

Macroeconomic Forecasting for the G7 countries under Uncertainty Shocks

Shovon Sengupta^{1,2,3}, Sunny Kumar Singh², Tanujit Chakraborty^{1,4}

¹ SAFIR, Sorbonne University Abu Dhabi, United Arab Emirates.

² BITS Pilani, Hyderabad Campus, India.

³ Fidelity Investments, Boston, USA.

⁴Sorbonne Center for Artificial Intelligence, Sorbonne University, Paris, France

Abstract

Accurate macroeconomic forecasting has become harder amid geopolitical disruptions, policy reversals, and volatile financial markets. Conventional vector autoregressions (VARs) overfit in high-dimensional settings, while threshold VARs struggle with time-varying interdependencies and complex parameter structures. We address these limitations by extending the Sims–Zha Bayesian VAR with exogenous variables (SZBVARx) to incorporate domain-informed shrinkage and four newspaper-based uncertainty shocks—economic policy uncertainty, geopolitical risk, US equity market volatility, and US monetary policy uncertainty. The framework improves structural interpretability, mitigates dimensionality, and imposes empirically guided regularization. Using G7 data, we study spillovers from uncertainty shocks to five core variables—unemployment, real broad effective exchange rates, short-term rates, oil prices, and CPI inflation—combining wavelet coherence (time–frequency dynamics) with nonlinear local projections (state-dependent impulse responses). Out-of-sample results at 12- and 24-month horizons show that SZBVARx outperforms 14 benchmarks, including classical VARs and leading machine-learning models, as confirmed by Murphy difference diagrams, multivariate Diebold–Mariano tests, and Giacomini–White predictability tests. Credible Bayesian prediction intervals deliver robust uncertainty quantification for scenario analysis and risk management. The proposed SZBVARx offers G7 policymakers a transparent, well-calibrated tool for modern macroeconomic forecasting under pervasive uncertainty.

Keywords: G7 countries, Bayesian VAR, Uncertainty shocks, Local projections, Credible intervals.

JEL Classification: C11, C32, C53, C55, E37.

1. Introduction

Uncertainty is deeply ingrained in the growth and development process of the G7 economies: Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States. The economic importance of accurate macroeconomic forecasting for these countries is self-evident. Collectively, the G7 economies account for about 28–30% of global GDP, more than 40% of global household wealth, and are home to the deepest and most liquid financial markets in the world ([International](#)

Monetary Fund, 2024b; Credit Suisse Research Institute, 2023). The G7 also host four of the world’s five largest currencies, with the US dollar, euro, yen, and pound sterling together accounting for about 89% of disclosed official foreign-exchange reserves. For these reasons, the G7 economies are by far the largest anchors of the international monetary system (International Monetary Fund, 2024a; Gopinath, Boz, Casas, Díez, Gourinchas & Plagborg-Møller, 2020). The G7 is also home to the majority of systemically important financial institutions and is the world’s largest source and recipient of cross-border capital flows (Rey, 2015; Miranda-Agrippino & Ricco, 2021). Given this significant footprint in world trade, finance, and monetary conditions, shocks that originate in the G7 or are transmitted through it have sizable spillover effects on emerging markets and the global economy more broadly (Bloom, Bunn, Mizen, Smietanka & Thwaites, 2019). The G7’s growth outlook continues to be buffeted by a perfect storm of adverse global shocks (including financial crises, political instability, technological change, wars, pandemics, and more), of which uncertainty, shocks, and their transmission are key (Baker, Bloom & Davis, 2016; Caldara & Iacoviello, 2022; Bloom, 2009). At the heart of this storm lie the network interactions among five key macroeconomic variables (unemployment rate, real broad effective exchange rate (REER), short-term interest rates (SIR), oil prices, and consumer price index (CPI) inflation), which are not only highly interrelated but also extremely sensitive to exogenous uncertainty shocks of all types (e.g., economic and monetary policy uncertainties, geopolitical risk, financial market volatility etc.). The shocks have been becoming more frequent and larger in magnitude, which has severe and measurable adverse impacts on expectations, financial stability, inflation, and real activity in the G7 economies. Forecasting frameworks can no longer be based only on fundamentals but must explicitly account for high-frequency information on uncertainty, which continues to be generated at an unprecedented speed and scale. Empirical studies in recent literature indicate that uncertainty shocks feed into each other (i.e., so-called “uncertainty multipliers”), which change the well-documented causal relationships among key macroeconomic variables, destabilize labor markets and investment, and result in volatile swings in exchange rates, flows, and prices (Caggiano, Castelnuovo & Groshenny, 2014; Smales, 2022; Rossi, 2013; Yang, Zhang, Liu, Liu, Tao, Lai & Huang, 2024). We treat them as core drivers of G7 forecasts.

In this paper, we examine the response of five macroeconomic variables—unemployment rate, REER, SIR, the oil price, and CPI inflation—to four sources of uncertainty shocks: economic policy uncertainty (EPU), geopolitical risk (GPR), US equity market volatility (USEMV), and US monetary policy uncertainty (USMPU). Such shocks, which encompass a range of uncertainty from policy-related disagreements to geopolitical tensions, are recognized as major driving forces of the outlook in advanced economies, and particularly G7 countries (Jones & Olson, 2013; Baker et al., 2016; Caldara & Iacoviello, 2022). Our work, therefore, is timely, as their increasing prominence in the economy suggests that they should be included in the forecasting system to understand and monitor the macroeconomic implications of global uncertainty and, potentially, to inform policy. The ongoing process of deepening interdependence across countries is a major motivation for multivariate forecasting as a key tool for G7 policymakers to monitor the interdependent co-movements of unemployment, inflation, exchange rates, interest rates, and commodity prices.

This article presents a structured platform to link the different uncertainty shocks to outcomes by emphasizing the G7 economy. The labor market, foreign exchange market, monetary policy, oil market, and CPI inflation outcomes are observed to be highly responsive to the jointly moving EPU,

GPR, USEMV, and USMPU variables in the G7 (Baker et al., 2016; Caldara & Iacoviello, 2022; Bloom, 2009; Caggiano et al., 2014; Smales, 2022; Jurado, Ludvigson & Ng, 2015). At the same time, pronounced spikes in EPU and GPR are linked to deteriorations in employment, volatile inflation and interest rates, shifting risk premia, and volatile exchange rates, particularly in the context of systemic cross-country shocks (Baker et al., 2016; Caldara & Iacoviello, 2022; Bloom, 2009; Caggiano et al., 2014; Smales, 2022; Jurado et al., 2015). These can propagate quickly through financial and trade linkages, inducing capital flows and currency realignments (Forbes & Warnock, 2012; Rossi, 2013; Ahir, Bloom & Furceri, 2022; Bekaert, Hoerova & Duca, 2013). Labor market responses are also particularly important, with hiring and investment often retrenching in the face of rising uncertainty (Bloom, 2009; Caggiano et al., 2014). By contrast, exchange rates tend to track sudden, uncertainty-induced capital flows, (Forbes & Warnock, 2012; Smales, 2022). The recent USMPU can also transmit the interest rate, and particularly so when mixed with EPU and GPR (Husted, Rogers & Sun, 2020; Tillmann, 2020). Oil price shocks are another potentially strong source of inflation and growth uncertainty, which have been found to respond strongly to policy and geopolitical news and interact with financial speculation (Kilian, 2009; Baumeister & Kilian, 2016; Kilian & Murphy, 2014; Kang, Ratti & Vespignani, 2017). In addition, the latest inflation evidence in the G7 shows it is now more sensitive to external uncertainty shocks and policy changes than to domestic slack (Coibion, Gorodnichenko & Ulate, 2019; Fajgelbaum, Goldberg, Kennedy & Khandelwal, 2020). Recent developments in the cross-country modeling frontier are therefore especially relevant for modeling the global nature of the interactions and support the use of multiple sources of uncertainty when forecasting the policy-relevant variables (Bai, Carriero, Clark & Marcellino, 2022; Baker et al., 2016; Caldara & Iacoviello, 2022).

In a data-rich environment like the G7, Bayesian VAR models, proposed by Litterman (1986); Sims, Doan & Litterman (1983); Sims & Zha (1998) have long been recognized as a well-performing benchmark. Recent developments, including hierarchical shrinkage priors (Baíbura, Giannone & Reichlin, 2010), stochastic volatility priors (Koop, 2013), and extensions to multi-country or high-dimensional macro systems (Miranda-Agrippino & Ricco, 2019; Bai et al., 2022; Besher, Sengupta & Chakraborty, 2025) continue to improve the accuracy and flexibility of these models (Carriero, Clark & Marcellino, 2019a; Feldkircher, Huber, Koop & Pfarrhofer, 2022; Gefang, Koop & Poon, 2023). On the empirical front, there is the need to include external drivers into the forecasting models, such as EPU (Baker et al., 2016), GPR (Caldara & Iacoviello, 2022), and financial or monetary shocks (Husted et al., 2020; Tang, 2015), as they have a significant impact on inflation, exchange rates, and other macro dynamics (Carriero, Galvao & Kapetanios, 2019b; Sengupta, Chakraborty & Singh, 2025; Bojer, 2022; Chakraborty, Besher, Panja & Sengupta, 2025). At the same time, a rich set of methodological advances, such as Bayesian Global VARs (Canova & Ciccarelli, 2004; Pesaran, Schuermann & Smith, 2009; Cuaresma, Feldkircher & Huber, 2016), time-varying parameter VARs (Koop & Korobilis, 2013; Bekiros, 2014), and deep learning models for multi-horizon prediction (Laborda, Ruano & Zamanillo, 2023), are being used to capture a range of complex features in the data, such as nonlinearities, regime changes, and spillovers.

Despite these advancements, classical and Bayesian VARs remain prominent forecasting workhorses but face limitations in capturing nonlinear dynamics, structural breaks, and uncertainty shock propagation (Sims & Zha, 1998; Koop & Korobilis, 2013; Canova, 2011; Kilian & Lütkepohl, 2017). Threshold and Markov-switching VARs introduce regime shifts, yet their complexity con-

strains practical use in high-dimensional G7 environments (Balke & Fomby, 1997; Lo & Zivot, 2001). Even recent time-varying parameter Bayesian VARs with structural identification can overlook key external policy and financial shocks (Koop & Korobilis, 2013; Del Negro & Schorfheide, 2008; Carriero et al., 2019a). The Sims-Zha Bayesian VAR with exogenous variables (SZBVARx) framework directly incorporates major uncertainty sources (Baker et al., 2016; Caldara & Iacoviello, 2022; Husted et al., 2020). Unlike previous VAR approaches to uncertainty, SZBVARx allows these shocks to have unrestricted effects on all of the core macroeconomic variables. In contrast to some large-scale multi-country models, SZBVARx allows for a good degree of country-level granularity but also explicitly introduces cross-country spillovers using a common set of exogenous variables (Pesaran et al., 2009; Cuaresma et al., 2016). Bayesian shrinkage in the SZBVARx model helps to ensure reliable estimation, especially in high-dimensional environments (Belmonte, Koop & Korobilis, 2014).

Against this background, the present paper contributes to the existing literature on three main fronts. *First*, on methods and theory, we propose a framework in which uncertainty is explicitly modeled as an autonomous and structural driver of the G7 macroeconomy. We directly incorporate four externally measured uncertainty indices in a Sims–Zha Bayesian VAR with exogenous variables (Jurado et al., 2015; Ludvigson, Ma & Ng, 2021; Baker et al., 2016; Caldara & Iacoviello, 2022), and treat uncertainty as a predictive state variable that augments more standard shocks. We couple nonlinear local projections (Carriero et al., 2019b) to recover regime-dependent impulse responses with shrinkage-based Bayesian estimation for robust inference in moderately large systems. We use Bayesian posterior sampling to construct predictive credible intervals that account for forecast uncertainty, providing empirically calibrated credible intervals (Baíbura et al., 2010; Carriero et al., 2019b; Koop, 2013). *Second*, on evidence, we consider five major macroeconomic variables for G7 countries. Long-run wavelet coherence confirms that uncertainty indices systematically lead macroeconomic variables along the business cycle and during major events like the global financial crisis and COVID-19, highlighting their value as truly exogenous drivers. State-dependent impulse responses reveal a number of regime asymmetries, with uncertainty shocks being both more impactful on prices and real quantities under tight monetary conditions and more persistent in easy regimes, especially through external channels for open economies. A thorough forecast evaluation against fourteen classical, machine-learning, and deep-learning benchmarks at 12- and 24-month horizons further highlights consistent gains in accuracy across a number of standard loss functions, validated by Multiple Comparisons with the Best, multivariate Diebold–Mariano and Giacomini–White tests, as well as Murphy-difference diagrams. *Third*, on application, our approach delivers output in a form that can be directly utilized by central banks and other policy institutions, including regime-conditioned impulse responses, coherent medium-term projections, and thus ready to be used in real time to assess, monitor, and report measures of forecast uncertainty, aiding in risk management, forward guidance, and scenario planning, as needed in environments of geopolitical or financial uncertainty (Bernanke, 2015; Adrian, Boyarchenko & Giannone, 2019; Fernández-Villaverde, Guerrón-Quintana, Rubio-Ramírez & Uribe, 2011; Huber, Koop, Onorante, Pfarrhofer & Schreiner, 2023). This paper bridges the gap between state-of-the-art econometric methods and the real-world need for uncertainty-aware forecasting of advanced economies.

The article is organized as follows. Section 2 presents the data and empirical setup. Section 3 provides wavelet-coherence evidence on G7 uncertainty transmissions and an impulse-response anal-

ysis under different monetary regimes. Section 4 details the SZBVARx framework, assumptions, and treatment of exogenous shocks. Section 5 reports the forecasting study such as baselines, performance metrics, results and experimental comparisons, significance and robustness, and uncertainty quantification via credible intervals. Section 6 discusses policy implications. Section 7 concludes with contributions, limitations, and future work.

2. Data and Preliminary Analysis

This study is concerned with the macroeconomic dynamics of the G7 countries. In particular, we jointly analyze five macro variables—Unemployment Rate, Real Broad Effective Exchange Rate (REER), Short-term Interest Rate (SIR), Oil Prices proxied by the benchmark West Texas Intermediate (WTI), and consumer price index (CPI) Inflation—together with four newspaper-based uncertainty measures: log-transformed country-wise Economic Policy Uncertainty (EPU), country-wise Geopolitical Risk (GPR), US Equity Market Volatility (USEMV), and US Monetary Policy Uncertainty (USMPU). The sample spans January 1995 to March 2024, a window that includes multiple business-cycle phases, a Global Financial Crisis, a zero lower bound era, and the COVID-19 pandemic, a global pandemic shock. We assess forecast performance using a rolling window design at two horizons: 12-month-ahead (“semi-long-term”) and 24-month-ahead (“long-term”). This allows us to contrast medium-run accuracy with longer-run stability in timeframes relevant for policy and investment decisions. The Section 2.1 documents the key data features through summary statistics and global characteristics, followed by tests for endogeneity among key macroeconomic variables and structural break detection.

2.1. Global Characteristics

This section provides a broad description of the data organization involved in our multivariate analysis for the G7 countries. We concentrate on five major macroeconomic variables, which are the unemployment rate, REER, SIR, oil price (WTI), and CPI inflation. These five macroeconomic indicators were chosen as they are key factors for major advanced economies. The unemployment rate is a measure of slack in the labor market and cyclical variation in employment. It is obtained from the Infra-Annual Labor Statistics. The REER is a weighted average of bilateral exchange rates and is normalized by relative consumer prices. It is used as a measure of each economy’s external competitiveness. SIR is used as a proxy for the stance of monetary policy and liquidity conditions. We use the three-month interbank rate for this purpose. Oil price (WTI) is included as a summary of global commodity market conditions that feed into both production costs and inflationary pressures. CPI inflation is the year-on-year change in consumer prices, and thus it measures the purchasing power of households and is the most direct indicator of price stability. All the macroeconomic series are sourced from the Federal Reserve Economic Data (FRED) database to ensure methodological consistency and cross-country comparability. A first pass through the trends and variability - plotted for each country-variable pair and presented in Figure A.1.1 of Appendix A.1.1 of the supplementary material, reveals substantial heterogeneity. Pronounced cyclical swings and crisis-related spikes characterize unemployment and CPI inflation, while exchange rates and oil prices exhibit more structural volatility. SIRs appear to be impacted by distinct monetary

policy regimes across the entire sample period, including the extended era of the zero lower bound (a prolonged period, particularly since the global financial crisis, when central banks kept short-term interest rates near zero, limiting conventional policy and leading to widespread adoption of tools such as quantitative easing and forward guidance). The autocorrelation diagnostics (see Appendix A.1.2, Figure A.1.2) indicate that most series display strong persistence through a gradually decaying autocorrelation structure. This widespread persistence motivates the use of a Bayesian VAR framework, which naturally accommodates persistent dynamics through its autoregressive structure while simultaneously capturing joint dynamics and cross-country spillover effects through the multivariate specification (Hyndman & Athanasopoulos, 2018; Diebold & Yilmaz, 2009; Pesararan, Schuermann & Weiner, 2004). Additionally, four newspaper-based high-frequency measures of uncertainty that are designed to capture different dimensions of global risk and policy uncertainty are utilized in this study. We start with the EPU index of Baker et al. (2016), which is constructed as the frequency of terms related to policy and uncertainty in leading newspapers. We follow Bloom (2009) and take the natural logarithm of the series (EPU) to reduce heteroscedasticity and the effect of spikes in uncertainty. The country-wise GPR index of Caldara & Iacoviello (2022) is designed to measure geopolitical risks to the global economy that stem from geopolitical events such as terrorism, armed conflict, and political and economic instability, using a news dataset hand-coded at the international level. It includes both sudden spikes in uncertainty and background levels of geopolitical tensions. USEMV index of Baker, Bloom, Davis & Kost (2019) is the share of news articles that connect economic and financial volatility to U.S. stock market returns and is thus a forward-looking measure of financial market stress. USMPU index of Husted et al. (2020) is based on a set of terms relating to monetary policy and uncertainty and constructed as the frequency with which articles in leading financial newspapers use those terms to identify uncertainty in Federal Open Market Committee decisions. These four indices are designed to provide complementary coverage of policy, geopolitical, financial-market, and monetary uncertainty, respectively.

Standard summary statistics, describing the distribution of the series, namely (minimum, quartiles, median, mean, standard deviation, and maximum) and two complexity measures: the coefficient of variation (CoV) and entropy, to describe the distributions of both macro series and uncertainty indices, are reported in Table 1. The CoV, standard deviation divided by the mean, renders volatility comparable across units of observation. A higher CoV implies that observations are more dispersed, making point forecasting more difficult. Entropy is a measure of randomness in distributions. It can be viewed as a proxy for informational complexity: higher entropy implies distributions are more dispersed and harder to predict. The picture is not uniform across countries. Some labor-market series are characterized by high CoV and entropy consistent with their erratic dynamics. In a number of countries, inflation also displays relatively high complexity. USEMV and USMPU, along with oil prices, appear to be among the more idiosyncratic exogenous drivers, as they display high CoV and entropy. This is also a manifestation of their episodic, shock-driven behavior that also transmits more forecast uncertainty. These findings demonstrate the need for a flexible, robust forecasting framework for the G7 economies.

Table 1: Summary statistics of the datasets utilized in this analysis for G7 countries. For each country, statistics are reported for unemployment rate, REER, SIR, CPI Inflation, EPU, and GPR. Global variables include Oil Price (WTI), USEMV, and USMPU.

Country	Series	Min	Value	Q1	Median	Mean	SD	Q3	Max	Value	CoV	Entropy
Canada	Unemployment Rate	4.50	6.30	7.20	7.30	1.39	8.10	14.30	19.17	5.91		
	REER	88.92	99.63	102.85	106.36	10.74	115.29	132.19	10.10	5.86		
	SIR	0.18	1.16	2.40	2.73	1.88	4.38	8.31	68.83	5.80		
	CPI Inflation	-0.95	1.27	1.99	2.13	1.32	2.57	8.13	62.30	5.94		
	EPU	1.48	1.92	2.16	2.15	0.28	2.38	2.83	13.15	6.08		
	GPR	0.06	0.13	0.17	0.21	0.16	0.24	1.72	75.64	8.03		
US	Unemployment Rate	3.10	4.30	5.10	5.61	1.85	6.20	14.40	32.96	5.90		
	REER	78.43	86.79	95.02	93.99	8.25	99.25	112.99	8.79	5.86		
	SIR	0.09	0.32	1.87	2.62	2.24	5.24	6.73	85.55	5.64		
	CPI Inflation	-2.10	1.63	2.29	2.54	1.66	3.18	9.06	65.25	5.91		
	EPU	1.65	1.94	2.05	2.07	0.18	2.21	2.70	8.92	6.10		
	GPR	0.75	1.63	1.99	2.28	1.29	2.58	13.23	56.87	5.97		
France	Unemployment Rate	6.50	8.30	9.10	9.41	1.52	10.30	13.20	16.20	5.90		
	REER	92.98	100.05	104.10	104.85	5.79	109.64	118.15	5.52	5.86		
	SIR	-0.58	0.03	2.12	1.98	1.98	3.56	8.06	99.91	5.73		
	CPI Inflation	-0.73	0.87	1.54	1.63	1.24	2.07	6.28	76.24	5.92		
	EPU	1.05	1.92	2.21	2.15	0.32	2.41	2.76	14.79	6.08		
	GPR	0.17	0.34	0.45	0.52	0.32	0.60	2.80	60.58	6.68		
Germany	Unemployment Rate	2.80	4.00	7.20	6.62	2.68	8.80	12.20	40.45	5.85		
	REER	94.46	99.89	102.85	105.03	6.68	109.39	127.17	6.37	5.86		
	SIR	-0.58	0.03	2.12	1.89	1.82	3.45	5.16	96.43	5.77		
	CPI Inflation	-1.04	1.12	1.51	1.80	1.55	2.01	8.82	86.49	5.89		
	EPU	1.45	1.94	2.10	2.14	0.29	2.30	2.93	13.46	6.08		
	GPR	0.08	0.23	0.32	0.40	0.29	0.48	2.66	74.32	6.92		
Japan	Unemployment Rate	2.10	3.00	3.90	3.84	0.95	4.70	5.80	24.73	5.96		
	REER	70.72	97.85	123.75	119.24	25.43	136.97	193.97	21.33	5.84		
	SIR	-0.08	-0.02	0.04	0.14	0.29	0.11	2.25	206.35	7.15		
	CPI Inflation	-2.56	-0.42	0.10	0.36	1.21	0.72	4.39	341.67	8.04		
	EPU	1.68	1.94	2.01	2.02	0.12	2.09	2.38	6.30	6.11		
	GPR	0.05	0.13	0.18	0.23	0.16	0.27	1.24	69.39	7.86		
UK	Unemployment Rate	3.50	4.70	5.30	5.76	1.47	7.10	9.00	25.53	5.92		
	REER	94.28	101.92	109.08	113.26	12.30	126.50	135.22	10.86	5.86		
	SIR	0.03	0.59	3.75	3.21	2.55	5.55	7.81	79.41	5.65		
	CPI Inflation	0.20	1.50	2.09	2.41	1.65	2.69	9.61	68.36	5.88		
	EPU	1.68	1.98	2.16	2.16	0.21	2.33	2.64	9.84	6.09		
	GPR	0.40	0.71	0.93	1.06	0.66	1.16	5.99	61.56	6.20		
Italy	Unemployment Rate	5.30	8.10	9.60	9.59	1.93	11.20	14.40	20.21	5.89		
	REER	86.42	100.05	102.85	103.00	4.04	106.10	110.63	3.92	5.86		
	SIR	-0.58	0.03	2.12	2.45	2.82	3.93	11.01	114.99	5.57		
	CPI Inflation	-0.58	1.07	2.00	2.23	1.95	2.73	11.84	87.66	5.77		
	EPU	1.68	1.98	2.16	2.16	0.21	2.33	2.64	9.84	6.09		
	GPR	0.03	0.08	0.12	0.14	0.09	0.17	0.65	67.34	9.25		
Global	Oil Price (WTI)	11.27	28.83	54.09	55.71	28.73	77.03	133.96	51.57	5.73		
	USEMV	9.57	16.24	18.93	21.18	8.02	23.97	69.83	37.90	5.82		
	USMPU	16.57	48.20	75.04	90.62	58.69	120.16	407.94	64.76	5.68		

To provide a comprehensive overview of the properties of the series, considered in this study, we report seven key statistical properties: skewness, kurtosis, nonlinearity, seasonality, stationarity, long-range dependence, and presence of outliers - in Table 2. The nonlinearity is captured by Tsay's and Keenan's test (Tsay, 1986; Keenan, 1985), which tests for departure from linear autoregressive structure and can accommodate threshold or regime-switching dynamics. The seasonality is captured by the Ollech–Webel combined procedure (Ollech & Webel, 2023), which is a robust diagnostic that bundles up several different tests to reduce misclassification. The stationarity is checked by the KPSS test (Kwiatkowski, Phillips, Schmidt & Shin, 1992), which takes stationarity as the null and is used in conjunction with standard unit-root tests. The long-range dependence is captured by the Hurst exponent (Shang, 2020), which differentiates between mean-reverting, memoryless, and trend-reinforcing behavior. The outliers are detected by Bonferroni-adjusted tests based on studentized residuals (Weisberg & Cook, 1982), which filter the rare but influential shocks. The diagnostics suggest that most of the G7 macroeconomic series are non-stationary, nonlinear, and exhibit long memory. Table 2 also indicates the presence of seasonality, mainly evident in unemployment rates, EPU, and USEMV. Statistically significant outliers cluster around the global financial crisis and the pandemic. In sum, the evidence suggests complex, persistent, and sometimes volatile dynamics, which call for uncertainty-resilient multivariate models that can accommodate nonlinearity, regime shifts, and persistent shocks.

To investigate potential endogeneity within a block of macroeconomic variables, we implement the Wu–Hausman (WH) test for the five principal variables in each of the G7 economies (Wu, 1973; Hausman, 1978; Heinisch, Scaramella & Schult, 2025). The Wu-Hausman (WH) test for endogeneity compares the Ordinary Least Squares (OLS) and Instrumental Variable (IV) estimators: OLS is efficient under exogeneity, but becomes inconsistent if regressors are endogenous, while IV remains consistent in both cases. The WH statistic measures the difference between these estimators, which should be asymptotically zero if all regressors are exogenous, and the null hypothesis follows a chi-squared distribution. Using a 10% significance level, p-values below 0.10 lead to rejection of exogeneity—indicating possible feedback or simultaneity and confirming the need for IV estimation (Hausman, 1978; Wu, 1973). Country-specific statistics and p-values (presented in Table 3) provide evidence of endogeneity at this threshold, supporting treatment of the block of these macroeconomic variables as a jointly determined system and the use of the four uncertainty indices as external drivers in the forecasting framework.

Table 2: Global characteristics of the economic time series under study for G7 countries. For each country, statistics are reported for Unemployment Rate, REER, SIR, CPI Inflation, EPU, and GPR. Global variables include Oil Price (WTI), USEMV, and USMPU.

Country	Series	Skewness	Kurtosis	Nonlinearity	Seasonality	Stationarity	Long-Range Dependence	Outlier(s) Detected
Canada	Unemployment Rate	1.03	2.67	Non-linear	Seasonal	Non-stationary	0.81	4
	REER	0.57	-0.83	Linear	Non-seasonal	Non-stationary	1.02	0
	SIR	0.53	-0.67	Non-linear	Non-seasonal	Non-stationary	0.76	0
	CPI Inflation	1.63	4.43	Non-linear	Non-seasonal	Non-stationary	0.84	3
	EPU	-0.07	-0.88	Non-linear	Seasonal	Non-stationary	0.87	0
	GPR	4.92	36.34	Non-linear	Non-seasonal	Trend Stationary	0.91	4
US	Unemployment Rate	1.35	1.84	Non-linear	Seasonal	Non-stationary	0.76	2
	REER	0.01	-0.94	Linear	Non-seasonal	Non-stationary	1.00	0
	SIR	0.34	-1.51	Non-linear	Non-seasonal	Non-stationary	0.86	0
	CPI Inflation	1.19	3.17	Non-linear	Non-seasonal	Trend Stationary	0.67	0
	EPU	0.38	0.01	Linear	Seasonal	Non-stationary	0.78	0
	GPR	4.45	30.07	Non-linear	Non-seasonal	Trend Stationary	0.87	4
France	Unemployment Rate	0.59	-0.33	Non-linear	Seasonal	Non-stationary	1.17	0
	REER	0.21	-0.93	Linear	Seasonal	Non-stationary	0.93	0
	SIR	0.37	-0.77	Non-linear	Non-seasonal	Non-stationary	0.72	0
	CPI Inflation	1.51	3.34	Linear	Non-seasonal	Trend Stationary	0.71	0
	EPU	-0.61	-0.24	Non-linear	Seasonal	Non-stationary	0.88	1
	GPR	3.09	13.99	Linear	Non-seasonal	Trend Stationary	0.79	6
Germany	Unemployment Rate	0.06	-1.40	Non-linear	Seasonal	Non-stationary	0.88	0
	REER	0.95	0.58	Linear	Seasonal	Non-stationary	0.87	0
	SIR	0.10	-1.43	Non-linear	Seasonal	Non-stationary	0.75	0
	CPI Inflation	2.40	7.12	Non-linear	Non-seasonal	Non-stationary	0.66	5
	EPU	0.65	0.19	Non-linear	Non-seasonal	Non-stationary	0.85	0
	GPR	3.51	19.45	Linear	Non-seasonal	Non-stationary	0.87	5
Japan	Unemployment Rate	0.04	-1.18	Linear	Seasonal	Non-stationary	0.99	0
	REER	0.21	-0.51	Non-linear	Non-seasonal	Non-stationary	0.82	0
	SIR	3.39	16.94	Non-linear	Non-seasonal	Non-stationary	1.10	3
	CPI Inflation	1.04	0.94	Linear	Non-seasonal	Non-stationary	0.71	0
	EPU	0.34	0.15	Non-linear	Non-seasonal	Trend Stationary	0.62	0
	GPR	2.35	7.72	Linear	Non-seasonal	Trend Stationary	1.03	5
UK	Unemployment Rate	0.54	-0.93	Non-linear	Seasonal	Non-stationary	1.13	0
	REER	0.29	-1.52	Non-linear	Non-seasonal	Non-stationary	0.89	0
	SIR	0.14	-1.62	Non-linear	Non-seasonal	Non-stationary	0.84	0
	CPI Inflation	2.35	6.45	Non-linear	Non-seasonal	Non-stationary	0.87	4
	EPU	0.06	-0.91	Non-linear	Seasonal	Non-stationary	0.76	0
	GPR	4.19	24.36	Non-linear	Non-seasonal	Trend Stationary	0.95	6
Italy	Unemployment Rate	0.00	-0.89	Non-linear	Seasonal	Non-stationary	1.22	0
	REER	-0.24	0.22	Non-linear	Seasonal	Non-stationary	1.01	2
	SIR	1.14	0.91	Non-linear	Non-seasonal	Non-stationary	0.84	0
	CPI Inflation	1.92	5.58	Non-linear	Non-seasonal	Trend Stationary	0.69	4
	EPU	0.06	-0.91	Non-linear	Seasonal	Non-stationary	0.76	0
	GPR	2.22	6.73	Non-linear	Non-seasonal	Non-stationary	0.	6
Global	Oil Price (WTI)	0.32	-0.92	Non-linear	Non-seasonal	Non-stationary	1.06	0
	USEMV	2.34	8.29	Non-linear	Seasonal	Trend Stationary	0.79	4
	USMPU	1.66	3.93	Linear	Non-seasonal	Trend Stationary	0.90	2

Finally, we assess the structural stability of the VAR model using the cumulative sum (CUSUM) based on the OLS procedure (Ploberger & Krämer, 1992) over the five macroeconomic variables, with potential implications for both specification and evaluation (Stock & Watson, 1996; Hansen, 2001). Appendix A.1.3 presents Figure A.1.3, which reveals heterogeneous patterns. The unemployment rate clearly presents episodes of instability for several economies around some targeted episodes (early 2000s and post-2012 in parts of the sample). REER registers breaks in a few countries in the early 2000s. Oil prices exhibit common breaks across all G7 in 2008–2009 and 2014–2015, consistent with global commodity shocks (although none of the breaks are found to be statistically significant). Most critically, SIRs exhibit systematic structural breaks in Canada, the United States, France, and Japan, consistent with the zero lower bound, quantitative-easing phases, and subsequent normalizations (Bernanke, 2020; Yellen, 2017). By contrast, Germany, the United Kingdom, and Italy display CUSUM paths within the confidence bands, consistent with

relative parameter stability. CPI inflation is comparatively more stable across the G7, and the few near-significant breakpoints are consistent with credible inflation-targeting frameworks ([Mishkin & Schmidt-Hebbel, 2007](#); [Bernanke, Laubach, Mishkin & Posen, 2000](#)).

Table 3: Wu-Hausman test results for endogeneity of the five key macroeconomic variables across G7 countries. The test statistic and p-values are reported for each country. A small p-value (<0.10) indicates rejection of the null hypothesis (H_0) of exogeneity, providing evidence of endogeneity among regressors.

Country	WH Test Statistic	WH Test (p-value)	Insights
Canada	6.655	0.00004	Evidence against exogeneity (endogeneity detected)
US	3.044	0.01735	Evidence against exogeneity (endogeneity detected)
Germany	35.893	0.00000	Evidence against exogeneity (endogeneity detected)
France	12.491	0.00000	Evidence against exogeneity (endogeneity detected)
Japan	2.064	0.08514	Evidence against exogeneity (endogeneity detected)
UK	14.620	0.00000	Evidence against exogeneity (endogeneity detected)
Italy	13.297	0.00000	Evidence against exogeneity (endogeneity detected)

Overall, the descriptive statistics, structural diagnostics, endogeneity tests, and stability checks jointly motivate our modeling choices. The macro series are persistent, often nonlinear, and subject to structural change; the uncertainty indices are high-frequency and shock-driven; there is evidence of feedback within the block of macroeconomic variables and of regime shifts - particularly in SIRs. These features argue for multivariate, shrinkage-based methods that can handle high-dimensional systems; for state-dependent tools to trace transmission under different monetary environments; and for forecasting designs that evaluate both medium- and long-run horizons. This empirical groundwork justifies the inclusion and treatment of variables in the forecasting system and guides the specification of the multivariate framework applied to the data.

3. Causality Analysis

The time–frequency causality analysis to map how uncertainty relates to the macro indicators is presented in Section 3.1. Then, transmission paths are assessed using impulse response analysis in Section 3.2. These diagnostics provide the empirical basis for variable selection, clarify propagation channels, and motivate the multivariate forecasting framework applied in subsequent sections.

3.1. Wavelet Coherence Evidence for Uncertainty Transmission in the G7

In this section, we examine the causal relationships between our focal macroeconomic variables and exogenous uncertainty measures in the G7 economies in the time-frequency space using wavelet coherence analysis (WCA), which was originally proposed by [Grinsted, Moore & Jevrejeva \(2004\)](#) and subsequently extended with scale-wise False Discovery Rate (FDR) correction by [Aguiar-Conraria & Soares \(2014\)](#). WCA is used as a descriptive time-frequency diagnostic to identify co-movement and lead-lag patterns whose strength and direction may differ across horizons and over time; as such, the resulting maps are interpreted as evidence on transmission timing and relevant frequency bands, rather than as structural tests of causality. Wavelet coherence plots visualize time-frequency information, with information on the lead-lag (phase) between the two variables

being analyzed represented by the direction of the arrows in the plot. Arrows pointing to the right (left) represent an in-phase (anti-phase) relation between the pair of variables, whereas upward (downward) pointing arrows represent a leading (lagging) relation in time. The cone of influence boundaries, by identifying the areas where edge effects could undermine statistical reliability, ensure the reliable interpretation of the coherence patterns throughout the analysis. The scale-wise application of FDR correction, using the Benjamini-Hochberg procedure, systematically removes statistically unreliable coherence regions at each temporal scale independently.

As a result, only the segments of the estimated coherence function that are unlikely to be affected by spurious inferences of significant deviations from zero will be apparent in the resulting maps. In these maps, shades of red (coherence values near 1 on a 0-1 scale) indicate the time points and frequencies at which the lead-lag relationship between uncertainty shocks and fundamentals is most likely to be statistically significant. The scale-wise FDR screen removes visually attractive but fragile patches, thereby increasing confidence that remaining shaded regions reflect genuine time-frequency associations rather than multiple-testing artifacts.

We compute bivariate coherence between each uncertainty index and macro series for each country using the complex Morlet wavelet ($\omega_o = 6$), which is implemented through the **biwavelet R** package. To prevent look-ahead bias, the WCA is calculated only in the in-sample period (Jan, 1995 – Mar, 2022), which is distinct from the 24 months reserved for out-of-sample forecast validation. Statistical significance is determined using 1,000 Monte Carlo replications and controlled at each scale with a Benjamini–Hochberg FDR adjustment set at $\alpha = 10\%$; only regions that remain significant after FDR correction are shaded (see Appendix A.1.4 for a more technical discussion of the FDR-corrected WCA).

Figure 1 shows a canonical example (Canada), and similar grids of WCA for the remaining G7 appear in the Appendix A.1.4 (Figures A.1.4 – A.1.9). Three patterns stand out in the case of Canada. First, coherence is strongest at intermediate periodicities (multi-quarter to multi-year bands) and exhibits pronounced clustering around 2008–2009 and 2020–2021 (reflecting the global financial crisis and the COVID-19 pandemic, respectively). This dominance of the medium term is consistent with the shock-driven nature of the uncertainty indices as well as the stylized volatility documented for oil prices and exchange rates. Second, coherence is stronger in times of stress relative to tranquil periods, suggesting that uncertainty and macroeconomic aggregates co-move more closely when policy and financial conditions are tight. Phase information frequently reveals the uncertainty series leading the macroeconomic variables in these frequency bands. Third, frequency localization varies across variables: SIR and REER co-move with USEMV and USMPU at short-to-medium periodicities (consistent with policy and financial-expectations channels), whereas unemployment rate and CPI inflation exhibit more medium-to-long periodicities vis-à-vis EPU and GPR, suggesting a slower speed of real-side adjustment. Cross-country heterogeneity is nontrivial. For the United States, coherence between oil prices and uncertainty is especially strong; for the European economies, inflation–uncertainty coherence is pronounced; and for Japan, low-frequency REER synchronization is observed with US-centered uncertainty measures.

The existence of FDR-robust regions in which uncertainty drives macroeconomic variables justifies treating EPU, GPR, USEMV, and USMPU as exogenous drivers in the multivariate forecasting

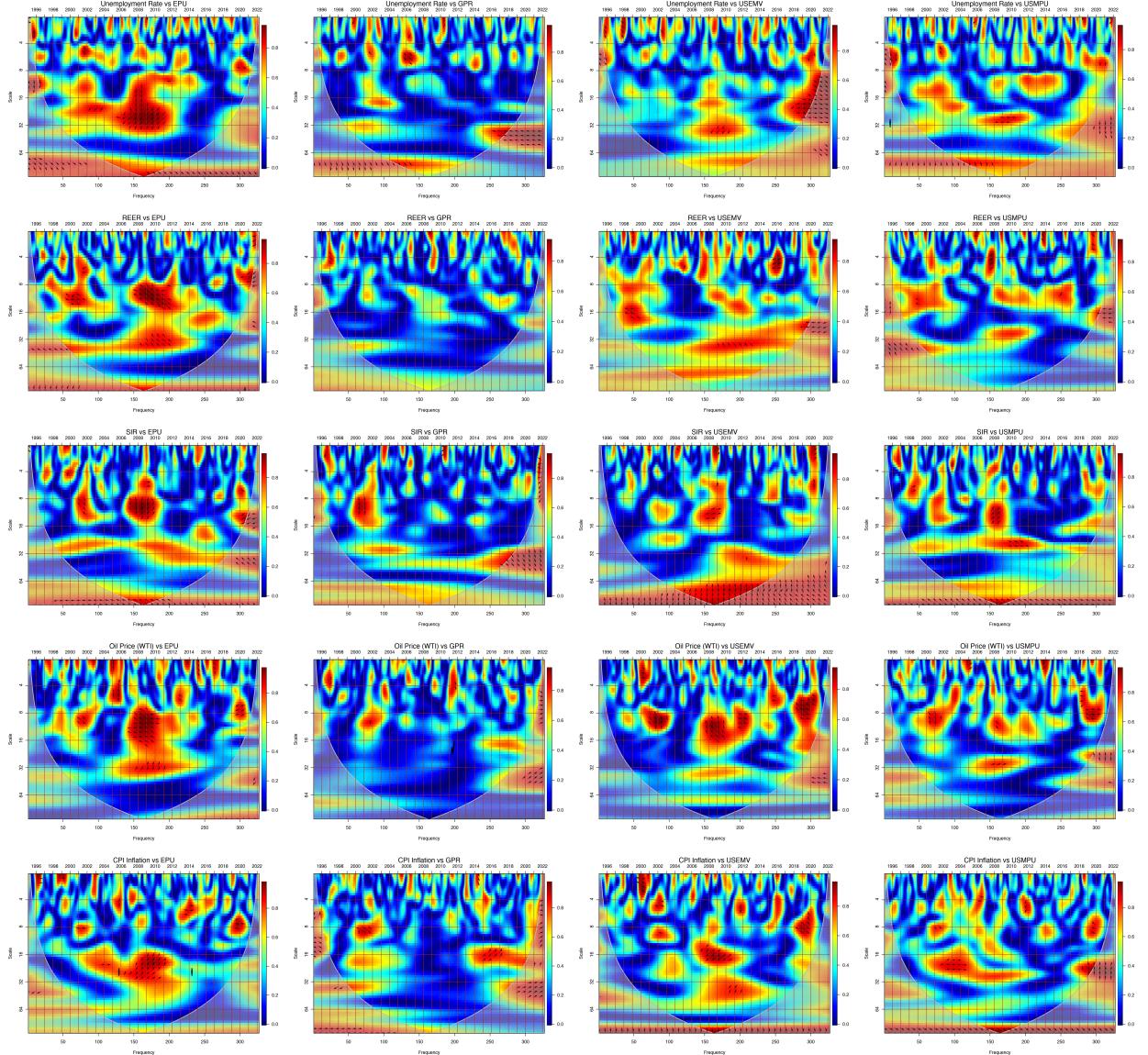


Figure 1: Scale-wise FDR-corrected wavelet coherence spectra for Canada (Jan, 1995 – Mar, 2022), displaying Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation against EPU, GPR, USEMV, and USMPU. Each subplot visualizes the time-frequency coherence between a macroeconomic variable (y-axis: Scale; x-axis: Frequency) and an uncertainty index, with warm colors (red/yellow) indicating intervals of high, statistically significant co-movement after FDR correction. The direction of phase arrows indicates lead-lag relationships: rightward arrows mean the series move in phase; leftward, in anti-phase; upward arrows signify the uncertainty index leads; downward, that it lags. Shaded regions inside the cone of influence reflect reliable coherence, while grid consistency enables robust cross-variable and cross-period comparison. The level of significance (α) for the FDR adjustment is set at 10%.

framework and thus complements the endogeneity evidence in the block of macroeconomic variables documented in Section 2.1 (Table 3). Overall, FDR-corrected WCA offers evidence on when and at what frequencies uncertainty shocks are associated with G7 macroeconomic aggregates, thus motivating an exogenous uncertainty block, stance conditioning, and horizon choice in the forecasting framework.

3.2. Impulse response analysis

This section focuses on the responses of the five key macroeconomic variables to the four newspaper-based uncertainty measures. The choice of the uncertainty drivers is based on the evidence presented in Section 3.1. This identification of exogenous uncertainty drivers informs the design of our empirical impulse-response analysis. We estimate impulse response functions using nonlinear local projections, as in Jordà (2005) (See Appendix A.1.5 for details). This methodology runs horizon-specific regressions that directly map innovations in the uncertainty measures to future values of each macro variable, without the need to invert a full VAR system or to impose strong parametric restrictions. We use the nonlinear extension described in Adämmer (2019), which employs a logistic transition function to generate state-contingent responses; observations are smoothly weighted by proximity to each regime, allowing both the size and timing of effects to depend on the monetary environment. As discussed in Section 3.1, the predominance of statistically significant coherence at medium-to-long periodicities suggests that dynamic interactions between uncertainty indices and financial or policy variables are most salient at these horizons. Guided by this frequency-localized coherence, we set both our key forecast horizons (12 and 24 months, corresponding to semi-long-term and long-term windows) and the maximal horizon for our impulse response analysis to align with the bands in which the strongest and most persistent co-movements and lead-lag relationships are identified. This ensures that our forecasting and IRF exercises are empirically grounded and tied directly to cycles supported by the underlying time-frequency evidence.

Lag orders of the endogenous macroeconomic variables (6 months) and exogenous uncertainty indicators (4 months) are chosen by using the Bayesian Information Criterion to account for persistence and delayed effects, respectively. Newey-West robust standard errors (Newey & West, 1987a) are used to compute the 95% confidence intervals, which are able to cope with time-varying volatility, especially during periods of crisis. Regime switching is lagged (`lag-switching = TRUE`) to ensure that the switching coincides with the pre-shock monetary environment. The IRF plots distinguish high-rate (blue) and low-rate (orange) regimes, while shaded regions represent 95% confidence bands. We compute horizon-by-horizon impulse responses up to 24 months. The smooth regime-switching is done by setting the regime probability following a logistic function, with curvature parameter $\gamma = 3$ to have a compromise between smooth transition and separation of the two regimes. Deterministic trends are turned off (`trend = 0`) to focus on cyclical features, and standardization of the short-term interest rate is done for each country to compare the magnitude across countries and over time. The country-specific IRFs are documented for Canada in Figures 2 - 3 and for the other six G7 economies in the Appendix A.1.6 (Figures A.1.10 - A.1.21), the United States (Figures A.1.10 and A.1.11), France (Figures A.1.12 and A.1.13), Germany (Figures A.1.14 and A.1.15), Japan (Figures A.1.16 and A.1.17), the United Kingdom (Figures A.1.18 and A.1.19) and Italy (Figures A.1.20 and A.1.21).

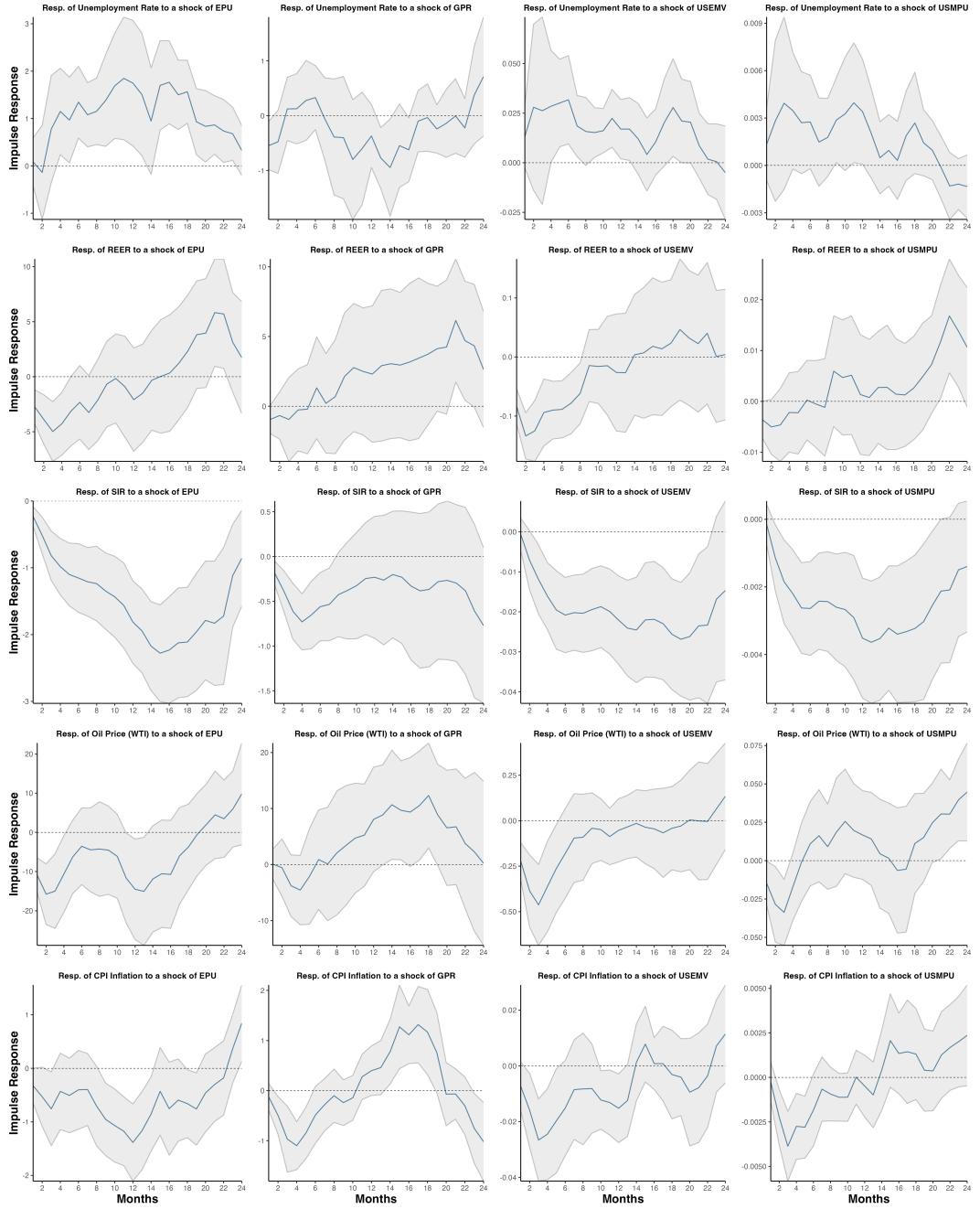


Figure 2: Impulse response functions (nonlinear local projections) of Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation to EPU, GPR, USEMV, and USMPU shocks for Canada (high-rate regime). Each row shows responses of a macro variable; each column, a distinct uncertainty shock. Solid blue lines depict IRFs with 95% confidence bands (grey). The x-axis is horizon (months), the y-axis the response magnitude. Compare responses by reading across rows (by variable) or down columns (by shock); zero response is marked by the dotted line.

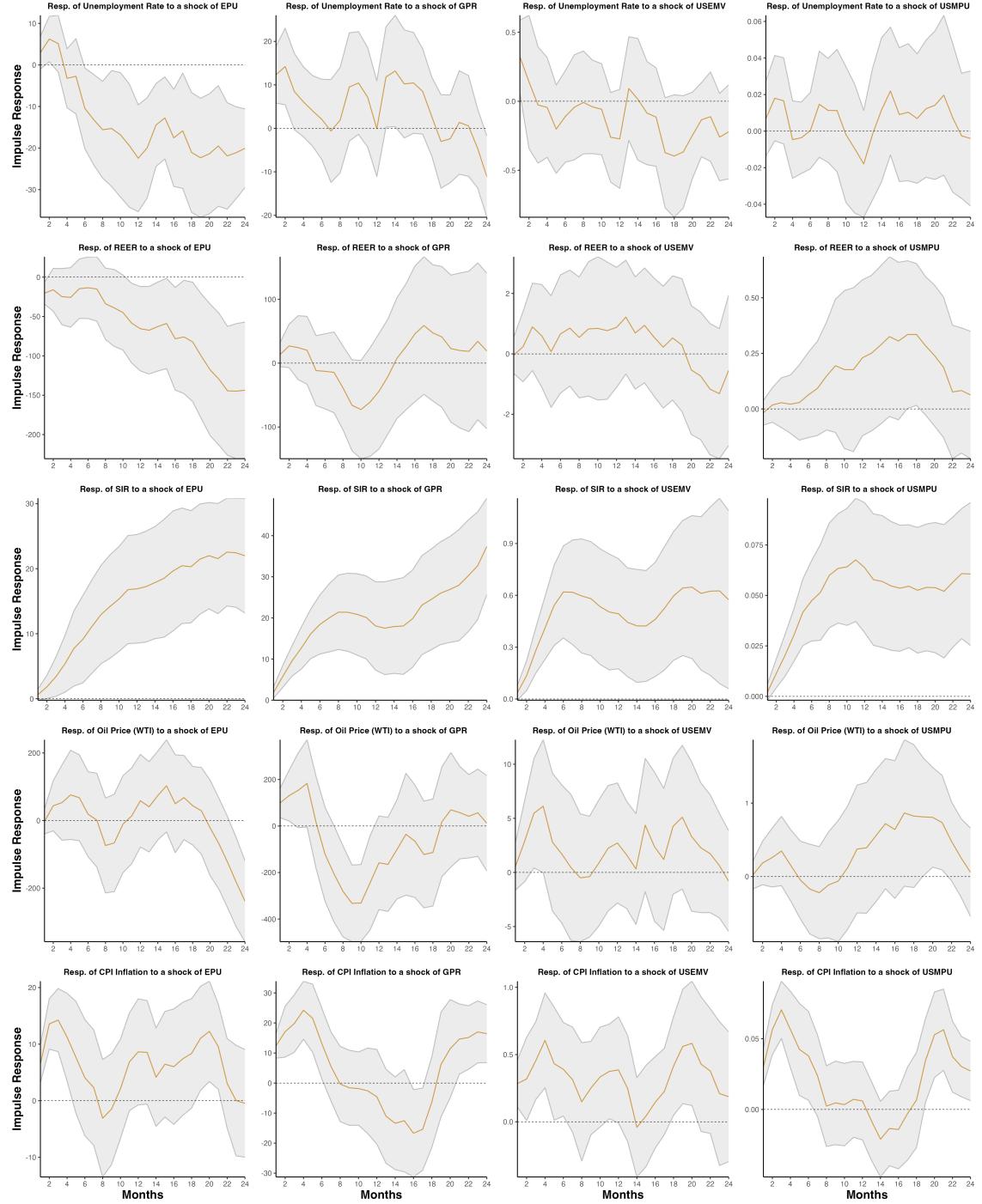


Figure 3: Impulse response functions (nonlinear local projections) of Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation to EPU, GPR, USEMV, and USMPU shocks for Canada (low-rate regime). Each row shows responses of a macro variable; each column, a distinct uncertainty shock. Solid orange lines depict IRFs with 95% confidence bands (grey). The x-axis is horizon (months), the y-axis the response magnitude. Compare responses by reading across rows (by variable) or down columns (by shock); zero response is marked by the dotted line.

The country-specific plots show responses (EPU, GPR, USEMV, USMPU) of uncertainty shocks for five macroeconomic variables. Dotted horizontal lines represent zero baselines, and shaded bands are 95% confidence intervals. The sample period is January, 1995 –Mar, 2024, where gradual changes in policy transmission are also captured through state-contingent responses across different monetary regimes ([Jordà, 2005](#)). The country grids display the impulse responses of the unemployment rate, REER, SIR, oil price (WTI), and CPI inflation to EPU, GPR, USEMV, and USMPU under the high- and low-rate regimes. We use the SIR as the regime-switching variable in the local projections framework for three reasons. First, following the documented evidence on the G7 in aggregate, the main episodes of monetary policy actions in each country before and after 2020 have also been associated with clear structural breaks in each country’s short rate dynamics, both in the mean-reversion and in the volatility properties of short rates. The same is true for both the COVID shock and recovery episodes, where reallocation in rate dynamics and regime-dependent volatility are particularly large (think, e.g., of the US taper tantrum episode or the ECB PEPP announcement) ([Darracq Pariès, Kornprobst & Priftis, 2024](#)). Second, the available evidence on nonlinear rate adjustments, and in particular on the documented asymmetries in short-term rate responses to inflation, output, and the term spread, are better accounted for by regime-switching specifications rather than linear models ([Haug & Siklos, 2006](#); [Ang & Bekaert, 2002](#)). Third, the OLS-CUSUM diagnostic suggests the presence of clear structural breaks in several G7 countries’ short rates. The zero-lower-bound and quantitative easing episodes, as well as the subsequent policy normalizations, broadly line up with those. Overall, these pieces of evidence jointly provide strong support for the use of the short rate as the regime threshold in the nonlinear local projections design.

For Canada (Figures [2](#) and [3](#)), IRFs are characterized by pronounced regime asymmetry. In high-rate states, EPU leads to transitory unemployment spikes alongside REER depreciation and a tightening bias in SIR, whereas GPR reinforces this pattern with sizable rate hikes and more volatile CPI inflation. In low-rate environments, similar shocks are met with accommodation and currency appreciation, but unemployment increases. USEMV effects are more muted, whereas USMPU highlights Canada’s financial integration with the United States through divergent REER and CPI inflation paths. The sign of CPI inflation responses also reverses with stance—positive in high-rate states and negative in low-rate states—pointing to regime-dependent expectations around imported prices and domestic slack; this is in line with asymmetric transmission in Canada ([Bhuiyan & Lucas, 2007](#)). Stance dependence is salient for the United States (Figures [A.1.10](#) and [A.1.11](#)). At high rates, EPU raises unemployment at its peak in month 14, coincides with large negative oil price responses during months 14–16, and drives near-term increases in CPI inflation with peaks at months 5–7 and again near month 10. GPR appreciates the REER, particularly via safe-haven flows. Low rates lead to amplified and more persistent real effects of EPU, keeping unemployment higher through month 16 while also generating persistent negative responses in CPI inflation. Interest rates respond contractionarily to GPR at high rates but accommodate at low rates—consistent with [Lhuissier & Tripler \(2021\)](#) and the “toxic interaction” between uncertainty and tight financial conditions ([Caldara, Fuentes-Albero, Gilchrist & Zakrjšek, 2016](#)). The impulse responses for France (Figures [A.1.12](#) and [A.1.13](#)) indicate that in the high-rate states, EPU leads to higher unemployment (peaking around months 18–20), higher REER (appreciation), and lower oil prices (peaking around month 12). In addition, CPI inflation disinflates initially but turns modestly positive after month 16. GPR shocks result in higher policy rates and lower CPI inflation, with the peak (negative) inflation effects materializing in months 2–4. In contrast, in the low-rate

regime, unemployment's response to EPU shocks is more gradual, yet it remains elevated through month 20. Policy rates fall more noticeably during months 12–15. Furthermore, CPI inflation responses are consistently negative throughout the forecast horizon. These results, which echo those for the euro area as a whole, are consistent with those reported by [Castelnuovo \(2019\)](#) for European countries, where they observe that uncertainty appears to have different transmission channels depending on the current policy rate environment. The impulse responses for the German economy (Figures A.1.14 and A.1.15) are also consistent with safe-haven effects. An increase in GPR raises unemployment, and this effect is persistent (remaining positive until month 20). The REER appreciates, peaking at about month 10. The response of interest rates to GPR turns positive after an initial period of accommodation and persists for the whole sample period. The response of CPI inflation is negative at first but becomes positive after month 10. All these features are consistent with a safe-haven role, also characterized by a credible commitment to price stability. The impulse responses for the low-rate state are instead different: unemployment jumps up following a shock to EPU, with the largest response at its peak, and then declining. The response of the REER to EPU is hump-shaped, peaking at months 10–12. The response of oil prices to EPU turns positive after an initial period of negative values. The response of CPI inflation is mostly countercyclical in the low-rate state, but it turns procyclical in the high-rate state ([Berg, 2019](#); [Grimme & Stöckli, 2018](#)). The IRFs for Japan (Figures A.1.16 and A.1.17) show state-dependent transmission and strong exchange-rate channels. High-rate episodes are characterized by EPU-induced oscillatory unemployment dynamics, a peak effect around month 18, and sizeable delayed REER appreciation after month 12, all in line with the safe-haven properties of the yen and the Bank of Japan's gradualist policy stance. Low-rate episodes are instead associated with more muted unemployment dynamics, more contained USEMV responses, and a switch from positive to negative CPI inflation responses—aligning with Japan's deflationary bias and the structural labor-market frictions highlighted by [Miyamoto \(2016\)](#). The larger responses in size and volatility under tight policy conditions are also consistent with [Lhuissier & Tripier \(2021\)](#)'s result that financial frictions amplify the effects of uncertainty when monetary policy has little room to maneuver. In the United Kingdom (Figures A.1.18 and A.1.19), an exchange-rate–inflation nexus is quite evident. In the high-rate regime, EPU increases unemployment (highest response around month 14) and is associated with persistent REER depreciation; CPI inflation displays an initial inflationary pulse followed by persistent disinflation (lasting to month 12). Interest rates act in a counter-cyclical way, with persistent negative responses, which can be interpreted as consistent with a “stabilization first” rather than a pure inflation-targeting policy in the face of geopolitical tensions. In the low-rate regime, unemployment responds to EPU more persistently, USEMV elicits a hump-shaped oil price response (with the maximum effect around month 16), and CPI inflation disinflates throughout; GPR displays persistent negative pressure on prices. These asymmetries are also consistent with the United Kingdom's nonlinear projections in ([Hauzenberger, Huber, Marcellino & Petz, 2025](#)) and with a heightened sensitivity of the United Kingdom's financial conditions, and their impact on the economy, in situations where policy space is more limited ([Cesa-Bianchi, Thwaites & Vicondoa, 2020](#)). For Italy (Figures A.1.20 and A.1.21), the difference between regimes is more in the nature of real-side adjustment: under high rates, EPU causes unemployment to increase moderately and persistently over months 12–16, the REER to appreciate, and oil prices to fall with large negative responses. There is some negative response for CPI inflation initially before a mild recovery. Low-rate effects of the same shocks have more persistent unemployment (running through to month 16), REER depreciation, more muted negative interest rate responses, and larger oil price declines; CPI

inflation is negative throughout with only a tentative recovery at the end. This is in line with recent euro-area evidence on real effects of uncertainty and limited offset from accommodation ([Bobasu, Quaglietti & Ricci, 2024](#)).

To summarize, the IRF grids reveal three broad patterns that are consistent across G7 economies. First, monetary stance matters: conditional on a given uncertainty shock, responses are typically larger and faster in the high-rate regime than in the low-rate regime, a pattern that WCA analysis also suggests by revealing clusters of statistically significant coherence around episodes with sharp movements in the SIR. This finding further justifies the use of the SIR as a transition variable in the state-dependent impulse-response design. Second, the policy-distinct measures that incorporate external risks (GPR, USEMV) have strong transmission through REER and oil price (WTI) with secondary effects on CPI inflation and unemployment at medium horizons; the policy-focused measures (EPU, USMPU) exhibit clearer connections to SIR, CPI inflation, and unemployment rate. Third, cross-country heterogeneity is significant - exchange rate channels are particularly important where financial-market or energy exposures are large, while real-side adjustments predominate elsewhere. These patterns motivate the forecasting design in what follows: incorporating the same set of uncertainty measures as exogenous drivers in the multivariate model and conditioning the evaluations on the monetary environment to capture the state-dependent transmission documented here.

4. SZBVARx Model: Methodology

SZBVARx provides a flexible structure for multivariate forecasting, extending the classical VAR approach by integrating Bayesian inference and allowing for exogenous variables. The model jointly estimates parameters and latent components, ensuring robust forecast performance in the presence of complex economic dynamics and uncertainty. The methodology builds upon the seminal work of [Sims & Zha \(1998\)](#) and extends the traditional Bayesian VAR framework to incorporate exogenous uncertainty shocks, following recent advances in [Carriero, Clark & Marcellino \(2015\)](#) and [Chan \(2020\)](#). We employ the SZBVARx framework for multivariate macroeconomic forecasting across G7 economies.

We specify the SZBVARx model as a reduced-form vector autoregression that jointly models m endogenous macroeconomic variables while incorporating k exogenous uncertainty measures. The model captures dynamic interdependencies among macroeconomic variables through autoregressive lags of order p , while allowing uncertainty shocks to contemporaneously affect all endogenous variables. The general specification takes the form:

$$y_t = \mu + \sum_{i=1}^p \Phi_i y_{t-i} + \Gamma x_t + u_t, \quad u_t \sim \mathcal{N}(0, \Sigma). \quad (1)$$

In this context, $y_t = [y_{1,t}, \dots, y_{m,t}]^\top \in \mathbb{R}^{m \times 1}$ denotes the m -dimensional vector of endogenous macroeconomic variables, $\mu \in \mathbb{R}^m$ represents the m -dimensional vector of intercept terms, $x_t = [x_{1,t}, \dots, x_{k,t}]^\top \in \mathbb{R}^{k \times 1}$ is a k -dimensional vector of exogenous uncertainty measures, and $u_t \in \mathbb{R}^m$ represents a white noise vector. The autoregressive coefficient matrices $\Phi_i \in \mathbb{R}^{m \times m}$, $\forall i = \{1, \dots, p\}$

capture cross-variable spillovers within the macroeconomic system. The matrix $\Gamma \in \mathbb{R}^{m \times k}$ measures the impact of macroeconomic uncertainty shocks on each endogenous variable, such that $\Gamma x_t \in \mathbb{R}^{m \times 1}$ represents the contemporaneous effect of the k exogenous uncertainty measures on the m endogenous variables. The covariance matrix $\Sigma \in \mathbb{R}^{m \times m}$ is positive definite and captures the contemporaneous covariance structure of the reduced-form errors. This specification can be expressed compactly by defining the regressor vector $Z_t = [1, y_{t-1}^\top, \dots, y_{t-p}^\top, x_t^\top]^\top \in \mathbb{R}^{d \times 1}$ and coefficient matrix $B = [\mu^\top, \Phi_1^\top, \dots, \Phi_p^\top, \Gamma^\top]^\top \in \mathbb{R}^{d \times m}$, yielding:

$$y_t = B^\top Z_t + u_t. \quad (2)$$

The regressor dimension $d = 1 + mp + k$ reflects the intercept term, mp autoregressive coefficients, and k exogenous variable coefficients. For estimation purposes, we stack observations over $t = 1, \dots, T$ to obtain the matrix form $Y = ZB + U$, where the sum of squared errors $\text{SSE}(B) = (Y - ZB)^\top(Y - ZB)$ forms the basis for likelihood construction. The SZBVARx framework is particularly suited for policy analysis, as it distinguishes between endogenous macroeconomic responses and exogenous uncertainty-driven fluctuations, enabling policymakers to assess the relative importance of domestic versus external uncertainty sources in driving macroeconomic volatility.

4.1. Prior Specifications

Bayesian inference in high-dimensional macroeconomic forecasting critically depends on the specification of prior distributions for model parameters. Following Sims & Zha (1998), we place a matrix-normal distribution on the coefficient matrix B conditional on the covariance matrix Σ , and an inverse-Wishart prior on Σ . Let $\beta = \text{vec}(B) \in \mathbb{R}^{md \times 1}$ denote the vectorized coefficient matrix, where $d = 1 + mp + k$ is the regressor dimension and $B \in \mathbb{R}^{d \times m}$. The prior specification takes the form:

$$B | \Sigma \sim \mathcal{MN}_{d,m}(B_0, \lambda_0^2 \Omega_0, \Sigma) \quad (3)$$

$$\Sigma \sim \mathcal{IW}(\Psi_0, \nu_0) \quad (4)$$

which is equivalent to the multivariate-normal vectorized form:

$$\beta | \Sigma \sim \mathcal{N}(\beta_0, \Sigma \otimes \lambda_0^2 \Omega_0) \quad (5)$$

where $\beta_0 = \text{vec}(B_0) \in \mathbb{R}^{md \times 1}$ represents the prior mean for the vectorized coefficients, $\Omega_0 \in \mathbb{R}^{d \times d}$ is a positive definite matrix encoding the prior covariance over regressors, $\Psi_0 \in \mathbb{R}^{m \times m}$ is the prior scale matrix for the covariance structure, $\nu_0 > m + 1$ denotes the prior degrees of freedom (ensuring finite prior mean $\mathbb{E}[\Sigma] = \Psi_0 / (\nu_0 - m - 1)$), and $\lambda_0 > 0$ is the overall tightness parameter controlling the strength of shrinkage toward the prior mean.

The SZBVARx framework adopts this conjugate *Matrix-Normal-Inverse-Wishart prior* as it provides computational efficiency through closed-form posterior distributions while allowing flexible control over parameter shrinkage. This prior specification is particularly well-suited for high-dimensional macroeconomic forecasting, as it naturally accommodates the structure of the reduced-form VAR system defined in equation (2). The Kronecker product structure $\Sigma \otimes \lambda_0^2 \Omega_0$ in equation (5) propagates cross-equation dependence via Σ , while structure in Ω_0 delivers lag-/regressor-specific

tuning; equation-specific tuning can be obtained by column-side scaling (e.g., equation-specific λ_j or diagonal rescaling of Σ), all while preserving conjugacy (Baíbura et al., 2010). Smaller values of λ_0 impose tighter shrinkage toward β_0 , effectively regularizing the model in high-dimensional settings. This conjugate prior structure directly shapes the model's ability to balance signal extraction and shrinkage, which is critical for reliable multivariate forecasting in the presence of complex economic interdependencies and uncertainty shocks. Informative, structured priors such as the Matrix-Normal–Inverse-Wishart allow for nuanced modeling of heterogeneity and regularization, enhancing stability and forecast reliability in complex, multi-country analyses. Given its relative strength in balancing computational tractability, forecast accuracy, and robustness to model misspecification, we incorporate this prior specification in the flexible SZBVARx framework for G7 macroeconomic forecasting¹

4.2. Hyperparameter Specification

The SZBVARx framework employs a systematic hyperparameter specification following Sims & Zha (1998) to operationalize the prior distributions described above, with parameters selected through grid search optimization detailed in Section 5.3. To accommodate the heterogeneous dynamics observed across G7 economies and forecast horizons, all hyperparameters are specified with country and horizon-specific scaling, denoted by superscripts (c, h) where c indexes countries and h denotes forecast horizons. The hyperparameter tuple $(p, \lambda_0, \lambda_1, \lambda_3, \lambda_4, \lambda_5, \mu_5, \mu_6)$ controls key aspects of model specification and prior shrinkage²: the lag length $p^{(c,h)} \in \{1, 2, \dots, p_{\max}\}$ determines the autoregressive order, while $\lambda_0^{(c,h)} \geq 0$ governs overall prior tightness as introduced in the Matrix-Normal–Inverse-Wishart specification above. The relative shrinkage parameters $\lambda_1^{(c,h)} \geq 0$ (own versus cross-variable lag tightness), $\lambda_3^{(c,h)} \geq 0$ (lag decay with harmonic decay when $\lambda_3 = 1$), $\lambda_4^{(c,h)} \geq 0$ (intercept shrinkage), and $\lambda_5^{(c,h)} \geq 0$ (exogenous coefficient shrinkage) provide granular control over different coefficient groups, reflecting the economic intuition that own lags typically matter more than cross-variable effects and that coefficient importance generally declines with lag length. The dummy observation weights $\mu_5^{(c,h)} \geq 0$ and $\mu_6^{(c,h)} \geq 0$ implement additional shrinkage through sum-of-coefficients and initial-conditions constraints, with the former imposing the economically motivated restriction $\sum_{i=1}^p \Phi_i \approx I_m$ that shrinks the system toward unit root behavior rather than enforcing row-sum constraints (Sims & Zha, 1998; Waggoner & Zha, 2003). The Matrix-Normal–Inverse-Wishart prior was employed uniformly across all G7 countries and forecast horizons (semi-long-term and long-term), as it emerged as the superior choice in preliminary grid-search experiments. This prior's dominance reflects its ability to balance computational tractability through conjugacy, flexible shrinkage control via the Kronecker structure $\Sigma \otimes \lambda_0^2 \Omega_0$, and proper uncertainty

¹The SZBVARx framework supports three prior specifications: (i) the Matrix-Normal–Inverse-Wishart (MN-IW) prior described above, (ii) a semi-conjugate flat-Gaussian prior, and (iii) a non-informative flat-flat prior. In our empirical implementation for G7 macroeconomic forecasting, the MN-IW prior was unanimously selected via grid-search optimization across all countries and forecast horizons, demonstrating its superior performance for this application. Consequently, the prior type is not treated as a tunable hyperparameter in Section 4.2, and all results reported herein employ the MN-IW specification.

²Note that λ_2 , which appears in some Minnesota prior specifications to control cross-equation shrinkage, is omitted in the Sims-Zha Normal-Wishart framework as cross-equation dependencies are captured through the Kronecker structure $\Sigma \otimes \lambda_0^2 \Omega_0$ of the prior covariance matrix.

quantification through the inverse-Wishart specification on Σ —features particularly valuable for capturing the complex interdependencies and heterogeneous dynamics characteristic of G7 macroeconomic systems.

4.3. Posterior Distributions and MCMC Inference

In Bayesian macroeconomic forecasting, the posterior distribution reflects updated prior beliefs about model parameters after incorporating observed data, providing a natural framework for uncertainty quantification that is particularly valuable in policy applications where parameter uncertainty can significantly affect forecast intervals and decision-making. Unlike classical approaches that treat parameters as fixed unknowns, the Bayesian framework acknowledges that macroeconomic relationships evolve over time and that our knowledge of these relationships is inherently clouded by uncertainty, making posterior distributions essential for robust inference in multivariate forecasting, where the curse of dimensionality amplifies parameter uncertainty. Given data (Y, Z) where $Z \in \mathbb{R}^{T \times d}$ is the regressor matrix and a chosen prior, Bayesian inference for the SZBVARx model proceeds by updating prior beliefs about model parameters using the reduced-form Gaussian likelihood:

$$\mathcal{L}(B, \Sigma | Y, Z) \propto |\Sigma|^{-T/2} \exp \left[-\frac{1}{2} \text{tr} \left(\Sigma^{-1} (Y - ZB)^\top (Y - ZB) \right) \right], \quad (6)$$

where $B \in \mathbb{R}^{d \times m}$ and $\Sigma \in \mathbb{R}^{m \times m}$. Under the conjugate Matrix-Normal–Inverse-Wishart prior specified in Equations (3)–(4), the posterior distributions retain the same functional form. Because of full conjugacy, the joint posterior factorizes as

$$p(\Sigma, B | Y, Z) = p(\Sigma | Y, Z)p(B | \Sigma, Y, Z), \quad (7)$$

yielding closed-form conditional posteriors:

$$B | \Sigma, Y, Z \sim \mathcal{MN}_{d,m}(\bar{B}, \bar{\Omega}, \Sigma), \quad (8)$$

$$\Sigma | Y, Z \sim \mathcal{IW}(\bar{\Psi}, \bar{\nu}), \quad (9)$$

with posterior hyperparameters:

$$\bar{\Omega}^{-1} = (\lambda_0^2 \Omega_0)^{-1} + Z^\top Z, \quad (10)$$

representing the posterior precision matrix for the coefficient covariance that combines prior information with the observed regressor cross-product;

$$\bar{B} = \bar{\Omega} \left[(\lambda_0^2 \Omega_0)^{-1} B_0 + Z^\top Y \right], \quad (11)$$

the posterior mean coefficient matrix that optimally weights the prior mean B_0 and the data-driven OLS estimate;

$$\bar{\Psi} = \Psi_0 + (Y - Z\bar{B})^\top (Y - Z\bar{B}) + (\bar{B} - B_0)^\top (\lambda_0^2 \Omega_0)^{-1} (\bar{B} - B_0), \quad (12)$$

the posterior scale matrix for Σ that accumulates prior scale, observed sum-of-squares, and adjustment terms reflecting the updating from prior to posterior mean; and

$$\bar{\nu} = \nu_0 + T, \quad (13)$$

the posterior degrees of freedom that increase with sample size T . Equivalently, in vectorized form:

$$\beta \mid \Sigma, Y, Z \sim \mathcal{N}(\bar{\beta}, \Sigma \otimes \bar{\Omega}), \quad (14)$$

where $\beta = \text{vec}(B)$ and $\bar{\beta} = \text{vec}(\bar{B})$.

Markov Chain Monte Carlo (MCMC) inference can proceed via two equivalent approaches. For direct sampling, we draw independently and identically distributed (i.i.d.) posterior samples: for $s = 1, \dots, S$, draw $\Sigma^{(s)} \sim \mathcal{IW}(\bar{\Psi}, \bar{\nu})$ and then $B^{(s)} \sim \mathcal{MN}_{d,m}(\bar{B}, \bar{\Omega}, \Sigma^{(s)})$, yielding i.i.d. posterior draws. Alternatively, for a Gibbs sampler with conditional dependence on the current draw, we alternate $B^{(s)} \sim \mathcal{MN}_{d,m}(\bar{B}, \bar{\Omega}, \Sigma^{(s-1)})$ and $\Sigma^{(s)} \sim \mathcal{IW}(\Psi_0 + (Y - ZB^{(s)})^\top(Y - ZB^{(s)}), \nu_0 + T)$. Both approaches are correct and yield a sequence of parameter draws $\{(B^{(s)}, \Sigma^{(s)})\}_{s=1}^S$ whose empirical distribution approximates the joint posterior, enabling Bayesian credible prediction intervals and posterior predictive inference essential for uncertainty quantification in high-dimensional G7 macroeconomic applications with country and horizon-specific hyperparameters configurations.

4.4. Forecast Generation and Stability

The SZBVARx framework generates forecasts through posterior predictive simulations that naturally incorporate the future shock uncertainty, providing comprehensive forecast distributions essential for policy analysis under uncertainty. This approach leverages the full posterior distribution $\{B^{(s)}, \Sigma^{(s)}\}_{s=1}^S$ obtained from the Gibbs sampling algorithm to produce forecasts that reflect all sources of uncertainty inherent in multivariate macroeconomic systems. For each MCMC draw s and forecast horizon h , the recursive forecast equation is:

$$y_{T+h|T}^{(s)} = \mu^{(s)} + \sum_{i=1}^p \Phi_i^{(s)} y_{T+h-i|T}^{(s)} + \Gamma^{(s)} x_{T+h} + u_{T+h}^{(s)}, \quad (15)$$

where $u_{T+h}^{(s)} \sim \mathcal{N}(0, \Sigma^{(s)})$ represents future innovations and $y_{T+j|T}^{(s)} = y_{T+j}$ for $j \leq 0$. When future uncertainty measures x_{T+h} are unobserved, we condition on policy scenarios or employ auxiliary forecasting models with uncertainty propagation. The point forecast is computed as the posterior mean:

$$\hat{y}_{T+h|y_T} = \frac{1}{S} \sum_{s=1}^S y_{T+h|y_T}^{(s)}, \quad (16)$$

while credible intervals utilize empirical quantiles of the forecast distribution $\{y_{T+h|y_T}^{(s)}\}_{s=1}^S$. System stability is ensured by verifying that all eigenvalues of the companion matrix

$$\mathbf{C} = \begin{bmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_{p-1} & \Phi_p \\ I_m & 0 & \cdots & 0 & 0 \\ 0 & I_m & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_m & 0 \end{bmatrix}$$

satisfy $|\rho_i| < 1$, $\forall i = \{1, \dots, mp\}$, with forecasts retained only for parameter draws satisfying this stability condition to ensure well-defined long-run dynamics ([Lütkepohl, 2005](#)). This stability check is particularly important in high-dimensional G7 applications where the interaction between domestic variables and uncertainty shocks can potentially generate explosive behavior without proper regularization through the country and horizon-specific prior specifications.

5. Experimental Evaluation and Analysis of Results

This section empirically benchmarks the proposed SZBVARx against a wide variety of baseline models, including classical econometric models, machine learning models, and deep learning models. The experimental setup aims to evaluate the generalization capability and effectiveness of each model in a more realistic macroeconomic environment. The target prediction task is to forecast the conditional changes in the next h months ($h \in \{12, 24\}$) of the five endogenous variables (Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation) in the presence of the four exogenous uncertainty shocks (EPU, GPR, USEMV, and USMPU). The goal is to evaluate the forecasting capability in both the semi-long-term (12 months) and long-term (24 months) horizons to provide a comprehensive evaluation of the model's performance in macroeconomic prediction tasks. The rest of this section is structured as follows. Section 5.1 introduces the baseline algorithms. Section 5.2 provides details on the evaluation metrics. Section 5.3 presents and discusses the experimental results. Section 5.4 explores the statistical significance of the forecasting performance gaps between different models, while Section 5.5 quantifies the uncertainty in predictive distributions through credible intervals (see Appendix A.1.9 for credible intervals of the 12-month ahead forecast horizon).

5.1. Baseline Models

To properly evaluate SZBVARx, we select an extended set of baseline models, representing traditional and contemporary approaches used in macroeconomic forecasting that can accommodate exogenous variables. We use *VARx*, *Threshold VAR (TVar)*, and *Vector exponential smoothing (VES)* as our baselines from the classical econometric literature. To benchmark against modern machine learning and deep learning, we consider *Neural Basis Expansion Analysis for Time Series with exogenous variables (NBEATSx)*, *Neural Hierarchical Interpolation for Time Series Forecasting with exogenous covariates (NHITSx)*, *CatBoost with exogenous covariates (CatBoostx)* and *LightGBM with exogenous covariates (LGBMx)*, *Xtreme Gradient Boosting with exogenous covariates (XGBx)*, *Temporal Fusion Transformer with exogenous covariates (TFTx)*, *Time-Series Dense Encoder with exogenous covariates (TiDEx)*, *Block recurrent neural networks with exogenous covariates (BRNNx)*, *Time-Series Mixer with exogenous covariates (TSMIXERx)*, and *DishTS ensembles augmented by CatBoostx and XGBx (DTS_CBx, DTS_XGBx)*. All the machine learning and neural models can take uncertainty shocks as covariates. We provide additional details and exogenous feature processing for each model in Appendix A.1.7.

5.2. Evaluation Metrics

To rigorously assess the forecasting accuracy and robustness of the SZBVARx, we use five standard metrics: Root Mean Square Error (RMSE), Symmetric Mean Absolute Percentage Error (SMAPE), Mean Absolute Scaled Error (MASE), Theil's U_1 (denoted as TU_1), and Median Absolute Percentage Error (MDAPE). RMSE is scale-dependent; SMAPE and MDAPE are percentage-based; MASE rescales errors by the in-sample seasonal naïve benchmark; and TU_1 is a scale-normalized index in $[0, 1]$. The mathematical formulations are:

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{1}{h} \sum_{t=1}^h (y_t - \hat{y}_t)^2}; & \text{MASE} &= \frac{\sum_{t=D+1}^{D+h} |\hat{y}_t - y_t|}{\frac{h}{D-S} \sum_{t=S+1}^D |y_t - y_{t-S}|}; \\ \text{SMAPE} &= \frac{1}{h} \sum_{t=1}^h \frac{|\hat{y}_t - y_t|}{(|\hat{y}_t| + |y_t|)/2} \times 100\%; & TU_1 &= \frac{\sqrt{\frac{1}{h} \sum_{t=1}^h (y_t - \hat{y}_t)^2}}{\sqrt{\frac{1}{h} \sum_{t=1}^h y_t^2} + \sqrt{\frac{1}{h} \sum_{t=1}^h \hat{y}_t^2}}; \\ \text{MDAPE} &= \text{median}_{t=1,\dots,h} \left(\frac{|y_t - \hat{y}_t|}{|y_t|} \right) \times 100\%. \end{aligned}$$

Here, y_t denotes the observed value at time t , \hat{y}_t the forecast, h the forecast horizon, and D, S the in-sample length and seasonal period used in MASE. Because no single metric suffices in macrocasting settings with persistence, breaks, and exogenous shocks, this suite jointly assesses magnitude accuracy (RMSE), percentage accuracy (SMAPE, MDAPE), scale-free performance relative to a seasonal naïve benchmark (MASE), and scale-normalized error independent of units (TU_1). Conventionally, the model minimizing a given error metric is regarded as the top performer ([Härdle & Athanasopoulos, 2018](#); [Sengupta et al., 2025](#); [Hewamalage, Ackermann & Bergmeir, 2023](#); [Komunjer & Owyang, 2012](#)).

5.3. Experimental results and baseline comparison

We evaluate SZBVARx on the joint dynamics of five stated endogenous indicators augmented with four exogenous uncertainty shocks. Hyperparameters are tuned by an explicit grid search in **R** using MSBVAR's `szbvar` routine and custom orchestration, selecting the tuple $(p, \lambda_0, \lambda_1, \lambda_3, \lambda_4, \lambda_5, \mu_5, \mu_6)$ that minimizes a rolling multivariate RMSE on the training sample under an expanding-window evaluation for each forecast horizon $h \in \{12, 24\}$. For a candidate hyperparameter vector θ , denote $e_{t+h|t}(\theta) = y_{t+h} - \hat{y}_{t+h|t}(\theta)$ and \mathcal{T}_h the set of training-origin dates used in the grid search. We define the Multivariate Root Mean Square Error (MRMSE) as:

$$\text{MRMSE}_h(\theta) = \sqrt{\frac{1}{m |\mathcal{T}_h|} \sum_{t \in \mathcal{T}_h} \|e_{t+h|t}(\theta)\|_2^2}, \quad (17)$$

where $m = 5$ is the number of endogenous variables and $\|\cdot\|_2$ denotes the Euclidean norm. We select $\hat{\theta}_h = \arg \min_{\theta} \text{MRMSE}_h(\theta)$. This procedure respects the time-series ordering and avoids peek-ahead bias ([Hamilton, 1994](#); [Lütkepohl, 2005](#)). The lag length p captures the persistence structure

of macroeconomic relationships, while shrinkage parameters $(\lambda_0, \lambda_1, \lambda_3, \lambda_4, \lambda_5)$ control the degree of regularization applied to blocks of coefficients in the reduced-form specification (see Equation (1)).

The dummy-observation weights (μ_5, μ_6) implement the sum-of-coefficients (long-run) and initial-conditions/drift controls in the Sims–Zha prior, respectively, which help stabilize medium- and long-horizon dynamics (Sims & Zha, 1998; Baínura et al., 2010; Giannone, Lenza & Primiceri, 2015). This optimization framework allows the Matrix-Normal–Inverse-Wishart (MN–IW) structure to adapt to country-specific features and horizon-specific persistence. Table 4 reveals three economically intuitive patterns: *(i) Medium-horizon parsimony (12-month ahead horizon)*: Most G7 countries favor $p = 1$ with moderate overall tightness $\lambda_0 = 0.2$ and light intercept shrinkage $\lambda_4 = 0.1$, consistent with first-order autoregressive sufficiency at medium horizons. The United States is a notable exception, selecting $p = 4$ and stronger lag decay $\lambda_3 = 3$; it also uses $\lambda_5 = 0.5$ and nonzero dummy weights $(\mu_5, \mu_6) = (0.5, 0.5)$ at 12-month ahead horizon. *(ii) Longer-horizon persistence (24-month ahead horizon)*: Several economies transition to higher lag orders and looser overall tightness: France adopts $p = 3$ with $\lambda_0 = 0.8$, while Germany, the United Kingdom, and Italy use $p = 4$ (Germany with $\lambda_0 = 0.4$). Intercept control tightens to $\lambda_4 = 0.5$ for France, Germany, the United Kingdom; Italy, Canada and the United States remain at $\lambda_4 = 0.1$. Lag decay increases to $\lambda_3 = 2$ for Germany (others mostly remain at 1). Relative own-vs.-cross tightness λ_1 becomes less stringent for the United States (0.1) and France (0.2) at 24-month horizon, compared with 0.05 elsewhere. *(iii) Long-run dummies and exogenous channels at 24-month ahead horizon*: The initial-conditions/drift weight μ_6 becomes active at 24-month horizon for Canada (1), France (0.5), Germany (0.5), the United Kingdom (1), and Italy (0.5), while remaining 0 for the United States and Japan. The sum-of-coefficients weight μ_5 is already active at 12-month ahead horizon for most countries and at 24-month ahead horizon increases for the United States (to 1), decreases for Japan (to 0.5), and remains 1 elsewhere except France ($\mu_5 = 0$ at 24-month ahead horizon). Shrinkage on exogenous coefficients is relaxed at 24-month ahead horizon for Canada ($\lambda_5 = 1$), and Italy ($\lambda_5 = 0.5$), allowing stronger transmission from uncertainty shocks; other countries retain $\lambda_5 = 0$ at 24-month ahead horizon. The grid search evaluated three reduced-form prior families: Normal–Wishart (conjugate MN–IW), Flat–Gaussian (Gaussian on $\text{vec}(B)$ with Jeffreys prior on Σ), and Flat–Flat. The MN–IW specification was unanimously selected across all G7 countries and both horizons (Table 4). Conjugacy yields closed-form posteriors and exact Gibbs updates for (B, Σ) , promotes coherent cross-equation shrinkage, and tends to improve forecast robustness in medium-dimensional systems (Sims & Zha, 1998; Baínura et al., 2010; Giannone et al., 2015; Carriero et al., 2015).

Against the baselines in Section 5.1 SZBVARx achieves one of the best and most consistent performances across both horizons (Section 5.4). The SZBVARx model was implemented in **R** via the **MSBVAR** package, while baseline models used **tsDyn** and **legion** in **R** and **Darts** in Python. The effectiveness of the proposed SZBVARx model lies in its ability to harness the benefits of hierarchical shrinkage to regularize the VARx system. This hierarchical shrinkage penalizes both variance inflation when increasing the dimension of the system (p), and allows for cross-equation coherence and structural flexibility to capture near-unit-root processes and level shifts with hyperparameters (μ_5, μ_6) . The competing VARx baseline has no regularization and is subject to overfitting in higher dimensions. TVAR alternatives require very large datasets for stable estimation of regimes, but underperform at long horizons. Alternative state-of-the-art machine learning and deep learning

models can potentially achieve high performance on long stationary panels, but in general, are over-parameterized in short macroeconomic samples, leading to overfitting. Overall, principled regularization of SZBVARx results in robust, interpretable, and superior forecasting performance under various macroeconomic environments.

Table 4: Trained hyper-parameters (*shrinkage, and model controls*) for SZBVARx models across G7 countries and both forecast horizons: 12-month-ahead (12) and 24-month-ahead (24). Parameters are estimated for optimal multivariate macroeconomic forecasting with exogenous uncertainty shocks.

Country	Horizon	p	λ_0	λ_1	λ_3	λ_4	λ_5	μ_5	μ_6
Canada	12	1	0.2	0.05	1	0.1	0	1	0
	24	1	0.6	0.05	1	0.1	1	1	1
US	12	4	0.2	0.05	3	0.1	0.5	0.5	0.5
	24	4	0.2	0.1	1	0.1	0	1	0
France	12	1	0.2	0.05	1	0.1	0	1	0
	24	3	0.8	0.2	1	0.5	0	0	0.5
Germany	12	1	0.2	0.05	1	0.1	0	1	0
	24	4	0.4	0.05	2	0.5	0	1	0.5
Japan	12	1	0.2	0.05	1	0.1	0	1	0
	24	2	0.2	0.05	1	0.1	0	0.5	0
UK	12	1	0.2	0.05	1	0.1	0	1	0
	24	4	0.2	0.05	1	0.5	0	1	1
Italy	12	1	0.2	0.05	1	0.1	0	1	0
	24	4	0.2	0.05	1	0.1	0.5	1	0.5

The country-horizon patterns in Table 4 thus reflect an adaptive, economically disciplined prior: tight where data is informative, permissive where persistence and breaks dominate, a configuration that underwrites SZBVARx’s robust, low-variance forecasts for the G7. Across Tables 5–11, SZBVARx delivers the most consistent top-tier performance across the five targets, both horizons, and scoring metrics. On the Unemployment Rate, SZBVARx is best at a 24-month-ahead horizon in France, the United Kingdom, and Italy, and is highly competitive elsewhere, while VARx leads in Canada and the United States, and TVAR in Germany; Japan’s best is VES. For CPI Inflation at the 24-month-ahead horizon, SZBVARx tops the list in Canada and in the United States, whereas VARx or TVAR dominate elsewhere (France: VARx; Germany: TVAR; United Kingdom: VARx; Italy: TVAR; Japan: VARx). For REER at a 24-month-ahead horizon, SZBVARx ranks first in Canada, France, and Italy, TVAR leading in the United States and Japan, XGBx in Germany, and CatBoostx in the UK. In Oil Price (WTI) at a 24-month-ahead horizon, SZBVARx leads in four countries - Canada, the United States, France, and Japan - with NHITSx winning in Germany and Italy and NBEATSx in the United Kingdom. For SIR at a 24-month-ahead horizon, deep nets generally dominate (e.g., BRNNx: Canada, France, Germany, United Kingdom, Italy); Japan’s best is DTS_CBx, while SZBVARx is often strongest at a 12-month-ahead horizon (e.g., Canada; France; Italy). Scale-normalized diagnostics corroborate these patterns: unemployment rate at 24-month-ahead horizon posts low TU₁ for SZBVARx in France/United Kingdom and Italy; CPI

Inflation at 24-month-ahead horizon yields efficient TU₁ for Canada, United States; and oil price (WTI) at 24-month-ahead horizon records uniformly small TU₁ across countries, indicating strong *scale-adjusted* accuracy.

Table 5: Evaluation of the SZBVARx model's performance relative to 14 baseline forecasters across all forecast horizons (12-month-ahead & 24-month-ahead) for Canada (**best** and **second-best** results are highlighted).

Variables	Horizons	Metrics	VARx	TVAR	VES	NBEATSx	LGBMx	DTS_CBx	DTS_XGBx	NHITSx	TSMIXERx	BRNNx	XGBx	CatBoostx	TiDEx	TFTx	SZBVARx
Unemployment Rate	12M	RMSE	0.37	0.43	0.58	2.04	2.43	2.91	4.54	3.38	2.50	1.79	2.82	1.93	1.70	1.70	0.39
		SMAPE	0.06	0.06	0.09	0.26	0.32	0.37	0.47	0.44	0.35	0.27	0.26	0.24	0.23	0.22	0.05
		MDAPE	5.81	4.08	11.40	29.84	37.29	41.85	51.80	59.19	39.28	38.16	16.06	20.76	31.04	24.60	4.11
	24M	TU ₁	0.03	0.04	0.10	0.16	0.18	0.21	0.29	0.33	0.18	0.14	0.11	0.13	0.14	0.14	0.06
		MASE	1.07	1.00	1.51	5.27	7.77	8.04	11.82	9.83	7.33	5.27	6.01	4.90	4.72	4.40	0.90
		RMSE	1.20	1.37	2.61	2.41	2.12	1.83	1.85	1.82	2.46	2.02	2.52	2.03	1.74	1.58	1.32
REER	12M	SMAPE	0.18	0.21	0.38	0.34	0.32	0.28	0.28	0.27	0.37	0.31	0.35	0.30	0.24	0.22	0.20
		MDAPE	23.30	27.33	47.95	37.46	37.91	30.88	34.85	31.38	48.26	38.00	40.22	33.79	28.39	20.21	23.59
		TU ₁	0.10	0.11	0.20	0.20	0.17	0.15	0.15	0.19	0.16	0.19	0.16	0.14	0.13	0.11	
	24M	MASE	2.79	3.35	6.66	5.41	5.30	4.58	4.53	4.51	6.37	5.17	6.00	4.94	3.95	3.46	3.09
		RMSE	2.22	5.49	4.95	2.37	3.27	0.91	4.32	5.26	13.93	63.50	4.08	1.55	8.63	10.04	2.10
		SMAPE	0.02	0.05	0.05	0.02	0.02	0.01	0.04	0.05	0.12	0.92	0.03	0.01	0.08	0.09	0.02
SIR	12M	MDAPE	1.59	5.62	5.06	1.49	0.86	0.71	3.33	4.51	11.49	63.03	2.53	1.50	7.11	9.83	1.71
		TU ₁	0.01	0.00	0.02	0.01	0.02	0.00	0.02	0.03	0.07	0.46	0.02	0.01	0.04	0.05	0.01
		MASE	2.55	7.18	6.62	2.48	3.05	1.01	5.04	6.19	17.47	86.23	4.36	1.85	10.10	12.50	2.48
	24M	RMSE	3.00	3.58	5.91	3.41	3.27	7.14	10.30	10.48	22.00	67.00	0.03	7.03	4.54	18.08	0.01
		SMAPE	0.04	0.03	0.03	0.03	0.03	0.06	0.10	0.20	1.00	0.05	0.06	0.04	0.16	0.04	0.03
		MDAPE	3.80	3.54	4.54	2.47	2.85	6.94	6.98	7.99	22.19	66.88	3.82	7.69	3.45	18.61	2.52
Oil Price (WTI)	12M	TU ₁	0.02	0.02	0.03	0.02	0.02	0.03	0.05	0.05	0.10	0.50	0.03	0.04	0.02	0.08	0.01
		MASE	4.21	3.81	6.31	3.22	3.07	7.43	9.55	10.81	26.59	79.03	5.66	7.85	4.45	20.84	3.08
		RMSE	0.46	2.00	3.74	3.75	1.69	2.50	3.02	2.26	3.48	2.71	1.21	2.00	2.96	3.80	0.34
	24M	SMAPE	0.08	0.43	1.16	1.11	0.37	0.63	0.78	0.54	1.02	0.72	0.22	0.50	0.76	1.20	0.06
		MDAPE	9.66	33.36	73.30	70.29	33.38	50.14	61.65	44.66	67.55	53.68	24.15	43.00	46.45	73.44	7.00
		TU ₁	0.05	0.23	0.58	0.51	0.19	0.32	0.38	0.28	0.50	0.36	0.13	0.25	0.36	0.54	0.03
CPI Inflation	12M	MASE	5.59	22.98	50.17	45.90	20.94	32.71	35.83	28.31	45.77	36.22	13.19	27.01	33.16	48.71	4.10
		RMSE	4.35	3.44	3.99	4.48	2.24	3.24	3.06	2.76	3.65	2.21	3.41	3.29	4.31	3.20	3.33
		SMAPE	1.65	1.34	1.47	1.36	0.65	0.95	0.95	0.44	1.18	0.58	1.11	1.01	1.42	1.03	1.12
	24M	MDAPE	96.53	7.57	80.50	91.29	50.03	70.66	67.79	2.03	80.53	44.57	72.22	75.01	88.78	74.70	7.79
		TU ₁	0.87	0.61	0.79	0.70	0.33	0.55	0.51	0.28	0.64	0.32	0.60	0.56	0.75	0.47	0.59
		MASE	21.26	16.86	19.72	20.76	11.15	15.50	15.04	10.87	17.35	10.75	16.65	15.82	20.18	14.69	16.60

Table 6: Evaluation of the SZBVARx model's performance relative to 14 baseline forecasters across all forecast horizons (12-month-ahead & 24-month-ahead) for the US (**best** and **second-best** results are highlighted).

Variables	Horizons	Metrics	VARx	TVar	VES	NBEATsx	LGBMx	DTS_CBx	DTS_XGBx	NHITSx	TSMIXERx	BRNNx	XGBx	CatBoostx	TlDEx	TFTx	SZBVARx	
Unemployment Rate	12M	RMSE	0.56	0.51	0.94	3.74	3.44	2.78	4.10	6.17	4.53	2.13	1.39	1.38	3.64	3.42	0.65	
		SMAPE	0.13	0.11	0.22	0.56	0.60	0.52	0.60	0.88	0.72	0.44	0.28	0.28	0.56	0.51	0.13	
		MDAPE	13.48	8.93	23.92	71.06	88.36	71.33	89.55	161.99	124.93	57.19	36.48	34.37	96.66	102.27	16.05	
	24M	RMSE	0.42	0.39	0.94	5.04	4.70	2.20	5.04	2.60	2.15	2.15	5.04	2.57	3.30	1.47	1.31	
		SMAPE	0.09	0.06	0.23	0.80	0.35	0.45	0.51	0.51	0.43	0.46	0.77	0.53	0.57	0.33	0.29	
		MDAPE	5.61	44.04	27.27	137.87	46.05	51.95	76.02	71.29	50.83	60.73	120.16	71.19	80.71	41.01	31.21	
REER	12M	RMSE	0.06	0.07	0.11	0.34	0.32	0.27	0.36	0.46	0.38	0.22	0.16	0.16	0.33	0.33	0.08	
		SMAPE	0.02	0.01	0.02	0.01	0.02	0.03	0.01	0.01	0.01	0.01	0.02	0.02	0.03	0.03	0.02	
		MDAPE	3.12	1.05	4.09	2.15	1.94	4.37	1.12	0.71	2.54	65.76	0.76	3.86	5.25	15.41	1.46	
	24M	RMSE	0.02	0.01	0.02	0.01	0.02	0.03	0.01	0.01	0.01	0.01	0.02	0.02	0.03	0.03	0.01	
		SMAPE	0.02	0.01	0.02	0.01	0.02	0.03	0.01	0.01	0.01	0.01	0.02	0.02	0.03	0.03	0.02	
		MDAPE	2.73	2.30	3.83	10.92	11.59	11.07	11.81	11.84	14.41	68.89	10.59	11.55	3.75	9.93	3.95	
SIR	12M	RMSE	3.98	2.67	4.00	11.98	12.00	12.00	14.00	12.67	15.72	74.35	12.63	12.50	5.45	11.08	4.50	
		SMAPE	0.03	0.02	0.04	0.02	0.03	0.05	0.02	0.01	0.02	0.98	0.01	0.04	0.05	0.18	0.02	
		MDAPE	16.88	4.40	80.18	63.50	29.45	51.86	52.52	44.84	70.55	56.49	12.34	43.25	60.13	22.18	11.12	
	24M	RMSE	11.05	3.98	58.21	45.84	20.29	36.50	36.80	30.37	49.74	41.03	9.19	30.72	45.03	21.61	7.24	
		SMAPE	1.51	0.74	1.14	1.29	1.08	0.96	1.08	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	
		MDAPE	5.61	50.81	80.40	83.56	5.50	70.42	71.52	45.60	53.44	54.44	7.57	81.50	77.84	65.83	65.83	
Oil Price (WTI)	12M	RMSE	17.96	1.68	1.21	8.21	8.38	8.52	9.67	8.86	10.84	53.07	8.52	8.73	3.22	7.61	3.00	
		SMAPE	0.02	0.01	0.02	0.06	0.06	0.06	0.07	0.06	0.08	0.06	0.06	0.06	0.06	0.06	0.06	
		MDAPE	11.05	16.5	16.1	16.96	10.76	13.79	14.75	10.73	10.29	15.96	13.69	14.85	13.83	13.53	13.53	
	24M	RMSE	7.01	2.75	29.77	6.08	11.10	19.87	12.38	23.70	32.03	44.30	13.67	12.39	36.04	24.10	6.50	
		SMAPE	0.07	0.37	0.32	0.07	0.12	0.24	0.15	0.31	0.45	0.75	0.15	0.14	0.53	0.27	0.07	
		MDAPE	8.70	38.74	38.79	4.05	23.49	23.48	34.24	43.25	55.10	50.01	11.70	48.23	21.71	4.26		
CPI Inflation	12M	RMSE	0.05	0.20	0.16	0.04	0.07	0.14	0.08	0.17	0.24	0.40	0.09	0.08	0.28	0.15	0.04	
		SMAPE	1.19	4.67	1.84	1.03	1.84	3.27	2.15	4.18	5.49	8.90	2.38	2.04	6.28	3.98	1.03	
		MDAPE	30.47	129.76	59.63	22.14	19.89	31.46	18.55	13.77	25.62	60.94	24.64	31.13	29.89	27.20	13.31	
	24M	RMSE	0.12	0.40	0.47	0.10	0.14	0.22	0.16	0.10	0.15	0.46	0.19	0.21	0.15	0.20	0.07	
		SMAPE	3.89	17.62	9.83	3.02	3.33	5.08	3.54	4.17	9.83	4.11	4.90	4.19	4.47	4.47	2.00	
		MDAPE	4.46	8.21	4.46	4.17	4.55	4.26	4.29	3.35	3.70	5.37	4.24	4.17	3.79	4.52	3.08	
SIR	12M	RMSE	0.88	3.72	1.58	0.61	0.58	0.49	1.98	2.57	1.25	0.85	0.81	1.97	2.09	1.33	0.31	
		SMAPE	0.33	0.25	0.69	0.35	0.16	0.16	0.75	1.21	0.37	0.20	0.20	0.68	0.65	0.31	0.08	
		MDAPE	46.41	20.35	103.41	32.34	14.08	9.86	51.76	86.16	27.14	18.60	23.32	44.29	50.86	42.39	5.00	
	24M	RMSE	3.16	1.13	20.69	12.73	16.78	19.69	10.62	29.30	30.72	14.35	17.68	27.66	26.94	3.06	0.06	
		SMAPE	0.17	0.13	0.56	0.21	0.09	0.08	0.34	0.46	0.21	0.11	0.11	0.28	0.30	0.30	0.16	
		MDAPE	3.34	1.65	0.73	0.31	0.35	1.56	4.57	2.99	1.98	4.72	4.62	4.33	4.33	4.33	0.06	
Oil Price (WTI)	12M	RMSE	9.04	0.88	0.83	4.37	3.50	4.05	3.98	1.68	11.74	6.06	4.83	4.48	3.66	11.73	0.92	
		SMAPE	0.04	0.06	0.04	0.03	0.04	0.04	0.02	0.11	0.89	0.04	0.04	0.03	0.11	0.01	0.01	
		MDAPE	0.70	5.84	6.00	4.15	3.50	3.29	3.52	10.79	61.79	3.49	3.73	3.23	12.01	0.73	0.00	
	24M	RMSE	1.82	1.16	1.97	10.48	9.07	9.99	9.64	4.02	4.02	16.73	10.61	11.02	8.60	31.40	2.03	
		SMAPE	0.09	0.04	0.03	0.04	0.05	0.03	0.03	0.06	0.96	0.03	0.05	0.04	0.03	0.01	0.01	
		MDAPE	85.84	153.94	64.32	59.45	72.69	69.89	72.45	52.60	59.49	56.64	63.16	67.06	58.82	63.10	55.10	
CPI Inflation	12M	RMSE	3.77	17.29	7.59	5.88	8.82	10.05	6.19	5.39	12.14	6.54	10.73	9.40	6.58	2.75		
		SMAPE	0.92	1.56	4.24	3.11	2.06	1.89	1.66	1.67	1.56	2.04	1.30	1.63	2.35	2.42	0.62	
		MDAPE	1.81	1.77	1.78	1.43	0.72	0.68	0.50	0.48	0.44	0.73	0.35	0.55	0.44	0.91	0.17	
	24M	RMSE	1.04	1.08	10.75	11.18	10.71	10.52	6.62	11.15	9.68	56.52	23.88	45.20	43.05	69.21	16.38	
		SMAPE	0.99	0.99	0.98	0.66	0.91	0.78	1.00	0.61	0.73	0.67	0.76	0.91	0.84	0.74	0.89	
		MDAPE	14.74	19.76	14.83	12.87	13.14	11.69	14.48	13.80	11.98	12.14	14.58	11.95	14.91	12.81	12.25	
Oil Price (WTI)	12M	RMSE	7.31	21.87	24.28	5.89	6.81	2.89	2.91	3.23	14.08	26.29	24.90	44.28	17.02	9.18	51.76	2.96
		SMAPE	0.07	0.05	0.06	0.07	0.10	0.17	0.17	0.17	0.17	0.47	0.92	0.49	0.75	0.82	0.40	
		MDAPE	5.26	28.67	28.27	4.49	6.97	9.40	13.66	38.15	17.73	56.07	11.88	9.21	53.16	25.38	4.45	
	24M	RMSE	0.05	0.16	0.14	0.04	0.04	0.06	0.09	0.19	0.15	0.39	0.12	0.06	0.38	0.16	0.05	
		SMAPE	1.12	4.02	4.64	0.93	1.15	1.49	2.57	4.58	3.84	2.75	1.58	8.93	4.40	1.09		
		MDAPE	15.79	47.94	52.49	9.71	19.97	23.39	20.56	11.68	19.66	54.05	24.55	21.14	10.89	29.05	6.12	
CPI Inflation	12M	RMSE	2.23	0.69	0.75	2.96	2.52	2.31	2.69	3.03	1.73	2.85	1.63	2.26	2.65	2.87	2.16	
		SMAPE	0.40	0.15	0.15	0.73	1.03	0.78	1.05	1.21	0.47	0.92	0.49	0.75	0.82	0.92	0.40	
		MDAPE	44.06	15.55	10.86	48.90	54.92	58.76	71.74	72.22	36.47	65.50	39.79	54.86	56.10	59.70	43.21	
	24M	RMSE	0.22	0.08	0.07	0.38	0.41	0.37	0.43	0.54</								

Table 8: Evaluation of the SZBVARx model's performance relative to 14 baseline forecasters across all forecast horizons (12-month-ahead & 24-month-ahead) for Germany (**best** and **second-best** results are highlighted).

Variables	Horizons	Metrics	VARx	TVar	VES	NBEATSx	LGBMx	DTS_CBx	DTS_XGBx	NHITSx	TSMIXERx	BRNNx	XGBx	CatBoostx	TIDEIx	TFTx	SZBVARx
Unemployment Rate	12M	RMSE	0.28	0.37	0.26	2.59	1.65	2.05	2.00	5.46	1.48	3.61	2.73	2.39	2.27	3.08	0.22
		SMAPE	0.07	0.08	0.06	0.62	0.37	0.49	0.46	0.92	0.37	0.73	0.57	0.55	0.46	0.64	0.06
		MDAPE	5.84	6.34	4.91	65.68	41.26	66.25	61.67	166.07	42.19	118.44	93.71	76.44	35.57	100.79	3.76
	24M	RMSE	0.15	0.20	0.13	0.32	0.32	0.25	0.24	0.47	0.19	0.37	0.31	0.28	0.29	1.89	0.04
		SMAPE	0.15	0.14	0.13	11.04	7.53	10.17	9.59	26.94	7.01	13.00	11.91	8.12	1.14	14.89	0.03
		MDAPE	17.38	9.78	23.59	71.44	14.96	38.58	27.75	40.32	51.30	128.46	23.46	42.23	42.50	18.79	14.44
REER	12M	RMSE	0.14	0.16	0.19	0.52	0.14	0.32	0.24	0.34	0.42	0.77	0.19	0.33	0.34	0.17	0.21
		SMAPE	0.24	0.10	0.19	0.52	0.14	0.32	0.24	0.34	0.42	0.77	0.19	0.33	0.34	0.17	0.21
		MDAPE	0.46	0.85	1.75	3.24	1.66	1.85	2.90	2.57	4.03	63.09	2.95	1.43	3.79	5.24	1.07
	24M	TU ₁	0.00	0.01	0.01	0.02	0.01	0.01	0.02	0.01	0.02	0.01	0.01	0.01	0.02	0.02	0.01
		MASE	1.51	3.25	5.11	9.58	6.62	4.84	8.18	6.79	9.60	178.52	7.32	4.64	10.82	13.06	3.12
		MAE	3.47	1.59	3.42	12.38	2.47	6.26	4.54	7.01	8.85	20.58	3.50	6.45	7.22	3.10	3.02
SIR	12M	RMSE	0.16	1.46	1.86	3.87	2.93	1.90	3.30	2.77	4.29	63.46	2.93	1.89	4.31	5.02	1.26
		SMAPE	0.01	0.01	0.02	0.03	0.02	0.02	0.03	0.02	0.03	0.02	0.02	0.02	0.04	0.05	0.01
		MDAPE	0.46	0.85	1.75	3.24	1.66	1.85	2.90	2.57	4.03	63.09	2.95	1.43	3.79	5.24	1.07
	24M	TU ₁	0.00	0.01	0.01	0.02	0.01	0.01	0.02	0.01	0.02	0.01	0.01	0.01	0.02	0.02	0.01
		MASE	5.72	2.71	54.39	32.60	30.84	28.13	18.94	6.69	16.51	26.62	17.62	26.00	26.92	40.54	5.81
		MAE	3.36	2.98	2.99	3.00	3.12	2.75	3.20	2.96	2.98	1.76	3.10	2.79	3.88	3.03	3.20
Oil Price (WTI)	12M	RMSE	0.16	0.16	4.21	2.81	2.50	2.19	1.55	0.65	1.55	2.05	1.49	2.04	2.31	3.55	0.47
		SMAPE	0.02	0.02	2.00	1.11	0.95	0.82	0.49	0.15	0.43	0.75	0.45	0.74	0.69	1.20	0.13
		MDAPE	12.44	1.69	111.31	74.23	69.45	57.40	38.33	13.22	37.69	56.94	40.11	53.75	39.91	74.13	12.26
	24M	TU ₁	0.07	0.05	1.00	0.51	0.48	0.41	0.26	0.09	0.24	0.38	0.40	0.64	0.40	0.64	0.07
		MASE	5.72	2.71	54.39	32.60	30.84	28.13	18.94	6.69	16.51	26.62	17.62	26.00	26.92	40.54	5.81
		MAE	3.36	2.98	2.99	3.00	3.12	2.75	3.20	2.96	2.98	1.76	3.10	2.79	3.88	3.03	3.20
CPI Inflation	12M	RMSE	10.96	89.07	28.72	6.10	6.39	14.14	13.89	21.79	31.16	44.29	14.48	11.13	43.79	24.84	5.32
		SMAPE	0.05	0.05	0.05	0.07	0.07	0.14	0.14	0.28	0.29	0.75	0.15	0.13	0.72	0.28	0.06
		MDAPE	10.65	9.51	34.08	4.51	5.74	17.39	10.34	29.30	30.22	56.08	13.01	10.46	51.21	24.78	5.07
	24M	TU ₁	0.07	0.72	0.16	0.07	0.04	0.09	0.09	0.16	0.18	0.39	0.09	0.07	0.32	0.16	0.03
		MASE	1.85	15.42	5.63	0.87	1.07	2.50	2.05	3.81	5.06	8.89	2.19	1.89	7.80	4.15	0.90
		MAE	3.68	3.14	9.72	1.50	13.75	13.87	12.44	14.33	13.53	13.41	13.82	12.52	16.91	13.08	14.56
CPI Inflation	12M	RMSE	4.48	6.37	1.63	3.33	2.38	2.73	2.64	3.16	0.81	3.47	1.51	2.50	2.67	3.75	4.00
		SMAPE	0.64	1.23	0.33	0.67	0.63	0.65	0.74	0.92	0.14	0.89	0.32	0.54	0.94	1.04	0.60
		MDAPE	107.86	93.74	29.37	38.35	45.81	44.06	52.43	63.44	11.20	62.38	26.26	39.89	51.34	70.84	97.69
	24M	TU ₁	0.33	0.61	0.16	0.40	0.32	0.39	0.36	0.48	0.09	0.54	0.16	0.34	0.33	0.47	0.31
		MASE	6.52	8.29	2.46	3.45	3.51	3.59	3.99	4.51	4.96	5.36	3.46	3.46	5.31	5.50	
		MAE	2.35	0.94	0.31	0.46	0.75	0.52	0.53	0.84	1.58	5.60	0.51	0.59	5.73	2.11	
SIR	12M	RMSE	21.61	18.67	54.13	9.78	23.03	27.30	20.36	8.45	15.37	54.04	25.60	24.31	16.27	27.89	12.66
		SMAPE	0.22	0.21	0.88	0.09	0.31	0.35	0.24	0.07	0.15	0.90	0.30	0.29	0.15	0.31	0.13
		MDAPE	27.72	20.70	58.64	8.29	29.74	31.13	20.11	5.55	15.52	60.87	24.25	25.59	12.49	22.61	14.99
	24M	TU ₁	0.11	0.11	0.46	0.08	0.15	0.19	0.13	0.05	0.09	0.46	0.17	0.16	0.09	0.18	0.07
		MASE	3.68	3.14	9.72	1.50	4.08	4.73	3.26	1.20	2.41	9.83	4.04	4.02	2.40	4.03	2.11
		MAE	3.36	2.98	2.99	3.00	3.12	2.75	3.20	2.96	2.98	1.76	3.10	2.79	3.88	3.03	3.20
Oil Price (WTI)	12M	RMSE	2.02	14.06	10.29	14.66	13.16	15.41	7.03	27.75	23.20	35.98	5.43	11.62	37.97	7.25	3.45
		SMAPE	0.02	0.16	0.13	0.18	0.16	0.19	0.09	0.26	0.27	0.65	0.06	0.14	0.36	0.08	0.04
		MDAPE	3.00	19.24	14.46	20.31	17.13	20.49	9.26	32.43	32.80	48.51	6.46	15.94	50.48	9.49	4.53
	24M	TU ₁	0.01	0.09	0.07	0.09	0.08	0.10	0.05	0.16	0.14	0.33	0.08	0.07	0.21	0.05	0.02
		MASE	1.78	12.40	9.90	13.99	12.76	14.98	6.59	22.44	23.59	35.19	4.76	11.19	32.96	6.35	3.13
		MAE	1.46	14.06	12.03	20.93	20.74	23.70	13.00	24.00	24.09	41.69	2.11	24.39	20.74	22.57	3.01
SIR	12M	RMSE	0.04	0.03	0.04	1.59	0.04	0.05	0.05	2.25	0.07	0.16	0.04	0.04	2.34	0.26	0.04
		SMAPE	0.60	0.58	0.90	1.99	0.66	1.45	1.56	1.99	1.88	1.87	0.88	1.15	1.91	1.64	0.66
		MDAPE	28.78	43.38	75.71	31.83	71.31	91.76	104.07	5529.54	144.34	294.66	62.67	82.26	3038.36	365.62	50.54
	24M	TU ₁	0.33	0.37	0.67	0.98	0.38	0.73	0.67	0.99	0.88	0.93	0.40	0.64	0.99	0.91	0.42
		MASE	1.34	1.32	2.00	71.33	1.57	2.40	2.65	11.08	3.56	8.21	1.92	2.02	9.62	11.22	1.50
		MAE	0.09	0.13	0.05	2.10	0.04	0.03	0.05	1.77	0.15	1.93	1.93	1.76	1.08	1.74	0.64
Oil Price (WTI)	12M	RMSE	5.70	18.91	30.27	10.95	21.28	19.05	27.03	38.62	27.85	44.29	19.82	14.19	66.64	25.57	12.08
		SMAPE	0.02	0.23	0.12	0.29	0.23	0.29	0.30	0.40	0.50	0.76	0.17	0.32	0.32	0.14	
		MDAPE	6.23	24.04	39.13	8.83	23.85	24.45	29.35	42.03	36.77	56.08	21.03	17.34	88.70	23.55	12.64
	24M	TU ₁	0.04	0.14	0.16	0.07	0.16	0.13	0.20	0.28	0.21	0.39	0.14	0.10	0.61	0.18	0.08
		MASE	0.95	3.43	5.98	1.77	4.03	3.29	4.91	6.66	5.00	8.89	3.56	2.49	12.29	4.10	2.

Table 10: Evaluation of the SZBVARx model's performance relative to 14 baseline forecasters across all forecast horizons (12-month-ahead & 24-month-ahead) for the UK (**best** and **second-best** results are highlighted).

Variables	Horizons	Metrics	VARx	TVar	VES	NBEATsx	LGBMx	DTS_CBx	DTS_XGBx	NHITSx	TSMIXERx	BRNNx	XGBx	CatBoostx	TlDEx	TFTx	SZBVARx	
Unemployment Rate	12M	RMSE	0.68	0.22	0.69	1.75	0.71	0.58	0.39	3.89	0.69	1.72	0.63	0.42	1.93	3.04	0.47	
		SMAPE	0.17	0.05	0.17	0.35	0.13	0.11	0.08	0.63	0.13	0.34	0.11	0.09	0.31	0.47	0.11	
		MDAPE	16.33	4.77	16.18	31.68	12.98	10.26	6.68	94.56	12.62	40.92	9.53	7.64	34.38	75.49	10.97	
	24M	RMSE	0.71	1.36	0.34	3.14	0.72	1.49	1.48	2.94	2.37	1.88	1.82	1.42	2.06	2.47	0.22	
		SMAPE	0.18	0.35	0.07	0.49	0.29	0.32	0.26	0.53	0.45	0.38	0.35	0.35	0.34	0.47	0.05	
		MDAPE	14.19	35.91	6.44	70.22	35.37	37.99	26.29	73.13	58.25	49.37	44.19	43.32	41.65	61.88	4.10	
REER	12M	RMSE	0.11	0.19	0.04	0.29	0.15	0.16	0.16	0.27	0.23	0.19	0.19	0.18	0.21	0.26	0.03	
		SMAPE	5.51	10.85	2.30	22.61	11.19	12.12	10.27	23.71	18.87	15.21	14.10	13.96	14.28	20.86	1.47	
		MDAPE	4.20	4.36	9.56	3.81	3.13	2.09	2.57	3.84	11.08	3.22	2.36	10.49	17.64	3.80		
	24M	RMSE	0.71	1.36	0.34	3.14	0.72	1.49	1.48	2.94	2.37	1.88	1.82	1.42	2.06	2.47	0.22	
		SMAPE	0.08	0.10	0.04	0.03	0.06	0.04	0.05	0.11	0.04	0.97	0.06	0.04	0.04	0.09	0.06	
		MDAPE	7.18	8.98	4.03	2.41	5.43	3.58	5.39	10.57	4.28	65.56	6.07	4.06	2.52	8.58	5.91	
SIR	12M	RMSE	0.04	0.05	0.02	0.03	0.02	0.03	0.06	0.09	0.03	0.49	0.03	0.02	0.03	0.05	0.03	
		SMAPE	5.79	16.54	8.26	4.80	2.77	2.75	4.40	7.16	2.90	71.98	5.36	2.55	5.04	8.69	5.44	
		MDAPE	0.05	0.11	0.04	0.02	0.02	0.03	0.06	0.09	0.02	1.02	0.04	0.04	0.07	0.04		
	24M	RMSE	5.79	21.09	74.25	75.54	33.85	40.32	63.28	24.15	55.79	45.61	11.12	42.44	31.51	73.24	6.45	
		SMAPE	0.11	0.15	0.59	0.65	0.21	0.27	0.44	0.14	0.33	0.28	0.06	0.26	0.22	0.58	0.05	
		MDAPE	8.82	11.26	34.40	39.14	15.51	19.79	27.74	10.77	22.66	20.45	4.89	18.91	16.50	34.47	3.67	
Oil Price (WTI)	12M	RMSE	1.05	1.42	3.88	4.89	1.80	2.55	3.37	1.36	2.66	0.60	2.17	2.28	4.07	0.49		
		SMAPE	0.21	0.28	1.17	1.29	0.40	0.54	0.09	0.26	0.65	0.56	0.11	0.51	0.15	1.25	0.08	
		MDAPE	21.15	45.08	39.37	3.47	34.34	17.30	14.74	20.14	24.11	50.07	11.03	17.85	52.48	20.68	6.45	
	24M	RMSE	0.06	0.43	0.16	0.04	0.06	0.09	0.10	0.14	0.15	0.39	0.08	0.06	0.34	0.08	0.05	
		SMAPE	8.82	12.16	34.40	39.14	15.51	19.79	27.74	10.77	22.66	20.45	4.89	18.91	16.50	34.47	3.67	
		MDAPE	8.30	30.19	95.40	85.80	58.50	61.57	77.37	75.92	72.86	45.45	84.68	62.56	81.31	76.81	9.11	
CPI Inflation	12M	RMSE	0.70	0.71	0.90	0.64	0.70	0.53	0.67	0.49	0.57	0.25	0.76	0.49	0.71	0.47	0.83	
		SMAPE	16.04	18.86	18.97	16.50	16.50	16.15	13.09	15.54	14.05	14.06	8.36	18.17	12.55	16.50	13.36	18.02
		MDAPE	2.16	4.08	4.56	4.03	2.58	2.88	3.28	2.70	2.39	5.79	3.18	3.77	2.17	2.26	1.99	3.88
	24M	RMSE	18.25	136.30	53.53	10.41	26.20	24.43	26.23	10.81	12.76	54.04	23.23	22.35	10.85	29.61	14.18	
		SMAPE	0.19	1.26	0.87	0.11	0.34	0.32	0.36	0.12	0.12	0.90	0.31	0.28	0.11	0.33	0.15	
		MDAPE	21.74	111.03	57.40	9.67	27.58	26.74	26.73	10.96	12.22	60.88	28.53	25.05	9.59	25.65	15.95	
REER	12M	RMSE	3.99	1.83	3.45	3.72	2.49	2.48	1.48	3.76	1.49	3.75	3.37	1.59	3.24	3.33	4.22	
		SMAPE	0.51	0.48	0.34	0.89	0.52	0.55	0.32	0.97	0.83	0.43	0.26	0.81	0.73		0.53	
		MDAPE	68.70	7.60	34.74	53.71	42.47	43.24	26.34	67.55	59.68	59.40	31.15	57.72	53.81	73.52		
	24M	RMSE	0.27	0.17	0.18	0.39	0.27	0.27	0.14	0.47	0.15	0.47	0.24	0.14	0.31	0.37	0.28	
		SMAPE	8.02	9.42	5.45	5.31	5.31	5.56	3.50	5.00	5.08	5.08	3.45	6.96	7.24	7.49		
		MDAPE	2.08	14.84	8.82	5.63	3.09	5.42	5.20	3.55	4.99	5.56	5.02	5.43	5.20	5.54	2.27	
SIR	12M	RMSE	1.91	8.15	2.08	6.24	4.51	2.45	2.69	4.57	63.18	4.28	2.51	7.66	4.10	1.09		
		SMAPE	0.02	0.07	0.02	0.06	0.03	0.02	0.02	0.04	0.92	0.04	0.02	0.06	0.04	0.01		
		MDAPE	1.68	7.97	2.00	6.00	3.32	1.44	1.89	1.91	4.64	62.95	4.20	1.73	5.65	4.25	0.82	
	24M	RMSE	6.03	24.99	7.03	20.33	12.11	6.70	8.21	7.09	14.04	220.63	12.75	7.08	22.69	13.36	3.35	
		SMAPE	0.03	0.07	0.01	0.04	0.03	0.02	0.02	0.04	0.98	0.03	0.03	0.09	0.05	0.01		
		MDAPE	3.71	8.29	0.65	3.62	2.65	3.31	1.91	1.89	2.90	65.99	2.36	3.55	8.88	5.35	1.48	
Oil Price (WTI)	12M	RMSE	1.20	1.94	4.23	2.50	0.89	1.83	1.30	0.92	1.67	1.69	0.85	1.72	1.77	3.19	0.54	
		SMAPE	0.34	0.66	2.00	0.84	0.23	0.65	0.40	0.22	0.34	0.56	0.21	0.58	0.34	1.32	0.15	
		MDAPE	31.93	51.97	111.75	49.70	20.78	50.07	35.46	48.91	46.83	21.65	45.58	35.63	78.54	14.19		
	24M	RMSE	14.22	23.97	54.59	27.06	10.42	23.67	16.02	10.20	21.45	9.94	21.87	18.45	39.26	6.79		
		SMAPE	2.83	4.75	3.00	2.78	3.13	2.42	3.18	2.58	1.80	3.20	2.39	3.39	2.82	3.51		
		MDAPE	1.63	1.80	1.78	1.46	1.40	1.29	1.80	1.80	1.82	1.82	1.79	1.27	1.50	1.51	1.81	
CPI Inflation	12M	RMSE	10.92	33.53	26.17	6.03	11.51	6.01	7.60	21.16	35.27	44.28	10.45	6.00	34.40	22.96	5.41	
		SMAPE	0.13	1.13	0.26	0.09	0.11	0.07	0.08	0.27	0.36	0.78	0.27	0.07	0.26	0.26	0.06	
		MDAPE	7.03	44.90	31.48	4.77	13.64	6.71	6.66	57.56	59.07	59.07	11.34	7.18	35.22	22.07	2.98	
	24M	RMSE	0.07	0.27	0.14	0.04	0.07	0.04	0.05	0.15	0.20	0.39	0.07	0.04	0.25	0.15	0.03	
		SMAPE	1.73	6.36	5.06	0.93	1.91	1.11	1.31	3.70	8.89	1.81	1.09	5.82	3.91	0.91		
		MDAPE	3.15	21.50	14.94	12.52	14.02	11.14	11.58	8.29	14.37	10.97	14.73	12.28	16.05			
REER	12M	RMSE	4.42	1.88	5.77	4.11	2.47	2.63	2.39	2.74	2.61	3.43	2.88	2.58	3.38	3.92	5.64	
		SMAPE	0.83	0.85	0.94	0.98	0.69	0.80	0.77	0.77	0.91	0.75	0.80	0.88	1.16	0.93		
		MDAPE	158.13	45.51	233.61	74.52	52.77	67.14	57.56	62.41	82.34	73.26	72.98	70.03	95.18	96.01	232.15	
	24M	RMSE	0.39	0.24	0.14	0.48	0.33	0.34	0.30	0.38	0.29	0.52	0.30	0.33	0.32	0.53	0.43	
		SMAPE	4.78	2.09	7.38	6.64												

plex, shock-prone macroeconomic environments. SZBVARx outperforms the classic VARx and VES in the nonstationary regimes and, in particular, provides much lower forecast variance than the machine learning and deep learning approaches when the sample sizes are small or moderately large. In particular, CatBoostx and VARx appear to be sensible baselines for the semi-long-term forecast (exchange rate and interest rate series), but quickly deteriorate in volatile and persistent environments. The broad evidence suggests the benefits of Bayesian shrinkage, explicitly modeling uncertainty, and carefully calibrating the prior for multivariate macroeconomic forecasting. The model’s performance evaluation endorses SZBVARx as a principled framework that provides a robust, stable, and interpretable platform for policy analysis across different G7 regimes.

5.4. Statistical significance of the results

As a further robustness check on the stability of our findings, we also conduct the multiple comparisons with the best (MCB) procedure (Lee & Lee, 2018). The MCB is a non-parametric multiple testing procedure that simultaneously compares multiple algorithms and settings with regard to their predictive accuracy. In the MCB test, we rank all algorithms according to the RMSE metric across all five target macroeconomic variables, and two forecast horizons (“semi-long-term” and “long-term”) for all G7 economies. The MCB test is based on the Nemenyi post-hoc procedure that corrects for the increase in Type I error arising from the multiple simultaneous comparisons (Koning, Franses, Hibon & Stekler, 2005; Svetunkov, 2023). For each algorithm i evaluated across D datasets, the mean rank \bar{R}_i is computed as:

$$\bar{R}_i = \frac{1}{D} \sum_{j=1}^D R_{ij} \quad (18)$$

where R_{ij} denotes the rank of algorithm i on dataset j . The critical distance at significance level α follows the Tukey distribution:

$$CD_\alpha = q_\alpha \sqrt{\frac{\mathcal{A}(\mathcal{A}+1)}{6D}} \quad (19)$$

In this expression, q_α is the critical value from the Tukey distribution at level α , \mathcal{A} is the number of algorithms, and D is the evaluation datasets. Two algorithms are statistically significantly different if $|R_i - R_j| > CD_\alpha$. As shown in Figure 4, the MCB analysis reveals a non-negligible degree of performance heterogeneity across the forecasting algorithms for the five macroeconomic indicators. SZBVARx consistently achieves the lowest mean ranks and its dominance in terms of RMSE performance over the full set of key variables is statistically confirmed. SZBVARx has mean ranks of 2.71, 3.43, 5.14, 2.00, and 5.14 for the unemployment rate, REER, SIR, oil price (WTI), and CPI inflation, respectively. The statistical significance is derived from the critical intervals resulting from the MCB procedure with the critical distance value of 5.34 (the upper part of the critical distance for the SZBVARx model, represented by the grey shaded regions in Figure 4, is used as the reference value for the test and is compared to other competing algorithms), where a number of top methods such as NBEATSx, NHITSx, and TSMIXERx have mean ranks that are significantly higher

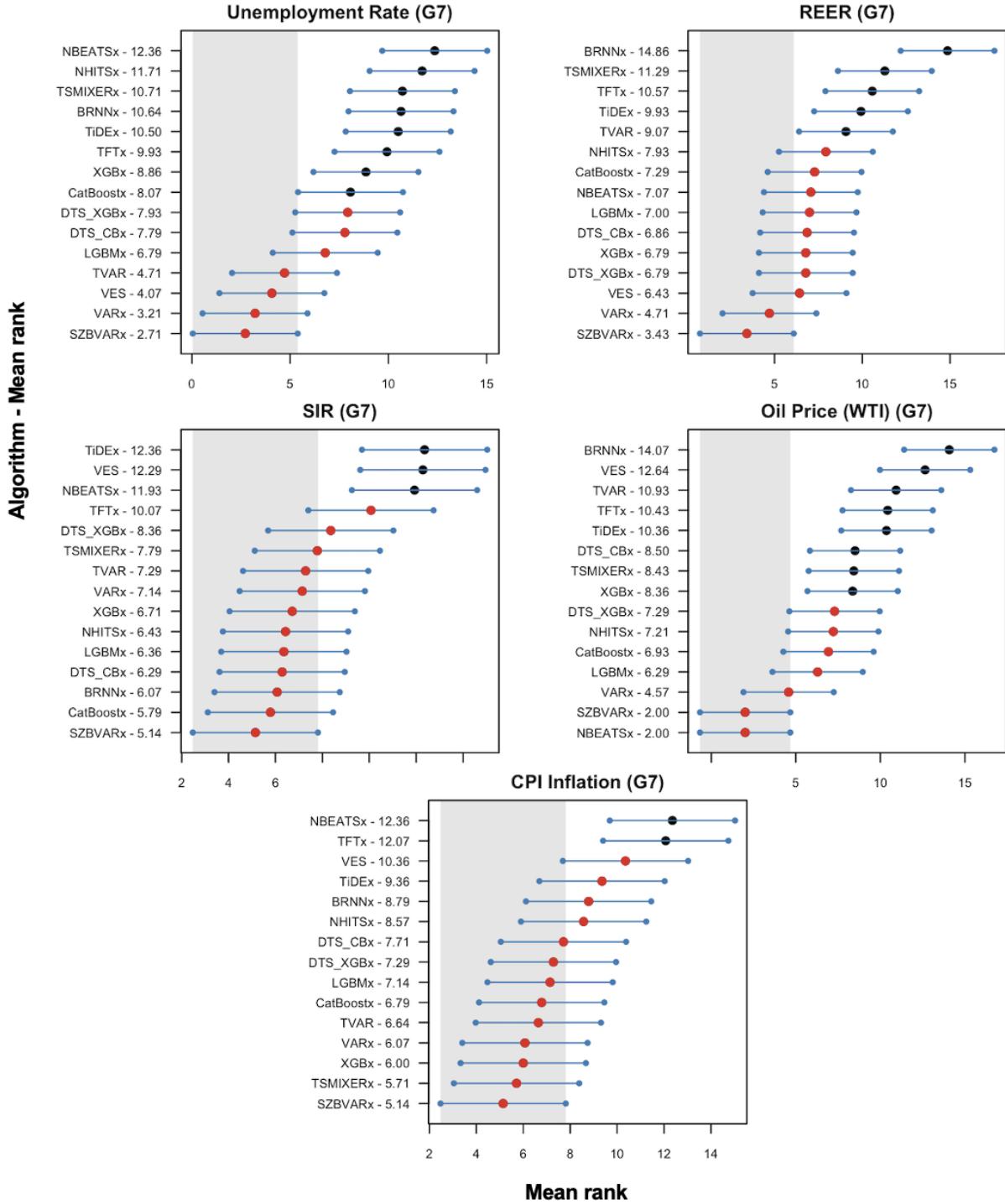


Figure 4: Multiple comparisons with the best (MCB) plots for G7 economies across 12-month (semi-long-term) ahead and 24-month (long-term) ahead forecasting horizons, ranking algorithms by RMSE metric. Key indicators: Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation. Labels (e.g., SZBVARx-2.71) denote model identifiers and their mean ranks, where lower values indicate superior out-of-sample predictive accuracy.

with non-overlapping critical intervals. In the case of deep learning and transformer-based models, TiDEx and TFTx exhibit a moderate level of performance for most of the variables, whereas BRNNx is conspicuously placed at the bottom for REER and oil price (WTI), with mean ranks that are above 10. At the same time, there also exists a group of competitive alternatives that occupy a second-tier position across several variables. In particular, VARx often is on par with top-performing models, and CatBoostx, outperforms most of the machine learning alternatives. CatBoostx demonstrates competitive performance for specific macroeconomic variables such as SIR, where it is the second-best (5.79) method only after SZBVARx, thus supporting its potential for practical applicability to monetary policy-related time series. NBEATSx is also able to show promising performance for oil price forecasting, with a mean rank that is the same as SZBVARx (2.00). The MCB test also suggests that the results are robust across the two forecast horizons. The fact that this pattern of algorithmic performance is observed across variables with fundamentally different statistical characteristics and volatility patterns further reinforces the evidence in favor of SZBVARx and its ability to provide consistently accurate forecasts for a range of multivariate macroeconomic time series for the G7 economies. The most strategically relevant baseline algorithms for comparison with SZBVARx across G7 countries, according to the results of the MCB test, are VARx and CatBoostx. The former is able to demonstrate competitive performance for unemployment rate and REER forecasting, where it consistently achieves second place in these two indicators. The latter exhibits strong performance for SIR, where it ranks second after SZBVARx. This aspect can be seen as a sign of its specialized applicability to time series that are directly relevant for monetary policy. Thus, their combination as complementary baselines is an ideal way to assess the relative performance of SZBVARx across the full set of variables and forecasting horizons.

Expanding on the MCB test described above, we apply a Murphy diagram (MD) difference analysis to further stress-test the statistical significance of the forecasting performances across G7 nations. The MD difference approach considered in this section operationalizes the extremal score framework of [Ehm, Gneiting, Jordan & Krüger \(2016\)](#) for rigorous pairwise comparison of forecasting methods on a large class of consistent scoring functions. The MD difference approach specifies extremal scoring functions that evaluate forecast performance on account of threshold-dependent decision problems ([Ziegel, Krüger, Jordan & Faschiati, 2017](#)). For two competing forecasts from different forecasting methods with corresponding observations, where forecasts are point predictions, the extremal score at threshold parameter θ is defined as:

$$\mathcal{S}_\theta(x, y) = \begin{cases} (\alpha - 1)(y - \theta) & \text{if } y \leq \theta < x \\ \alpha(y - \theta) & \text{if } x \leq \theta < y \\ 0 & \text{otherwise} \end{cases} \quad (20)$$

where x represents the forecast value, y represents the actual observed value at time t , $\alpha \in (0, 1)$ is the expectile level parameter controlling asymmetric penalization (with $\alpha = 0.5$ providing symmetric scoring), and $\theta \in \mathbb{R}$ is the threshold parameter ([Ehm et al., 2016](#)). This formulation implements the elementary scoring function for expectiles, representing a special case of the consistent scoring function family for expectile functionals as established in [Gneiting \(2011\)](#). Following the scoring convention, lower extremal scores indicate superior forecasting performance. The Murphy diagram difference function computes the performance differential between two forecasting methods across

the spectrum of threshold values:

$$D(\theta) = \bar{\mathcal{S}}_\theta(\hat{y}_{\text{SZBVARx}}, y) - \bar{\mathcal{S}}_\theta(\hat{y}_{\text{baseline}}, y) = \frac{1}{N} \sum_{t=1}^N [\mathcal{S}_\theta(\hat{y}_{\text{SZBVARx},t}, y_t) - \mathcal{S}_\theta(\hat{y}_{\text{baseline},t}, y_t)] \quad (21)$$

where $\bar{\mathcal{S}}_\theta(\hat{y}_j, y)$ represents the mean extremal score for forecasting method j at threshold θ , and N is the total number of forecast-observation pairs. Since lower extremal scores indicate better performance, negative values of $D(\theta)$ indicate superior performance of SZBVARx, while positive values suggest the baseline method (VARx or CatBoost) performs better at the given threshold. To account for potential heteroskedasticity and autocorrelation in the score differences, we employ the Newey-West heteroskedasticity- and autocorrelation-consistent (HAC) variance estimator with Bartlett kernel and lag truncation parameter ℓ (Newey & West, 1987b):

$$\widehat{\text{Var}}[D(\theta)] = \frac{1}{N} \left(\hat{\gamma}_0 + 2 \sum_{j=1}^{\ell} w_j \hat{\gamma}_j \right) \quad (22)$$

where $w_j = 1 - \frac{j}{\ell+1}$ are Bartlett kernel weights that ensure positive semi-definiteness of the covariance matrix estimator, and $\hat{\gamma}_j$ are the sample autocovariances of the mean-demeaned score differences:

$$\hat{\gamma}_j = \frac{1}{N} \sum_{t=j+1}^N (d_t(\theta) - \bar{d}(\theta)) (d_{t-j}(\theta) - \bar{d}(\theta)) \quad (23)$$

where $d_t(\theta) = \mathcal{S}_\theta(\hat{y}_{\text{SZBVARx},t}, y_t) - \mathcal{S}_\theta(\hat{y}_{\text{baseline},t}, y_t)$ denotes the score difference at time t , and $\bar{d}(\theta) = \frac{1}{N} \sum_{t=1}^N d_t(\theta)$ is the sample mean of score differences (Andrews, 1991). The lag truncation parameter is chosen as $\ell = \lfloor 4(N/100)^{2/9} \rfloor$ following the automatic bandwidth selection rule of Newey & West (1994). Pointwise confidence bands are constructed as:

$$D(\theta) \pm z_{\alpha/2} \sqrt{\widehat{\text{Var}}[D(\theta)]} \quad (24)$$

where $z_{\alpha/2}$ is the critical value from the standard normal distribution for the desired confidence level (90% in our implementation). The threshold parameter θ orchestrates forecast evaluation by selectively emphasizing distinct dimensions of prediction error asymmetry (Holzmann & Klar, 2014). Lower θ values amplify sensitivity to systematic underestimation, while elevated values intensify penalties for consistent overforecasting. This parametric flexibility empowers Murphy difference diagrams to deliver distribution-agnostic performance benchmarking across competing forecasting architectures, revealing error distribution patterns and establishing dominance hierarchies under heterogeneous scoring regimes (Jordan, Krüger & Lerch, 2019).

This threshold-adaptive methodology complements our MCB analysis by revealing performance heterogeneity across diverse decision-theoretic contexts, providing a more nuanced understanding of forecasting dominance patterns beyond aggregate error metrics³.

³Implementation uses the `murphydiagram` R package's `murphydiagram_diff` function (Krueger & Jordan, 2019) with HAC-based confidence bands and automatic lag truncation.

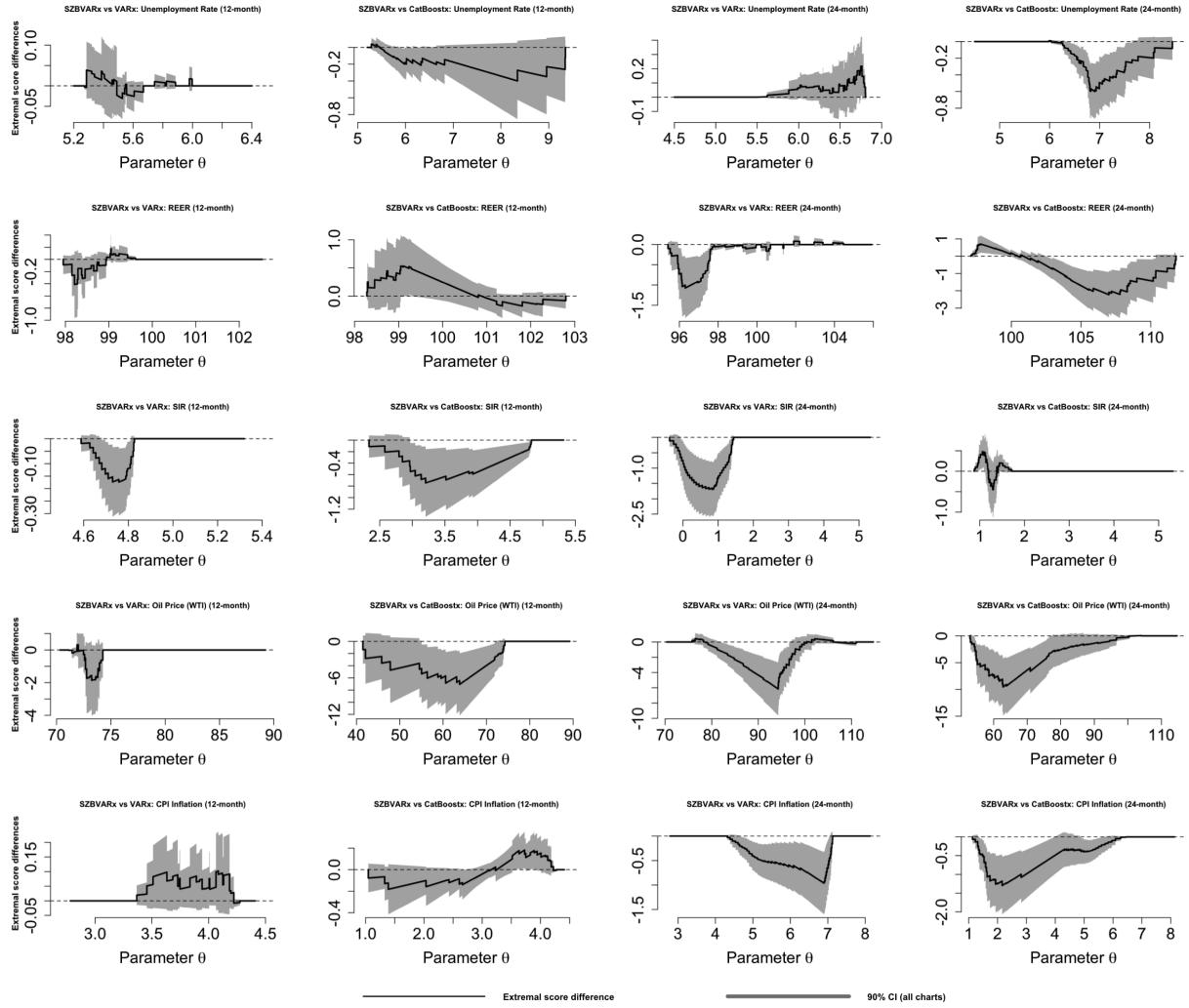


Figure 5: Murphy diagram difference plots comparing SZBVARx forecasting performance against VARx and CatBoostx baselines across 12-month and 24-month-ahead forecast horizons for Canada. Panels in the first and second columns correspond to the 12-month horizon, while panels in the third and fourth columns display the results for the 24-month horizon. The analysis covers five key macroeconomic indicators: Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation. Each subplot displays extremal score differences with 90% HAC-based confidence bands (gray shaded areas) across threshold parameter values. Negative differences indicate superior SZBVARx performance, while the magnitude reflects the strength of the performance advantage at each decision threshold.

Figure 5 presents the Murphy difference diagram for Canada. For the unemployment rate, SZBVARx outperforms the alternatives in most of the threshold ranges between 5.2 and 6.4, while in certain decision criteria, VARx also becomes competitive. For REER, the positive difference areas are generally assigned to CatBoostx in threshold ranges between 98 and 100, while the negative differences dominate overall, suggesting the superiority of SZBVARx. The oil price also displays negative differences almost everywhere, although there are certain decision criteria where the challenger models prevail. In the case of CPI inflation, the positive difference areas are mainly related to VARx ($\theta \approx 3.4 - 4.3$) and CatBoostx ($\theta \approx 3.5 - 4.1$) at the 12-month horizon, which may be interpreted as the superiority of the baseline models at those threshold values. At 24 months, the differences are negative for all decision criteria, which in turn implies the dominance of SZBVARx. In summary, there is no methodology that is clearly superior to the others for all of the considered threshold parameters and forecast horizons. However, each method displays its excellence in certain decision regimes. This complementarity across approaches and decision criteria for Canada paves the way to further our analysis with the other G7 countries.

A more nuanced view of the competitive relationships emerges when looking at the remaining G7 plots available in Appendix A.1.8 (Figures A.1.22 through A.1.27). These plots verify the general supremacy of SZBVARx but also feature zones where either the standard VARx or CatBoostx forecasts dominate across different threshold intervals. This result indicates that forecasting dominance is not unidirectional and that different methodologies present their own advantages depending on the economic variable and the decision-theoretic framework in focus. When considering the competitive landscape across countries, there are significant variations. In the United States, SZBVARx shows dominance in the unemployment indicator ($\theta \approx 3.5 - 7$) and in oil price, while baseline models are more competitive in CPI inflation decision regimes ($\theta \approx 3.5 - 5.0$ in the case of 12-month horizon, $\theta \approx 5.5 - 8.0$ in the case of 24-month horizon). The case of France presents a more marked supremacy of SZBVARx, mainly in unemployment ($\theta \approx 6.8 - 9.0$) and SIR ($\theta \approx 2.8 - 4.0$). The traditional models remain competitive in some specific regimes for REER ($\theta \approx 96.0 - 97.0$). The German case can be considered the one that most strongly confirms the superiority of SZBVARx, with negative differences across the whole spectrum of selected variables and decision thresholds of interest: unemployment rate ($\theta \approx 2.8 - 3.2$), REER ($\theta \approx 98.0 - 99.0$), SIR ($\theta \approx 3.3 - 4.0$), and CPI inflation ($\theta \approx 6.0 - 8.0$). In contrast, Japan shows the most even competitive situation, with the baseline approaches showing superiority in unemployment rate ($\theta \approx 2.7 - 2.9$) and in a few CPI inflation regimes ($\theta \approx 3.4 - 4.0$). SZBVARx models dominate the forecasting task in REER ($\theta \approx 75 - 85$, $\theta \approx 90 - 105$). The case of the United Kingdom illustrates well the complementarity of the methodological approaches. On the one hand, SZBVARx presents a clear superiority in the REER forecasting task ($\theta \approx 95 - 105$). On the other hand, the traditional approaches are more competitive in oil price (WTI) and in CPI inflation ($\theta \approx 70 - 85$, $\theta \approx 4.5 - 10$). The Italian situation is similar to the one just described, with baseline superiority in SIR ($\theta \approx 0 - 1$) and CPI inflation ($\theta \approx 2 - 8$), and SZBVARx advantage in unemployment rate and REER.

Augmenting the detailed findings of the Murphy difference analysis, we complement our evaluation of forecast performance differences with statistical inference on the (multivariate) Diebold-Mariano (DM) test statistic of Mariano & Preve (2012)⁴. The DM test is a Wald-type test that

⁴The multivariate Diebold-Mariano test is implemented using the `multDM` R package by Drachal (2025). The R document can be found here: <https://cran.r-project.org/web/packages/multDM/multDM.pdf>.

generalizes the original one model-versus-one to testing for equal predictive accuracy of two or more non-nested models over a set of target series. The multivariate framework allows accounting for serial correlation and heteroskedasticity of loss differentials while jointly testing the null hypothesis that the competing models are all equally accurate against the alternative hypothesis that they are not. The resulting finite-sample corrected test statistic S_c has an asymptotic chi-squared (χ^2) distribution, and the finite-sample correction is shown to improve size control in moderate samples. The test is invariant to the ordering of the models, and the testing procedure can be flexibly applied using various loss functions. We follow Mariano & Preve (2012) and use absolute scaled error loss, block length $q = 1$ (standard choice), and finite-sample correction in our implementation. Table 12 reports the p-values from pairwise comparison of SZBVARx versus VARx and CatBoostx across the G7 economies and 12-month and 24-month forecast horizons. Results that are statistically significant at 10% (reported in bold) indicate that, overall, SZBVARx exhibits superior predictive accuracy, especially at longer horizons and relative to CatBoostx. Notably, for Canada, the United States, France, Germany, Japan and the United Kingdom at 24-month horizon, the SZBVARx outperforms both baselines and exhibits several cases of significance at 12-month horizon. VARx and CatBoostx occasionally show parity or local advantage, but the multivariate DM test results largely confirm the relative superiority and robustness of the SZBVARx approach for multivariate macroeconomic forecasting in the presence of exogenous uncertainty shocks.

To augment the multivariate Diebold-Mariano test results further, we also employ the unconditional multivariate Giacomini-White (GW) test (Borup, Eriksen, Kjær & Thyrsgaard, 2024). The GW test framework has three main advantages: (i) HAC covariance estimation with Parzen kernel to account for heteroskedasticity and serial correlation in forecast errors⁵, (ii) an explicit accounting for estimation uncertainty in nested model comparisons relevant for rolling window evaluation exercises, and (iii) Wald-type of the loss differential vector, allowing for joint hypothesis testing and more powerful statistical inference (Giacomini & White, 2006; Mariano & Preve, 2012; Hansen, Lunde & Nason, 2011). The GW test evaluates $H_0 : \mathbb{E}[\Delta\mathcal{L}_{t+1}] = 0$ ("unconditional case" since no conditioning variables used) using a statistic $\mathcal{S} = T\bar{d}'\hat{\Sigma}^{-1}\bar{d}$, where \bar{d} is the mean loss differential vector, T is the number of forecast observations and $\hat{\Sigma}$ is the HAC covariance matrix. Rejection of the null hypothesis suggests that at least one of the models is superior in predictive performance.

Table A.1.1 in Appendix A.1.8 reports p-values from the unconditional multivariate GW test for SZBVARx versus VARx and CatBoostx, spanning G7 economies, five macroeconomic indicators, and both 12- and 24-month forecast horizons. The results provide compelling evidence of SZBVARx's forecasting edge, with 101 out of 140 tests (72.14%) significant at the 10% level. Significance rates are uniformly high across countries—ranging from 65% (Canada), 70% (France, United States) to 75% (United Kingdom, Germany, Italy, Japan)—and across variables, peaking at 78.57% for both unemployment rate and SIR, and remaining robust for REER (75.00%), oil price (WTI) (71.43%), and CPI inflation (57.14%). Horizon-specific results further underscore SZBVARx's dominance at longer horizons: the model outperforms CatBoostx in 80.00% and VARx in 77.14% of 24-month tests, compared to 65.71% (for both VARx and CatBoostx) at 12-month horizon, respectively. Notably, robust significance is observed for CPI inflation and SIR forecasts in Canada, the United States, France, Japan, and the United Kingdom at 24-month horizon, with several significant results at 12-month horizon as well. Although our implementation is unconditional, the

⁵Multivariate GW test has been implemented using `fEval` Python package (<https://github.com/ogrnz/fEval>).

remarkable consistency and breadth of significant results across countries, variables, and horizons provides robust statistical evidence for the forecasting supremacy of SZBVARx over both statistical and machine learning benchmarks, especially at extended horizons and under macroeconomic uncertainty.

Table 12: Forecasting performance comparison using the Multivariate Diebold-Mariano Test for SZBVARx, VARx, and CatBoostx across G7 countries at 12-month-ahead and 24-month-ahead forecast horizons. **Bold with *** values denote significance at the 10% level

Country	Forecast Horizon	SZBVARx Vs VARx (p-value)	SZBVARx Vs CatBoostx (p-value)
Canada	12	0.446	0.015*
	24	0.002*	0.000*
US	12	0.073*	0.010*
	24	0.010*	0.000*
France	12	0.464	0.002*
	24	0.019*	0.000*
Germany	12	0.114	0.068*
	24	0.005*	0.030*
Japan	12	0.031*	0.002*
	24	0.044*	0.000*
UK	12	0.004*	0.686
	24	0.011*	0.073*
Italy	12	0.123	0.606
	24	0.001*	0.378

5.5. Uncertainty Quantification through Predictive Credible Intervals

In addition to point forecasts, characterizing the predictive uncertainty is also of great interest for economic decision-making and risk management. Bayesian forecasting yields the *predictive* distribution for future observations *conditional on posterior-mean parameters*. We report credible prediction intervals that reflect *future shock uncertainty* conditional on posterior-mean parameters $(\bar{\mu}, \bar{\Phi}_1, \dots, \bar{\Phi}_p, \bar{\Gamma}, \bar{\Sigma})$ after hyperparameter tuning (Koop & Korobilis, 2010; Geweke & Whiteman, 2006) from Section 4.4⁶. A γ -level credible interval $[L_{T+h}, U_{T+h}]$ satisfies $\mathbb{P}(L_{T+h} \leq y_{T+h} \leq U_{T+h} | y_{1:T}) = \gamma$. We simulate future shocks for each draw s and horizon $h = 1, \dots, H$ as:⁷

$$\varepsilon_{T+h}^{(s)} \sim \mathcal{N}_m(0, \bar{\Sigma}), \quad (25)$$

and compute the forecast trajectory stepwise:

$$\tilde{y}_{T+h|y_T}^{(s)} = \bar{\mu} + \sum_{i=1}^p \bar{\Phi}_i \tilde{y}_{T+h-i|y_T}^{(s)} + \bar{\Gamma} x_{T+h} + \varepsilon_{T+h}^{(s)}, \quad (26)$$

⁶This ensures intervals are centered on optimal point forecasts from hyperparameter tuning rather than resimulating from different specifications. The snap-centering adjustment aligns the simulated distribution with the deterministic forecast path while preserving stochastic uncertainty. Intervals are constructed component-wise (marginal); joint credible regions are not reported.

⁷Here, $s = 1, \dots, S$ indexes Monte Carlo shock draws, with parameters fixed at posterior means.

initialized with $\hat{y}_{T+1-j|y_T}^{(s)} = y_{T+1-j}$ for $j = 1, \dots, p$.⁸ To ensure forecast consistency, we apply snap-centering: $\delta_h = \hat{y}_{T+h|y_T}^{\text{PF}} - \hat{y}_{T+h|y_T}$, where $\hat{y}_{T+h|y_T}$ is the deterministic forecast (with $\varepsilon_{T+h} \equiv 0$) and $\hat{y}_{T+h|y_T}^{\text{PF}}$ is the tuned point forecast,⁹ yielding:

$$\hat{y}_{T+h|y_T}^{(s)} = \tilde{y}_{T+h|y_T}^{(s)} + \delta_h. \quad (27)$$

We impose admissibility constraints by truncating draws to economically meaningful support bounds $[a_\ell, b_\ell]$ for each variable ℓ , where a_ℓ and b_ℓ are determined by the natural constraints of each macroeconomic variable.¹⁰ Let $\mathcal{S}_{\ell,h} = \{s : a_\ell \leq \hat{y}_{\ell,T+h|y_T}^{(s)} \leq b_\ell\}$ denote admissible draws. The γ -level credible interval is the shortest interval containing the point forecast with at least γ probability mass.¹¹

$$[L_{\ell,T+h}, U_{\ell,T+h}] = \arg \min_{[L,U]: \hat{y}_{\ell,T+h|y_T}^{\text{PF}} \in [L,U]} \left\{ U - L : \frac{1}{|\mathcal{S}_{\ell,h}|} \sum_{s \in \mathcal{S}_{\ell,h}} \mathbb{1}\{L \leq \hat{y}_{\ell,T+h|y_T}^{(s)} \leq U\} \geq \gamma \right\}. \quad (28)$$

When truncation binds, the effective mass covered is $\gamma_{\text{eff}} = \gamma \cdot \rho_{\ell,h}$. These PF-anchored shortest-mass (Highest Posterior Density-style) intervals guarantee $\hat{y}_{\ell,T+h}^{\text{PF}} \in [L_{\ell,T+h}, U_{\ell,T+h}]$ and are appropriate under truncation-induced skewness. We use $S = 1000$ draws and $\gamma = 0.50$.

Figure 6 reports the out-of-sample forecast performance of 24-month-ahead forecast of all G7 countries. We observe that the heterogeneity across the countries and across the variables in terms of the predictive uncertainty is remarkable. The SZBVARx model outperforms the VARx and CatBoostx baselines in comparison to the ground truth. We find that the credible intervals include the realized values in a majority of cases across countries and variables, providing empirical evidence that the uncertainty quantification procedure described in Equations (25)–(28) yields well-calibrated predictive distributions for the 24-month forecast horizon. It is worth noting that in all countries and for all variables, the credible intervals are small in the first forecast horizons and they widen at longer horizons. This occurs because forecast uncertainty accumulates over longer horizons as shock uncertainty compounds over time in Bayesian predictive intervals. We see relatively narrow intervals for unemployment rate forecasts in the United Kingdom and Italy, moderate uncertainty bands in Japan and Germany, and substantially wider uncertainty in Canada, France, and the United States. For REER predictions, we observe the tightest intervals in Germany and the United States, followed by the United Kingdom, France, and Italy, while Canada and Japan feature broader uncertainty. Tight intervals are also observed in the United States and Germany for SIR forecasts, while the uncertainty is moderate in the United Kingdom, Italy, and Japan, and fairly wide in France and Canada, in line with elevated monetary policy uncertainty in those countries. In contrast, the intervals for oil price predictions are consistently wide but relatively homogeneous across all G7 countries, highlighting the high volatility and geopolitical risk sensitivity in energy markets. CPI

⁸The exogenous path uses the last H training observations: $x_{T+h} = x_{T-H+h}$ for $h = 1, \dots, H$, ensuring no data leakage.

⁹The adjustment δ_h is a vector of length m .

¹⁰E.g., unemployment rate $\in [0, 100]$, real exchange rate $\in [0, \infty)$. Acceptance rate $\rho_{\ell,h}$ measures the proportion of admissible draws; low rates indicate potential misspecification.

¹¹When the point forecast lies outside the truncated empirical support, we anchor at the point forecast and extend to the nearest γ -mass window, ensuring $\hat{y}_{\ell,T+h|y_T}^{\text{PF}} \in [L_{\ell,T+h}, U_{\ell,T+h}]$.

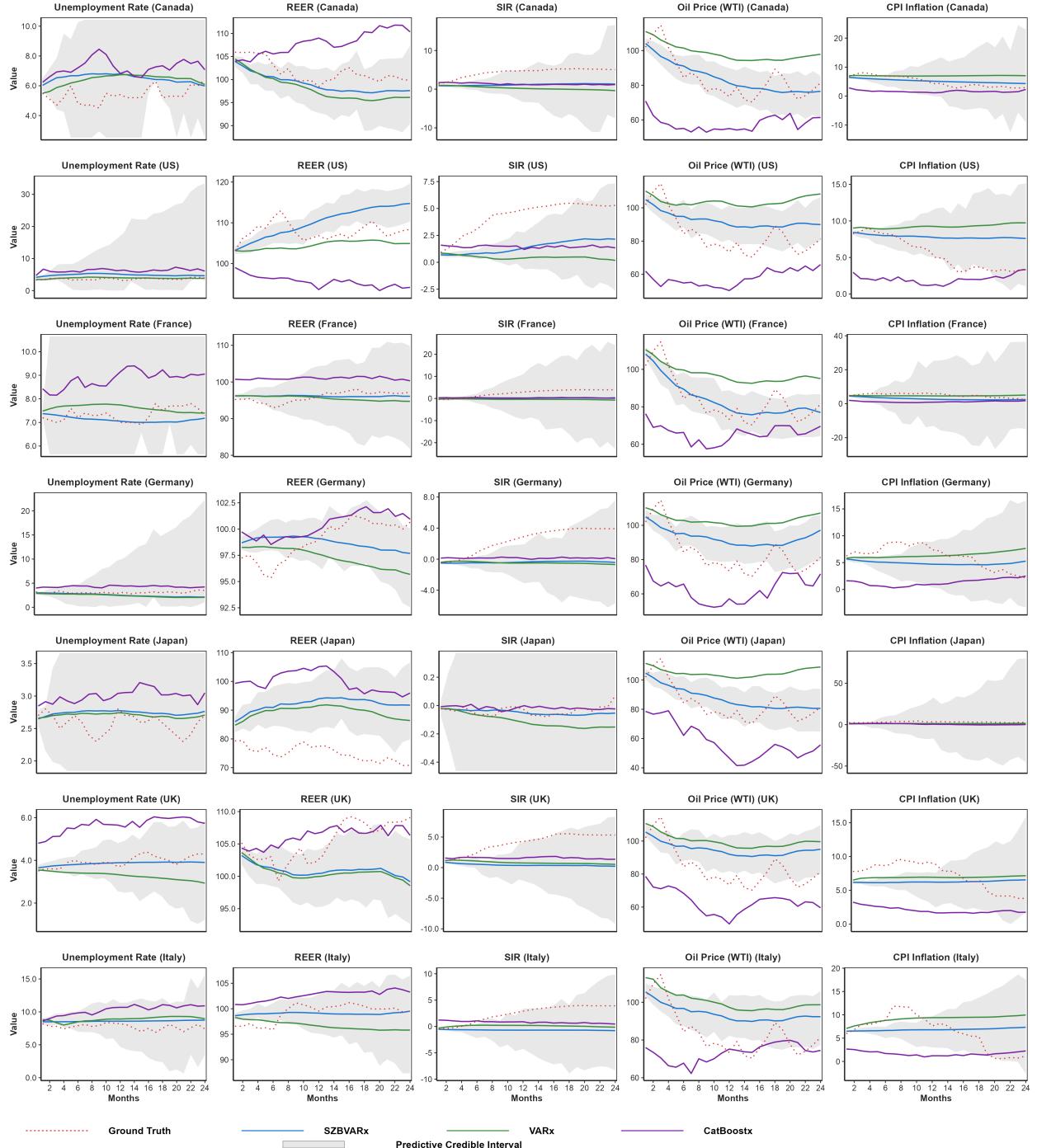


Figure 6: Visualization of 24-month-ahead forecasts for five macroeconomic indicators across G7 countries. Each panel displays the ground truth (red dotted line), point forecasts from SZBVARx (blue), VARx (green), and CatBoostx (purple) models, along with predictive intervals from SZBVARx (grey shaded regions) quantifying forecast uncertainty. Variables shown are: Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation Rate. Rows represent countries (Canada, US, France, Germany, Japan, UK, Italy from top to bottom), while columns represent the five macroeconomic variables.

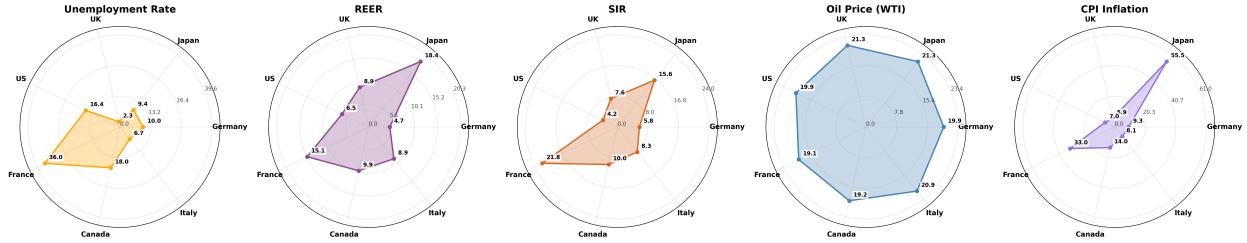


Figure 7: Visualization of average predictive credible interval widths for 24-month-ahead forecasts across five G7 macroeconomic indicators using SZBVARx model. Each subplot displays interval width distributions across G7 economies for: Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation. Lower interval widths indicate superior prediction precision and enhanced forecast reliability.

inflation forecasts stand out with the most cross-country heterogeneity: the United Kingdom, the United States, and Germany maintain relatively narrow intervals, while Italy and Canada exhibit moderate to wide uncertainty, and France and Japan display exceptionally wide intervals, reflecting high inflation volatility and policy uncertainty in the forecast period. The 12-month ahead forecast results, presented in Appendix A.1.9 and Figure A.1.28, exhibit qualitatively similar patterns with narrower credible intervals overall, reflecting reduced forecast uncertainty at shorter horizons.

To better understand the cross-country heterogeneity in the forecast uncertainty, Figure 7 plots the average width of the 50% predictive credible intervals across the five macroeconomic variables in each of the G7 countries¹². The radar chart shows that the unemployment rate intervals have the smallest range in the United Kingdom (2.3) and the widest range in France (36.0); the REER intervals are the narrowest in Germany (4.7) and the widest in Japan (18.4); the SIR intervals range from the smallest in the United States (4.2) to the largest in France (21.8) reflecting the cross-country heterogeneity in the monetary policy uncertainty; the oil price intervals are the most homogeneous across the G7 countries (19–21) with the narrowest range in Germany; and the CPI inflation intervals are the most dispersed ranging from the United Kingdom (5.9) to Japan (55.5), with France (33.0) and Japan exhibiting abnormally high inflation uncertainty during the forecast horizon. These point-wise summary measures are in line with the visual patterns in Figure 6 and highlight the importance of quantifying the forecast uncertainty on both the country and the variable dimensions.

6. Policy Implications

The empirical results of this paper point to several policy-relevant facts. First, uncertainty shocks are economically meaningful and transmit through identifiable channels. Geopolitical risk

¹²We report the average width of 50% predictive credible intervals to provide a standardized measure of forecast uncertainty that is directly comparable across countries and variables. Mathematically, the interval width $W_{0.50} = Q_{0.75} - Q_{0.25}$ captures the interquartile range of the predictive distribution, where Q_p denotes the p -th quantile. This choice is consistent with $\gamma = 0.50$ used in constructing the predictive credible intervals and provides a robust measure of central dispersion that is less sensitive to extreme tail behavior than wider intervals, making it particularly suitable for cross-country comparisons.

and equity-market volatility primarily operate via the external sector—appreciations (depreciations) of the REER and movements in oil prices; while economic policy uncertainty and monetary policy uncertainty affect the SIR, inflation, and unemployment. Second, transmission is state dependent: impulse responses are typically larger and faster in high-rate (tight) regimes and more persistent but diminished in low-rate (easy) regimes, indicating an interaction between uncertainty, financial frictions, and available policy space (Caldara et al., 2016; Lhuissier & Tripier, 2021). Third, the proposed SZBVARx delivers superior accuracy at 12–24 months, consistent with the advantages of Bayesian shrinkage and stochastic-volatility priors in high-dimensional settings (Bañbura et al., 2010; Carriero et al., 2019a). Finally, the Bayesian predictive credible intervals demonstrate reasonable empirical coverage and appropriately scale with forecast horizon, providing credible uncertainty quantification that is essential for policy communication, risk assessment, and scenario planning under macroeconomic uncertainty (Giannone et al., 2015; Clark, 2011; Knüppel, 2015).

These results have several policy implications. For monetary policy, the decision framework should condition explicitly on the rate regime and contemporaneous uncertainty indicators. As wavelet coherence indicates that the four uncertainty series are generally leading the macro variables, and the macro block is endogenous, it is reasonable to consider economic policy uncertainty, geopolitical risk, equity-market volatility, and monetary policy uncertainty as exogenous shocks for the purpose of scenario design. In practice, this means that central banks should generate regime-conditioned (high-rate vs. low-rate) parallel projections and evaluate the degree to which unemployment and inflation paths start to materially diverge; where divergence is sizeable, they should lean against amplification in tight conditions and tolerate slower real-side adjustment, recognizing asymmetric interactions between uncertainty and financial conditions (Caldara et al., 2016; Lhuissier & Tripier, 2021). Given its medium-horizon advantage, the SZBVARx is also well suited to forward guidance, where coherence across variables over 1–2 years matters for credibility (Bernanke, 2020; Yellen, 2017).

For exchange rate and external sector management, the high prominence of the external channel suggests that geopolitical risk and financial-market volatility may push risk premia and capital flows in ways that induce safe-haven appreciations for some currencies and depreciation for others, with first-round effects on competitiveness and imported inflation. In particular, small, open, and energy-intensive economies should accompany monetary policy moves with Forex liquidity backstops, reserve management, and transparent communication about how uncertainty shocks translate into exchange-rate movements and the inflation target (Kilian, 2009; Caldara & Iacoviello, 2022). In the euro area, cross-country heterogeneity in exchange-rate sensitivity further argues for the provision of national macro-stabilization tools to accompany area-wide policy in case of common shocks with asymmetric external impacts.

For energy and commodity risk, the forecast evaluation results indicate that oil-price projections exhibit the widest prediction intervals—particularly during global stress periods—suggesting that fiscal and regulatory authorities should incorporate energy-price stress tests into budget planning and sectoral oversight while supporting hedging strategies that reduce exposure to abrupt commodity swings (Kilian, 2009). Where energy import dependence is high, pre-specified contingencies (such as temporary energy-levy adjustments or targeted transfers) can be activated when prediction intervals for oil-price pass-through indicate elevated near-term inflation risk.

For fiscal policy, uncertainty shocks in low-rate regimes tend to raise unemployment while dampening inflation, strengthening the case for automatic stabilizers that are most active when monetary space is constrained ([Auerbach & Gorodnichenko, 2012](#)). Predictive credible intervals provide a basis for sizing fiscal buffers ex ante: wider unemployment or inflation intervals warrant larger precautionary space and more conservative debt-sustainability margins. The model indicates particularly wide prediction intervals for CPI inflation or short-term interest rates; thus, budget authorities should broaden the revenue and expenditure scenario ranges in their frameworks.

For macroprudential policy, equity-market volatility and monetary policy uncertainty propagate via financial conditions and rate expectations. The width and skew of credible intervals around interest rate and exchange rate forecasts can therefore guide countercyclical capital buffer calibration and liquidity coverage targets. When the model signals widening short rate and real effective exchange rate uncertainty under tight regimes, earlier activation and slower release of buffers can mitigate credit-channel amplification; when regimes are stable and credible intervals are narrow, more targeted borrower-based measures may be preferred to broad capital tightening ([Carstens, 2021](#)).

Finally, for forecast governance and communication, institutions should leverage the point-forecast gains of SZBVARx. Publishing fan charts based on predictive credible intervals can shift attention from single central paths to risk-weighted ranges, consistent with best practice in inflation-report communication and with the need to avoid overconfidence from parametric intervals that may be misspecified ([Britton, Fisher & Whitley, 1998](#)). A practical workflow is to (i) nowcast and monitor the exogenous uncertainty block with thresholds that trigger scenario refreshes; (ii) run and disclose parallel regime-conditioned projections, documenting policy choices against the implied amplification-versus-persistence trade-off; and (iii) communicate credible intervals to anchor expectations and guide contingency planning. While judgment is indispensable in rare regime shifts or institutional changes, the architecture developed here offers a practical and evidence-based template for resilient policy under uncertainty in the G7.

7. Conclusion and Discussion

This paper develops and evaluates an uncertainty-resilient Bayesian VAR framework with exogenous drivers (SZBVARx) for multivariate macroeconomic forecasting in the G7. The framework integrates a compact block of external uncertainty indicators - EPU, GPR, USEMV, and USMPU - with five core macroeconomic variables: unemployment rate, REER, SIR, oil price (WTI), and CPI inflation. The central premise is that policy and geopolitical uncertainty, financial-market volatility, and monetary policy uncertainty are systematic determinants of macroeconomic outcomes. Embedding these drivers directly in the forecasting system improves point predictions and clarifies the mechanisms through which external risks propagate across G7 economies.

Two features set the proposed method apart on a methodological level. First, non-linear local projections ([Jordà, 2005](#)) are used to estimate impulse responses, which are state-dependent and path-dependent, in contrast to tightly specified structural VARs that require strong parametric restrictions. Second, Bayesian point forecasts are combined with credible intervals derived from the

posterior predictive distribution to quantify forecast uncertainty across horizons ([Giannone et al., 2015](#); [Clark, 2011](#)). In large-scale benchmarks against 14 different methods, from classical time series to machine learning and deep learning, at multiple horizons, the SZBVARx systematically improves accuracy across a range of metrics. A suite of statistical tests confirms that improvements are economically meaningful and not driven by any single metric. Taken together, these results provide a new standard for macroeconomic forecasting in advanced economies that explicitly models uncertainty and regime dependence. Although the previous section already provides a detailed discussion of policy implications, it is worth emphasizing the immediate relevance of the framework for policy-makers. The SZBVARx produces coherent medium-term projections across variables, transparent regime-conditioned impulse responses that make the amplification–persistence trade-off explicit, and Bayesian credible intervals suitable for fan-chart communication and contingency planning. In an environment characterized by recurrent global shocks and evolving monetary regimes, these features constitute a practical decision-support toolkit and a sound methodological foundation for future, uncertainty-aware forecasting.

The findings in this paper open up several avenues for interesting extensions. Newspaper-based measures of uncertainty are noisy and cross-country comparable. Market-implied indicators, sentiment-based gauges, or measures of supply-chain and climate-policy risk, in particular, could be used to strengthen the robustness of our findings. Regime classification is based on the short-term policy rate. Composite state variables that aggregate information from the policy rate and financial-conditions indices or term-premium measures may improve classification. The parsimonious forecasting block could be expanded to more sectoral disaggregates (e.g., tradables vs. non-tradables inflation; vacancy–unemployment dynamics) to better isolate policy-relevant margins. On the uncertainty quantification side, while this paper uses Bayesian predictive credible intervals obtained from the posterior predictive distribution, an important avenue for future research is the use of multivariate conformal prediction intervals ([Vovk, Gammerman & Shafer, 2005](#); [Romano, Patterson & Candes, 2019](#)). This would be a particularly promising direction for macroeconomic forecasting, where model uncertainty is pervasive and coverage validity is critical for policy communication. A fully real-time assessment, in which the mapping from forecasts to nowcasts would account for publication lags and data revisions, would make the framework even more operationally amenable. We finally point out that the approach could be easily extended to emerging markets and developing economies, many of which are affected by different transmission channels, subject to higher volatility, and with typically less policy space to maneuver than G7 economies. We leave the exploration of the above directions for future work.

Competing interest

There are no competing interests to be declared.

Data and Code Availability Statement

The macroeconomic time series data for the G7 countries analyzed in this study were sourced from the Federal Reserve Economic Data (FRED) repository (<https://fred.stlouisfed.org/>). Measures of economic policy uncertainty, geopolitical risk, and related uncertainty indices were obtained from the Policy Uncertainty database (<https://www.policyuncertainty.com/>). The code and compiled datasets used to reproduce the results of this study are available at: <https://github.com/ctanujit/Macrocasting>.

References

- Adrian, T., Boyarchenko, N., & Giannone, D. (2019). Vulnerable growth. *American Economic Review*, 109, 1263–1289.
- Adämmer, P. (2019). lpirfs: An r package to estimate impulse response functions by local projections. *The R Journal*, 11, 421–438.
- Aguiar-Conraria, L., & Soares, M. J. (2014). The continuous wavelet transform: Moving beyond uni-and bivariate analysis. *Journal of economic surveys*, 28, 344–375.
- Ahir, H., Bloom, N., & Furceri, D. (2022). *The World Uncertainty Index*. Technical Report National Bureau of Economic Research.
- Andrews, D. W. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59, 817–858.
- Ang, A., & Bekaert, G. (2002). Regime switches in interest rates. *Journal of Business & Economic Statistics*, 20, 163–182.
- Auerbach, A. J., & Gorodnichenko, Y. (2012). Measuring the output responses to fiscal policy. *American Economic Journal: Economic Policy*, 4, 1–27.
- Bai, Y., Carriero, A., Clark, T. E., & Marcellino, M. (2022). Macroeconomic forecasting in a multi-country context. *Journal of Applied Econometrics*, 37, 1230–1255.
- Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131, 1593–1636.
- Baker, S. R., Bloom, N., Davis, S. J., & Kost, K. J. (2019). *Policy news and stock market volatility*. Technical Report National Bureau of Economic Research.
- Balke, N. S., & Fomby, T. B. (1997). Threshold cointegration. *International economic review*, (pp. 627–645).
- Bańbura, M., Giannone, D., & Reichlin, L. (2010). Large bayesian vector auto regressions. *Journal of applied Econometrics*, 25, 71–92.

- Baumeister, C., & Kilian, L. (2016). Forty years of oil price fluctuations: Why the price of oil may still surprise us. *Journal of Economic Perspectives*, 30, 139–160.
- Bekaert, G., Hoerova, M., & Duca, M. L. (2013). Risk, uncertainty and monetary policy. *Journal of Monetary Economics*, 60, 771–788.
- Bekiros, S. (2014). Forecasting with a state space time-varying parameter var model: Evidence from the euro area. *Economic Modelling*, 38, 619–626.
- Belmonte, M. A., Koop, G., & Korobilis, D. (2014). Hierarchical shrinkage in time-varying parameter models. *Journal of Forecasting*, 33, 80–94.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57, 289–300.
- Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, 29, 1165–1188.
- Berg, T. O. (2019). Business uncertainty and the effectiveness of fiscal policy in germany. *Macroeconomic Dynamics*, 23, 1442–1470.
- Bernanke, B. S. (2015). *The courage to act: A memoir of a crisis and its aftermath*. W. W. Norton & Company.
- Bernanke, B. S. (2020). The new tools of monetary policy. *American Economic Review*, 110, 943–983.
- Bernanke, B. S., Laubach, T., Mishkin, F. S., & Posen, A. S. (2000). Inflation targeting: Lessons from the international experience. princeton, nj: Princeton univ. press, 1999. inflation targeting has become a viable framework of monetary policy in a number of industrial economies in the 1990s. by applying this framework.
- Besher, D., Sengupta, A., & Chakraborty, T. (2025). Forecasting climate policy uncertainty: Evidence from the united states. *arXiv preprint arXiv:2507.12276*, .
- Bhuiyan, R., & Lucas, R. F. (2007). Real and nominal effects of monetary policy shocks. *Canadian Journal of Economics/Revue canadienne d'économique*, 40, 679–702.
- Bloom, N. (2009). The impact of uncertainty shocks. *econometrica*, 77, 623–685.
- Bloom, N., Bunn, P., Mizen, P., Smietanka, P., & Thwaites, G. (2019). The impact of brexit on uk firms. *NBER Working Paper Series*, .
- Bobasu, A., Quaglietti, L., & Ricci, M. (2024). Tracking global economic uncertainty: implications for the euro area. *IMF Economic Review*, 72, 820–857.
- Bojer, C. S. (2022). Understanding machine learning-based forecasting methods: A decomposition framework and research opportunities. *International Journal of Forecasting*, 38, 1555–1561.
- Borup, D., Eriksen, J. N., Kjær, M. M., & Thyrgaard, M. (2024). Predicting bond return predictability. *Management Science*, 70, 931–951.

- Britton, E., Fisher, P., & Whitley, J. (1998). The inflation report projections: understanding the fan chart. *Chart*, 8, 30.
- Caggiano, G., Castelnuovo, E., & Groshenny, N. (2014). Uncertainty shocks and unemployment dynamics in us recessions. *Journal of Monetary Economics*, 67, 78–92.
- Caldara, D., Fuentes-Albero, C., Gilchrist, S., & Zakrajšek, E. (2016). The macroeconomic impact of financial and uncertainty shocks. *European Economic Review*, 88, 185–207.
- Caldara, D., & Iacoviello, M. (2022). Measuring geopolitical risk. *American Economic Review*, 112, 1194–1225.
- Canova, F. (2011). *Methods for applied macroeconomic research*. Princeton university press.
- Canova, F., & Ciccarelli, M. (2004). Forecasting and turning point predictions in a bayesian panel var model. *Journal of Econometrics*, 120, 327–359.
- Carriero, A., Clark, T. E., & Marcellino, M. (2015). Bayesian vars: specification choices and forecast accuracy. *Journal of Applied Econometrics*, 30, 46–73.
- Carriero, A., Clark, T. E., & Marcellino, M. (2019a). Large bayesian vector autoregressions with stochastic volatility and non-conjugate priors. *Journal of Econometrics*, 212, 137–154.
- Carriero, A., Galvao, A. B., & Kapetanios, G. (2019b). A comprehensive evaluation of macroeconomic forecasting methods. *International Journal of Forecasting*, 35, 1226–1239.
- Carstens, A. (2021). Central banks and inequality. *Remarks at the Markus' Academy, Princeton University's Bendheim Center for Finance, May*, 6.
- Castelnuovo, E. (2019). *Domestic and Global Uncertainty: A Survey and Some New Results*. CAMA Working Paper 75/2019 Centre for Applied Macroeconomic Analysis (CAMA), Australian National University Canberra.
- Cesa-Bianchi, A., Thwaites, G., & Vicondoa, A. (2020). Monetary policy transmission in the united kingdom: A high frequency identification approach. *European Economic Review*, 123, 103375.
- Chakraborty, T., Besher, D., Panja, M., & Sengupta, S. (2025). Neural arfima model for forecasting bric exchange rates with long memory under oil shocks and policy uncertainties. *arXiv preprint arXiv:2509.06697*,
- Challu, C., Olivares, K. G., Oreshkin, B. N., Garza, F., Mergenthaler-Canseco, M., & Dubrawski, A. (2023). Nhits: Neural hierarchical interpolation for time series forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37.
- Chan, J. C. (2020). Large bayesian vars: A flexible kronecker error covariance structure. *Journal of Business & Economic Statistics*, 38, 68–79.
- Chen, S.-A., Li, C.-L., Yoder, N., Arik, S. Ö., & Pfister, T. (2023). Tsmixer: An all-mlp architecture for time series forecasting. [arXiv:2303.06053](https://arxiv.org/abs/2303.06053) arXiv preprint.
- Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 785–794).

- Clark, T. E. (2011). Real-time density forecasts from bayesian vector autoregressions with stochastic volatility. *Journal of Business & Economic Statistics*, 29, 327–341.
- Coibion, O., Gorodnichenko, Y., & Ulate, M. (2019). Is inflation just around the corner? the phillips curve and global inflationary pressures. In *AEA Papers and Proceedings* (pp. 465–469). volume 109.
- Credit Suisse Research Institute (2023). *Global Wealth Report 2023*. Technical Report Credit Suisse AG, a UBS Group company Zurich.
- Cuadrasma, J. C., Feldkircher, M., & Huber, F. (2016). Forecasting with global vector autoregressive models: A bayesian approach. *Journal of Applied Econometrics*, 31, 1371–1391.
- Darracq Pariès, M., Kornprobst, A., & Priftis, R. (2024). *Monetary policy strategies to navigate post-pandemic inflation: An assessment using the ECB's New Area-Wide Model*. 2935. ECB Working Paper.
- Das, A., Kong, W., Leach, A., Mathur, S., Sen, R., & Yu, R. (2023). Long-term forecasting with tide: Time-series dense encoder. *arXiv preprint arXiv:2304.08424*, .
- De Silva, A., Hyndman, R. J., & Snyder, R. (2010). The vector innovations structural time series framework: a simple approach to multivariate forecasting. *Statistical Modelling*, 10, 353–374.
- Del Negro, M., & Schorfheide, F. (2008). Forming priors for dsge models (and how it affects the assessment of nominal rigidities). *Journal of Monetary Economics*, 55, 1191–1208.
- Diebold, F. X., & Yilmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *The Economic Journal*, 119, 158–171.
- Ehm, W., Gneiting, T., Jordan, A., & Krüger, F. (2016). Of quantiles and expectiles: consistent scoring functions, choquet representations and forecast rankings. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 78, 505–562.
- Ekambaram, V., Jati, A., Nguyen, N., Sinthong, P., & Kalagnanam, J. (2023). Tsmixer: Lightweight mlp-mixer model for multivariate time series forecasting. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '23)*. ACM. Long Beach, CA, USA, August 6–10, 2023.
- Fajgelbaum, P. D., Goldberg, P. K., Kennedy, P. J., & Khandelwal, A. K. (2020). The return to protectionism. *The quarterly journal of economics*, 135, 1–55.
- Fan, W., Wang, P., Wang, D., Wang, D., Zhou, Y., & Fu, Y. (2023). Dish-ts: A general paradigm for alleviating distribution shift in time series forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37, 7522–7529.
- Feldkircher, M., Huber, F., Koop, G., & Pfarrhofer, M. (2022). Approximate bayesian inference and forecasting in huge-dimensional multicountry vars. *International Economic Review*, 63, 1625–1658.
- 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, 2530–2561.

- Forbes, K. J., & Warnock, F. E. (2012). Capital flow waves: Surges, stops, flight, and retrenchment. *Journal of International Economics*, 88, 235–251.
- Gefang, D., Koop, G., & Poon, A. (2023). Forecasting using variational bayesian inference in large vector autoregressions with hierarchical shrinkage. *International Journal of Forecasting*, 39, 346–363.
- Geweke, J., & Whiteman, C. (2006). Bayesian forecasting. *Handbook of economic forecasting*, 1, 3–80.
- Giacomini, R., & White, H. (2006). Tests of conditional predictive ability. *Econometrica*, 74, 1545–1578.
- Giannone, D., Lenza, M., & Primiceri, G. E. (2015). Prior selection for vector autoregressions. *The Review of Economics and Statistics*, 97, 436–451.
- Gneiting, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association*, 106, 746–762.
- Gopinath, G., Boz, E., Casas, C., Díez, F. J., Gourinchas, P.-O., & Plagborg-Møller, M. (2020). Dominant currency paradigm. *American Economic Review*, 110, 677–719.
- Grimme, C., & Stöckli, M. (2018). Measuring macroeconomic uncertainty in germany. In *CESifo Forum* (pp. 46–50). München: ifo Institut-Leibniz-Institut für Wirtschaftsforschung an der ... volume 19.
- Grinsted, A., Moore, J. C., & Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear processes in geophysics*, 11, 561–566.
- Hamilton, J. D. (1994). *Time Series Analysis*. Princeton, NJ: Princeton University Press.
- Hamilton, J. D. (2020). *Time series analysis*. Princeton university press.
- Hansen, B. E. (2001). The new econometrics of structural change: Dating breaks in us labor productivity. *Journal of Economic perspectives*, 15, 117–128.
- Hansen, B. E. (2011). Threshold autoregression in economics. *Statistics and its Interface*, 4, 123–127.
- Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The model confidence set. *Econometrica*, 79, 453–497.
- Haug, A. A., & Siklos, P. L. (2006). The behavior of short-term interest rates: international evidence of non-linear adjustment. *Studies in Nonlinear Dynamics & Econometrics*, 10.
- Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the econometric society*, (pp. 1251–1271).
- Hauzenberger, N., Huber, F., Marcellino, M., & Petz, N. (2025). Gaussian process vector autoregressions and macroeconomic uncertainty. *Journal of Business & Economic Statistics*, 43, 27–43.

- Heinisch, K., Scaramella, F., & Schult, C. (2025). *Assumption errors and forecast accuracy: A partial linear instrumental variable and double machine learning approach*. Technical Report IWH Discussion Papers.
- Herzen, J., Lässig, F., Piazzetta, S. G., Neuer, T., Tafti, L., Raille, G., Van Pottelbergh, T., Pasieka, M., Skrodzki, A., Huguenin, N. et al. (2022). Darts: User-friendly modern machine learning for time series. *Journal of Machine Learning Research*, 23, 1–6.
- Hewamalage, H., Ackermann, K., & Bergmeir, C. (2023). Forecast evaluation for data scientists: common pitfalls and best practices. *Data Mining and Knowledge Discovery*, 37, 788–832.
- Holzmann, H., & Klar, B. (2014). The role of the information set for forecasting—with applications to risk management. *Annals of Applied Statistics*, 8, 595–621.
- Huber, F., Koop, G., Onorante, L., Pfarrhofer, M., & Schreiner, J. (2023). Nowcasting in a pandemic using non-parametric mixed frequency vars. *Journal of Econometrics*, 232, 52–69.
- Husted, L., Rogers, J., & Sun, B. (2020). Monetary policy uncertainty. *Journal of Monetary Economics*, 115, 20–36.
- Hyndman, R., Koehler, A., Ord, K., & Snyder, R. (2008). *Forecasting with exponential smoothing: the state space approach*. Springer.
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.
- International Monetary Fund (2024a). Currency composition of official foreign exchange reserves (cofer). IMF Data.
- International Monetary Fund (2024b). *World Economic Outlook, October 2024: Policy Pivot, Rising Threats*. Technical Report International Monetary Fund Washington, DC.
- Jones, P. M., & Olson, E. (2013). The time-varying correlation between uncertainty, output, and inflation: Evidence from a dcc-garch model. *Economics Letters*, 118, 33–37.
- Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *American economic review*, 95, 161–182.
- Jordan, A., Krüger, F., & Lerch, S. (2019). Evaluating probabilistic forecasts with scoringrules. *Journal of Statistical Software*, 90, 1–37.
- Jurado, K., Ludvigson, S. C., & Ng, S. (2015). Measuring uncertainty. *American Economic Review*, 105, 1177–1216.
- Kang, W., Ratti, R. A., & Vespignani, J. L. (2017). Oil price shocks and policy uncertainty: New evidence on the effects of us and non-us oil production. *Energy Economics*, 66, 536–546.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30.
- Keenan, D. M. (1985). A tukey nonadditivity-type test for time series nonlinearity. *Biometrika*, 72, 39–44.

- Kilian, L. (2006). New introduction to multiple time series analysis, by helmut lütkepohl, springer, 2005. *Econometric theory*, 22, 961–967.
- Kilian, L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review*, 99, 1053–1069.
- Kilian, L., & Lütkepohl, H. (2017). *Structural vector autoregressive analysis*. Cambridge University Press.
- Kilian, L., & Murphy, D. P. (2014). The role of inventories and speculative trading in the global market for crude oil. *Journal of Applied Econometrics*, 29, 454–478.
- Knüppel, M. (2015). Evaluating the calibration of multi-step-ahead density forecasts using raw moments. *Journal of Business & Economic Statistics*, 33, 270–281.
- Komunjer, I., & Owyang, M. T. (2012). Multivariate forecast evaluation and rationality testing. *Review of Economics and Statistics*, 94, 1066–1080.
- Koning, A. J., Franses, P. H., Hibon, M., & Stekler, H. O. (2005). The m3 competition: Statistical tests of the results. *International Journal of Forecasting*, 21, 397–409.
- Koop, G., & Korobilis, D. (2010). Bayesian multivariate time series methods for empirical macroeconomics. *Foundations and Trends in Econometrics*, 3, 267–358.
- Koop, G., & Korobilis, D. (2013). Large time-varying parameter vars. *Journal of Econometrics*, 177, 185–198.
- Koop, G. M. (2013). Forecasting with medium and large bayesian vars. *Journal of Applied Econometrics*, 28, 177–203.
- Krueger, F., & Jordan, A. (2019). *murphydiagram: Murphy Diagrams for Forecast Comparisons*. R package version 0.12.2.
- Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of econometrics*, 54, 159–178.
- Laborda, J., Ruano, S., & Zamanillo, I. (2023). Multi-country and multi-horizon gdp forecasting using temporal fusion transformers. *Mathematics*, 11, 2625.
- Lee, S., & Lee, D. K. (2018). What is the proper way to apply the multiple comparison test? *Korean journal of anesthesiology*, 71, 353–360.
- Lhuissier, S., & Tripier, F. (2021). Regime-dependent effects of uncertainty shocks: A structural interpretation. *Quantitative Economics*, 12, 1139–1170.
- Lim, B., Arik, S. Ö., Loeff, N., & Pfister, T. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37, 1748–1764.
- Litterman, R. B. (1986). Forecasting with bayesian vector autoregressions—five years of experience. *Journal of Business & Economic Statistics*, 4, 25–38.

- Liu, Y., Liang, X. S., & Weisberg, R. H. (2007). Rectification of the bias in the wavelet power spectrum. *Journal of Atmospheric and Oceanic Technology*, 24, 2093–2102.
- Lo, M. C., & Zivot, E. (2001). Threshold cointegration and nonlinear adjustment to the law of one price. *Macroeconomic Dynamics*, 5, 533–576.
- Ludvigson, S. C., Ma, S., & Ng, S. (2021). Uncertainty and business cycles: exogenous impulse or endogenous response? *American Economic Journal: Macroeconomics*, 13, 369–410.
- Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Berlin: Springer.
- Mariano, R. S., & Preve, D. (2012). Statistical tests for multiple forecast comparison. *Journal of econometrics*, 169, 123–130.
- Miranda-Agrippino, S., & Ricco, G. (2019). Bayesian vector autoregressions: Estimation. In *Oxford Research Encyclopedia of Economics and Finance*.
- Miranda-Agrippino, S., & Ricco, G. (2021). The transmission of monetary policy shocks. *American Economic Journal: Macroeconomics*, 13, 74–107.
- Mishkin, F. S., & Schmidt-Hebbel, K. (2007). Does inflation targeting make a difference?
- Miyamoto, H. (2016). *Uncertainty Shocks and Labor Market Dynamics in Japan*. Working Paper SDES-2016-8 Kochi University of Technology, School of Economics and Management Kochi, Japan. Revised June 2016.
- Newey, W. K., & West, K. D. (1987a). Hypothesis testing with efficient method of moments estimation. *International Economic Review*, (pp. 777–787).
- Newey, W. K., & West, K. D. (1987b). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55, 703–708.
- Newey, W. K., & West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *The review of economic studies*, 61, 631–653.
- Ollech, D., & Webel, K. (2023). A random forest-based approach to combining and ranking seasonality tests. *Journal of Econometric Methods*, 12, 117–130.
- Oreshkin, B. N., Carpow, D., Chapados, N., & Bengio, Y. (2019). N-beats: Neural basis expansion analysis for interpretable time series forecasting. *arXiv preprint arXiv:1905.10437*, .
- Pesaran, M. H., Schuermann, T., & Smith, L. V. (2009). Forecasting economic and financial variables with global vars. *International journal of forecasting*, 25, 642–675.
- Pesaran, M. H., Schuermann, T., & Weiner, S. M. (2004). Modeling regional interdependencies using a global error-correcting macroeconomic model. *Journal of Business & Economic Statistics*, 22, 129–162.
- Ploberger, W., & Krämer, W. (1992). The cusum test with ols residuals. *Econometrica: Journal of the Econometric Society*, (pp. 271–285).

- Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). Catboost: unbiased boosting with categorical features. *Advances in neural information processing systems*, 31.
- Rey, H. (2015). *Dilemma not trilemma: the global financial cycle and monetary policy independence*. Technical Report National Bureau of Economic Research.
- Romano, Y., Patterson, E., & Candes, E. (2019). Conformalized quantile regression. *Advances in neural information processing systems*, 32.
- Rossi, B. (2013). Exchange rate predictability. *Journal of economic literature*, 51, 1063–1119.
- Sengupta, S., Chakraborty, T., & Singh, S. K. (2025). Forecasting cpi inflation under economic policy and geopolitical uncertainties. *International Journal of Forecasting*, 41, 953–981.
- Shang, H. L. (2020). A comparison of hurst exponent estimators in long-range dependent curve time series. *Journal of Time Series Econometrics*, 12, 20190009.
- Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: journal of the Econometric Society*, (pp. 1–48).
- Sims, C. A., Doan, T., & Litterman, R. B. (1983). *Forecasting and conditional projection using realistic prior distributions*. National Bureau of Economic Research.
- Sims, C. A., & Zha, T. (1998). Bayesian methods for dynamic multivariate models. *International Economic Review*, (pp. 949–968).
- Smales, L. A. (2022). The influence of policy uncertainty on exchange rate forecasting. *Journal of Forecasting*, 41, 997–1016.
- Stock, J. H., & Watson, M. W. (1996). Evidence on structural instability in macroeconomic time series relations. *Journal of Business & Economic Statistics*, 14, 11–30.
- Svetunkov, I. (2023). *Forecasting and analytics with the augmented dynamic adaptive model (ADAM)*. CRC Press.
- Svetunkov, I., Chen, H., & Boylan, J. E. (2023). A new taxonomy for vector exponential smoothing and its application to seasonal time series. *European Journal of Operational Research*, 304, 964–980.
- Tang, J. (2015). *Uncertainty and the signaling channel of monetary policy*. Technical Report Working Papers.
- Tillmann, P. (2020). Monetary policy uncertainty and the response of the yield curve to policy shocks. *Journal of Money, Credit and Banking*, 52, 803–833.
- Torrence, C., & Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological society*, 79, 61–78.
- Tsay, R. S. (1986). Nonlinearity tests for time series. *Biometrika*, 73, 461–466.

- Tsay, R. S. (1998). Testing and modeling multivariate threshold models. *Journal of the American Statistical Association*, 93, 1188–1202.
- Tsay, R. S. (2005). *Analysis of financial time series*. John Wiley & Sons.
- Veleda, D., Montagne, R., & Araujo, M. (2012). Cross-wavelet bias corrected by normalizing scales. *Journal of Atmospheric and Oceanic Technology*, 29, 1401–1408.
- Vovk, V., Gammerman, A., & Shafer, G. (2005). *Algorithmic learning in a random world* volume 29. Springer.
- Waggoner, D. F., & Zha, T. (2003). Likelihood preserving normalization in multiple equation models. *Journal of Econometrics*, 114, 329–347.
- Weisberg, S., & Cook, R. D. (1982). *Residuals and influence in regression*. Chapman & Hall.
- Wu, D.-M. (1973). Alternative tests of independence between stochastic regressors and disturbances. *Econometrica: journal of the Econometric Society*, (pp. 733–750).
- Yang, X., Zhang, Q., Liu, H., Liu, Z., Tao, Q., Lai, Y., & Huang, L. (2024). Economic policy uncertainty, macroeconomic shocks, and systemic risk: Evidence from China. *The North American Journal of Economics and Finance*, 69, 102032.
- Yellen, J. L. (2017). Inflation, uncertainty, and monetary policy. *Business Economics*, 52, 194–207.
- Ziegel, J. F., Krüger, F., Jordan, A., & Fasciati, F. (2017). Murphy diagrams: Forecast evaluation of expected shortfall. *arXiv preprint arXiv:1705.04537*, .

A.1. Appendix

A.1.1. Trend Plots

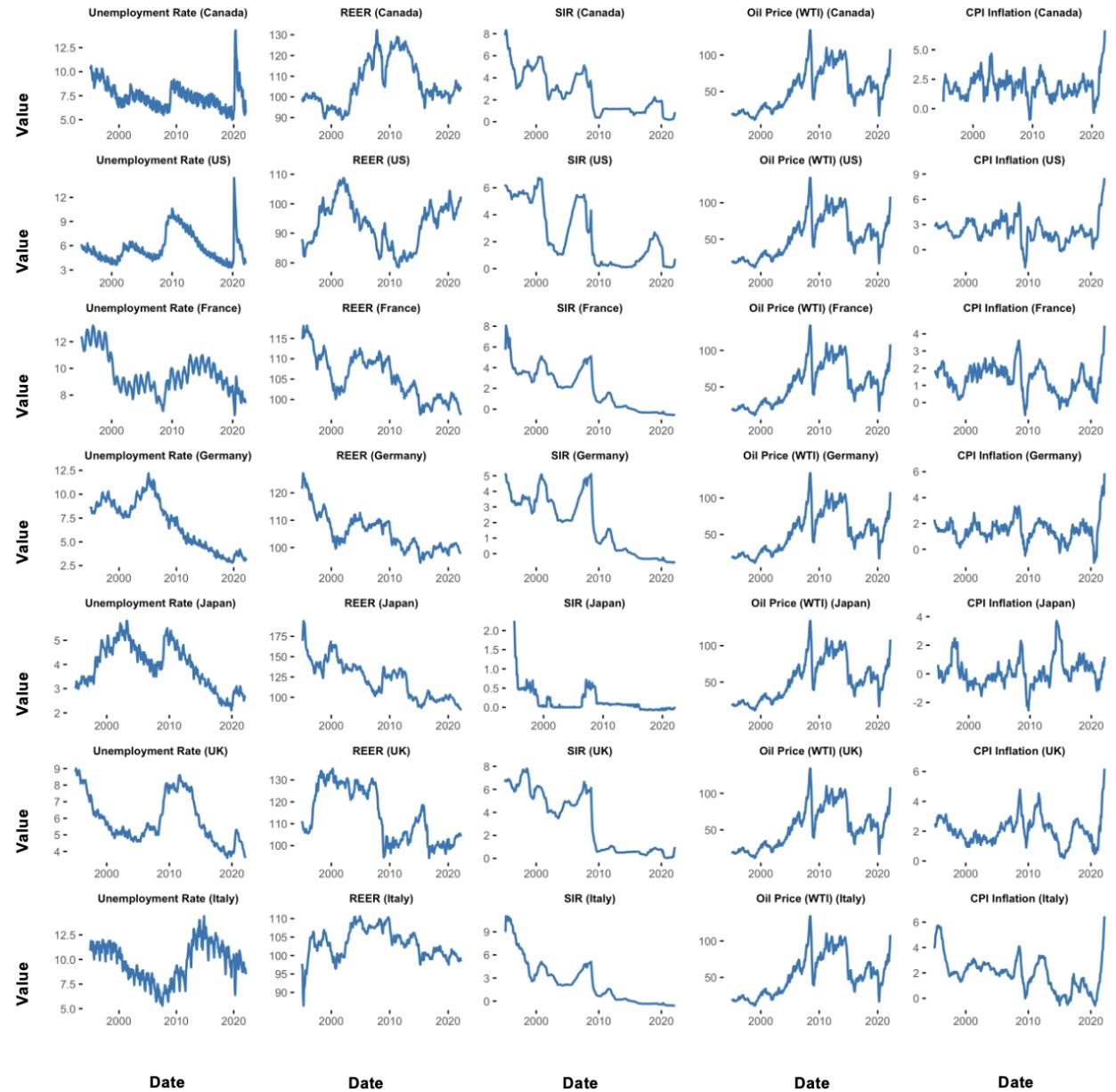


Figure A.1.1: Trend plots for Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation across G7 countries. Each subplot visualizes the time series evolution of a specific macroeconomic variable for a given country over the training period (Jan, 1995 – Mar, 2022).

A.1.2. ACF Plots

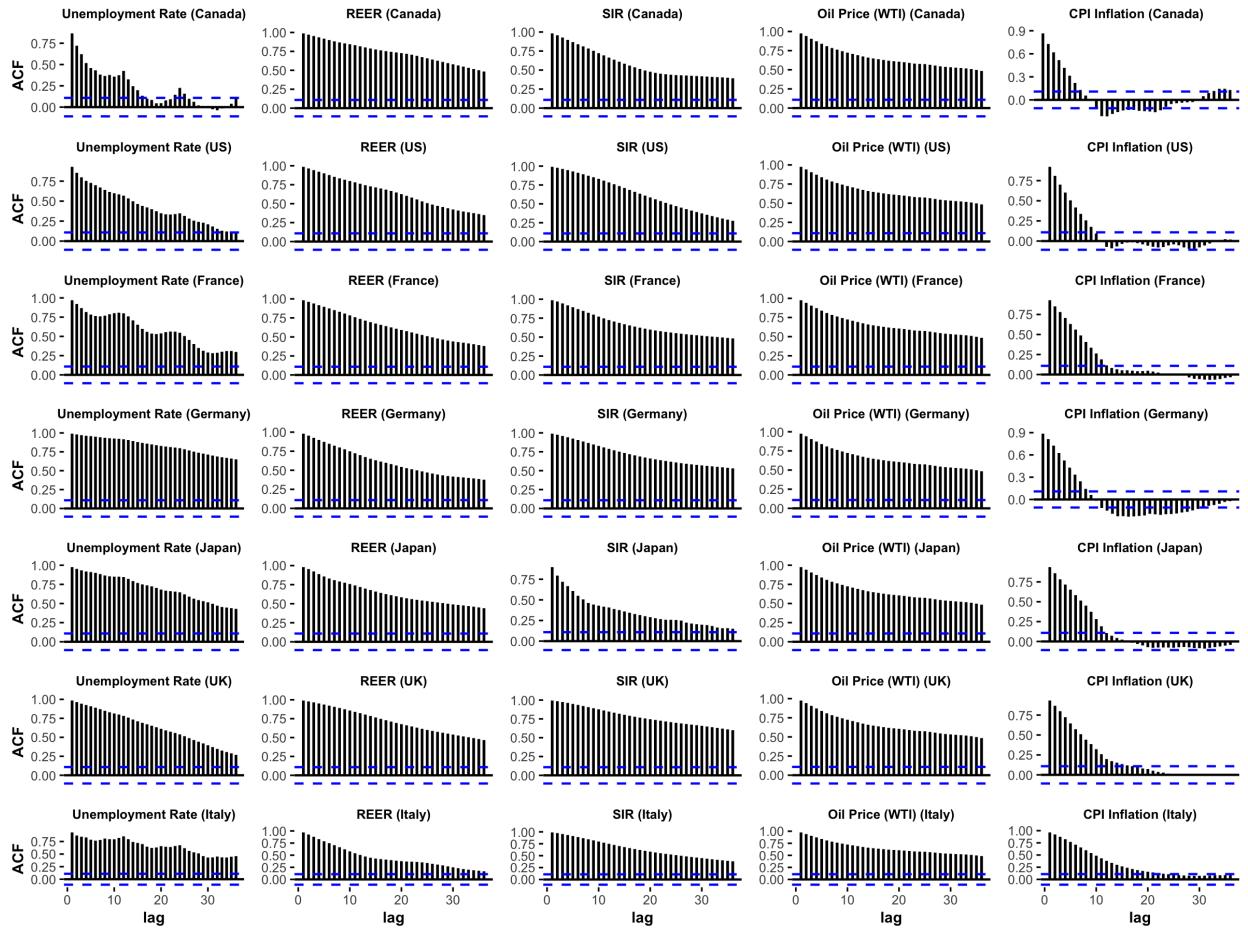


Figure A.1.2: ACF plots for Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation across G7 countries. Each subplot displays the autocorrelation structure for a specific macroeconomic variable and country over the training period (Jan, 1995 – Mar, 2022).

A.1.3. OLS-CUSUM test for structural breakpoint analysis

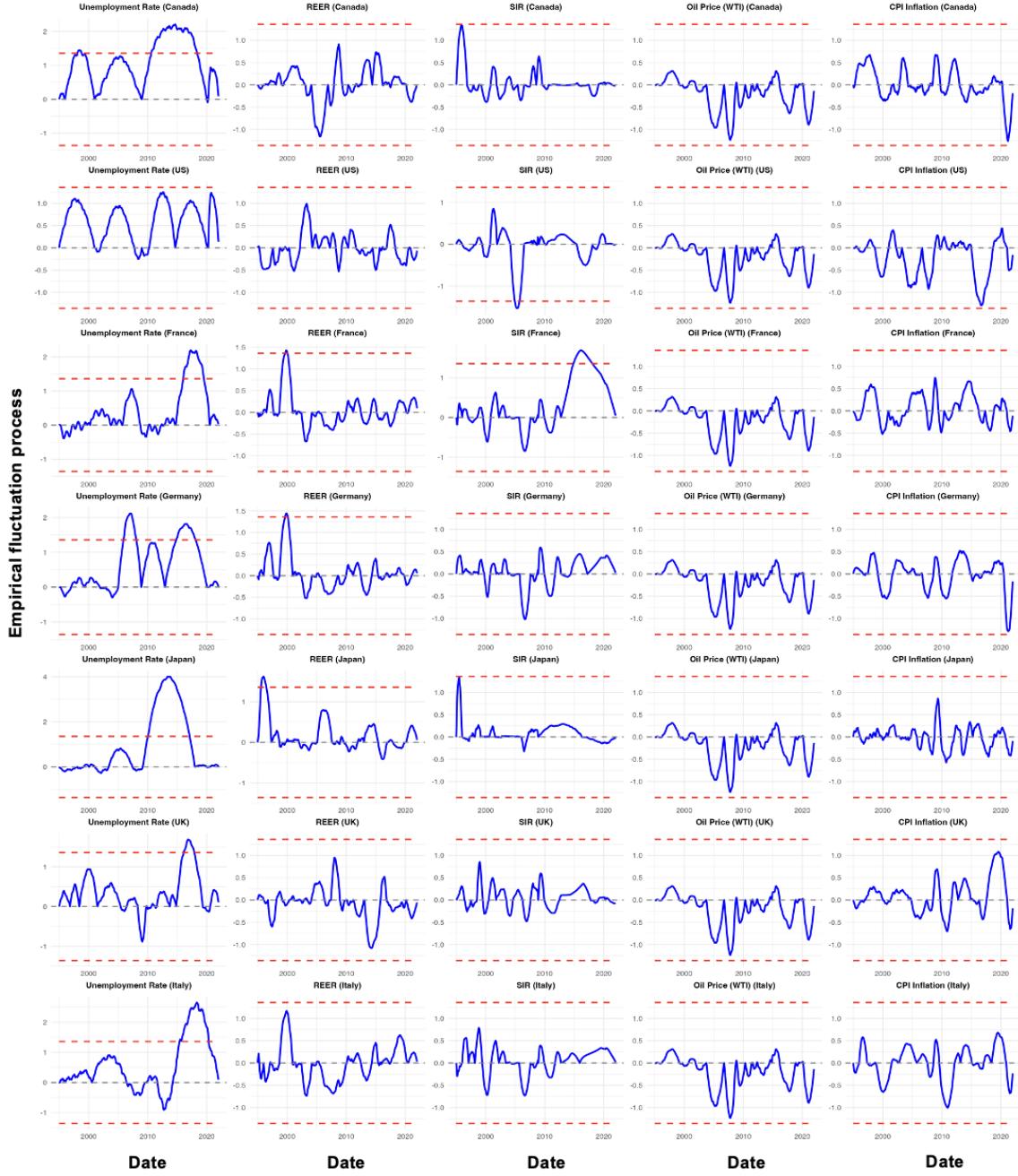


Figure A.1.3: OLS-CUSUM test plots for structural breakpoint detection in Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation across G7 countries. Blue lines show the CUSUM process; red dashed lines indicate 95% confidence boundaries. Excursions beyond the boundaries signal structural breaks in the respective macroeconomic series.

A.1.4. Causal Analysis

This appendix presents the formal mathematical details of the False Discovery Rate (FDR) corrected Wavelet Coherence Analysis (WCA) workflow used for finding time-frequency-localized associations and lead-lag relations between five key macroeconomic variables and four exogenous uncertainty shocks over the G7 sample. The implementation via the `biwavelet` package in **R** leverages the most recent advances in signal decomposition, bias correction, and valid multiple testing control to guide valid discovery of statistically significant findings.

Let $x(t)$ and $y(t)$ be two discrete-time processes sampled at interval dt , with t indexing observation times. The continuous wavelet transform (CWT) of $x(t)$, using a complex Morlet mother wavelet $\psi(\cdot)$ with $\omega_0 = 6$ and unit-energy normalization, is given by:

$$\mathcal{W}_x(s, \tau) = \sqrt{\frac{dt}{s}} \sum_{t'} x(t') \bar{\psi} \left(\frac{(t' - \tau)dt}{s} \right),$$

where s is the scale (inverse frequency), τ the translation (in time), and $\bar{\psi}$ the complex conjugate of the mother wavelet (Torrence & Compo, 1998). The cross-wavelet transform between $x(t)$ and $y(t)$ is:

$$\mathcal{W}_{xy}(s, \tau) = \mathcal{W}_x(s, \tau) \cdot \overline{\mathcal{W}_y(s, \tau)},$$

where $\overline{\mathcal{W}_y(s, \tau)}$ is the complex conjugate of the CWT of $y(t)$.

The localized wavelet coherence, bounded by $\mathcal{R}^2(s, \tau) \in [0, 1]$, is computed as:

$$\mathcal{R}^2(s, \tau) = \frac{|\mathcal{S}(s^{-1}\mathcal{W}_{xy}(s, \tau))|^2}{\mathcal{S}(s^{-1}|\mathcal{W}_x(s, \tau)|^2) \cdot \mathcal{S}(s^{-1}|\mathcal{W}_y(s, \tau)|^2)},$$

where $\mathcal{S}(\cdot)$ represents a two-dimensional smoothing operator applied in both time and scale dimensions using a Morlet-consistent time-scale smoothing kernel (Torrence & Compo, 1998; Grinsted et al., 2004). The phase difference, providing lead-lag information, is given by:

$$\phi_{xy}(s, \tau) = \arg(\mathcal{S}(s^{-1}\mathcal{W}_{xy}(s, \tau))),$$

where $\arg(\cdot)$ denotes the argument of the complex number. Lead-lag is interpreted via the phase-arrow convention: right-pointing arrows indicate in-phase relationships, left-pointing arrows indicate anti-phase relationships; arrows tilting upward indicate x leads y , and downward-tilting arrows indicate y leads x ; the angle magnitude encodes the lag duration.

To address low-frequency bias in cross-power, we apply the scale normalization of Veleda, Montagne & Araujo (2012), premultiplying each transform by \sqrt{s} before forming cross-power. This correction chiefly affects cross-power estimates, while the normalized coherence $\mathcal{R}^2(s, \tau)$ remains essentially invariant after smoothing operations. For auto-power bias background, see Liu, Liang & Weisberg (2007).

Statistical significance is assessed via Monte Carlo simulation using AR(1) surrogate time series that preserve the lag-1 autocorrelation structure of the original data, generating empirical null

distributions for $\mathcal{R}^2(s, \tau)$ at each (s, τ) point with $B = 1000$ realizations and significance level $\alpha = 0.05$. All testing excludes regions within the cone of influence (COI) where edge effects contaminate the wavelet transform ([Torrence & Compo, 1998](#)).

As the method involves multiple simultaneous hypothesis tests across the time-frequency grid, control for false discoveries is ensured through a two-stage approach. First, the Benjamini-Hochberg FDR procedure ([Benjamini & Hochberg, 1995](#)) is applied at each scale under positive regression dependency on subsets conditions:

- For a given scale s , let m denote the number of time points under consideration outside the COI. Compute a set of p-values $\{p_{s,\tau_1}, p_{s,\tau_2}, \dots, p_{s,\tau_m}\}$ corresponding to the test for statistical significance of coherence at each temporal location τ_j .
- Sort the p-values in ascending order, denoted as $p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(m)}$, where $p_{(k)}$ is the k^{th} smallest p-value among the m hypotheses at scale s .
- For a chosen false discovery rate (FDR) level α , determine the largest integer k^* such that $p_{(k)} \leq \frac{k}{m}\alpha$.
- All points for which the p-value $p_{s,\tau} \leq p_{(k^*)}$ are then declared statistically significant at FDR level α .

Second, to conservatively perform a global sensitivity analysis that allows for possible dependence in the time-frequency domain, we compute pooled Benjamini-Yekutieli q-values ([Benjamini & Yekutieli, 2001](#)) over all (s, τ) points that control FDR at any pre-specified level under any possible dependence structure. Cluster-wise contiguity on the time-frequency plane can also be optionally used to respect the geometry of spatial dependence. Here k indexes through the hypotheses ranked at a particular scale. s, m is the number of time points tested at that scale outside of the COI, and α is the chosen FDR level (e.g., 0.10).

The proposed correction method rigorously limits the anticipated rate of false positives amongst all detected time-frequency points, thereby boosting the confidence in findings from high-resolution WCA investigations in macroeconomic frameworks. The findings denote correlation trends and lead-lag interconnections, as opposed to structural causation, with phase variations delivering directional clues concerning the temporal order between shocks in uncertainty and macroeconomic metrics.

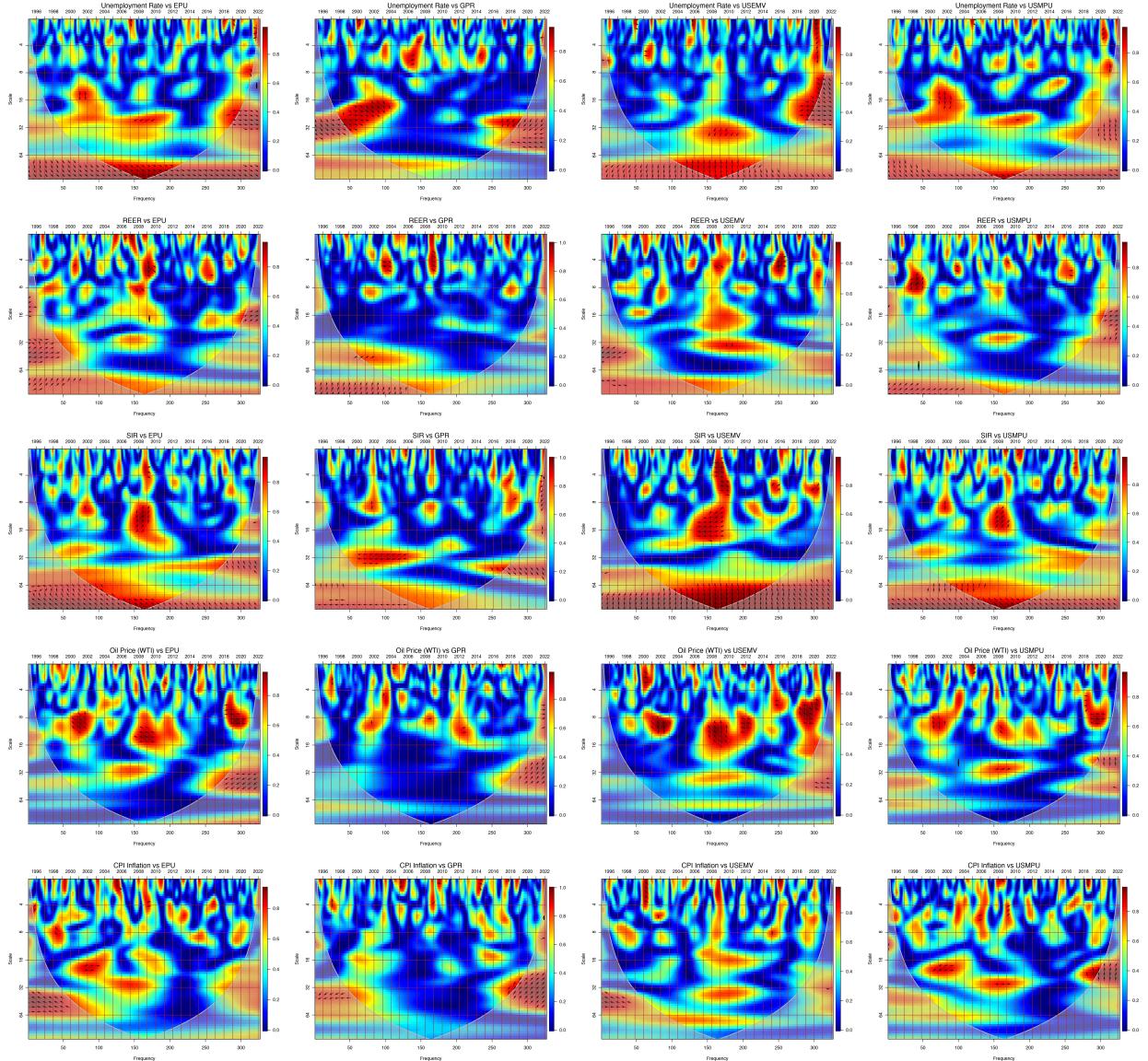


Figure A.1.4: Scale-wise FDR-corrected wavelet coherence spectra for the US (Jan, 1995 – Mar, 2022), displaying Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation against EPU, GPR, USEMV, and USMPU. Each subplot visualizes the time-frequency coherence between a macroeconomic variable (y-axis: Scale; x-axis: Frequency) and an uncertainty index, with warm colors (red/yellow) indicating intervals of high, statistically significant co-movement after FDR correction. The direction of phase arrows indicates lead-lag relationships: rightward arrows mean the series move in phase; leftward, in anti-phase; upward arrows signify the uncertainty index leads; downward, that it lags. Shaded regions inside the cone of influence reflect reliable coherence, while grid consistency enables robust cross-variable and cross-period comparison. The level of significance (α) for the FDR adjustment is set at 10%.

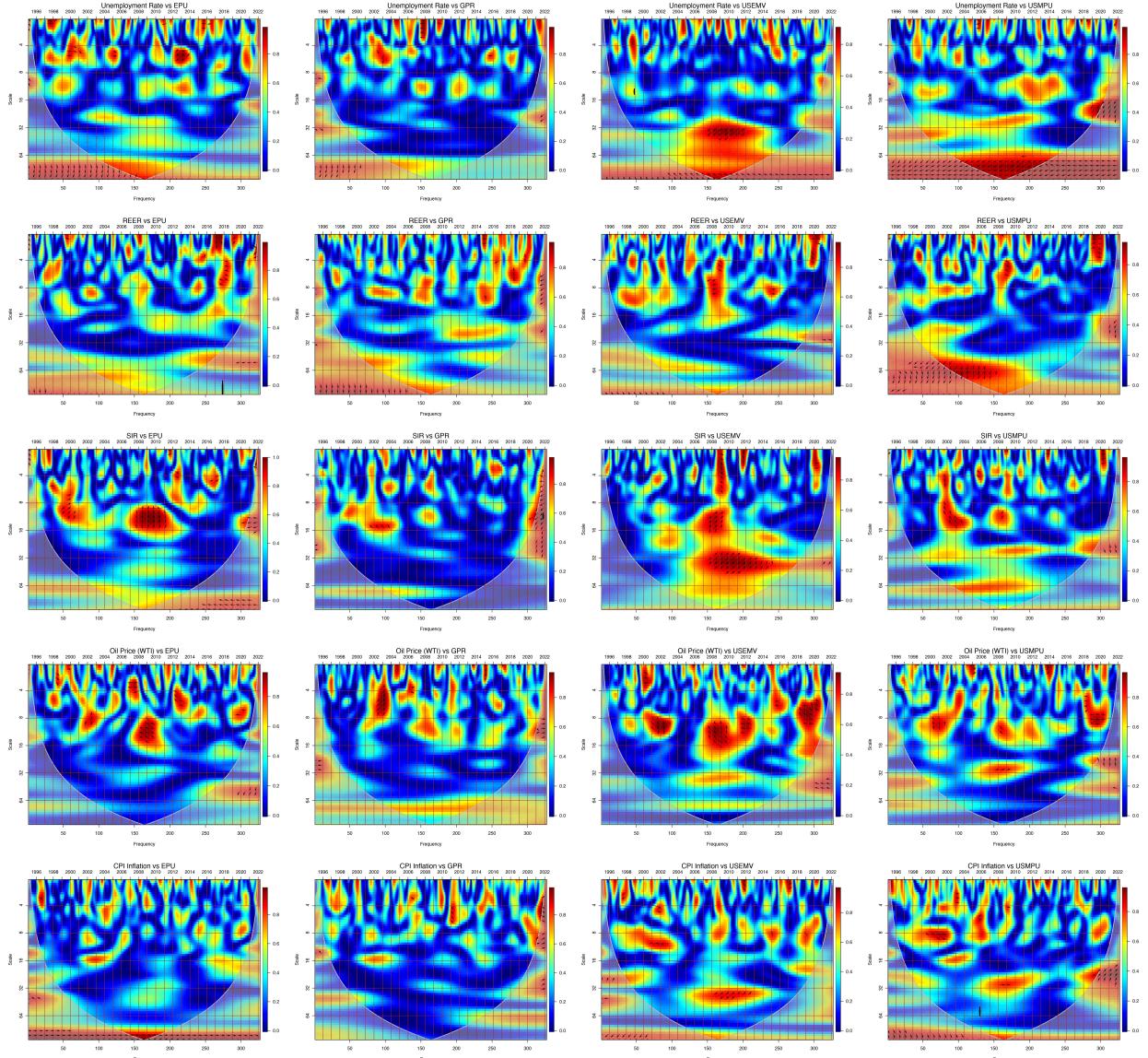


Figure A.1.5: Scale-wise FDR-corrected wavelet coherence spectra for France (Jan, 1995 – Mar, 2022), displaying Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation against EPU, GPR, USEMV, and USMPU. Each subplot visualizes the time-frequency coherence between a macroeconomic variable (y-axis: Scale; x-axis: Frequency) and an uncertainty index, with warm colors (red/yellow) indicating intervals of high, statistically significant co-movement after FDR correction. The direction of phase arrows indicates lead-lag relationships: rightward arrows mean the series move in phase; leftward, in anti-phase; upward arrows signify the uncertainty index leads; downward, that it lags. Shaded regions inside the cone of influence reflect reliable coherence, while grid consistency enables robust cross-variable and cross-period comparison. The level of significance (α) for the FDR adjustment is set at 10%.

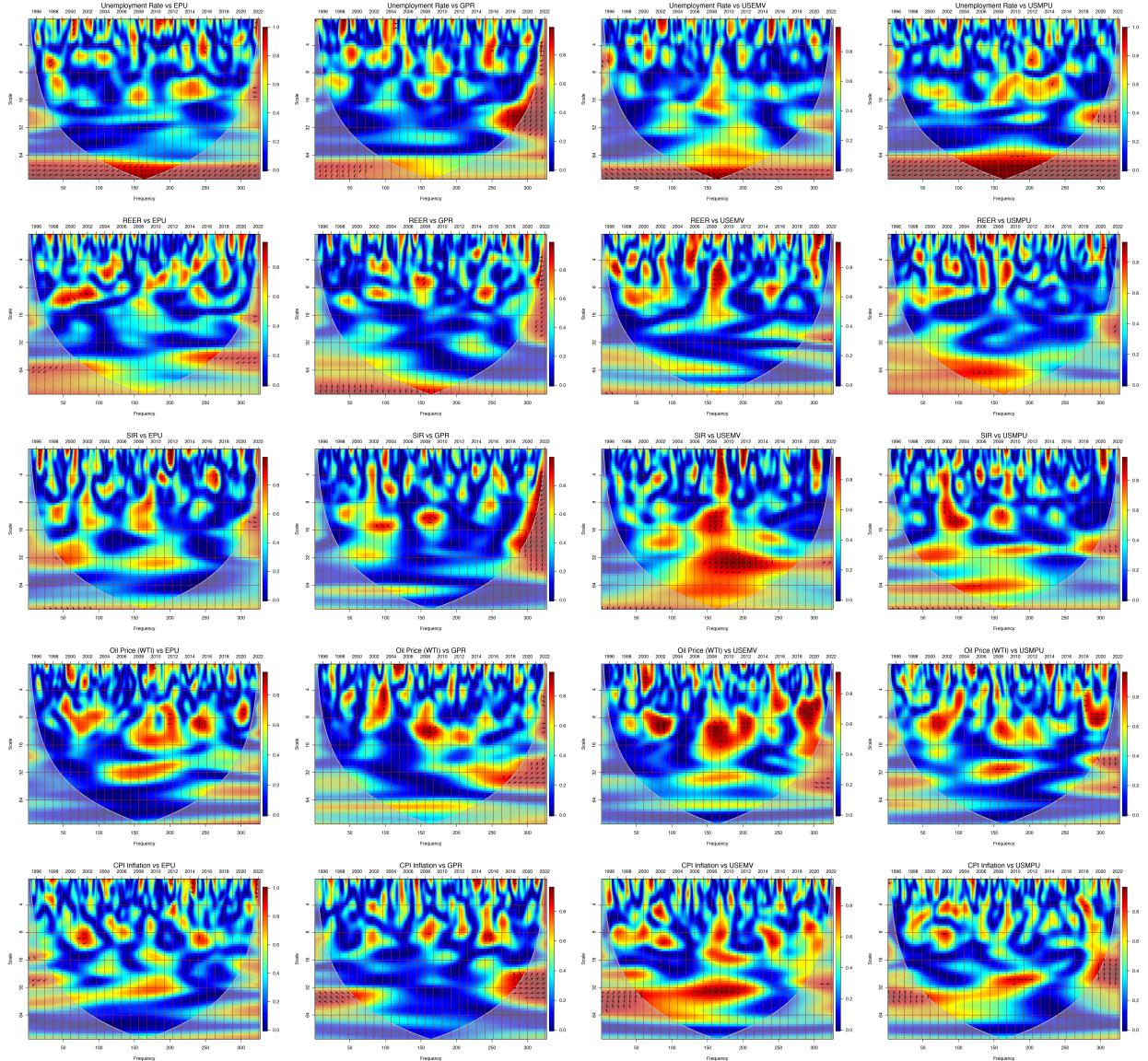


Figure A.1.6: Scale-wise FDR-corrected wavelet coherence spectra for Germany (Jan, 1995 – Mar, 2022), displaying Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation against EPU, GPR, USEMV, and USMPU. Each subplot visualizes the time-frequency coherence between a macroeconomic variable (y-axis: Scale; x-axis: Frequency) and an uncertainty index, with warm colors (red/yellow) indicating intervals of high, statistically significant co-movement after FDR correction. The direction of phase arrows indicates lead-lag relationships: rightward arrows mean the series move in phase; leftward, in anti-phase; upward arrows signify the uncertainty index leads; downward, that it lags. Shaded regions inside the cone of influence reflect reliable coherence, while grid consistency enables robust cross-variable and cross-period comparison. The level of significance (α) for the FDR adjustment is set at 10%.

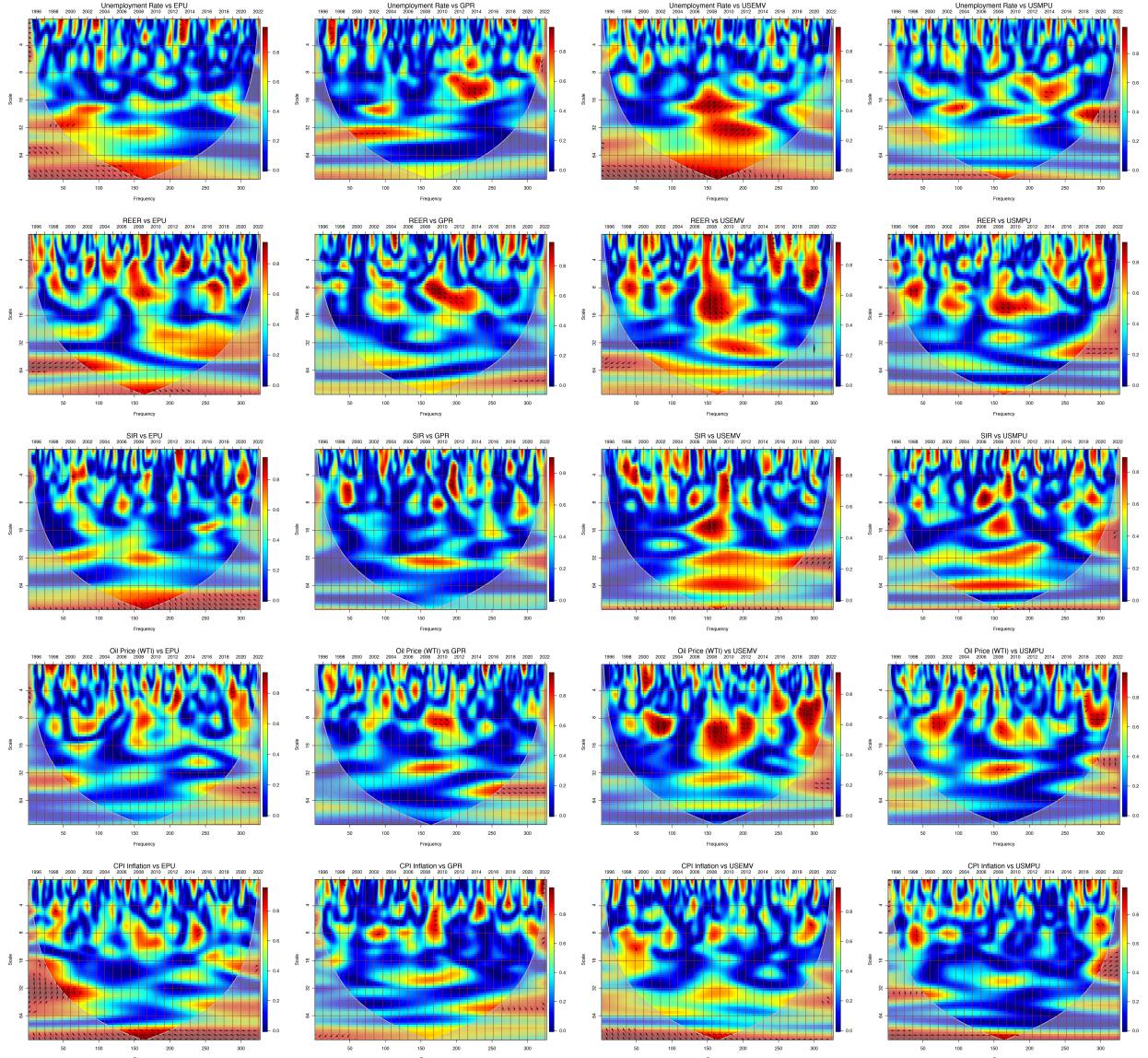


Figure A.1.7: Scale-wise FDR-corrected wavelet coherence spectra for Japan (Jan, 1995 – Mar, 2022), displaying Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation against EPU, GPR, USEMV, and USMPU. Each subplot visualizes the time-frequency coherence between a macroeconomic variable (y-axis: Scale; x-axis: Frequency) and an uncertainty index, with warm colors (red/yellow) indicating intervals of high, statistically significant co-movement after FDR correction. The direction of phase arrows indicates lead-lag relationships: rightward arrows mean the series move in phase; leftward, in anti-phase; upward arrows signify the uncertainty index leads; downward, that it lags. Shaded regions inside the cone of influence reflect reliable coherence, while grid consistency enables robust cross-variable and cross-period comparison. The level of significance (α) for the FDR adjustment is set at 10%.

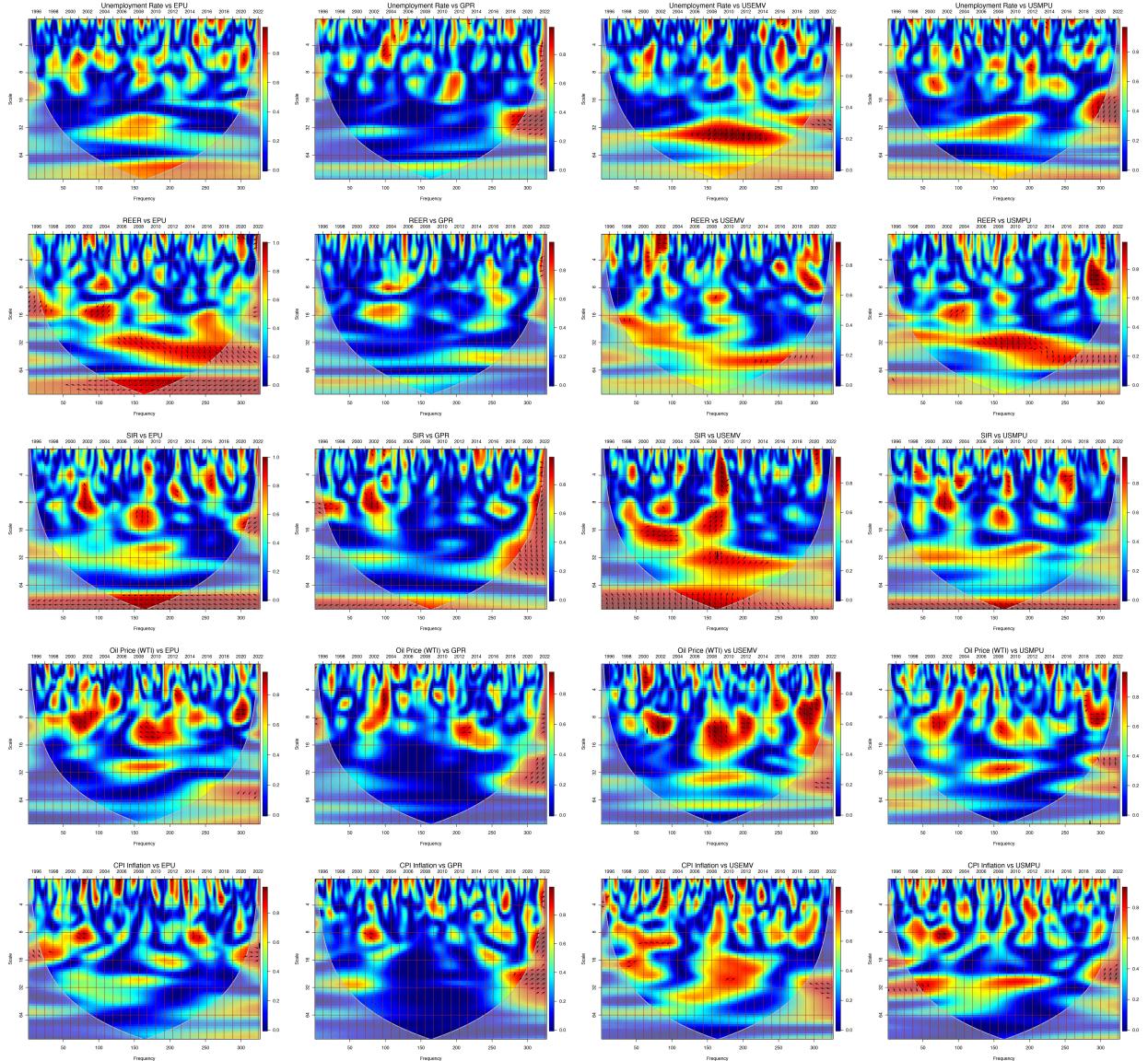


Figure A.1.8: Scale-wise FDR-corrected wavelet coherence spectra for the UK (Jan, 1995 – Mar, 2022), displaying Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation against EPU, GPR, USEMV, and USMPU. Each subplot visualizes the time-frequency coherence between a macroeconomic variable (y-axis: Scale; x-axis: Frequency) and an uncertainty index, with warm colors (red/yellow) indicating intervals of high, statistically significant co-movement after FDR correction. The direction of phase arrows indicates lead-lag relationships: rightward arrows mean the series move in phase; leftward, in anti-phase; upward arrows signify the uncertainty index leads; downward, that it lags. Shaded regions inside the cone of influence reflect reliable coherence, while grid consistency enables robust cross-variable and cross-period comparison. The level of significance (α) for the FDR adjustment is set at 10%.

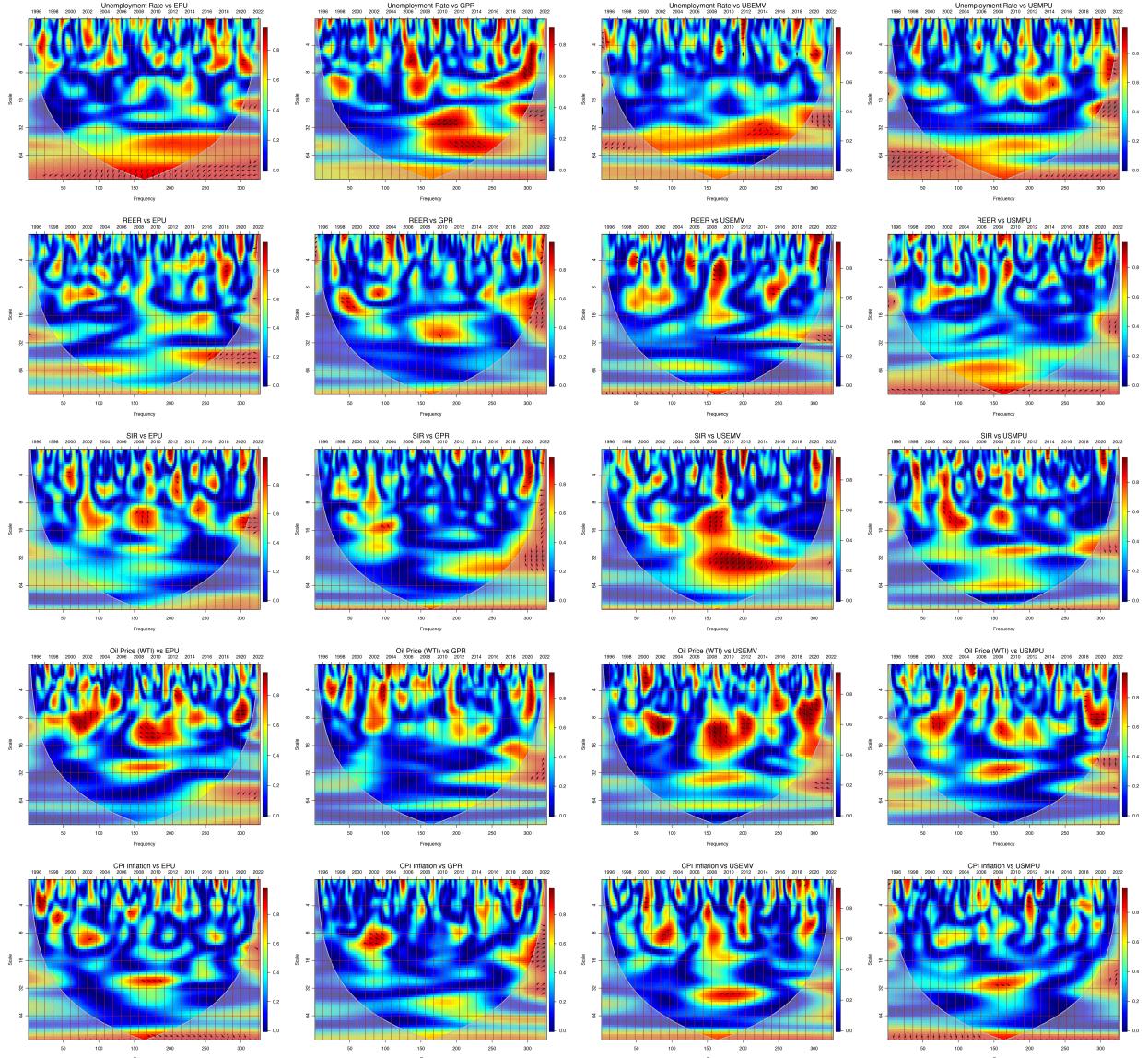


Figure A.1.9: Scale-wise FDR-corrected wavelet coherence spectra for Italy (Jan, 1995 – Mar, 2022), displaying Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation against EPU, GPR, USEMV, and USMPU. Each subplot visualizes the time-frequency coherence between a macroeconomic variable (y-axis: Scale; x-axis: Frequency) and an uncertainty index, with warm colors (red/yellow) indicating intervals of high, statistically significant co-movement after FDR correction. The direction of phase arrows indicates lead-lag relationships: rightward arrows mean the series move in phase; leftward, in anti-phase; upward arrows signify the uncertainty index leads; downward, that it lags. Shaded regions inside the cone of influence reflect reliable coherence, while grid consistency enables robust cross-variable and cross-period comparison. The level of significance (α) for the FDR adjustment is set at 10%.

A.1.5. Methodological framework for non-linear local projections

Impulse responses are estimated horizon-wise using nonlinear local projections as implemented in the `lpirfs` package (Adämmer, 2019), extending Jordà (2005) to smoothly switch regimes via a

logistic function. For each G7 country c and forecast horizon h , the system

$$\begin{aligned} \mathbf{y}_{c,t+h} = & \boldsymbol{\alpha}_c^h + \sum_{j=1}^p \mathbf{B}_{j,1,c}^h \left(\mathbf{y}_{c,t-j} [1 - F(z_{c,t-1})] \right) + \sum_{j=1}^p \mathbf{B}_{j,2,c}^h \left(\mathbf{y}_{c,t-j} F(z_{c,t-1}) \right) \\ & + \sum_{m=1}^M \boldsymbol{\Gamma}_{m,1,c}^h (u_{m,c,t} [1 - F(z_{c,t-1})]) + \sum_{m=1}^M \boldsymbol{\Gamma}_{m,2,c}^h (u_{m,c,t} F(z_{c,t-1})) \\ & + \delta_c^h \text{trend}_t + \boldsymbol{\varepsilon}_{c,t+h}^h, \end{aligned} \quad (\text{A.1})$$

is estimated by OLS for each $h = 0, \dots, H$, where $\mathbf{y}_{c,t}$ is a $K \times 1$ vector of macro variables (unemployment rate, REER, SIR, Oil Price (WTI), and, CPI inflation; $K = 5$), $u_{m,c,t}$ denotes uncertainty shocks (EPU, GPR, USEMV, USMPU; $M = 4$), and $F(z_{c,t-1}) = [1 + \exp(\gamma z_{c,t-1})]^{-1}$ is the transition function of the standardized short-term rate $z_{c,t-1}$. Lag order p is selected via Bayesian Information Criterion (BIC) and held fixed across h . The regime index is driven by the standardized short-term interest rate (SIR)

$$z_{c,t-1} = \frac{\text{SIR}_{c,t-1} - \mu_c}{\sigma_c},$$

where μ_c and σ_c are the within-country sample mean and standard deviation computed once and held fixed across horizons. The (decreasing) logistic transition function is

$$F(z_{c,t-1}) = \frac{1}{1 + \exp(\gamma z_{c,t-1})}, \quad \gamma > 0,$$

so that higher values of $z_{c,t-1}$ (tight/high-rate periods) imply $F(z_{c,t-1}) \approx 0$ and the $(1-F)$ -weighted regime (regime 1, “high IR”) dominates, whereas lower values of $z_{c,t-1}$ (loose/low-rate periods) imply $F(z_{c,t-1}) \approx 1$ and the F -weighted regime (regime 2, “low IR”) dominates. We lag the transition function to avoid contemporaneous feedback. We consider two complementary implementations consistent with [Adämmer \(2019\)](#): (i) Identified-shock nonlinear LPs (`1p_nl_iv`). In practice, we estimate a separate model for each uncertainty proxy m , treating $u_{m,c,t}$ as the identified shock series. In this case, the regime- r impulse response of variable i at horizon h to a unit innovation in u_m is the i th element of $\boldsymbol{\Gamma}_{m,r,c}^h$, extracted using the selection vector \mathbf{e}'_i , where \mathbf{e}_i is a $K \times 1$ unit vector with unity in the i th position and zeros elsewhere, enabling isolation of the specific variable’s response from the multivariate coefficient vector:

$$\text{IRF}_{i,m,r}^h = \mathbf{e}'_i \boldsymbol{\Gamma}_{m,r,c}^h, \quad r \in \{1, 2\}.$$

(ii) VAR-based nonlinear IRFs (`1p_nl`). When no external shock series is supplied, `1pirfs` estimates a reduced VAR to obtain the residual covariance $\boldsymbol{\Sigma}$, applies a Cholesky factorization to construct orthogonal shocks, and uses the local projection coefficients in (A.1) to compute state-dependent IRFs by horizon, as detailed in [Adämmer \(2019\)](#). HAC Newey–West standard errors yield confidence intervals:

$$\text{CI}_{95\%} = [\hat{\theta} - 1.96 \hat{\sigma}_{NW}(\hat{\theta}), \hat{\theta} + 1.96 \hat{\sigma}_{NW}(\hat{\theta})].$$

Model reliability requires sufficient variation in $F(z_{c,t-1})$ so both regimes are well-represented. This framework enables flexible, state-dependent IRF estimation while accounting for serial correlation and smooth monetary regime shifts, following ([Adämmer, 2019](#)).

A.1.6. IRF Plots for G7

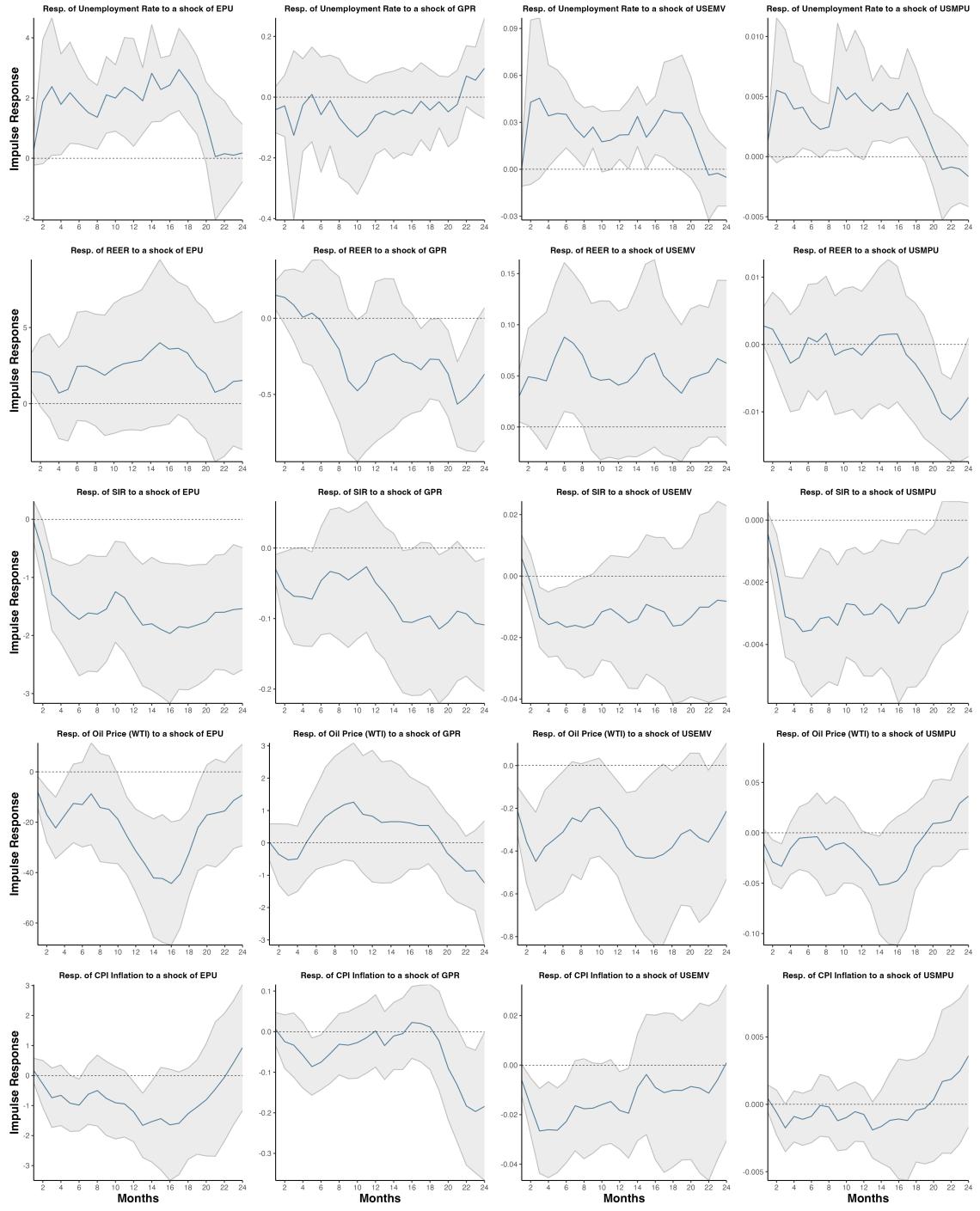


Figure A.1.10: Impulse response functions (nonlinear local projections) of Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation to EPU, GPR, USEMV, and USMPU shocks for the US (high-rate regime). Each row shows responses of a macro variable; each column, a distinct uncertainty shock. Solid **blue** lines depict IRFs with 95% confidence bands (grey). The x-axis is horizon (months), the y-axis the response magnitude. Compare responses by reading across rows (by variable) or down columns (by shock); zero response is marked by the dotted line.

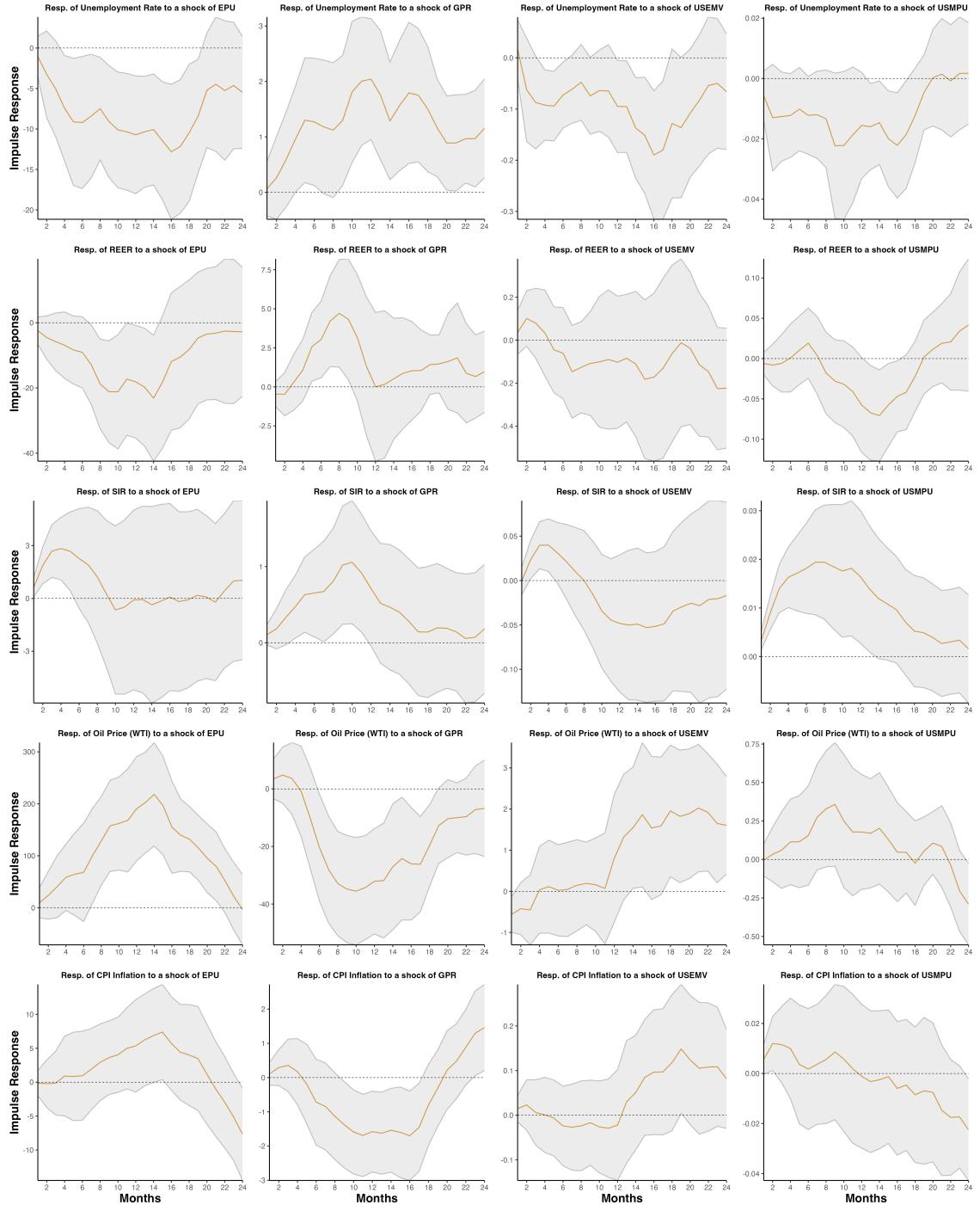


Figure A.1.11: Impulse response functions (nonlinear local projections) of Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation to EPU, GPR, USEMV, and USMPU shocks for the US (low-rate regime). Each row shows responses of a macro variable; each column, a distinct uncertainty shock. Solid orange lines depict IRFs with 95% confidence bands (grey). The x-axis is horizon (months), the y-axis the response magnitude. Compare responses by reading across rows (by variable) or down columns (by shock); zero response is marked by the dotted line.

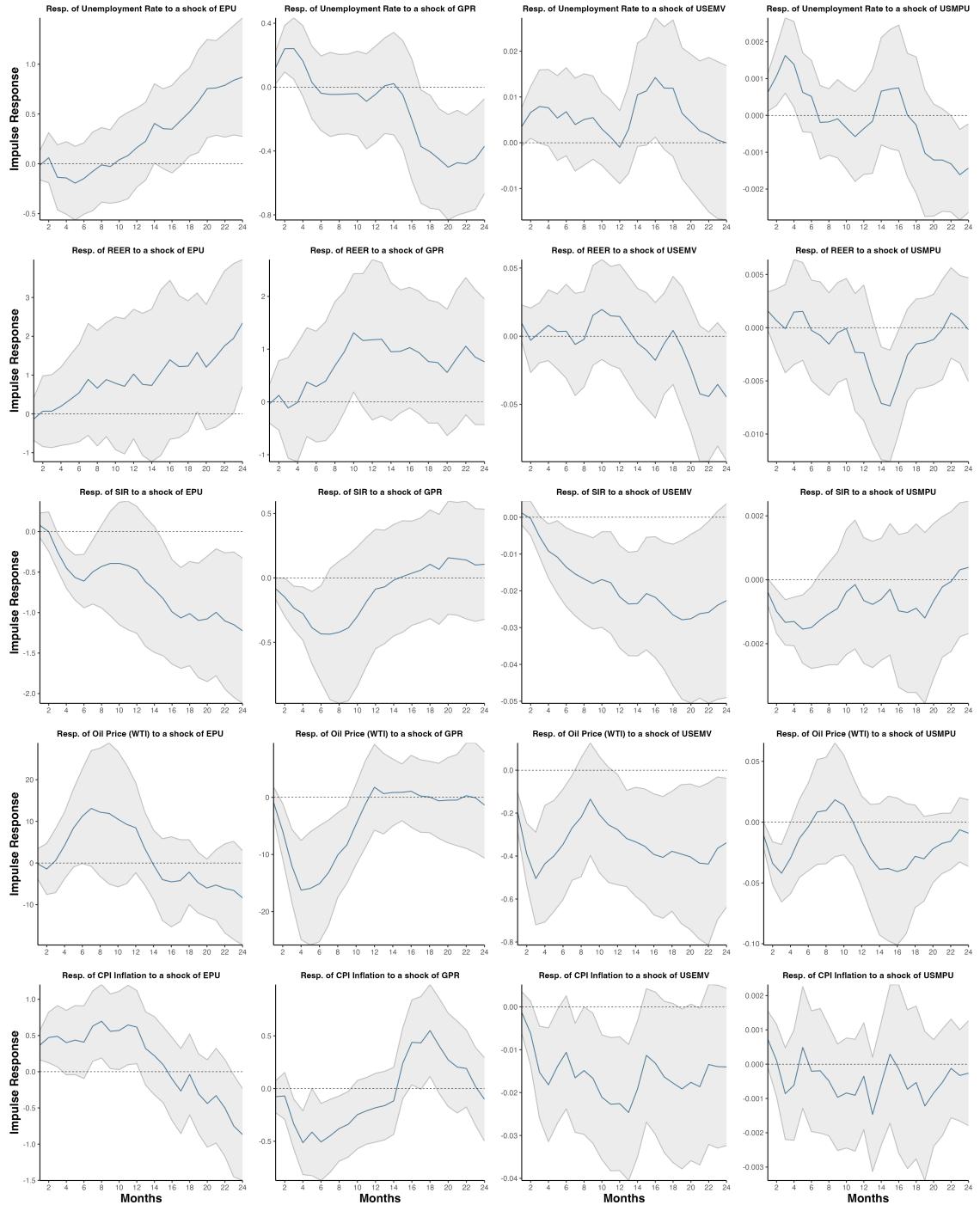


Figure A.1.12: Impulse response functions (nonlinear local projections) of Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation to EPU, GPR, USEMV, and USMPU shocks for France (high-rate regime). Each row shows responses of a macro variable; each column, a distinct uncertainty shock. Solid blue lines depict IRFs with 95% confidence bands (grey). The x-axis is horizon (months), the y-axis the response magnitude. Compare responses by reading across rows (by variable) or down columns (by shock); zero response is marked by the dotted line.

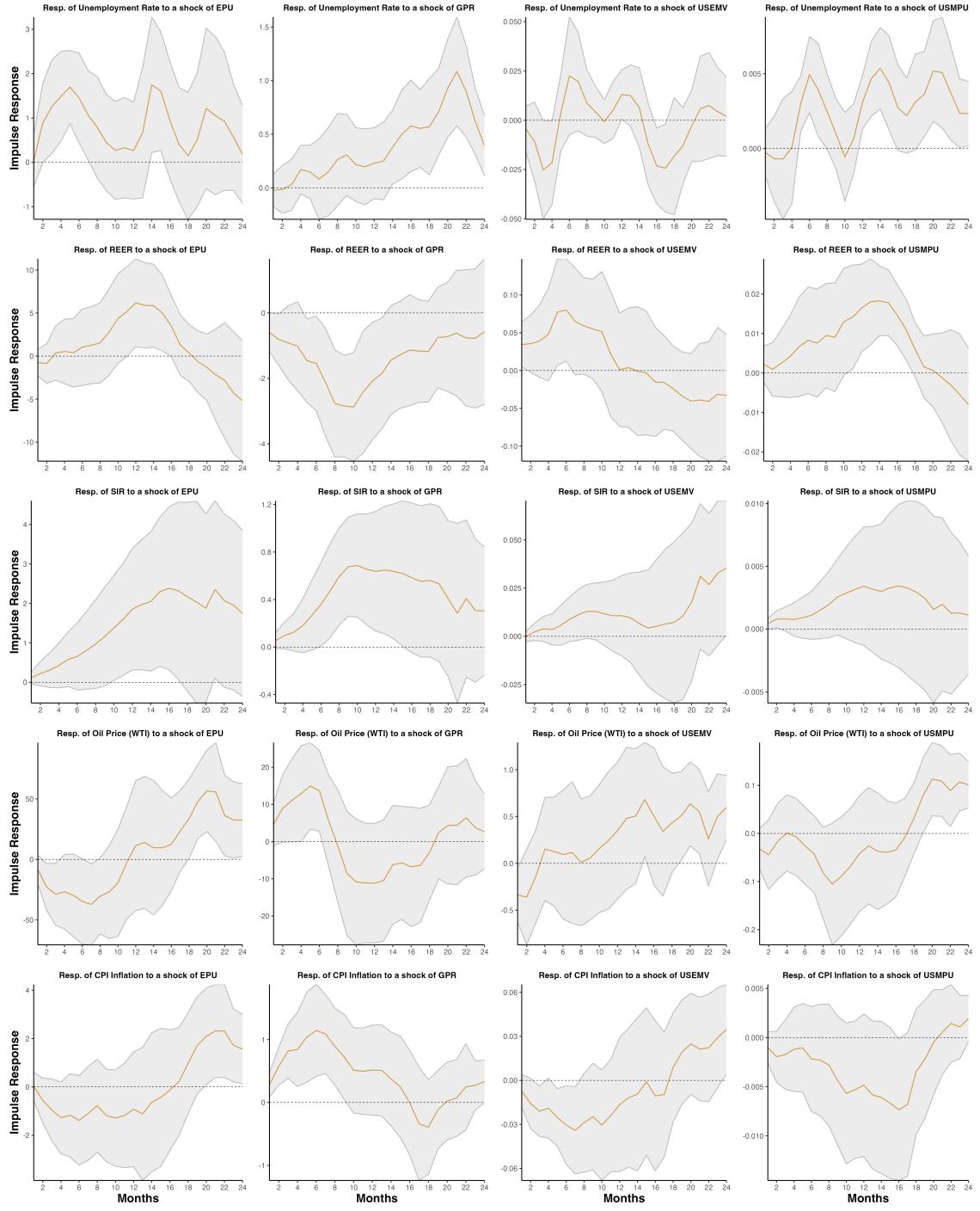


Figure A.1.13: Impulse response functions (nonlinear local projections) of Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation to EPU, GPR, USEMV, and USMPU shocks for France (low-rate regime). Each row shows responses of a macro variable; each column, a distinct uncertainty shock. Solid orange lines depict IRFs with 95% confidence bands (grey). The x-axis is horizon (months), the y-axis the response magnitude. Compare responses by reading across rows (by variable) or down columns (by shock); zero response is marked by the dotted line.

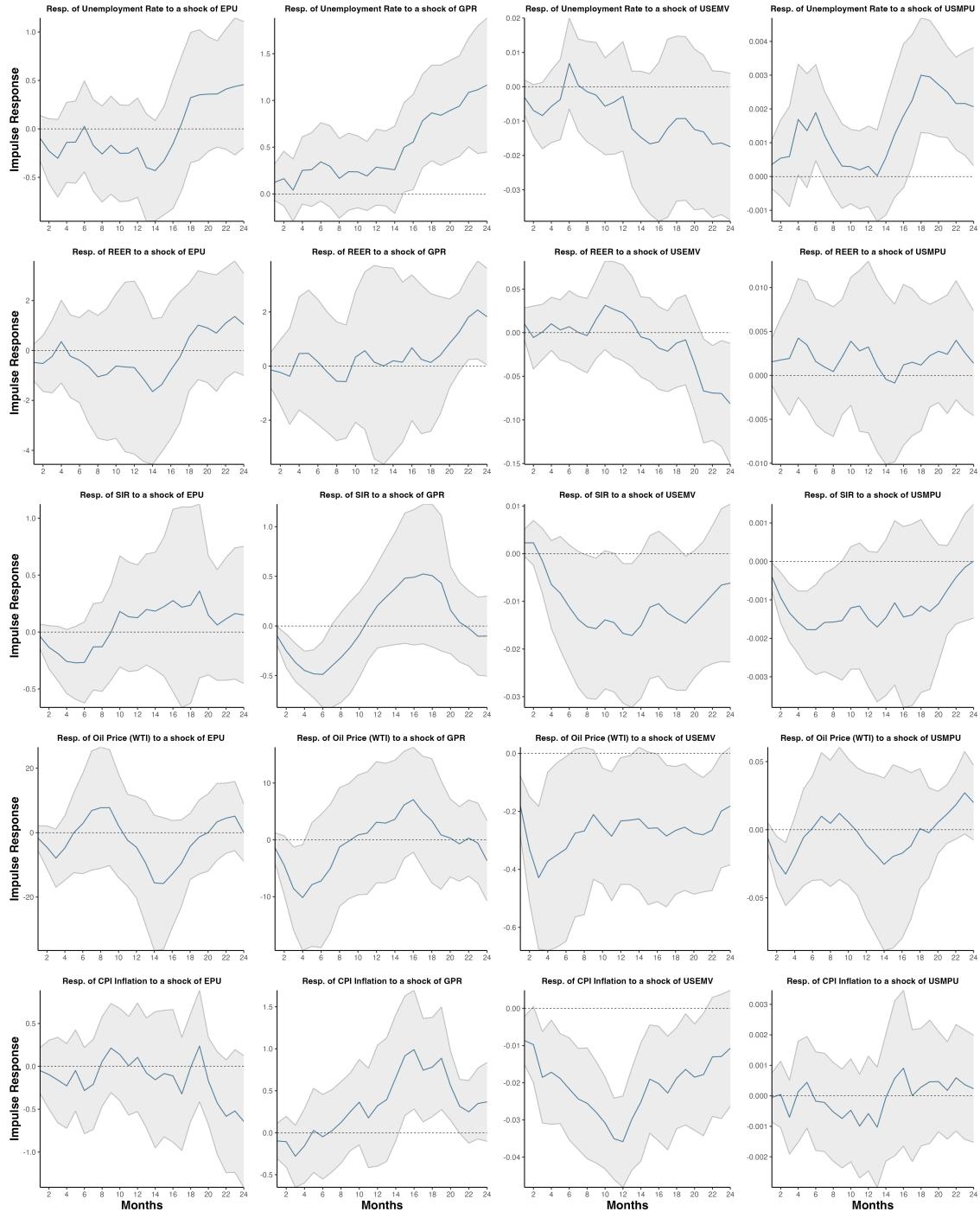


Figure A.1.14: Impulse response functions (nonlinear local projections) of Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation to EPU, GPR, USEMV, and USMPU shocks for Germany (high-rate regime). Each row shows responses of a macro variable; each column, a distinct uncertainty shock. Solid **blue** lines depict IRFs with 95% confidence bands (grey). The x-axis is horizon (months), the y-axis the response magnitude. Compare responses by reading across rows (by variable) or down columns (by shock); zero response is marked by the dotted line.

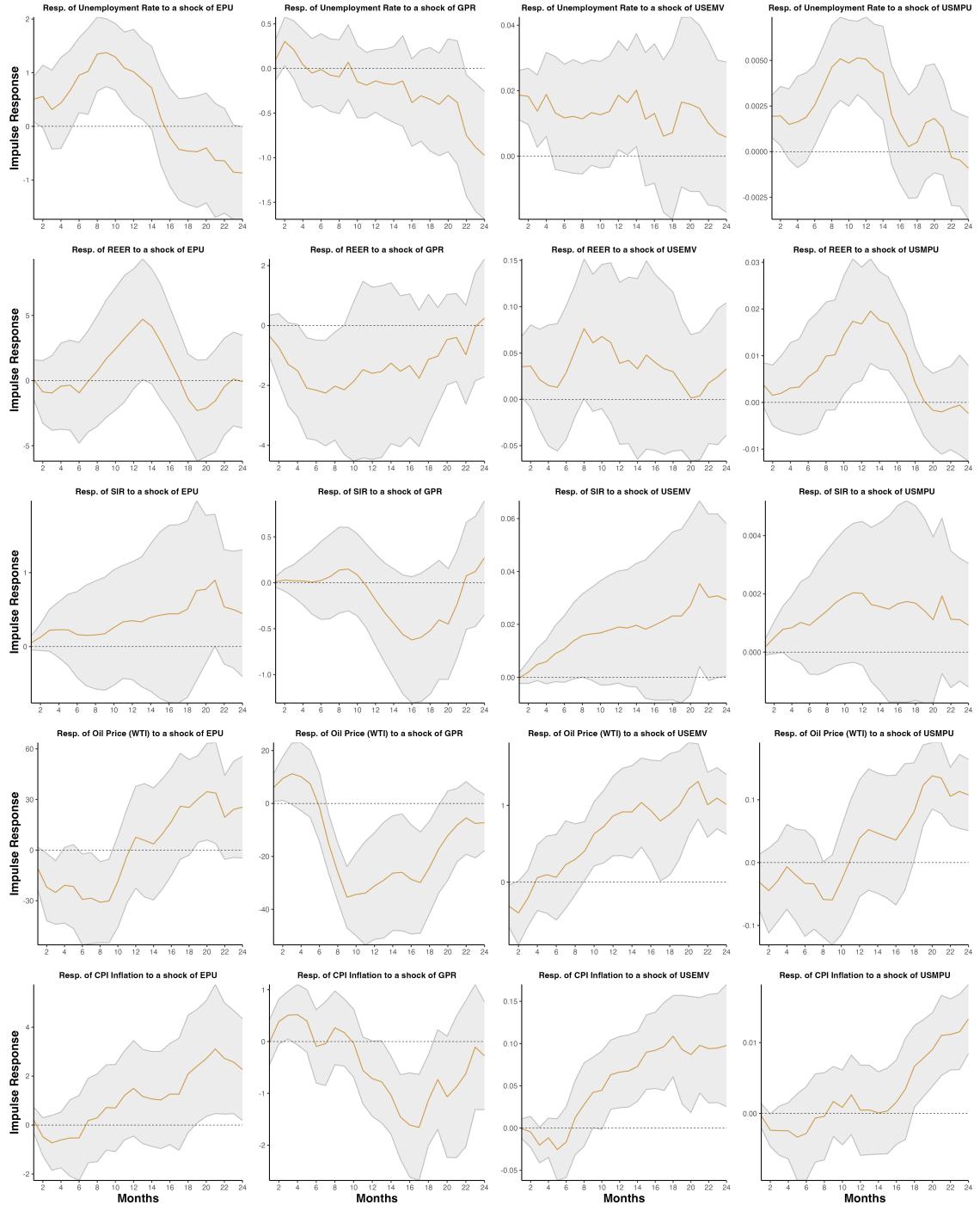


Figure A.1.15: Impulse response functions (nonlinear local projections) of Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation to EPU, GPR, USEMV, and USMPU shocks for Germany (low-rate regime). Each row shows responses of a macro variable; each column, a distinct uncertainty shock. Solid orange lines depict IRFs with 95% confidence bands (grey). The x-axis is horizon (months), the y-axis the response magnitude. Compare responses by reading across rows (by variable) or down columns (by shock); zero response is marked by the dotted line.

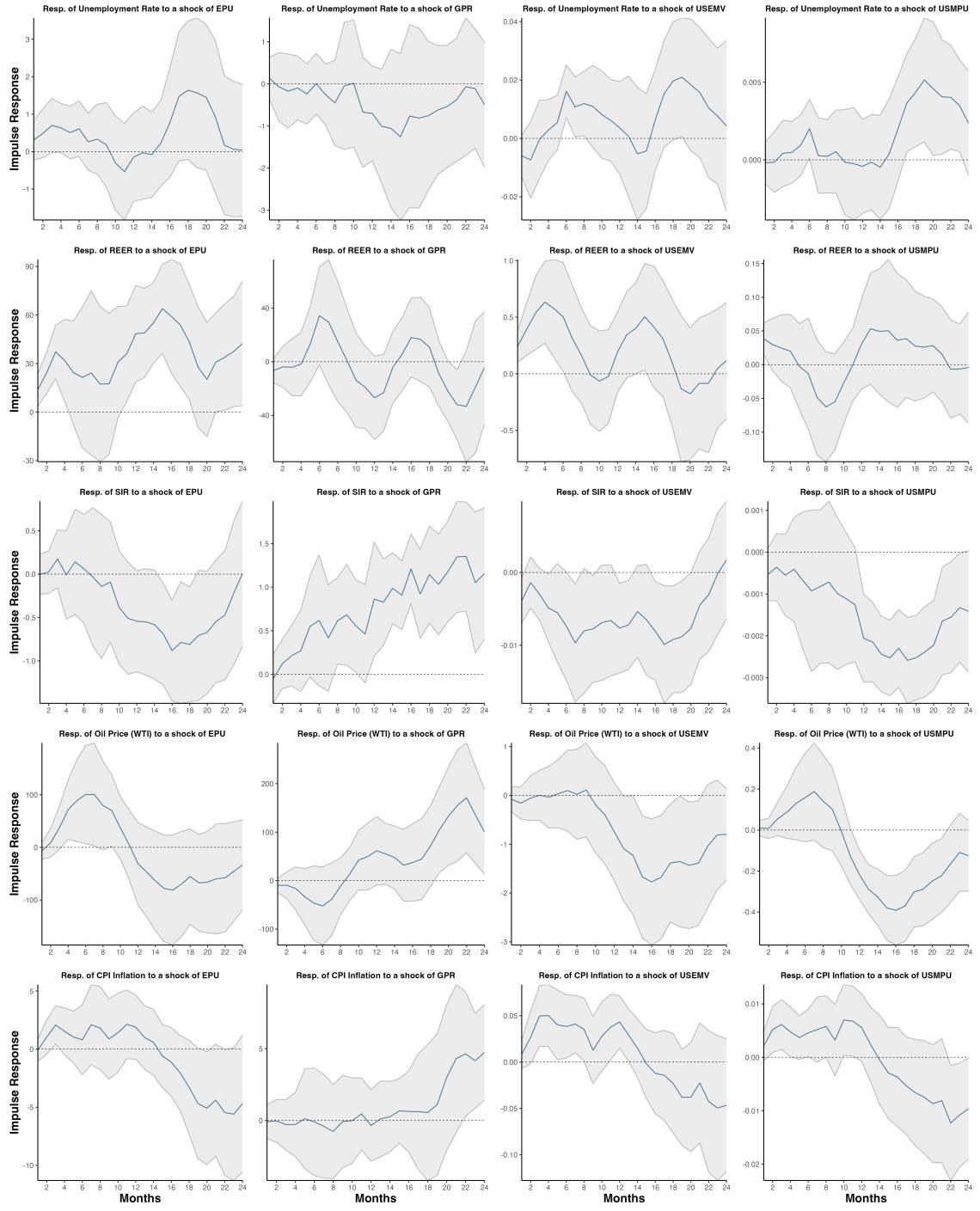


Figure A.1.16: Impulse response functions (nonlinear local projections) of Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation to EPU, GPR, USEMV, and USMPU shocks for Japan (high-rate regime). Each row shows responses of a macro variable; each column, a distinct uncertainty shock. Solid **blue** lines depict IRFs with 95% confidence bands (grey). The x-axis is horizon (months), the y-axis the response magnitude. Compare responses by reading across rows (by variable) or down columns (by shock); zero response is marked by the dotted line.

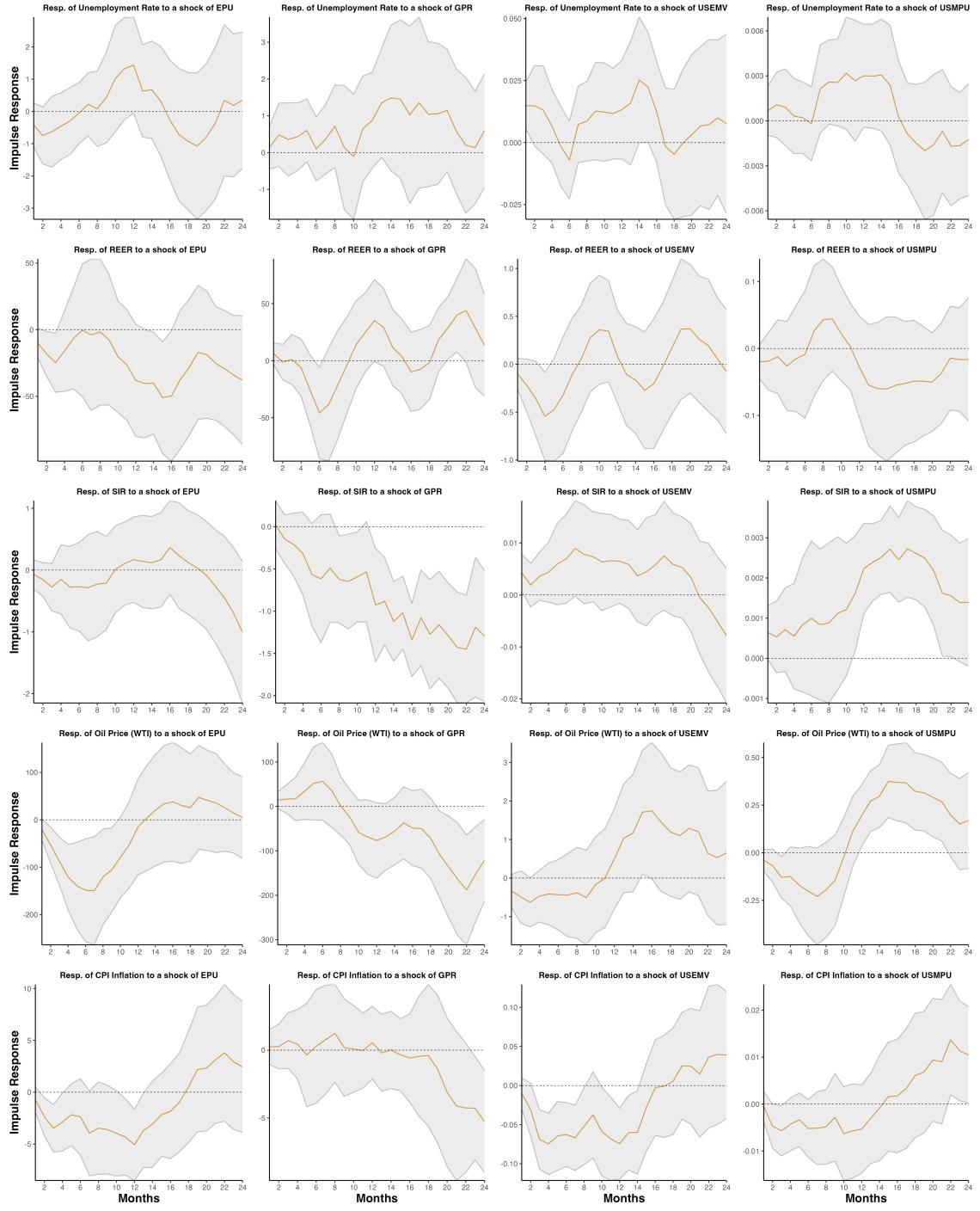


Figure A.1.17: Impulse response functions (nonlinear local projections) of Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation to EPU, GPR, USEMV, and USMPU shocks for Japan (low-rate regime). Each row shows responses of a macro variable; each column, a distinct uncertainty shock. Solid orange lines depict IRFs with 95% confidence bands (grey). The x-axis is horizon (months), the y-axis the response magnitude. Compare responses by reading across rows (by variable) or down columns (by shock); zero response is marked by the dotted line.



Figure A.1.18: Impulse response functions (nonlinear local projections) of Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation to EPU, GPR, USEMV, and USMPU shocks for the UK (high-rate regime). Each row shows responses of a macro variable; each column, a distinct uncertainty shock. Solid blue lines depict IRFs with 95% confidence bands (grey). The x-axis is horizon (months), the y-axis the response magnitude. Compare responses by reading across rows (by variable) or down columns (by shock); zero response is marked by the dotted line.

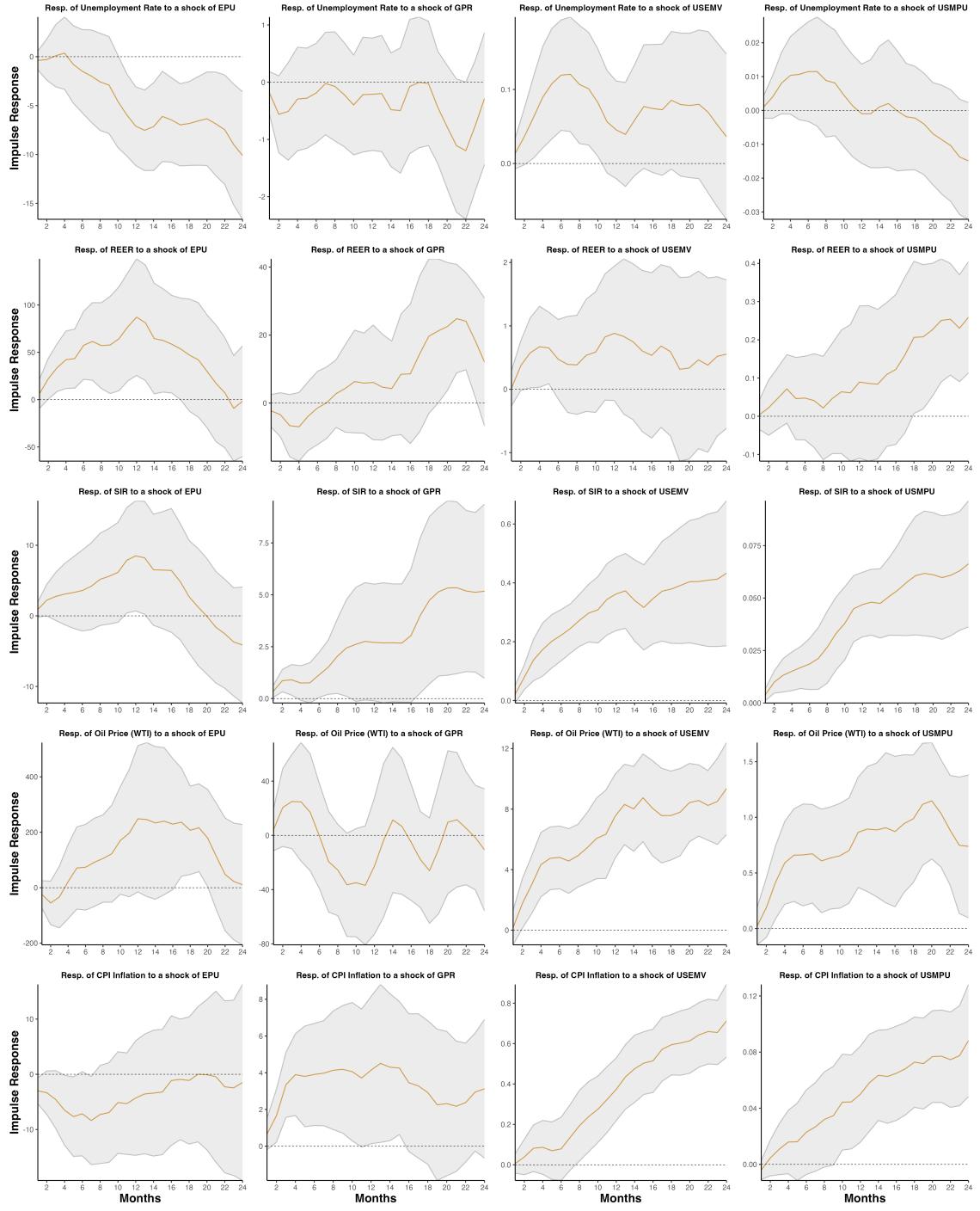


Figure A.1.19: Impulse response functions (nonlinear local projections) of Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation to EPU, GPR, USEMV, and USMPU shocks for the UK (low-rate regime). Each row shows responses of a macro variable; each column, a distinct uncertainty shock. Solid orange lines depict IRFs with 95% confidence bands (grey). The x-axis is horizon (months), the y-axis the response magnitude. Compare responses by reading across rows (by variable) or down columns (by shock); zero response is marked by the dotted line.

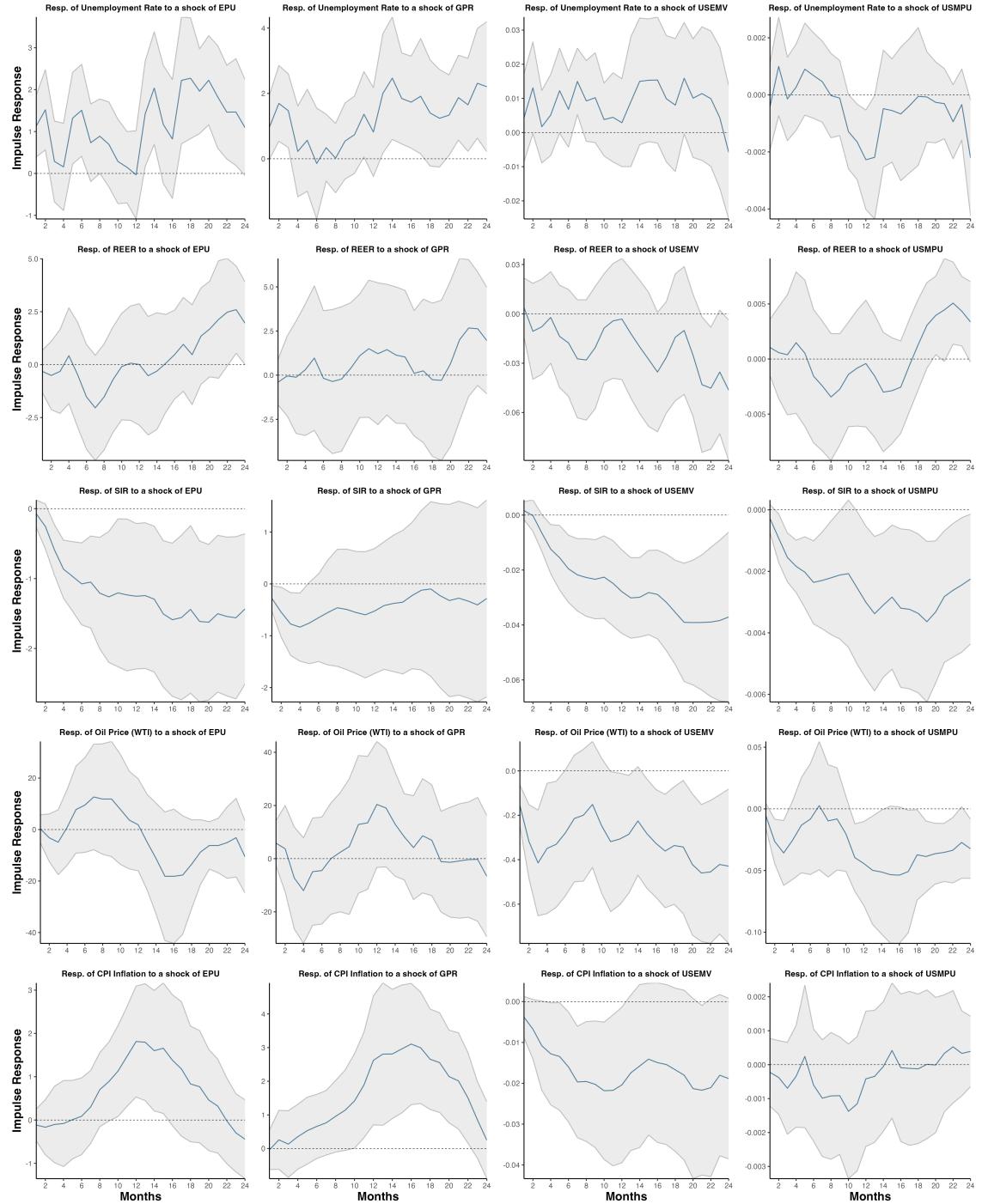


Figure A.1.20: Impulse response functions (nonlinear local projections) of Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation to EPU, GPR, USEMV, and USMPU shocks for Italy (high-rate regime). Each row shows responses of a macro variable; each column, a distinct uncertainty shock. Solid blue lines depict IRFs with 95% confidence bands (grey). The x-axis is horizon (months), the y-axis the response magnitude. Compare responses by reading across rows (by variable) or down columns (by shock); zero response is marked by the dotted line.

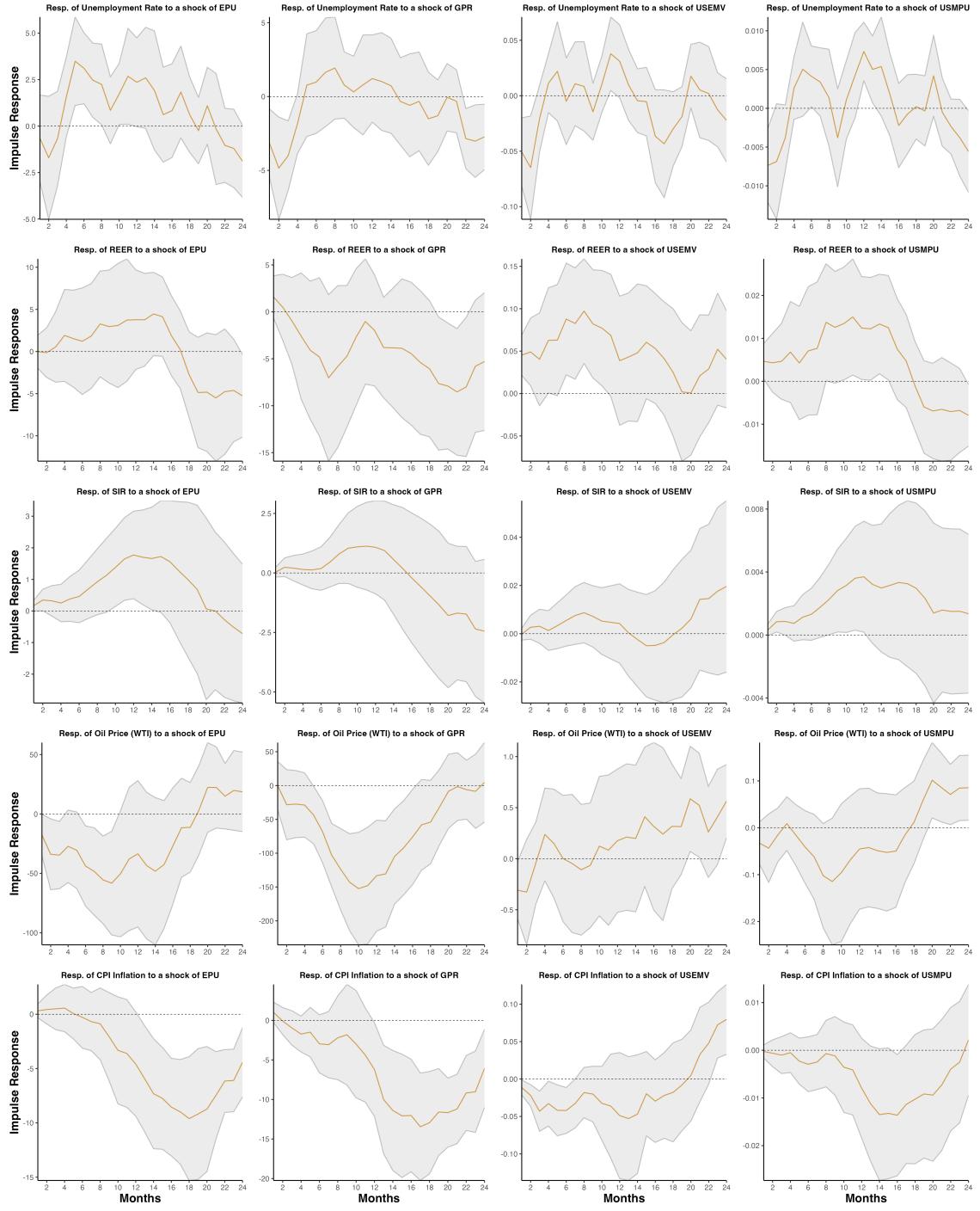


Figure A.1.21: Impulse response functions (nonlinear local projections) of Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation to EPU, GPR, USEMV, and USMPU shocks for Italy (low-rate regime). Each row shows responses of a macro variable; each column, a distinct uncertainty shock. Solid orange lines depict IRFs with 95% confidence bands (grey). The x-axis is horizon (months), the y-axis the response magnitude. Compare responses by reading across rows (by variable) or down columns (by shock); zero response is marked by the dotted line.

A.1.7. Baseline Models: An overview

Our baseline models include:

- *Vector Autoregressive model with exogenous covariates (VARx)*: The Vector Autoregression with exogenous variables model is a system of equations describing the joint dynamics of a set of time series with external predictors, and thereby facilitates rich cross-variable and shock transmission channels([Sims, 1980](#); [Kilian, 2006](#); [Hamilton, 2020](#); [Tsay, 2005](#)).
- *Threshold Vector Autoregressive model (TVAR)*: TVAR models permit the coefficients to switch between different regimes specified by the value of a threshold variable, and so can be used to capture nonlinearities and regime shifts in time series data ([Tsay, 1998](#); [Lo & Zivot, 2001](#); [Hansen, 2011](#)). A downside to TVAR models is that they do not allow for exogenous covariates. In practice, a p -lag TVAR model with two regimes can be specified as:

$$\mathbf{y}_t = \begin{cases} \mu^{(1)} + \sum_{i=1}^p \mathbf{A}_i^{(1)} \mathbf{y}_{t-i} + \varepsilon_t^{(1)}, & \text{if } q_{t-d} \leq r, \\ \mu^{(2)} + \sum_{i=1}^p \mathbf{A}_i^{(2)} \mathbf{y}_{t-i} + \varepsilon_t^{(2)}, & \text{if } q_{t-d} > r, \end{cases}$$

where q_{t-d} is the threshold variable (typically a lagged endogenous series), r is the estimated threshold, d is the delay, and the $\mathbf{A}_i^{(j)}$'s are regime-specific coefficients. The threshold is selected by searching for the optimal value in the data. Forecasts are then generated by fitting a TVAR over the training data, and then predicting forward to obtain the required horizons.

- *Vector Exponential Smoothing (VES)*: VES jointly smooths and forecasts multiple time series, capturing trend and seasonality ([Svetunkov, Chen & Boylan, 2023](#); [Hyndman, Koehler, Ord & Snyder, 2008](#); [De Silva, Hyndman & Snyder, 2010](#)). VES does not natively incorporate exogenous variables.
- *Neural Basis Expansion Analysis for Time Series with exogenous variables (NBEATSx)*: A deep neural model using stacked fully-connected blocks and backward/forward basis expansions, for flexible and highly accurate forecasting, that allows for exogenous regressors, excels at pure time series signal extraction ([Oreshkin, Carpor, Chapados & Bengio, 2019](#)).
- *Neural Hierarchical Interpolation for Time Series Forecasting with exogenous covariates (NHITSx)*: An extended version of NBEATSx that factorizes a time series with a set of hierarchical interpolation modules. Hierarchical interpolation can better model short- and long-term patterns simultaneously. NHITSx can also effectively support long input and output sequences by directly learning the hierarchical basis and interpolation weights. It improves model accuracy and computational efficiency ([Challu, Olivares, Oreshkin, Garza, Mergenthaler-Canseco & Dubrawski, 2023](#)).
- *Block Recurrent Neural Network with exogenous covariates (BRNNx)*: BlockRNN refers to recurrent neural network models (including LSTM and GRU blocks) structured for multivariate forecasting. These models are adept at learning sequential dependencies, providing strong performance for time series with significant temporal structure or autocorrelation, and can also handle exogenous covariates ([Herzen, Lässig, Piazzetta, Neuer, Tafti, Raille, Van Pottelbergh, Pasieka, Skrodzki, Huguenin et al., 2022](#)).

- *Time-Series Mixer with exogenous covariates (TSMIXERx)*: TSMIXERx is an all-MLP (multi-layer perceptron) neural architecture without recurrent or convolutional components. It models time series by mixing information both along the time axis (temporal mixing) and variable axis (feature mixing) through specialized mixer layers, offering memory and computational scalability. TSMIXERx is adept at capturing cross-variable dependencies and can process exogenous variables efficiently ([Chen, Li, Yoder, Arik & Pfister, 2023](#); [Ekambaram, Jati, Nguyen, Sinhong & Kalagnanam, 2023](#)).
- *CatBoostx, Light Gradient Boosting Machine (LGBMx), & Xtreme Gradient Boosting Machine (XGBx) with exogenous covairates*: These are tree ensemble methods that accept both categorical and numerical features, can model nonlinearities, and can ingest exogenous variables as additional features ([Prokhorenkova, Gusev, Vorobev, Dorogush & Gulin, 2018](#); [Ke, Meng, Finley, Wang, Chen, Ma, Ye & Liu, 2017](#); [Chen & Guestrin, 2016](#); [Herzen et al., 2022](#)).
- *Temporal Fusion Transformer with exogenous covariates (TFTx)*: TFTx incorporates attention and gating mechanisms to balance variable selection, interpretability, and model expressiveness for multi-horizon time series forecasting with exogenous data ([Lim, Arik, Loeff & Pfister, 2021](#)).
- *Time-Series Dense Encoder with exogenous covariates (TiDEX)*: A recently developed deep neural architecture with dense encoders and decoders, designed for long-term time series prediction, offering a good trade-off between model expressivity and efficiency([Das, Kong, Leach, Mathur, Sen & Yu, 2023](#)).
- *DishTS with CatBoostx (DTS_CBx) & DishTS with XGBx (DTS_XGBx) with exogenous covariates*: These two hybrid models use DishTS (Distribution Shift in Time Series Forecasting approach as proposed by [Fan, Wang, Wang, Wang, Zhou & Fu \(2023\)](#)) to normalize the time series and extract robust feature representations to tackle nonstationarity and distribution shifts. The time series transformed with DishTS, along with the covariates, are fed to CatBoostx or XGBoostx ensemble regressors. These models combine the bias-correction and normalization from DishTS with modern gradient boosting ensembles for forecasting strength and regularization, resulting in an improved robustness to structural shifts for macroeconomic data ([Fan et al., 2023](#)).

It is worth emphasizing that, as in their univariate versions, TVAR and VES models in their multivariate forms do not allow for exogenous variables to enter the model specification. This rules out the possibility to fully account for the presence of external shocks in a multivariate setting. Overall, this very comprehensive set of baselines allow us to robustly benchmark SZBVARx against both conventional econometric approaches and the frontiers of machine learning.

A.1.8. Statistical significance of the results

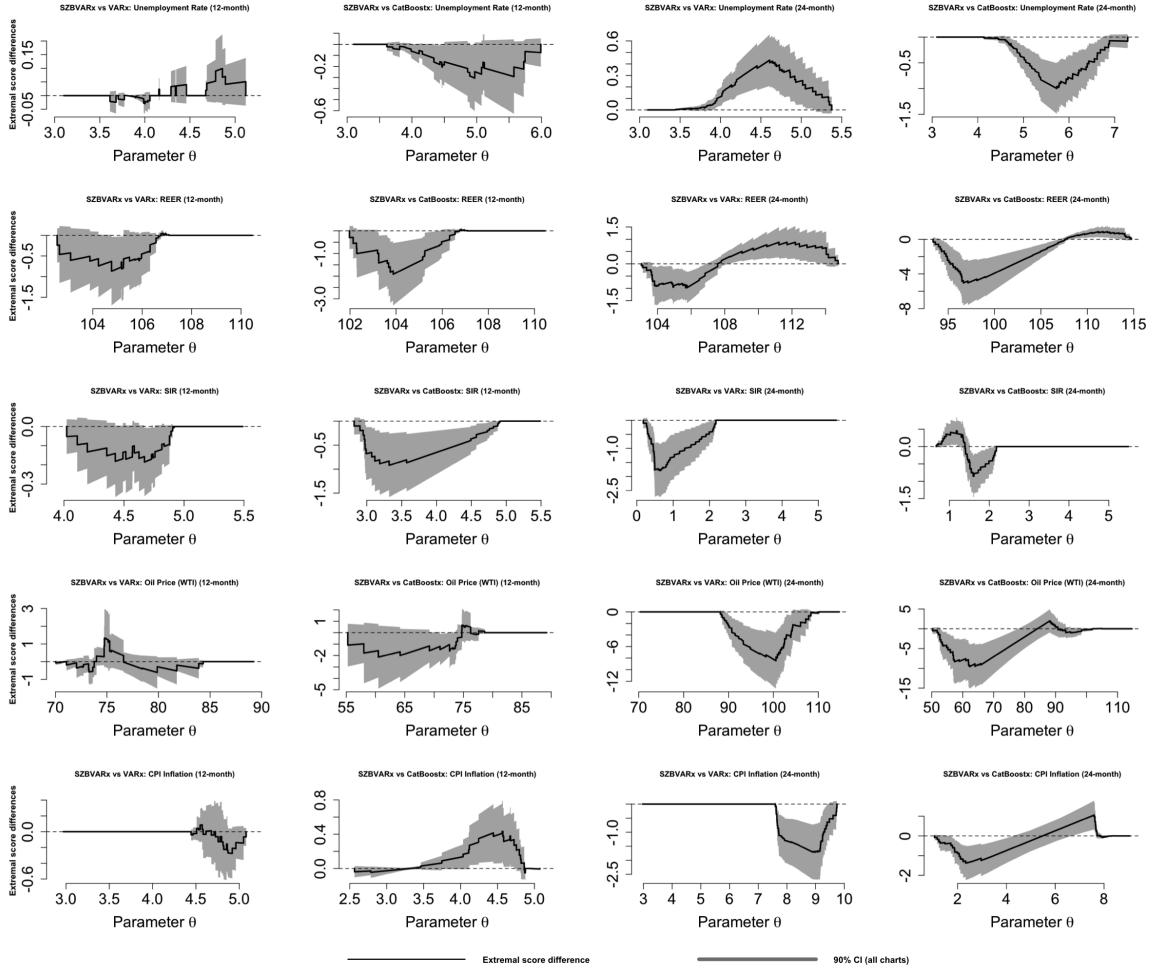


Figure A.1.22: Murphy diagram difference plots comparing SZBVARx forecasting performance against VARx and CatBoostx baselines across 12-month-and 24-month-ahead forecast horizons for the US. Panels in the first and second columns correspond to the 12-month horizon, while panels in the third and fourth columns display the results for the 24-month horizon. The analysis covers five key macroeconomic indicators: Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation. Each subplot displays extremal score differences with 90% HAC-based confidence bands (gray shaded areas) across threshold parameter values. Negative differences indicate superior SZBVARx performance, while the magnitude reflects the strength of performance advantage at each decision threshold.

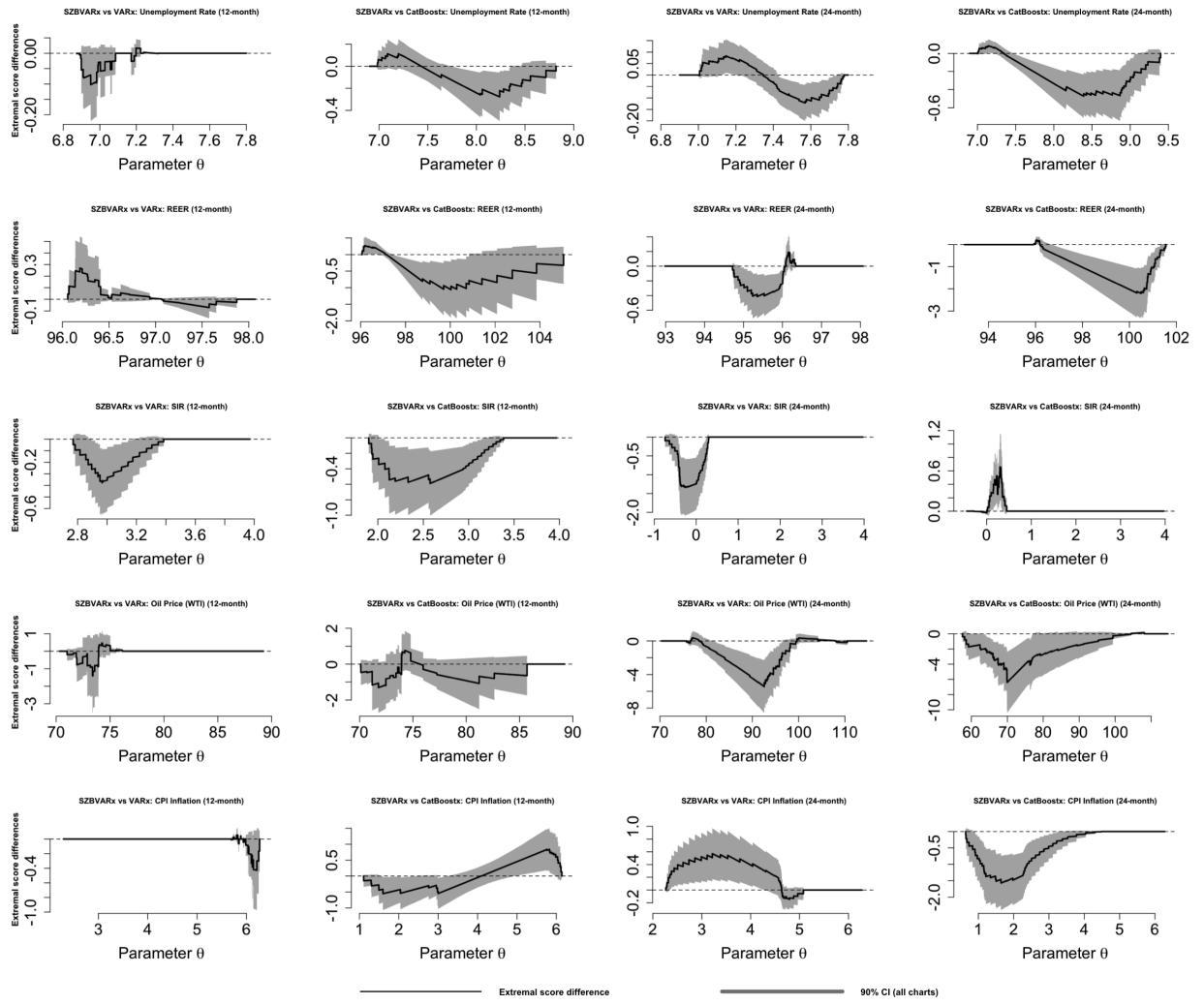


Figure A.1.23: Murphy diagram difference plots comparing SZBVARx forecasting performance against VARx and CatBoostx baselines across 12-month-and 24-month-ahead forecast horizons for France. Panels in the first and second columns correspond to the 12-month horizon, while panels in the third and fourth columns display the results for the 24-month horizon. The analysis covers five key macroeconomic indicators: Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation. Each subplot displays extremal score differences with 90% HAC-based confidence bands (gray shaded areas) across threshold parameter values. Negative differences indicate superior SZBVARx performance, while the magnitude reflects the strength of the performance advantage at each decision threshold.

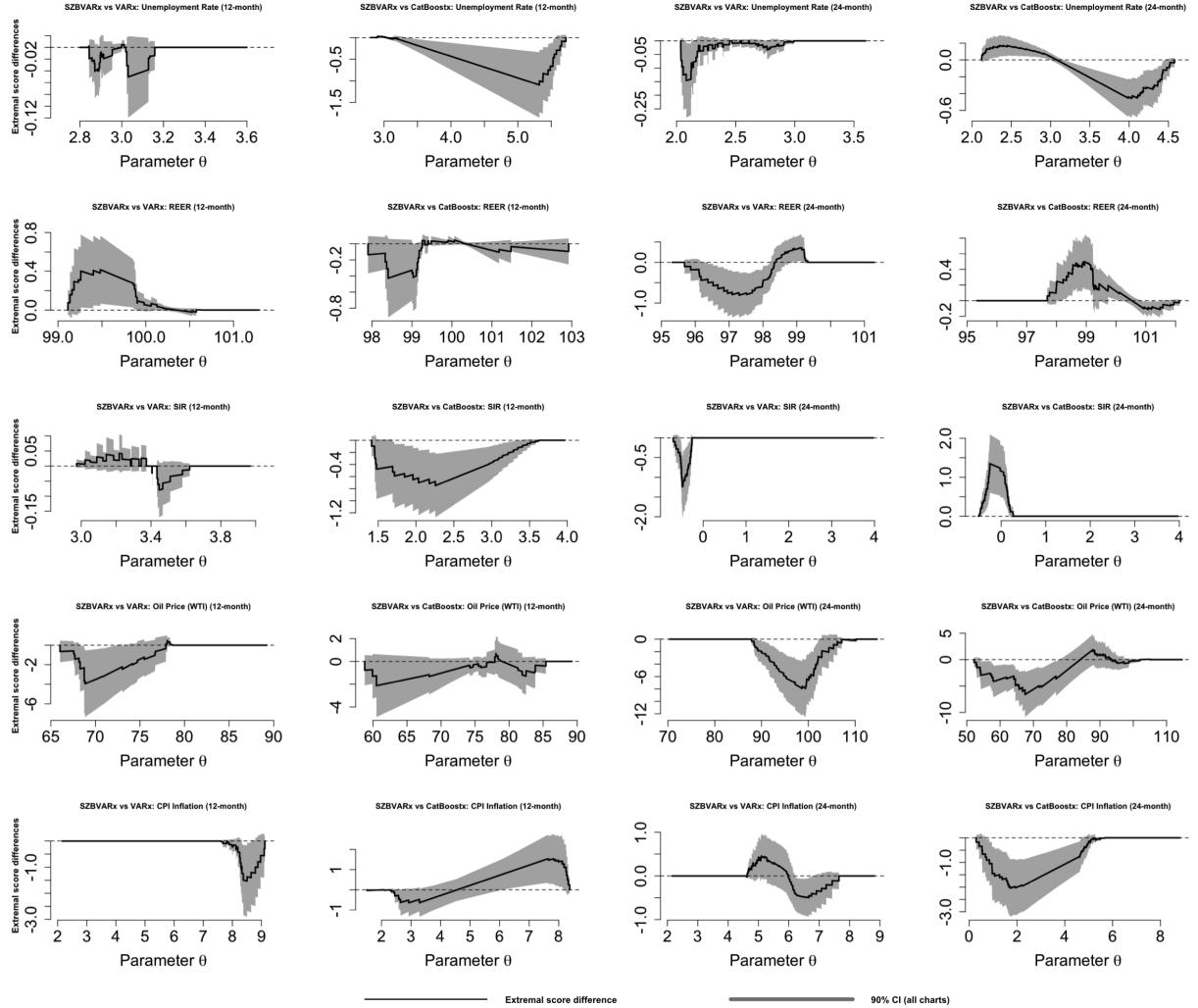


Figure A.1.24: Murphy diagram difference plots comparing SZBVARx forecasting performance against VARx and CatBoostx baselines across 12-month-and 24-month-ahead forecast horizons for Germany. Panels in the first and second columns correspond to the 12-month horizon, while panels in the third and fourth columns display the results for the 24-month horizon. The analysis covers five key macroeconomic indicators: Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation. Each subplot displays extremal score differences with 90% HAC-based confidence bands (gray shaded areas) across threshold parameter values. Negative differences indicate superior SZBVARx performance, while the magnitude reflects the strength of the performance advantage at each decision threshold.

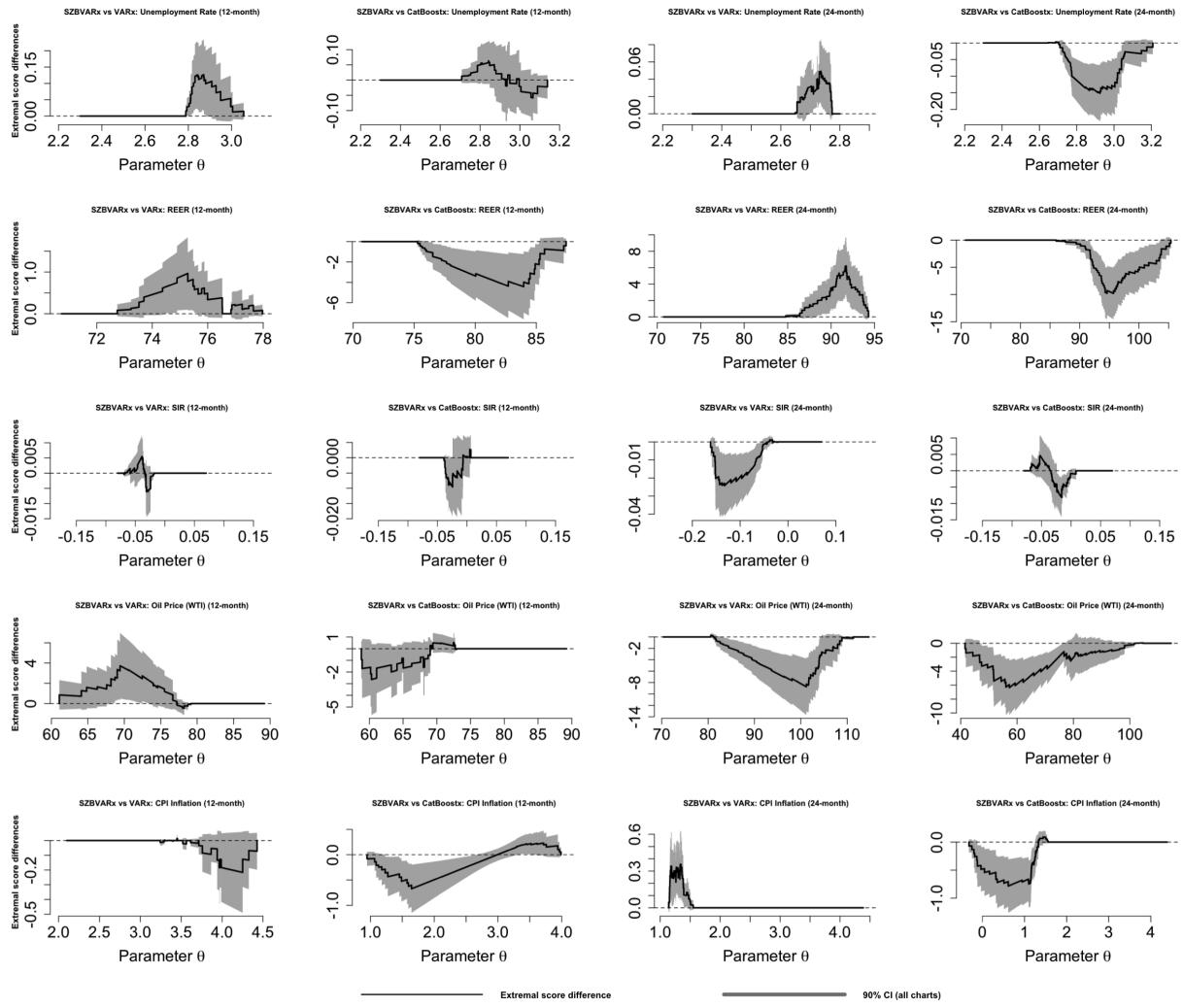


Figure A.1.25: Murphy diagram difference plots comparing SZBVARx forecasting performance against VARx and CatBoostx baselines across 12-month-and 24-month-ahead forecast horizons for Japan. Panels in the first and second columns correspond to the 12-month horizon, while panels in the third and fourth columns display the results for the 24-month horizon. The analysis covers five key macroeconomic indicators: Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation. Each subplot displays extremal score differences with 90% HAC-based confidence bands (gray shaded areas) across threshold parameter values. Negative differences indicate superior SZBVARx performance, while the magnitude reflects the strength of the performance advantage at each decision threshold.

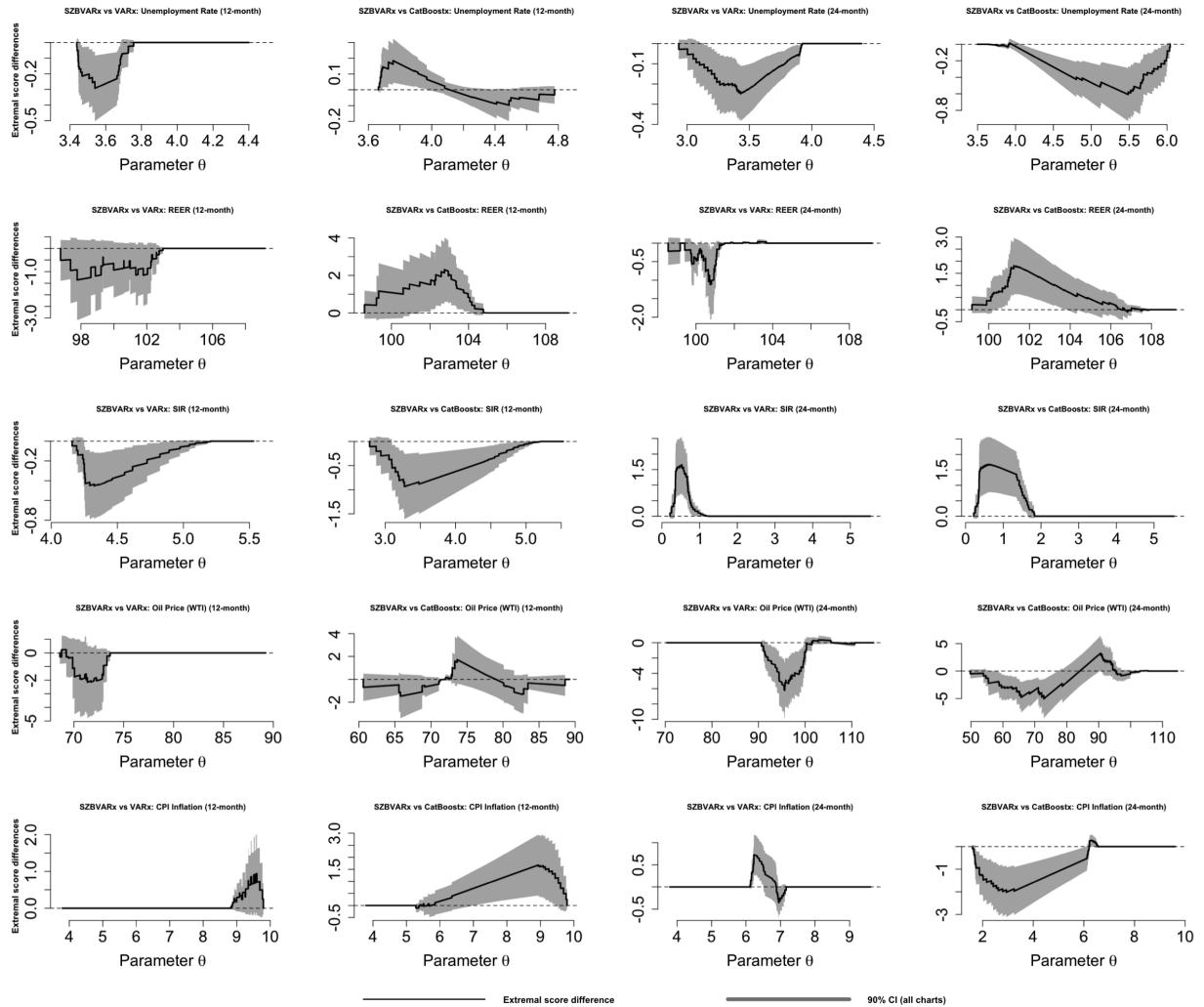


Figure A.1.26: Murphy diagram difference plots comparing SZBVARx forecasting performance against VARx and CatBoostx baselines across 12-month-and 24-month-ahead forecast horizons for the UK. Panels in the first and second columns correspond to the 12-month horizon, while panels in the third and fourth columns display the results for the 24-month horizon. The analysis covers five key macroeconomic indicators: Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation. Each subplot displays extremal score differences with 90% HAC-based confidence bands (gray shaded areas) across threshold parameter values. Negative differences indicate superior SZBVARx performance, while the magnitude reflects the strength of the performance advantage at each decision threshold.

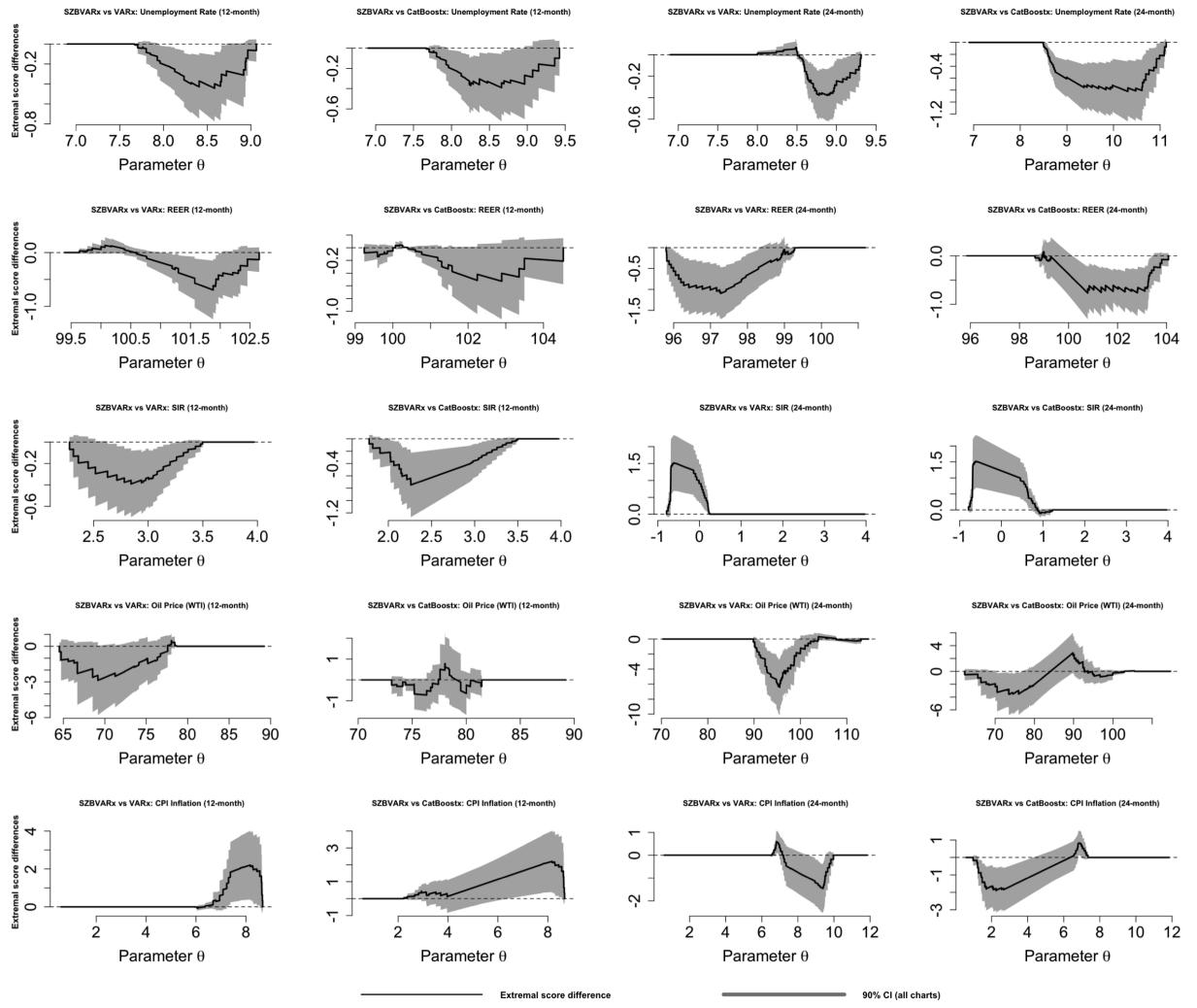


Figure A.1.27: Murphy diagram difference plots comparing SZBVARx forecasting performance against VARx and CatBoostx baselines across 12-month-and 24-month-ahead forecast horizons for Italy. Panels in the first and second columns correspond to the 12-month horizon, while panels in the third and fourth columns display the results for the 24-month horizon. The analysis covers five key macroeconomic indicators: Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation. Each subplot displays extremal score differences with 90% HAC-based confidence bands (gray shaded areas) across threshold parameter values. Negative differences indicate superior SZBVARx performance, while the magnitude reflects the strength of the performance advantage at each decision threshold.

Table A.1.1: Multivariate Giacomini-White test comparing SZBVARx with VARx and CatBoostx across G7 countries, five macroeconomic indicators (Unemployment Rate, REER, SIR, Oil Price (WTI), CPI Inflation), and 12-month-and 24-month-ahead forecast horizons. **Bold with *** denotes statistical significance at the 10% level, indicating rejection of equal conditional predictive ability and highlighting SZBVARx's robust gains.

Country	Variable	SZBVARx vs VARx (12M) (p-value)	SZBVARx vs CatBoostx (12M) (p-value)	SZBVARx vs VARx (24M) (p-value)	SZBVARx vs CatBoostx (24M) (p-value)
Canada	Unemployment Rate	0.4471	0.0821*	0.3416	0.0000*
	REER	0.2992	0.3922	0.0242*	0.0120*
	SIR	0.0015*	0.0000*	0.0041*	0.7414
	Oil Price (WTI)	0.1163	0.0004*	0.0009*	0.0072*
	CPI Inflation	0.0000*	0.3139	0.0791*	0.0396*
US	Unemployment Rate	0.2658	0.0099*	0.0000*	0.0000*
	REER	0.0158*	0.0000*	0.6372	0.0000*
	SIR	0.0024*	0.0000*	0.0046*	0.6034
	Oil Price (WTI)	0.0724*	0.1616	0.0005*	0.0528*
	CPI Inflation	0.6139	0.0000*	0.0121*	0.5240
France	Unemployment Rate	0.0174*	0.0071*	0.6266	0.0000*
	REER	0.2960	0.0324*	0.1162	0.0000*
	SIR	0.0183*	0.0000*	0.0059*	0.0020*
	Oil Price (WTI)	0.4691	0.0739*	0.0005*	0.0356*
	CPI Inflation	0.1068	0.8804	0.0351*	0.0011*
Germany	Unemployment Rate	0.0324*	0.0000*	0.0025*	0.0063*
	REER	0.0036*	0.0000*	0.1992	0.0530*
	SIR	0.9062	0.0000*	0.0608*	0.0001*
	Oil Price (WTI)	0.0000*	0.1119	0.0003*	0.0645*
	CPI Inflation	0.0233*	0.3276	0.8558	0.0101*
Japan	Unemployment Rate	0.0110*	0.9039	0.0000*	0.0000*
	REER	0.0048*	0.0000*	0.0007*	0.0000*
	SIR	0.9373	0.4240	0.0349*	0.6121
	Oil Price (WTI)	0.0010*	0.0841*	0.0003*	0.0000*
	CPI Inflation	0.1240	0.0003*	0.0000*	0.0131*
UK	Unemployment Rate	0.0000*	0.6738	0.0068*	0.0000*
	REER	0.0005*	0.0234*	0.0074*	0.1159
	SIR	0.0000*	0.0000*	0.0000*	0.0001*
	Oil Price (WTI)	0.1846	0.3239	0.0001*	0.0595*
	CPI Inflation	0.0008*	0.0375*	0.4879	0.0013*
Italy	Unemployment Rate	0.0000*	0.0069*	0.0557*	0.0002*
	REER	0.0176*	0.0913*	0.0269*	0.0000*
	SIR	0.0199*	0.0000*	0.0001*	0.0028*
	Oil Price (WTI)	0.0401*	0.4465	0.0000*	0.5644
	CPI Inflation	0.0123*	0.1351	0.3587	0.2495

A.1.9. Empirical Results - Uncertainty Quantification through Predictive Credible Intervals

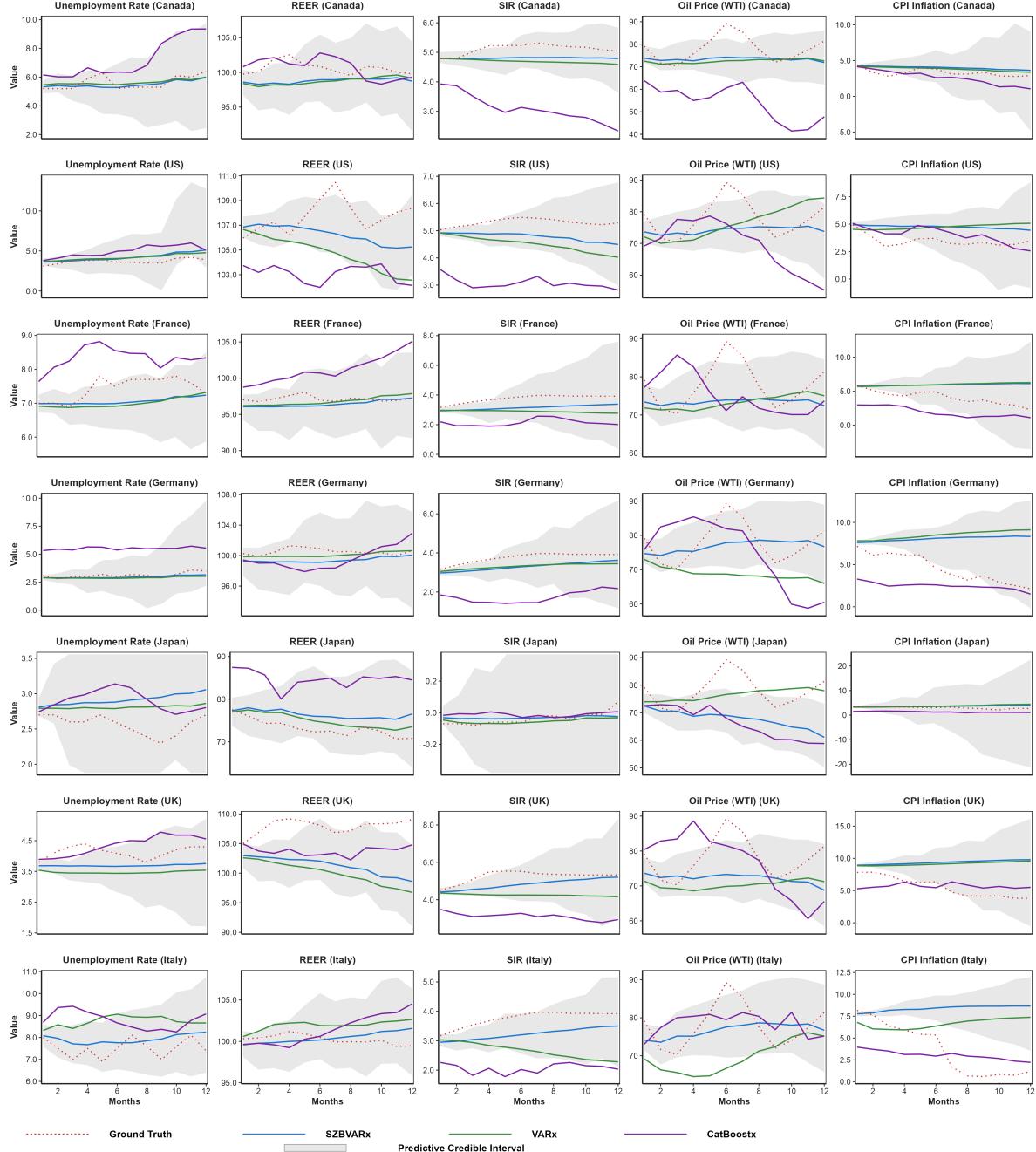


Figure A.1.28: Visualization of 12-month-ahead forecasts for five macroeconomic indicators across G7 countries. Each panel displays the ground truth (red dotted line), point forecasts from SZBVARx (blue), VARx (green), and CatBoostx (purple) models, along with predictive credible intervals from SZBVARx (grey shaded regions) quantifying forecast uncertainty. Variables shown are: Unemployment Rate, REER, SIR, Oil Price (WTI), and CPI Inflation Rate. Rows represent countries (Canada, US, France, Germany, Japan, UK, Italy from top to bottom), while columns represent the five macroeconomic variables.

Figure A.1.28 reports the 12-month-ahead forecast performance for all G7 countries. Overall,

results are qualitatively similar to the 24-months horizon, with some notable quantitative differences. As it is standard, the credible intervals are narrower. In particular, the SZBVARx model exhibits, once again, a strong alignment with ground truth as compared to VARx and CatBoostx baselines and the predictive credible intervals cover most of the realized values. In general, the intervals are widening over time and this pattern is even more accentuated than for the 24 months ahead horizon. The information transmission through exchange rates has the overall same features as in the 24-month-ahead horizon. In particular, the unemployment rate displays the narrowest intervals in the United Kingdom and Italy and the widest in France and the United States. For the REER, the tightest intervals are estimated for Germany and the United States, while the widest are estimated for Japan and Canada. As for the 24-month-ahead horizon, the forecasted SIR exhibits the narrowest uncertainty bands in the United States and Germany, while the wider intervals are obtained for France and Canada. This again points to important heterogeneity in monetary policy across countries, which is materializing in economic activity dynamics. Predictions of oil prices are instead fairly homogeneous across countries, with moderate uncertainty bands. Finally, the CPI inflation forecasts show large cross-country dispersion in predictive uncertainty, with the United Kingdom, the United States, and Germany being the countries with the narrowest intervals. France and Japan are instead the two countries with the largest predictive uncertainty, although the level of dispersion is smaller than for the 24-months horizon.