant at the 5 percent level with a value of -0.229. The results show that stronger economic growth decreases overoptimistic forecasting after the financial crisis and pandemic have been controlled. Robust growth will enhance fiscal forecasts while it will decrease unexpected expenditures, which will simplify the process of maintaining budget forecasts. The relationship only appears after research isolate crises because they need to differentiate between normal economic fluctuations and major disruptive events.

The change in CAD, which became essential for two-year models, lost its importance when outer-year specifications were tested. The model shows a negative coefficient between -0.044 and -0.063 in Models 2 to 5, but the standard errors remain high enough to prevent any results from reaching statistical significance. The results indicate that external imbalances affect fiscal forecasts primarily during brief periods, which shows that immediate financial weaknesses exist. The transitory signals over three years will become less distinct because of the ongoing effects of more permanent factors.

The DBT changes no longer show significance through any of the models which produce results that extend beyond the current year. The full specification shows that the value of the coefficient maintains its negative direction while the forecasting accuracy decreases because standard errors have increased. This evidence demonstrates that advanced economies experience no regular relationship between their annual debt variations and their forecasting performance. The level of debt or its long-term trajectory through time remains important because these

impacts will either emerge through country fixed effects or need an alternative model to be shown.

The MGR stays useless throughout all research models, which confirms the results from the two-year study. The point estimate shows stability between 0.16 and 0.17, but the standard errors show high variability. The research demonstrates that MGR influences educational and health and social services spending, but the research shows that its total spending effect is too spread out and delayed and diverse to detect using first difference analysis in a panel research design.

The outer-year models show that fiscal rules maintain their minimal impact because their coefficients stay close to zero while their standard errors remain large. The out-year analysis demonstrates that rules do not create forecasting accuracy changes after researchers control for visible economic conditions. The rules should not be considered ineffective because their effects extend beyond what a basic two-way system can show. Exploring such mechanisms requires richer institutional data.

The most significant discovery shows that crisis dummies have an extremely strong impact. The CRS and CVD show negative values which reach 1% significance with their coefficients showing –0.082 and –0.195. The pandemic effect exceeds the 2008 crisis by more than double its impact. The two crises demonstrate strong over-optimism because actual spending surpassed the three-year budget forecasts. The addition of these dummies increases R^2 from 0.055 in Model 5 to 0.246 in Model 6 which shows that extraordinary shocks serve as the main factor that explains forecast error variation throughout all time periods.

The constant term remains positive and significant across all models which shows that the model maintains over-optimism during non-crisis periods. The value of the effect decreases when crisis dummies are included in the analysis yet it maintains its economic significance.

The two-year findings receive confirmation through outer-year results, which also provide additional evidence. UNP acts as the primary economic driver that determines economic results because its effects become stronger throughout extended time frames. The short-term impact of INF remains negligible, but its three-year mark shows actual importance because price changes now affect total spending. Crises dominate explanatory power, accounting for most forecast error variation. The relationship between external balances, debt dynamics, migration, and fiscal rules shows no consistent pattern. The medium-term economic forecasting process depends on stable macroeconomic factors, which only recognize inflation as important when forecasting beyond traditional time frames.

6. ROBUSTNESS CHECKS

The study uses three different robustness tests which examine three separate types of potential estimation bias to prove that its empirical results do not depend on a single research design. The study conduct a two-way fixed effects estimation of the baseline model by including country and year dummies as basic model. The method handles unobserved country differences which remain constant through time while also managing common global financial patterns and synchronized policy actions which would otherwise distort the estimated coefficients.

Second, the Driscoll-Kraay estimator is re-estimated with a longer lag structure (lag=3). The baseline specification assumes that serial correlation in the residuals is adequately captured by a two-year lag. Extending this to three years provides a more stringent test of whether the results are sensitive to the assumed persistence of within-country dependencies. Third, the models are re-estimated using Panel-Corrected Standard Errors (PCSE; Beck & Katz, 1995). While the Driscoll-Kraay estimator is robust to cross-sectional dependence, heteroskedasticity, and serial correlation, PCSE offers an alternative nonparametric correction that has been widely applied in comparative political economy and fiscal policy research. Convergence of results across these two estimators would strengthen confidence in the reported standard errors.

It is worth noting that alternative validation approaches were considered but deemed less suitable for the present context. Out-of-sample testing, while valuable for purely predictive exercises, is less informative when the primary objective is structural inference, specifically, identifying the partial effects of macroeconomic variables on forecast errors within the observed sample period. Similarly, alternative estimators such as system GMM or Feasible Generalized Least Squares were not employed because the fixed-effects Driscoll-Kraay framework already addresses the key panel data complications identified without imposing additional moment conditions or parametric assumptions on the error structure. The chosen robustness strategy thus prioritizes validating the core findings against the most salient threats to inference in this dataset.

The complete model specification (Model 6) serves as the basis for all tests. The results for Out-year and Outer-year appear in Table 8.

Table 8 Robustness Results for Out-Year and Outer-Year Models

		Baseline Model		TWFE Model		DKL3 Model		PCSE Model	
									-0.158
GDP	**Coef.**	-0.101	-0.229**	-0.0240	-0.240	-0.101	-0.229***	-0.0306	0.111
	Std.Err	0.0743	0.0816	0.171	0.224	0.0785	0.0706	0.0958	0.140**
UNP	**Coef.**	0.0778**	0.128***	0.0879	0.127	0.0778**	0.128***	0.0794*	0.0558
	Std.Err	0.0284	0.0367	0.0534	0.0747	0.0299	0.0389	0.0462	0.126
INF	**Coef.**	0.0863	0.217**	-0.0333	0.130	0.0863	0.217**	0.0549	0.0807
	Std.Err	0.0739	0.103	0.0866	0.0830	0.0723	0.0957	0.0675	-0.0685
CAD	**Coef.**	-0.0581	0.0560	0.0153	0.200	-0.0581	0.0560	-0.152*	0.0999
	Std.Err	0.0594	0.164	0.0516	0.131	0.0571	0.159	0.0866	-0.0909
DBT	**Coef.**	-0.00850	-0.108	0.103	-0.0915	-0.0085	-0.108	0.0076	0.137
	Std.Err	0.0855	0.113	0.129	0.226	0.0941	0.120	0.113	0.202**
MGR	**Coef.**	0.164	0.170	0.0943	0.115	0.164*	0.170*	0.161**	0.100
	Std.Err	0.101	0.101	0.0858	0.135	0.0947	0.0879	0.0806	-0.00659
EXR	**Coef.**	-0.00844	-0.00133	-0.176	-0.0161	-0.0084	-0.00133	-0.0102	0.0226
	Std.Err	0.0123	0.0185	0.246	0.0274	0.0123	0.0188	0.0163	-0.00747
EXC	**Coef.**	-0.00658	-0.00941	0.0283	0.0621*	-0.0065	-0.00941	-0.0053	0.0266
	Std.Err	0.0139	0.0191	0.0281	0.0340	0.0131	0.0175	0.0183	-0.146***
CVD	**Coef.**	-0.155***	-0.195***	-0.0317	-0.131**	-0.155***	-0.195***	-0.140***	0.0268
	Std.Err	0.0192	0.0193	0.0607	0.0524	0.0168	0.0171	0.0218	-0.060**
CRS	**Coef.**	-0.0684***	-0.081***	-0.0412	-0.0968	-0.0684***	-0.08***	-0.0557**	0.0279
	Std.Err	0.0102	0.0196	0.0596	0.0593	0.0101	0.0199	0.0233	0.692***
cons	**Coef.**	0.661***	0.643***	0.610***	0.611***	0.661***	0.643***	0.697***	0.107
	Std.Err	0.106	0.156	0.121	0.125	0.105	0.149	0.0892	0.323
R^2		0.251	0.246	0.246	0.353	0.353	0.246	0.246	

(*) indicates MacKinnon (1996) one-sided p-values: (***) significance at 1%, (**) at 5% and (*) at the 10%.

Source: Author's own calculation using Stata 15.

The testing of robustness shows that essential research results maintain their consistency. UNP shows positive and significant results across most model specifications at both forecasting time periods. The change in INF shows different results based on the selected estimator. It appears positive yet lacks significance in the baseline test, it shows negative results in the two-way fixed effects model, and it never reaches significant levels. The internal world economy of a country shows limited changes which create this uncertainty. CAD and DBT show mainly non-significant results throughout all tests which proves that external imbalances and short-term debt changes do not impact forecast accuracy in any systematic way.

MGR shows modest, estimator-sensitive effects: insignificant in baseline, marginally significant under alternative lag, and significant at 5% under PCSE, which shows demographic pressures produce only minor effects. The results show that fiscal rule dummies stay insignificant in all tested models, which proves that rules by themselves cannot enhance accurate forecasting.

Crisis indicators serve as the strongest predictor for determining the outcome. CRS and CVD both show negative impacts which remain statistically

significant at 1% level across two testing methods and their various lag settings while retaining their PCSE validation, although their impact strength reduced. The two-way fixed effects system shows that variables become insignificant because year dummies capture all shared disturbances, which demonstrates that crises occur throughout all nations but show different effects across each country. The three-year period shows greater crisis effects because of increasing uncertainty throughout that timeframe.

The constant term shows positive results across all models because it tests fundamental baseline system which shows excessive optimism during non-crisis periods. The model fit demonstrates substantial improvement through crisis inclusion because R^2 values increase from less than 0.06 to approximately 0.25 which then advances to 0.35 with two-way fixed effects models.

7. DISCUSSION

The results contribute to the literature in three ways. First, MTEF errors show systematic links to macroeconomic conditions. Second, the importance of determinants varies with the forecast horizon, consistent with the finding that accuracy deteriorates as the horizon lengthens. Third, extraordinary shocks exert a dominant influence, which exceeds the impact of standard macroeconomic factors.

With a positive and significant coefficient across horizons and specifications, UNP stands out as the most reliable and stable determinant. This extends previous annual-budget studies to the medium-term (Gentry, 1989; Mayper et al., 1991; Brogan, 2012; Boukari & Veiga, 2018). Medium-term frameworks frequently assume smooth trajectories, making them susceptible to UNP shocks, especially over longer horizons; deteriorating labor markets increase social transfers and uncertainty.

GDP exhibits a complex trend. After accounting for crises, it becomes negative and significant in the three-year models. But it is negligible in two-year models. Growth is important for forecast accuracy, but its impact depends on the macroeconomic climate and is only apparent when crises are taken into consideration. This balances research that reports null results (Ríos et al., 2018) with studies that find significant growth effects (Brogan, 2012; Buettner & Kauder, 2015; Merola & Pérez, 2013; Boukari & Veiga, 2018).

INF becomes significant only at extended time periods. The two-year models show no impact of INF while the three-year models demonstrate positive results. The research shows that INF identifies as a determining factor (Gentry, 1989; Allan, 1965; Afonso & Silva, 2012; Sedmihradská & Čabla, 2013; Kara, 2024) while price assumptions and indexation mechanisms become increasingly important for longer time periods. Forecasts should consider INF as the primary factor that creates uncertainty instead of using as a standard technical component.

CAD shows a horizon-specific pattern. The relationship between forecast errors and two-year models shows negative results until the introduction of crises which causes three-year models to lose their predictive value. External imbalances may shape near-term fiscal conditions but are not robust medium-term drivers.

Changes in DBT and MGR show no significant effects in any specification. The debt dynamics exhibit too much uncertainty and their level-dependent behavior makes it difficult to assess debt dynamics while the migration effects extend beyond immediate time frames and create effects through channels that current methods cannot detect.

EXR and EXC are consistently negligible. This qualifies rather than contradicts the literature (Smith, 2007; Beetsma et al., 2009, 2011; Pina & Venes, 2011; Frankel, 2011; Heinemann, 2006). More important than mere existence are bindingness, design, and credibility (Harris et al., 2013; Raudla et al., 2020). Forecast accuracy is not increased by merely having a rule. Recent studies similarly report that formal institutional indicators, whether governance quality for FDI (Podvorica et al., 2025) or fiscal rules for debt sustainability (Rebić & Arčabić, 2023), fail to explain economic outcomes unless accompanied by effective implementation and a supportive macroeconomic context.

Beyond the direct effect of fiscal rules, the broader issue of transparency and accountability in public financial management appears to matter for fiscal outcomes. Prijaković (2023) discovered that Croatian cities which show greater online budget disclosure tend to have lower direct debt obligations. The research shows that budgeting transparency functions as a fiscal discipline tool because it enables real accountability systems to operate instead of existing as mere official reporting requirements. The present study examines forecast accuracy rather than debt levels, yet the study shows that institutional arrangements do not guarantee better fiscal results because their success depends on the wider governance system which controls their operations.

The most clear finding is that crisis variables are the strongest. Both CRS and CVD are harmful and highly significant, but CVD produces a greater impact. This finding supports research evidence which shows that extreme events disrupt normal operations (PSC, 2011; Calitz et al., 2013; Boukari & Veiga, 2018; Kara, 2024). The forecasts experience their highest instability during medium-term frameworks because crises cause most of the error variation that impacts them.

The results of the robustness tests provide evidence which confirms this interpretation. The main results of the study stay the same with different estimation methods and different lag patterns and two-way fixed effect models. The most stable factors affecting things are UNP and crisis variables. The variables CAD and MGR demonstrate estimator sensitivity so users need to interpret their results with caution. The main results stay intact because one estimator and one lag option do not produce any different outcomes.

There are multiple policy implications which arise from this situation. First, MTBs should more clearly include how the labor market works. Budget officials need to apply multiple baseline paths for labor market prediction because UNP serves as the most accurate forecasting method. Longer time periods should be assessed using INF as a principal factor which creates forecasting uncertainty. Expenditure forecasts should include sensitivity analysis which evaluates different price assumptions and their impact on inflation-related wage costs and procurement expenses and transfer payments and indexation costs.

The overwhelming effects of crises which arise during disaster situations demonstrate that contingency planning needs to be established as an essential requirement. The framework needs to establish three different types of scenarios which include baseline and upside and downside scenarios together with specific financial response plans because deterministic baselines do not provide adequate coverage. The null results for fiscal rules show that policymakers need to focus on designing and enforcing rules while creating trustworthy institutions instead of pursuing only ceremonial compliance. The binary rule indicators need to be replaced because they fail to show the vital aspects which exist in actual situations.

The study has certain limitations which researchers must observe. The study becomes more valid when researchers study developed economies but this approach limits their ability to apply findings from the study to developing countries. The researchers intentionally excluded political factors which include election cycles and ideology and the institutional elements of budget centralization and fiscal transparency and independent fiscal institutions to study macroeconomic impacts. Future studies could focus on these aspects or investigate whether factors vary depending on the type of expenditure.

The research investigates how macroeconomic elements affect MTEF errors during different time intervals in developed nations despite its existing limitations. The research shows that labor market conditions determine inflation effects which become more important as time goes on and external balances show minimal influence during crisis periods. The combination of economic instability and institutional design and major unexpected events creates problems for medium-term budgeting and technical forecasting accuracy. The research needs realistic labor market assumptions and explicit price-risk analysis and trustworthy fiscal institutions and capacity to handle unexpected shocks for better forecasting results.

8. CONCLUSION

The research aimed to identify macroeconomic factors which affected MTEF errors across thirteen advanced economies from 2005 until 2024 while examining two-year and three-year forecast periods. The researchers used an econometric framework together with multiple testing methods to establish that MTB forecast errors occurred due to specific technical factors rather than appearing as random events. The time forecasting period together with the existing

macroeconomic situation serve as the main factors which determine forecast accuracy. The research adds to existing knowledge by examining medium-term fiscal planning through a study which goes beyond traditional annual forecasting evaluation methods.

The research results show that labor market trends serve as the primary factor which determines how accurate forecasts will be. Across all model specifications and estimation techniques, unemployment emerges as the most consistent and robust determinant of expenditure forecast errors. The result demonstrates that medium-term fiscal frameworks require accurate labor market development tracking because fiscal effects of unemployment increase with time through automatic stabilizers and discretionary policy actions. Other macroeconomic variables such as GDP growth and inflation show different effects which depend on the time period under consideration. The growth becomes important when specific model specifications are used while inflation has greater impact on public spending at extended time periods because price changes affect capital expenditures.

The study reveals that extraordinary macroeconomic shocks serve as the main element which determines research results. The global financial crisis together with the COVID-19 pandemic generates strong effects which demonstrate statistical significance for forecast errors, which enables models to explain data better than before. The results demonstrate that medium-term budgeting systems face their most basic predictability assumption because of major system breakdowns which occur during operational disruptions. The presence of major exogenous disruptions demonstrates the fundamental restrictions which deterministic forecasting methods face because they cannot predict the complete fiscal results which their systems require.

The study provides critical knowledge about the functioning methods used by financial institutions. The fiscal rule indicators do not show any meaningful connection to forecast accuracy because established rules do not provide enough protection to achieve dependable medium-term projections. Three elements determine how framework systems function which include their design characteristics, their user value and their integration level with existing institutional frameworks. The current finding demonstrates that fiscal governance systems need to operate as mandatory rules together with strong institutional capacity to achieve substantial impacts on fiscal outcomes according to existing research.

The results reach a definitive conclusion which demonstrates that MTEF accuracy needs multiple methods which extend beyond technical model improvements. Policymakers require better economic scenario assessments which should include complete labor market and inflation behavior analysis and they must establish better institutional systems and emergency response strategies. Future research could build on these findings by integrating political and institutional variables, expanding the country sample to include developing economies, and exploring the heterogeneity of forecast errors across different expenditure

categories. The extensions would improve our understanding of how macroeconomic factors and institutional systems interact with medium-term budgeting fiscal prediction accuracy.

Author Contributions: The author confirms sole responsibility for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation.

Funding: The research presented in the manuscript did not receive any external funding.

Conflict of interest: None.

Acknowledgement of AI or AI-assisted tools use: In this study, artificial intelligence (ChatGPT) was used to translate the title and abstract into Croatian.

REFERENCES

Afonso, A., & Silva, J. (2012). *The Fiscal Forecasting Track Record of the European Commission and Portugal.* Lisbon: ISEG Economics Working Paper No. 37/2012/DE/UECE. https://doi.org/10.2139/ssrn.2181117

Alesina, A., & Perotti, R. (1996). Fiscal Discipline and the Budget Process. *The American Economic Review, 86*(2), 401-407.

Allan, C. M. (1965). Fiscal Marksmanship, 1951-1963. *Oxford Economic Papers, 17*(2), 317-327. https://doi.org/10.1093/oxfordjournals.oep.a040992

Allen, R., & Tommasi, D. (2001). *Managing Public Expenditure: A Reference Book for Transition Countries.* Paris: OECD. https://doi.org/10.1787/9789264192607-en

Allen, R., Chaponda, T., Fisher, L., & Ray, R. (2017). *Medium-Term Budget Frameworks in Sub-Saharan African Countries.* 203. IMF Working Paper. https://doi.org/10.5089/9781484314203.001

Atiyas, İ., & Sayın, Ş. (1997). *Siyasi Sorumluluk, Yönetsel Sorumluluk ve Bütçe Sistemi: Bir Yeniden Yapılanma Önerisine Doğru.* İstanbul: Türkiye Ekonomik ve Sosyal Etüdler Vakfı (TESEV).

Beck, N., & N. Katz, J. (1995). What to do (and not to do) with Time-Series Cross-Section Data. *The American Political Science Review, 89*(3), 634-647. https://doi.org/10.2307/2082979

Beetsma, R., Bluhm, B., Giuliodori, M., & Wierts, P. (2011). *From First-Release to Ex-Post Fiscal Data: Exploring the Sources of Revision Errors in the EU.* Amsterdam: Tinbergen Institute Discussion Paper. https://doi.org/10.2139/ssrn.1844173

Beetsma, R., Giuliodori , M., & Wierts, P. (2009). Planning to Cheat: Eu Fiscal Policy in Real Time. *Economic Policy, 24*(60), 753-804. https://doi.org/10.1111/j.1468-0327.2009.00230.x

Boex, J., Martinez-Vazquez, J., & Mcnab, R. (2000). Multi-Year Budgeting: A Review of International Practices and Lessons for Developing and Transitional Economies. *Public Budgeting & Finance, 20*(2), 91-112. https://doi.org/10.1111/0275-1100.00013

Boukari, M., & Veiga, F. J. (2018). Disentangling Political and Institutional Determinants of Budget Forecast Errors: A Comparative Approach. *Journal of Comparative Economics, 46*(4), 1030-1045. https://doi.org/10.1016/j.jce.2018.03.002

Breuer, C. (2015). On the Rationality of Medium-Term Tax Revenue Forecasts: Evidence from Germany. *ahrbücher für Nationalökonomie und Statistik, 235*(1), 22-40. https://doi.org/10.1515/jbnst-2015-0104

Brogan, M. (2012). The Politics of Budgeting: Evaluating the Effects of the Political Election Cycle on State-Level Budget Forecast Errors. *Public Administration Quarterly, 36*(1), 84-115.

Buettner, T., & Kauder, B. (2015). Political Biases Despite External Expert Participation: An Empirical Analysis of Tax Revenue Forecasts in Germany. *Public Choice, 164*(3), 287-307. https://doi.org/10.1007/s11127-015-0279-2

Calitz, E., Siebrits, K., & Stuart, I. (2013b). *Enhancing the Credibility of Fiscal Forecasts in South Africa: is a Fiscal Council the Only Way?* Working Papers 25. Stellenbosch: Stellenbosch University, Department of Economics.

Choi, J.-Y., & Park, N. (2013). *Achieving Medium Term Expenditure Framework Reform A Case Study of Korea.* The World Bank Policy Research Working Paper 6342. https://doi.org/10.1596/1813-9450-6342

Dorotinsky, W., & Watkins, J. (2013). *Common Practices in Setting Expenditure Ceilings within National Budgets.* Washington D.C.: World Bank.

Driscoll, J. C., & Kraay, A. C. (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *The Review of Economics and Statistics, 80*(4), 549–560. https://doi.org/10.1162/003465398557825

Filc, G., & Scartascini, C. (2010). Is Latin America on the Right Track?: An Analysis of Medium-Term Frameworks and the Budget Process. *IDB Working Paper Series No. IDB-WP-160.* Washington, DC: Inter-American Development Bank. https://doi.org/10.18235/0010825

Francesco, M. D., & Barroso, R. (2015). *Review of International Practices for Determining Medium-Term Resource Needs of Spending Agencies.* World Bank Policy Research Working Paper No. 7289. https://doi.org/10.1596/1813-9450-7289

Frankel, J. A. (2011). Over-optimism in Forecasts by Official Budget Agencies and Its Implications. *Oxford Review of Economic Policy, 27*(4), 536-562. https://doi.org/10.1093/oxrep/grr025

Gentry, W. M. (1989). Do State Revenue Forecasters Utilize Available Information? *National Tax Journal, 42*(4), 429-439. https://doi.org/10.1086/NTJ41788812

Harris, J., Hughes, R., Ljungman, G., & Sateriale, C. (2013). Medium-Term Budget Frameworks in Advanced Economies: Objectives, Design, and Performance. M. Cangiano, T. Curristine, & M. Lazare içinde, *Public Financial Management and Its Emerging Architecture* (s. 137-173). International Monetary Fund. https://doi.org/10.5089/9781475531091.071

Heinemann, F. (2006). Planning or Propaganda? An Evaluation of Germany's Medium-term Budgetary Planning. *FinanzArchiv / Public Finance Analysis, 62*(4), 551-578. https://doi.org/10.1628/001522106X172698

Holmes, M., & Evans, A. (2003). *A Review of Experience in Implementing Medium Term Expenditure Frameworks in a PRSP Context: A Synthesis of Eight Country Studies.* London: Overseas Development Institute .

Houerou, P. L., & Taliercio, R. (2002). Medium term expenditure frameworks : from concept to practice - preliminary lessons from Africa. *Working Paper 23651.* Washington, DC: The World Bank.

Jena, P. R. (2006). Fiscal Marksmanship: Link between Forecasting Central Tax Revenues and State Fiscal Management. *Economic and Political Weekly, 41*(37), 3971-3976.

Kara, B. (2024). The Performance of Medium-Term Budgeting in Türkiye: An Analysis of Budget Forecasts. *Yönetim Ve Ekonomi Dergisi, 31*(4), 659-676. https://doi.org/10.18657/yonveek.1521261

Kara, B. (2024). *Türkiye'de Kamu Bütçesi Tahminlerinin Gerçekliği.* Ankara: Gazi Kitabevi.

Kara, B., & Sarıoğlu, F. (2026). The Impact of Economic Factors on Medium-Term Budget Revenue Forecasts: Insights from an Ex Post Analysis of Advanced Economies. *Journal of Risk and Financial Management, 19*(1), 34. https://doi.org/10.3390/jrfm19010034

Kumar, M., & Ter-Minassian, T. (2007). Fiscal discipline: key issues and overview. M. Kumar, & T. Ter-Minassian içinde, *Promoting fiscal discipline* (s. 1-8). IMF. https://doi.org/10.5089/9781451871142.001

MacKinnon, J. G. (1996). Numerical Distribution Functions for Unit Root and Cointegration Tests. *Journal of Applied Econometrics, 11*(6), 601-618. https://doi.org/10.1002/(SICI)1099-1255(199611)11:6<601::AID-JAE417>3.0.CO;2-T

Mayper, A. G., Granof, M., & Giroux, G. (1991). An Analysis of Municipal Budget Variances. *Accounting, Auditing & Accountability Journal, 4*(1), 29-50. https://doi.org/10.1108/09513579110003358

Merola, R., & Pérez, J. (2013). Fiscal Forecast Errors: Governments versus Independent Agencies? *European Journal of Political Economy, 32*, 285–299. https://doi.org/10.1016/j.ejpoleco.2013.09.002

Muzychenko, G., Koliada, T., & Churkina, I. (2017). Impact Of European Integration Processes on The Necessity of Long-Term Budget Planning Introduction in Ukraine. *Baltic Journal of Economic Studies, 3*(5), 329-333. http://dx.doi.org/10.30525/2256-0742/2017-3-5-329-333

OECD. (2023). Spending Better: Post-Covid Budget Frameworks. *45th Annual Meeting of Senior Budget Officials.* Athens: OECD Publishing.

PCS, & RIG. (2011). *States' Revenue Estimating: Cracks in the Crystal Ball.* New York: Pew Center on the States and The Nelson A. Rockefeller Institute of Government.

Penner, R. G. (2001). *Errors in Budget Forecasting.* Washington D.C.: The Urban Institute.

Pesaran, M. H. (2004). *General Diagnostic Tests for Cross Section Dependence in Panels.* Bonn: IZA Discussion Paper No. 1240. https://doi.org/10.2139/ssrn.572504

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics, 22*(2), 265–312. https://doi.org/10.1002/jae.951

Pesaran, M. H. (2015). Testing Weak Cross-Sectional Dependence in Large Panels. *Econometric Reviews, 34*(6-10), 1089-1117. https://doi.org/10.1080/07474938.2014.956623

Pina, Á. M., & Venes, N. (2011). The Political Economy of EDP Fiscal Forecasts: An Empirical Assessment. *European Journal of Political Economy, 27*(3), 534-546. https://doi.org/10.1016/j.ejpoleco.2011.01.005

Plesko, G. A. (1988). The Accuracy of Government Forecasts and Budget Procejtions. *National Tax Journal*, 41(4), 483-501. https://doi.org/10.1086/NTJ41788755

Podvorica, N., Hoti, A., & Gashi, P. (2025). Fdi Drivers In Central Europe, The Baltics, And The Western Balkans: The Role Of Macroeconomic And Institutional Factors. *Economic Thought and Practice*, 34(1), 65-84. https://doi.org/10.17818/EMIP/2025/15

Prijaković, S. (2023). Budget Transparency and Debt of Cities in Crotia. *Economic Thought and Practice*, 32(2), 585-613. https://doi.org/10.17818/EMIP/2023/2.14

Radu, D. (2023). Domestic Medium-Term Budgetary Frameworks in the EU: Fit for Purpose and for the Future? Discussion Paper 189. Brussels: European Commission.

Raudla, R., Bur, S., & Keel, K. (2020). The Effects of Crises and European Fiscal Governance Reforms on the Budgetary Processes of Member States. Journal of Common Market Studies, 58(3), 740-756. https://doi.org/10.1111/jcms.12972

Rebic, P., & Arcabic, V. (2023). Sustainability And Character Of Fiscal Policy In The Republic Of Croatia. *Economic Thought and Practice*, 32(2), 455-472. https://doi.org/10.17818/EMIP/2023/2.8

Ríos, A.-M., Guillamón, M.-D., Benito, B., & Bastida, F. (2018). The Influence of Transparency on Budget Forecast Deviations in Municipal Governments. Journal of Forecasting, 37(4), 457–474. https://doi.org/10.1002/for.2513

Schiavo-Campo, S. (2009). Potemkin Villages: "The" Medium-Term Expenditure Framework in Developing Countries. *Public Budgeting & Finance, 29*(2), 1-26. https://doi.org/10.1111/j.1540-5850.2009.00926.x

Schick, A. (1998). *A Contemporary Approach to Public Expenditure Management.* World Bank Institute Governance, Regulation, and Finance Division.

Schroeder, L. (2007). Forecasting Local Revenues and Expenditures. A. Shah içinde, *Local Budgeting* (s. 53-79). Washington: The World Bank. https://doi.org/10.1596/978-0-8213-6945-6

Sedmihradská, L., & Čabla, A. (2013). Budget Accuracy in Czech Municipalities and the Determinants of Tax Revenue Forecasting Errors. *Ekonomická revue - Central European Review of Economic Issues, 16*(4), 197–206. https://doi.org/10.7327/cerei.2013.12.01

Smith, D. L. (2007). Rules, Participants, and Executive Politics in State Tax Revenue Forecasting. *Journal of Public Budgeting, 19*(4), 472-487. https://doi.org/10.1108/JPBAFM-19-04-2007-B004

Uralovich, P. U. (2025). Analysis of Interacting Factors in Medium-Term Forecasting and Planning of Government Budget Revenues. *Central Asian Journal of Innovations on Tourism Management and Finance, 7*(1), 280–288. doi:10.51699/cajitmf.v7i1.1100

Vlaicu, R., Verhoeven, M., Grigoli, F., & Mills, Z. (2014). Multiyear budgets and fiscal performance: Panel data evidence. *Journal of Public Economics, 111*, 79-95. https://doi.org/10.1016/j.jpubeco.2013.12.011

Wildawsky, A., & Zapico-Goñi, E. (1993). *National Budgeting for Economic and Monetary Union.* Leiden: Brill | Nijhoff. https://doi.org/10.1163/9789004640740

Dr. sc. Berat Kara

Docent
Sveučilište Istanbul Medeniyet, Turska
Fakultet političkih znanosti
E-mail: berat.kara@medeniyet.edu.tr
Orcid: https://orcid.org/0000-0002-6948-2197

ZAŠTO SREDNJOROČNE PROCJENE RASHODA POGAĐAJU? NOVI DOKAZI IZ PANEL PODATAKA

Sažetak

Ova studija istražuje makroekonomske odrednice pogrešaka u srednjoročnim procjenama rashoda u naprednim gospodarstvima od 2005. do 2024., s naglaskom na dvogodišnje i trogodišnje horizonte. Primjenom panelnih metoda analizira se uloga nezaposlenosti, gospodarskog rasta, inflacije, stanja tekućeg računa, javnog duga i fiskalnih pravila, uzimajući u obzir i krizna razdoblja. Rezultati pokazuju da je nezaposlenost najrobusnija i najdosljednija odrednica pogrešaka u svim modelima i horizontima, što ističe važnost tržišta rada u srednjoročnom fiskalnom planiranju. Inflacija je značajna uglavnom na duljim horizontima, što upućuje na jačanje njezinih kumulativnih učinaka tijekom vremena. Suprotno tome, rast BDP-a ima ograničenu i modelno ovisnu ulogu, dok javni dug, migracije i fiskalna pravila ne pokazuju sustavan utjecaj. Velike krize znatno povećavaju pogreške i dominiraju objašnjavajućom snagom modela.

Ključne riječi: srednjoročno proračunsko planiranje, proračunske procjene, srednjoročne procjene, procjene rashoda, točnost proračuna.

JEL klasifikacija: H61, C23, E62, E27.