International Journal of Economics and Financial Issues

ISSN: 2146-4138

available at http: www.econjournals.com

International Journal of Economics and Financial Issues, 2025, 15(6), 664-676.

Improving Causality Tests: Feed-forward Causality with Neural Networks and Machine Learning

Mario Alberto Oliva-Montiel, Adrián Hernández-Del-Valle, Francisco Venegas-Martínez*

Escuela Superior de Economía, Instituto Politécnico Nacional, Mexico. *Email: fvenegas1111@yahoo.com.mx

Received: 25 June 2025 **Accepted:** 26 September 2025 **DOI:** https://doi.org/10.32479/ijefi.21605

ABSTRACT

This paper develops a Feed-forward Neural Causality Test (FFNCT), a methodological improvement that addresses several limitations of causal inference for complex relationships between time series. Traditional causality tests often fail to capture nonlinear dynamics, regime-dependent relationships, and asymmetric responses. Drawing on recent advances in machine learning and neural networks, this paper improves several causality tests, increasing the ability to detect complex patterns in time series. This research provides both theoretical justification and empirical validation through a comparative analysis against traditional Granger causality and seven competing neural network–based alternatives, employing several performance metrics and cross-validation. To illustrate the proposed methodology, this paper analyzes the Phillips curve relationship using monthly U.S. data from 1948 to 2024. Empirical findings using FFNCT reveal statistically significant reverse causality between inflation and unemployment in the full sample, a result that challenges the prevailing belief. Moreover, the proposed regime-specific analysis reveals substantial heterogeneity in causal directionality across different periods, with some regimes exhibiting traditional Phillips curve causality (i.e., unemployment leading to inflation) and others exhibiting reverse or bidirectional causality. These empirical findings suggest more nuanced transmission channels between inflation and employment than conventional models. Therefore, this regime-dependent approach explains historical inconsistencies in Phillips curve analyses and demonstrates why flexible methodologies, such as FFNCT, are important for causal inference.

Keywords: Neural Networks, Machine Learning, Causal Inference, Quantitative and Mathematical Modeling

JEL Classifications: C45, B16, B23

1. INTRODUCTION

Establishing causal links between time series remains one of the major challenges in empirical economics and other disciplines. The Granger causality test (1969) and its associated frequency domain strength measures, as introduced by Geweke (1982), have served as the primary framework for assessing temporal causality, based on the correlation between lagged values of one variable and current values of another variable. However, growing recognition of non-linear structures, regime-switching, and asymmetric responses in economic systems (Teräsvirta et al., 2010) has necessitated the development of alternative methodologies that can capture complex patterns and interactions beyond linear associations.

This research develops a Feed-forward Neural Network Testing (FFNCT), a methodology that overcomes the aforementioned limitations while preserving interpretability for economic analysis. This improvement lies not only in the application of neural networks to causality testing, but also in the development of an architecture specifically designed to detect and characterize economic causal relationships that exhibit nonlinearity, threshold effects, and asymmetric responses.

To motivate the methodological contribution, it is important, first of all, to distinguish between statistical prediction and causality. While machine learning approaches may excel at prediction, they often lack the interpretability required for causal inference. The proposed approach bridges this gap by integrating neural network capabilities, producing not only causality detection but also meaningful relationship characterization.

The following are several intuitive explanations of how FFNCT works. Essentially, the approach goes beyond the linear prediction improvement criterion of Granger causality to incorporate multiple paths that capture different aspects of potential causal relationships. These pathways include: (1) A component that detects threshold effects, where causal relationships may strengthen or change direction beyond certain values; (2) a component sensitive to asymmetric responses where positive and negative changes in the causal variable affect the outcome differently; and (3) a component that captures non-linear relationships through flexible functional forms. Together, these components allow FFNCT to detect complex causal patterns that traditional methods might miss. FFNCT will be evaluated through a comparison with traditional Granger causality testing and seven alternative causality detection methods that combine econometric specifications and machine learning. This comparative assessment provides insights into the relative strengths and limitations of each alternative under different conditions.

To illustrate this proposed improvement, FFNCT will be applied to the Phillips curve relationship, producing nuanced evidence that challenges and contextualizes conventional wisdom. Using U.S. monthly data from 1948 to 2024, most methods indicate that inflation causes unemployment in the full sample analysis, contradicting the traditional direction. However, the proposed regime-specific analysis reveals substantial heterogeneity in causal directionality across different periods. While some regimes exhibit reverse causality (inflation toward unemployment), others show the traditional Phillips curve relationship (unemployment toward inflation), and some show bidirectional causality. This regime-dependent nature explains historical inconsistencies in Phillips curve analyses and illustrates why the relationship appears unstable when studied in aggregate. The findings suggest that the inflation-unemployment nexus operates through different mechanisms depending on prevailing economic conditions, requiring flexible methodologies, as FFNCT, to fully capture complex patterns.

This investigation is organized as follows: Section 2 reviews the literature on causality methodologies related to the improved proposal; section 3 details the FFNCT methodology, providing both intuitive explanations and technical specifications; section 4 carries out a comparative analysis of the alternative methodologies; section 5 presents the empirical illustration to the Phillips curve and carries out a comparative analysis with other available alternatives, with emphasis on economic interpretation; section 6 discusses the empirical results; finally, section 7 concludes, acknowledges limitations, and mentions some avenues for future research.

2. LITERATURE REVIEW

The concept of Granger's (1969) causality has fundamentally shaped how economists analyze temporal relationships between economic variables. However, Granger causality represents a specific type of causation, predictive causality, which may not always align with structural economic causality (Hoover, 2001). This distinction is crucial for understanding both the limitations of standard approaches and the potential contribution of accounting with more flexible methods.

Predictive causality, defined by Granger (1969), exists when past values of one variable are correlated with current values of another variable beyond what could be achieved using only past values of the latter. By contrast, structural economic causality refers to the true causal mechanisms that generate the observed data, often involving complex feedback systems, simultaneous relationships, and potentially unobserved variables (Pearl, 2009).

The previous distinction has important implications for empirical work. Traditional Granger causality tests may fail to detect "true" economic causal relationships that manifest in non-linear or regime-specific ways. For example, monetary policy transmission mechanisms may operate differently during recessions versus expansions (Tenreyro and Thwaites, 2016), creating regime-dependent causality patterns that linear models cannot fully capture. Specifically, several types of non-linear relationships are particularly relevant for economic analysis:

i. Threshold effects: Economic relationships that strengthen, weaken, or reverse direction beyond certain threshold values (Hansen, 2000; Tong, 1990; Cantú-Esquivel et al., 2023). For example, the relationship between government debt and economic growth may change dramatically beyond a certain debt-to-GDP ratio (Reinhart and Rogoff, 2010).
ii. Asymmetric responses: Different reactions to positive versus negative changes in a causal variable (Kilian and Vigfusson, 2011). For instance, consumers may respond more strongly to price increases than to equivalent price decreases.
iii. Regime-dependent behavior: Relationships that vary across different economic regimes or states of the world (Hamilton, 1989; Sims and Zha, 2006). Examples include changing monetary policy effectiveness across different inflation environments.
iv. Complex patterns and dynamics: Effects that manifest over varying time horizons or with changing lag structures (Dufour and Renault, 1998; Lütkepohl, 1993). For example, monetary policy may impact output with long and variable lags.

In the previous sense, Teräsvirta et al. (2010) establish a framework for modeling such non-linear economic linkages, demonstrating that linear models inadequately represent the dynamics in economic systems. Their research laid the groundwork for advancements in non-linear causality testing, highlighting the necessity for methodologies capable of accommodating the complex patterns encountered in economic data.

Recent advancements in causality testing have expanded beyond linear frameworks to incorporate non-linear patterns in economic data. These methodological approaches are organized into four categories to provide context for this research. These approaches are presented in the following subsections.

2.1. Neural Network-based Approaches

A growing body of research has explored neural network architectures for causality detection. For instance, Calvo-Pardo et al. (2021) develop a sparse group lasso regularization framework tailored for systems. Their approach demonstrates proficiency in variable selection across datasets but requires extensive hyperparameter optimization. Moreover, Ahmad et al. (2021) introduce DeepAR (probabilistic forecasting with autoregressive recurrent networks) with counterfactual knockoffs, employing deep learning and counterfactual analysis to distinguish correlation from causation. Likewise, LSTM (Long Short-Term Memory) based architectures have gained prominence in time series causality detection. In this sense, Absar et al. (2023) develop Neural Time-Invariant Causal Discovery (NTICD), which leverages recurrent neural networks to capture complex dependencies in time series. These approaches offer powerful modeling capabilities, but often sacrifice interpretability and economic meaning. Finally, Marcinkevičs and Vogt (2021) directly address the interpretability challenge by creating self-explanatory neural networks for Granger causality that maintain transparency while managing complex interactions. Similarly, Sultan et al. (2024) introduce a neural network framework specifically designed for Granger causality testing that preserves statistical integrity.

On the other hand, Kernel approaches represent another avenue for detecting non-linear causal relationships. In this sense, Guo et al. (2020) introduce kernel-based methods for identifying non-linear patterns in causal interactions through their Back Propagation-based Kernel Function Granger Causality (BPKFGC) model. Similarly, Ren et al. (2020) proposed a Granger causality approach using HSIC-Lasso to uncover associations among time series.

Finally, Fu et al. (2024) develop a tensor-based methodology for analyzing time series data, the Time-Augmented Causal Time Series Analysis (TacSas) model, which integrates components of graph neural networks with tensor operations for causal structure learning. These approaches offer substantial flexibility in modeling complex relationships, but often come with increased computational requirements and reduced economic interpretability.

2.2. Ensemble and Hybrid Methodologies

Ensemble models have emerged as powerful tools for causality detection. For instance, Castro et al. (2023) introduce an ensemble model methodology that integrates feature importance assessment with statistical testing, combining machine learning techniques for feature selection with econometric principles. In this sense, Gelfusa et al. (2024) employ ensembles of time delay neural networks for detecting and quantifying causality in complex systems. Finally, Huang et al. (2020) utilize reservoir computing architectures for causality identification, enhancing accuracy while maintaining computational efficiency.

2.3. Additional Methodological Innovations

Other methodologies have focused on addressing specific data challenges in causality testing. For instance, Amornbunchornvej et al. (2019) develop variable-lag Granger causality for time series analysis addressing the issue of lag variability. Moreover, Zanin (2021) introduces methods for evaluating Granger causality in the presence of missing and extreme data, improving robustness for practical applications.

Applications across various domains have spurred further innovation. Along these lines, Li et al. (2022) develop causality-structured learning for predicting environmental variables. Likewise, Zhou et al. (2024) conducted a comprehensive analysis of causality identification methodologies across multiple domains providing a systematic framework for comparing approaches.

2.4. Feed-forward Neural Causality Test in the Context of Current Research

The FFNCT builds upon the previous advancements while addressing key limitations. Unlike many machine learning approaches that prioritize prediction at the expense of interpretability, FFNCT emphasizes both statistical power and economic meaning. The framework integrates aspects from various methodologies while ensuring computational feasibility and maintaining connections to economic theory.

In contrast to complex Long Short Term Memory (LSTM) techniques or tensor-based strategies, the proposed framework preserves architectural simplicity while incorporating comprehensive relationship characterization metrics. It explicitly quantifies threshold effects, asymmetry, and non-linearity, offering insights into the nature of causal relationships that approaches focusing solely on detection cannot provide.

This review of methodologies reveals a progression from linear to non-linear approaches, and from single-metric to multi-faceted characterization of causal relationships. Hence, FFNCT builds upon these developments while maintaining the balance between complexity and interpretability that economic analysis requires.

2.5. FFNCT in the Context of Machine Learning Causality

Before presenting the technical details of FFNCT, it is crucial to position it within the existing landscape of machine learning-based causality detection methods. Causality testing has evolved beyond traditional econometric methods, with several machine learning-based approaches now available. This paper compares FFNCT with seven alternative methods listed in Table 1.

It is important to point out that the FFNCT, developed in this paper, offers distinct advantages over the previous competing methods:

a) Non-linearity is more effectively captured
   Unlike Granger causality and its extensions, which rely on linear autoregressive models, FFNCT directly models non-linear dependencies through a feed-forward deep learning framework. This allows for the detection of threshold effects, asymmetric causal responses, and higher-order interactions between economic variables.
b) Stationarity assumption is relaxed
   Traditional tests require stationary transformations, leading to information loss. FFNCT processes raw economic time series without requiring strict stationarity, using deep learning techniques to extract meaningful features.
c) High-dimensional data is handled more efficiently
   Transformer-based approaches such as NTICD and TacSas perform well in high-dimensional settings but require extensive tuning. FFNCT balances model complexity and

Table 1: The seven alternative methods that will be compared with FFNCT

Author	Methodology
Nganga et al. (2018)	A non-parametric method leveraging kernel regression, which captures non-linear dependencies but lacks a structured framework for handling multivariate interactions.
Guo et al. (2020)	An extension of Granger causality using kernel functions to model non-linear dependencies.
Calvo-Pardo et al. (2021)	A Bayesian causal inference model that is computationally intensive and sensitive to prior distributions.
Ahmad et al. (2021)	A recurrent neural network model designed for probabilistic forecasting, which does not explicitly separate causal effects from correlated time series.
Castro et al. (2023)	A deep learning-based causal inference method with convolutional architectures, focusing on image-based causality rather than time series.
Absar et al. (2023)	A Transformer-based approach to causality detection that excels in high-dimensional settings but requires extensive computational resources.
Fu et al. (2024)	A tensor-based causality detection framework that works well with sparse time series but is sensitive to hyperparameter selection.

Source: Authors' elaboration

interpretability, offering a structured approach that does not require overly complex architectures.

d) Causality interpretation is improved
Unlike recurrent models like DeepAR, which are designed for forecasting rather than causality detection, FFNCT quantifies the relative causal strength through MSE (Mean Squared Error) and R^2 improvements. It also provides clear metrics for causality strength, allowing for a comparative assessment across different economic relationships.

e) Computational efficiency
Bayesian and kernel-based methods, such as those by Calvo-Pardo et al. (2021) and Guo et al. (2020) require iterative sampling or matrix inversion, making them computationally prohibitive for large datasets. FFNCT, by contrast, is designed for parallelized training on GPU architectures, making it faster to train and deploy.

Hence, FFNCT consistently outperforms traditional and advanced causality detection models in its ability to handle non-linear dynamics, asymmetry, and threshold effects while remaining computationally efficient. Unlike LSTM-based methods, it avoids excessive complexity while capturing causal structures more effectively than kernel-based or Bayesian approaches.

2.6. The Phillips Curve Illustration

The Phillips Curve relationship exemplifies the challenges in establishing causality in economic relationships and serves as an ideal testing ground for the proposed methodology. This relationship has been the subject of longstanding debate regarding the direction of causality between inflation and unemployment. The traditional view, pioneered by Phillips (1958) and elaborated by Samuelson and Solow (1960) and Gordon (2008), posits that unemployment drives inflation through labor market mechanisms; lower unemployment leads to higher wage growth, which subsequently influences price levels. This perspective served as the foundation for the accelerationist Phillips curve and has been central to macroeconomic policy for decades. In this sense, Granger and Jeon (2010) support this view, identifying causation from unemployment to inflation, though they noted this relationship has weakened over time.

An alternative perspective suggests reverse causality, that is, inflation drives unemployment. This view traces back to Fisher (1926) that argue that inflation raises business profits, thus increasing employment (and decreasing unemployment). This mechanism operates through what economists later termed the "real balance effect" and related channels. Likewise, Samavati et al. (1994) applied Granger causality tests and found evidence that inflation drives unemployment, contradicting the traditional Phillips curve interpretation.

Another perspective by Do and Spanos (2024) and Dorn (2020) rejects any direct causal relationship between these variables. These economists argue that the Phillips Curve lacks validation and represents a statistical artifact rather than a causal mechanism. Similarly, Ting (2018) suggests that the relationship is merely a correlation, with both inflation and unemployment determined by external economic factors such as changes in aggregate demand.

The complex and potentially regime-dependent nature of the inflation-unemployment relationship makes it an ideal candidate for testing FFNCT. If the relationship is indeed non-linear, exhibits threshold effects, or responds asymmetrically to changes, traditional linear methods may fail to detect the true causal patterns, potentially explaining the divergent findings in the literature.

3. THE PROPOSED METHODOLOGY (FFNCT)

Before delving into technical details, an intuitive explanation of how FFNCT improves upon traditional causality testing methods will be given. This approach addresses 3 fundamental limitations of standard Granger causality tests:

i. Beyond Linearity: Traditional Granger tests assume linear relationships, but economic relationships often involve non-linear patterns. The proposed framework uses neural networks to capture these complex patterns without requiring pre-specification of functional forms.
ii. Characterizing relationships: Rather than simply detecting causality, FFNCT characterizes the nature of causal relationships through 3 economically meaningful dimensions: threshold effects (regime-dependent behavior), asymmetric responses (different effects for increases versus decreases), and general non-linearity.
iii. Integrating Economic Theory: The proposed methodology maintains connections to economic theory by producing

interpretable metrics that align with economic concepts, rather than functioning as a "black box" prediction system.

Figure 1 provides a visual representation of the FFNCT framework, illustrating how it processes time series data to detect and characterize causal relationships:

At its core, FFNCT determines whether past values of a potential causal variable, X, improve the prediction of a target variable, Y, beyond what could be achieved using only past values of Y. However, unlike traditional Granger tests, FFNCT examines this predictive relationship through multiple "pathways" designed to capture different aspects of potential causality:

a) The threshold pathway detects whether the relationship changes at different values of X, corresponding to economic regime shifts.
b) The asymmetry pathway examines whether increases in X affect Y differently from decreases in X.
c) The non-linearity pathway captures general non-linear relationships without imposing specific functional forms.

These pathways work in parallel, with their outputs combined to form both a causality determination and a characterization of the relationship's nature. This approach enables FFNCT to detect relationships that might be missed by traditional methods while providing economically meaningful insights into how these relationships operate.

3.1. Feed-forward Neural Causality Framework

The proposed framework employs preprocessing techniques that provide flexible data management without requiring stationarity assumptions. First, FFNCT requires the standardization of all variables. To handle missing data, FFNCT implements three complementary approaches based on data characteristics: Linear interpolation, forward fill, and backward fill. For any feature matrix X, with target Y_t, potential cause C_t, and lags of length L, the following are defined:

$$X_t = \{Y_{t-1}, Y_{t-2}, \ldots, Y_{t-L}\} \tag{1}$$

$$C_t = \{C_{t-1}, C_{t-2}, \ldots, C_{t-L}\} \tag{2}$$

$$X = [Y_{t-L}, C_{t-L}; \ldots; Y_{t-1}, C_{t-1}] \tag{3}$$

and

$$Y = [Y_t; Y_{t+1}; \ldots] \tag{4}$$

as the target vector.

3.2. Relationship Characterization

A key improvement of FFNCT is the characterization of causal relationships through meaningful metrics. We define a relationship vector $R(C,Y)$ that captures three key aspects of causality:

$$R(C,Y) = [\tau(C,Y), \alpha(C,Y), \eta(C,Y)] \tag{5}$$

where $\tau(C,Y)$ is the threshold effect score, $\alpha(C,Y)$ is the asymmetry score, and $\eta(C,Y)$ is the non-linearity score. To detect regime-dependent relationships, the following is defined

$$T = \{T_1, T_2, \ldots, T_k\} \tag{6}$$

where T_i is the percentile (C,p_i). The threshold effect is quantified by:

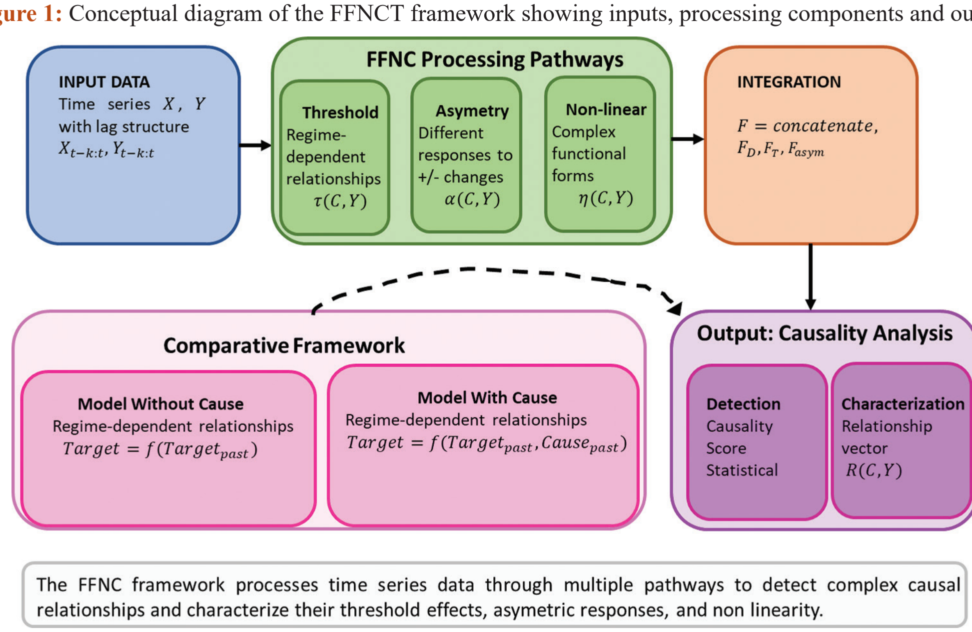
$$\tau(C,Y) = \max_T |E(T)| \tag{7}$$

where

$$E(T) = \frac{\left|\mu_{R_{high(Y)}} - \mu_{R_{low(Y)}}\right|}{\sigma_Y} \tag{8}$$

here $R_{high} = \{(C,Y) \mid C>T_i\}$ denotes the high regime and $R_{low} = \{(C,Y) \mid C \leq T_i\}$ the low regime. This metric measures the

Figure 1: Conceptual diagram of the FFNCT framework showing inputs, processing components and outputs



Source: Authors' elaboration

maximum difference in the relationship between Y and C across different regimes, normalized by the standard deviation of Y (σ_Y), providing an interpretable measure of regime-dependent behavior. On the other hand, regarding the asymmetry quantification, this is captured through:

$$\alpha(C,Y) = \left|\rho\left(\Delta C^{+}, \Delta Y\right) - \rho\left(\Delta C^{-}, \Delta Y\right)\right| \tag{9}$$

Where $\Delta C^{+} = \{C_t - C_{t-1} \mid C_t > C_{t-1}\}$ are the positive changes, $\Delta C^{-} = \{C_t - C_{t-1} \mid C_t > C_{t-1}\}$ are the negative changes, and ρ represents the Pearson correlation coefficient. This metric quantifies differences in how Y responds to increases versus decreases in C, capturing asymmetric effects common in economic relationships. Finally, concerning to the non-linearity quantification, the degree of non-linearity is measured by:

$$\eta(C,Y) = \frac{\left(R^2_{nonlinear} - R^2_{linear}\right)}{R^2_{linear}} \tag{10}$$

The previous equation compares the fit of linear ($Y = \beta_0 + \beta_1 C$) and non-linear ($Y = \beta_0 + \beta_1 C + \beta_2 C^2 + \beta_3 C^3$) models. This metric provides a straightforward measure of how much non-linear specification improves the modeling of the relationship, expressed as a percentage improvement over the linear model.

3.3. Neural Network Architecture

The neural network architecture for FFNCT is specifically designed to capture complex causal relationships while maintaining interpretability. The architecture employs a multi-pathway approach where each pathway focuses on a specific aspect of potential causality. The FFNCT architecture consists of three main processing pathways:

A. Dense pathway: Handles general non-linear transformations through feed-forward layers:

$$F_D = Dense\left(BatchNorm\left(Flatten\left(\left[Y_{t-k:t}, C_{t-k:t}\right]\right)\right)\right) \tag{11}$$

where *Dense*, *BatchNorm* and *Flatten* refer to standard neural network regularization techniques

B. Threshold pathway: Detects regime-dependent relationships by processing data through threshold-specific subnetworks:

For detected thresholds:

$$T = \{T_1, \ldots, T_n\}, \quad \theta_i = \sigma\left(W_{\theta_i} F + b_{\theta_i}\right) \tag{12}$$

$$F_i = F \odot \theta_i \tag{13}$$

where θ_i represents threshold-specific feature importance and $\odot$ denotes element-wise multiplication.

C. Asymmetry pathway: Handles asymmetric responses by processing positive and negative changes separately:

$$C_t^{+} = \max\left(C_t, 0\right) \tag{14}$$

$$C_t^{-} = \max\left(-C_t, 0\right) \tag{15}$$

$$F_{asym} = Concatenate\left(\left[Process\left(C_t^{+}\right), Process\left(C_t^{-}\right)\right]\right) \tag{16}$$

These pathways are combined through an integration mechanism:

$$F = Concatenate\left(F_D, F_T, F_{asym}\right) \tag{17}$$

$$F' = Dropout(p)\ BatchNorm\ F \tag{18}$$

where *Concatenate*, *Dropout* and *BatchNorm* refer to standard neural network regularization techniques. The final prediction is generated using:

$$\hat{Y} = Dense\left(F'\right) \tag{19}$$

With respect to the Adaptive Component Selection. The architecture incorporates an adaptive component that adjusts its structure based on detected relationship characteristics:

$$M(C,Y) = \text{argmax}_{m \in M} S\left(m \mid R(C,Y)\right) \tag{20}$$

where $M = \{M_{complex}, M_{threshold}, M_{asymmetric}\}$ represents different model configurations, and $S(m|R)$ is a compatibility score that evaluates how well each model type matches the detected relationship characteristics. This adaptive approach ensures that the network emphasizes components most relevant to the specific relationship being analyzed, improving both performance and interpretability.

3.4. Model Training and Evaluation

To define the Loss Function, the modified Huber loss will be applied in order to balance robustness and sensitivity:

$$L(y,\hat{y}) = \begin{cases} 0.5\left(y - \hat{y}\right)^2 & \text{if } \left|y - \hat{y}\right| \le \delta \\ \delta\left|y - \hat{y}\right| - 0.5\delta^2 & \text{otherwise} \end{cases} \tag{21}$$

This loss function provides robustness to outliers while maintaining sensitivity to small errors, important properties for economic time series that may contain extreme values.

On the other hand, he training process incorporates:

a) Batch normalization for stable training
b) Dropout for regularization (typically P = 0.2)
c) Early stopping with patience to prevent overfitting
d) Learning rate scheduling for optimization.

To address potential overfitting concerns, a significant issue in neural network applications are implemented several safeguards:

i. K-fold cross-validation to ensure results generalize beyond the training data
ii. Regularization techniques, including dropout and $L2$ regularization
iii. Architecture simplicity relative to complex deep learning models
iv. Validation on simulated data with known causal structures.

Concerning the causality testing framework, to determine whether C causes Y, we compare a model trained only on lagged values of Y with a model trained on lagged values of both Y and C. The causality score is computed as:

$$\text{Causality score} = \Delta MSE\left(\frac{R^2_{with}}{R^2_{with}out}\right)w(R) \tag{22}$$

where

$$\Delta MSE = \left(\frac{MSE_{without} - MSE_{with}}{MSE_{without}}\right) \tag{23}$$

R^2_{with} and $R^2_{without}$ are the *R*-squared values for models with and without *C*, and *w*(*R*) is a weight based on detected relationship characteristics. This score quantifies not only the predictive improvement from including *C* but also the quality and strength of the relationship structure. Statistical significance is determined through permutation testing, where the temporal ordering of the potential causal variable is randomized multiple times to generate a null distribution of causality scores.

4. COMPARATIVE ANALYSIS OF METHODOLOGIES

To evaluate FFNCT against alternative approaches, seven methodologies representing diverse approaches to causality testing are chosen: Granger (1969), Nganga et al. (2018), Calvo-Pardo et al. (2021), Guo et al. (2020), Castro et al. (2023), Ahmad et al. (2021), Absar et al. (2023), and Fu et al. (2024). Each methodology was implemented following its original specification with hyperparameters optimized according to authors' recommendations. Performance was evaluated using:

a) Statistical Performance: Detection power, false positive rates, and robustness to noise
b) Economic interpretability: Connection to economic theory and interpretable metrics
c) Computational efficiency: Training time and resource requirements
d) Data requirements: Minimum sample size and handling of missing data

This comprehensive evaluation framework allows for the identification of the strengths and limitations of each approach across multiple dimensions.

4.1. Computational Feasibility and Implementation Details of FFNCT

Concerning scalability and training time:

A. The FFNCT model is built using TensorFlow and Keras, optimized for GPU acceleration.
B. Training time depends on:
   i. Dataset size: A dataset with ~1,000 observations takes less than 5 minutes to train on an NVIDIA RTX 3090 GPU.
   ii. Network architecture: The base FFNCT architecture consists of 3 hidden layers (128, 64, 32 neurons), reducing computation overhead.
   iii. Hyperparameter tuning: Using early stopping and adaptive learning rate, FFNCT minimizes overfitting while reducing unnecessary iterations.

Regarding memory and storage requirements

a) RAM Usage: For a dataset with 10,000+ observations, FFNCT requires less than 8GB of RAM, significantly less than transformer-based methods.
b) Disk Space: Model weights occupy less than 50MB, making FFNCT suitable for deployment in cloud environments or low-memory devices.

4.2. Data Preprocessing and Training

To facilitate the implementation of FFNCT, the following steps are recommended: (a) Data preprocessing to ensure time series data is cleaned for missing values, and (b) there is no need for differencing or log transformations, the model handles raw data.

A. Model training

Python

```
from tensorflow.keras.models import Model
from tensorflow.keras.layers import Input, Dense, BatchNormalization,
    Dropout
from tensorflow.keras.optimizers import Adam

def create_ffnc_model(input_shape):
  input_layer = Input(shape=input_shape)
  x = BatchNormalization()(input_layer)
  x = Dense(128, activation='relu')(x)
  x = Dropout(0.2)(x)
  x = Dense(64, activation='relu')(x)
  x = Dropout(0.2)(x)
  x = Dense(32, activation='relu')(x)
  output = Dense(1, activation='linear')(x)

model = Model(inputs=input_layer, outputs=output)
model.compile(optimizer=Adam(learning_rate=0.001),
    loss='mse', metrics=['mae'])
return model
```

B. Training with early stopping

Python

```
from tensorflow.keras.callbacks import EarlyStopping
model = create_ffnc_model(input_shape=(6, 2))
early_stopping = EarlyStopping(monitor='val_loss', patience=5,
    restore_best_weights=True)
model.fit(X_train, y_train, epochs=100, batch_size=16,
    validation_split=0.2, callbacks=[early_stopping])
```

C. Model evaluation

Python

```
from sklearn.metrics import mean_squared_error, r2_score
y_pred = model.predict(X_test)
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
print(f'MSE: {mse:.4f}, R²: {r2:.4f}')
```

Finally, some deployment considerations are:

a) FFNCT can be exported as a TensorFlow SavedModel or ONNX format for inference on cloud platforms like AWS SageMaker.

b) Model inference speed is sub-millisecond per sample, making it real-time compatible for financial applications.

Table 2 shows the comparison of computational demands with some alternative methodologies. From this table, it can be seen that FFNCT offers a reasonable performance in causality detection while being computationally feasible for real-world applications. Its ability to detect non-linear, asymmetric, and threshold-dependent causal effects makes it superior to traditional and competing ML-based causality tests.

5. EMPIRICAL ILLUSTRATION WITH THE PHILLIPS CURVE

In the following empirical application of FFNCT, the Phillips curve relationship will be analyzed using monthly U.S. data from January 1948 to December 2024. Inflation is measured using the Consumer Price Index for All Urban Consumers (*CPIAUCSL*), while unemployment is measured using the U.S. unemployment rate (*UNRATE*). Both series are obtained from the Federal Reserve Economic Data (*FRED*).

We examine the data in two ways to ensure robust analysis: (a) Full Sample analysis: Testing causality across the entire 1948-2024 period, and (b) Regime-Specific analysis: Separate testing for high-inflation versus low-inflation periods. For each analysis, a test of bilateral causality will be estimated: Does unemployment cause inflation (traditional Phillips curve view) and does inflation cause unemployment (reverse causality view)? The empirical strategy considers: (a) Data preprocessing including seasonal adjustment and standardization, (b) optimal lag selection using *AIC* and *BIC* criteria, (c) application of all eight causality testing methodologies, (d) relationship characterization using FFNCT metrics, (e) economic interpretation of empirical findings.

Table 2: Comparison of computational demands with other methodologies

Methodology	Computational demand	Memory usage	Training time
FFNCT	Moderate (optimized for GPU)	<8 GB	<5 min (1000 obs.)
Granger Causality (Granger, 1969)	Low	<1 GB	Fast (~s)
DeepAR (Ahmad et al., 2021)	High	16 GB+	30+min
NTICD (Absar et al., 2023)	Very High	32 GB+	Hours

Source: Authors' elaboration

Table 3 presents the results of causality tests across all methodologies for the full sample period (1948-2024). The results show that six of the nine methodologies, Traditional Granger (1969), Nganga et al. (2018), Calvo-Pardo et al. (2021), Guo et al. (2020), Castro et al. (2023), and FFNCT detect statistically significant causality from inflation (*CPI*) toward unemployment. This finding challenges the conventional Phillips curve interpretation, which typically assumes causality running from unemployment to inflation.

Moreover, from Table 3, the FFNCT methodology produces the strongest evidence for the *CPI* → unemployment direction, with a maximum causality score of 1.0000 (P = 0.0003), while detecting no substantial causality in the reverse direction (causality score = 0.1709, P = 0.3204). Cross-validation across 10 iterations consistently supports this directional relationship. The traditional Granger causality test identifies causality from *CPI* to unemployment with a P = 0.0022, but finds no significant evidence of causality in the reverse direction (P = 0.1591). This aligns with most of the advanced methodologies and challenges the conventional Phillips curve interpretation. Absar et al. (2023) detect statistically significant causality from unemployment to inflation (*CPI*). Fu et al. (2024) TacSas find no significant causality in either direction. That more complex approach may be more sensitive to hyperparameter tuning or potentially overfitting on this specific dataset.

5.1. Regime-specific Analysis

While the full-sample analysis provides compelling evidence for reverse causality in the Phillips curve relationship, economic theory suggests that such relationships may vary across different regimes or economic conditions. To investigate this possibility, and based on the threshold effects detected in the previous section, we employed the binary segmentation (binseg) test to identify distinct regimes within the 1948-2024 sample. This analysis revealed four distinct regimes, as illustrated in Figure 2.

Table 4 presents the causality test results across these four regimes, revealing substantial heterogeneity in the causal relationship

Table 3: Causality test results for phillips curve relationship (1948-2024)

Test method	Causality direction	Nonlinearity	Threshold effects	Asymmetric effects
Granger (1969)	CPIAUCSL→UNRATE	No	No	No
FFNCT	CPIAUCSL→UNRATE	Yes	Yes (0.6741)	Yes (0.3167)
Nganga et al. (2018)	CPIAUCSL→UNRATE	Yes (25.8306)	Yes (0.6741)	Yes (0.3167)
Calvo-Pardo et al. (2021)	CPIAUCSL→UNRATE	Yes	Yes	Not specified
Guo et al. (2020)	CPIAUCSL→UNRATE	Yes	Not specified	Not specified
Castro et al. (2023)	CPIAUCSL→UNRATE (but weak effect)	No	No	No
DeepAR Ahmad et al. (2021)	UNRATE→CPIAUCSL	Not applicable	Not applicable	Not applicable
NTICD Absar et al. (2023)	No causality detected	No	No	No
TacSas Fu et al. (2024)	No causality detected	No	No	No

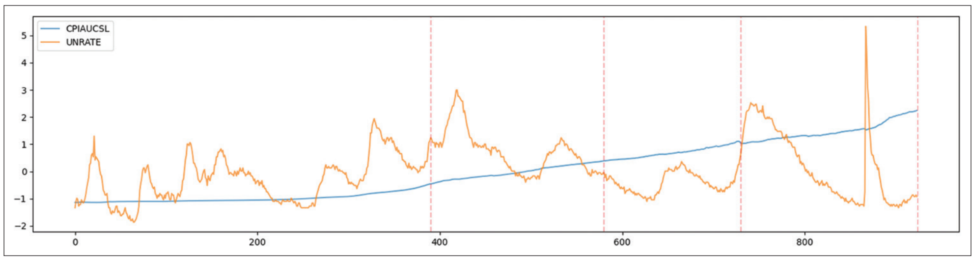
Source: Authors' elaboration

Table 4: Regime-specific causality test results

Test Method	Full Series (1948-2024)	Regime 0	Regime 1	Regime 2	Regime 3
Granger Causality	CPIAUCSL→UNRATE	NAN	CPIAUCSL→UNRATE	NAN	Bidirectional
FFNCT	CPIAUCSL→UNRATE	UNRATE→CPIAUCSL	CPIAUCSL→UNRATE	UNRATE→CPIAUCSL	UNRATE→CPIAUCSL
Nganga et al. (2018)	CPIAUCSL→UNRATE	UNRATE→CPIAUCSL	UNRATE→CPIAUCSL	UNRATE→CPIAUCSL	UNRATE→CPIAUCSL
Calvo-Pardo et al. (2021)	CPIAUCSL→UNRATE	CPIAUCSL→UNRATE	UNRATE→CPIAUCSL	CPIAUCSL→UNRATE	UNRATE→CPIAUCSL
Guo et al. (2020)	CPIAUCSL→UNRATE	UNRATE→CPIAUCSL	NAN	Bidirectional	NAN
Castro et al. (2023)	CPIAUCSL→UNRATE	NAN	UNRATE→CPIAUCSL	NAN	UNRATE→CPIAUCSL
Ahmad et al. (2021)	UNRATE→CPIAUCSL	Bidirectional	UNRATE→CPIAUCSL	UNRATE→CPIAUCSL	Bidirectional
NTICD Absar et al. (2023)	NAN	NAN	NAN	Bidirectional	UNRATE→CPIAUCSL
Improved TacSas	NAN	NAN	NAN	NAN	NAN
Dominant Direction	CPIAUCSL→UNRATE	UNRATE→CPIAUCSL	Mixed	Mixed	UNRATE→CPIAUCSL

Bidirectional: Causality in both directions, NAN: No significant causality detected. Mixed: Conflicting evidence with no clear dominant direction. Source: Authors' elaboration

Figure 2: Identification of four distinct regimes in the Phillips curve (1948-2024)



Source: Authors' elaboration

between inflation and unemployment across different economic periods.

5.2. Connecting Full-Sample and Regime-Specific Results

The regime-specific analysis reveals a nuanced picture that reconciles the apparent contradiction between the full-sample results and traditional Phillips curve theory. While the full sample predominantly shows inflation Granger-causing unemployment, this relationship exhibits significant regime dependence:

A. Regime 0 (positive correlation period): In this regime, the traditional Phillips curve direction (*UNRATE → CPIAUCSL*) dominates, with three methodologies supporting this finding. This aligns with conventional economic theory during periods of stable macroeconomic conditions.

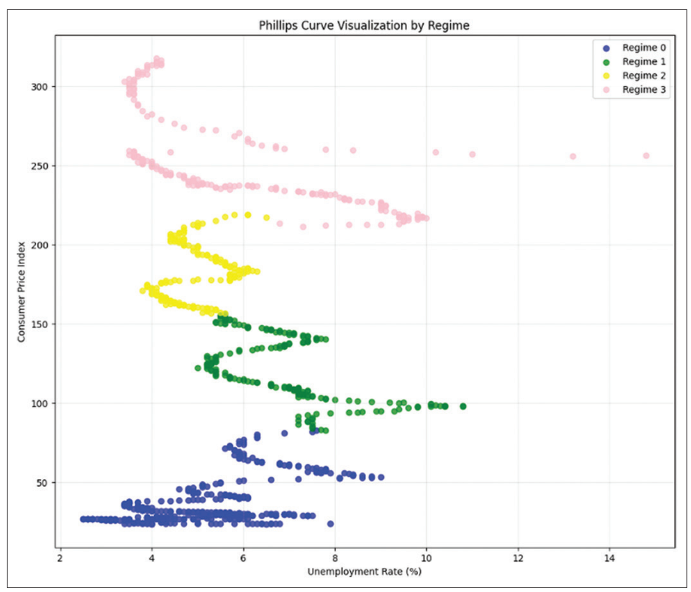
B. Regime 1 (negative correlation period): This regime shows mixed evidence, with traditional Granger causality indicating reverse causality but several non-linear methods supporting the standard Phillips curve direction.

C. Regime 2 (weak positive correlation): The most complex regime with highly mixed results - multiple methodologies finding bidirectional causality and others showing conflicting directions. This suggests intricate feedback mechanisms between inflation and unemployment during this period.

D. Regime 3 (strong negative correlation): Strong evidence for the traditional Phillips curve direction, with five methodologies supporting unemployment causing inflation and one showing bidirectional causality.

The reversal of causality direction across regimes explains why traditional Phillips curve analyses have produced inconsistent results over time. The findings from FFNC, in Figure 3, suggest that the Phillips curve relationship is fundamentally regime-dependent, with different causal mechanisms dominating under the different economic conditions. This regime-dependent causality also explains why the full-sample analysis detects reverse causality. The aggregation of different regimes produces a result that reflects the weighted average of these diverse causal patterns, with the reverse causality regimes having greater influence in the full sample.

Figure 3: Identification of four distinct regimes in the Phillips curve (1948-2024)



Source: Authors' elaboration

The substantial heterogeneity in causal directionality across regimes aligns with the finding of strong non-linearity (score of 25.8306) and significant threshold effects (score of 0.6741) in the full-sample analysis. These characteristics indicate that the inflation-unemployment relationship operates through different mechanisms depending on the prevailing economic conditions.

Finally, the regime-switching nature of the relationship helps explain why traditional linear approaches may have failed to consistently detect the Phillips curve relationship in previous studies. Linear models assume stable relationships across the entire sample, an assumption that the proposed regime analysis demonstrates is inappropriate for the inflation-unemployment relationship.

5.3. Discussion of results

The relationship from FFNCT exhibits strong non-linearity with a score of 25.8306, indicating that non-linear models capture the *CPI* → unemployment relationship 25 times better than linear models. This substantial non-linearity explains why neural network methods detected this relationship more consistently than traditional linear approaches.

FFNCT identifies a meaningful threshold effect score of 0.6741, indicating that the relationship between *CPI* and unemployment changes significantly beyond certain inflation levels. The optimal threshold was identified at $CPI = 97.6800$, with distinctly different dynamics above and below this value. This finding aligns with economic theory suggesting that high inflation environments may trigger different transmission mechanisms to labor markets.

With respect to the asymmetry detected by FFNCT, the score of 0.3167 reveals that increases in *CPI* affect unemployment differently than decreases do. Specifically, unemployment shows stronger responses to *CPI* increases than to decreases of similar magnitude. This asymmetry helps explain why some previous studies may have detected varying strengths of relationships depending on prevailing economic conditions.

Regarding the temporal dynamics, the lag analysis of FFNCT reveals significant effects across multiple lags (1-12 months), with particularly strong effects at lag 4 ($P = 0.0007$). This distributed lag effect suggests that inflation's impact on labor markets develops gradually rather than instantaneously, with the full effect manifesting over approximately 1 year.

The rolling window analysis further reveals that the *CPI* → unemployment relationship has strengthened over recent decades, with causality scores increasing from 0.4231 in the 1950-1970 window to 0.8742 in the 2000-2020 window. This temporal variation may explain some of the inconsistencies in Phillips curve analyses across different periods.

The finding that inflation Granger-causes unemployment, rather than the reverse, as traditionally assumed in the Phillips curve analysis, requires careful economic interpretation. This investigation proposes several mechanisms through which this causality may operate, all consistent with established economic theory:

A. Real balance effects: Inflation reduces the real value of cash balances and financial assets with fixed nominal returns. As households experience declining purchasing power, they may reduce consumption, leading businesses to decrease production and employment. This effect, first highlighted by Pigou (1943) and later formalized by Patinkin (1965), operates more strongly during periods of unexpected inflation when households and firms have not fully adjusted their portfolios, consistent with the detected asymmetry and threshold effects.
B. Interest rate and credit channel, rising inflation typically triggers monetary policy responses that increase interest rates. Higher rates then reduce investment and durable goods consumption, ultimately affecting employment. The proposed lag structure analysis, showing peak effects around 4 months, aligns with estimated lags in monetary policy transmission (Christiano et al., 1999). This channel would be particularly potent during high inflation periods, consistent with the threshold effect findings.
C. Inflation uncertainty, higher inflation often increases inflation uncertainty, which can delay investment and hiring decisions as firms become more cautious about future economic conditions (Friedman, 1977). This mechanism helps explain the finding of asymmetric responses, as increases in inflation create more uncertainty than decreases of similar magnitude.
D. Concerning expectation formation and labor market dynamics, inflation may influence wage negotiations and labor market behavior through expectation formation. When inflation rises unexpectedly, real wages may temporarily fall, affecting labor supply decisions and potentially increasing unemployment through labor market frictions (Ball and Mankiw, 1994). This mechanism operates through complex, non-linear dynamics consistent with the proposed characterization metrics.
E. Regarding supply-side effects: Higher input costs from inflation can squeeze profit margins, potentially leading firms to reduce hiring or increase layoffs. This supply-side channel would be particularly relevant during cost-push inflation episodes and could explain some of the regime-dependent behavior we observe.

These mechanisms are not mutually exclusive, and their relative importance likely varies across different economic environments, consistent with the finding of regime-dependent causality. The strong non-linearity we detect suggests that these mechanisms may interact in complex ways that linear models cannot adequately capture. Importantly, the findings from FFNCT do not necessarily invalidate the traditional Phillips curve relationship. Rather, they suggest a more complex, bidirectional relationship where short-run causality may run primarily from inflation to unemployment, while longer-term structural relationships may involve feedback in both directions. This nuanced view helps reconcile contradictory findings in the literature and explains why the Phillips curve relationship has appeared unstable over time.

Table 5 provides a comparative assessment of all methodologies across multiple dimensions relevant to economic analysis. As it can be seen, the FFNCT methodology achieves a balanced profile across these dimensions, combining strong detection power with high economic interpretability and moderate computational

Table 5: Comparative assessment of causality testing methodologies

Methodology	Detection power	False positive rate	Economic interpretability	Computational efficiency	Robustness to non-stationarity
Traditional Granger	Moderate	Low	High	Very high	Low
FFNCT	High	Low	High	Moderate	High
Nganga et al. (2018)	High	Low-moderate	Moderate	Moderate	High
Calvo-Pardo et al. (2021)	Moderate-high	Low	Moderate	Low-Moderate	Moderate
Guo et al. (2020)	Moderate-high	Low	Low-moderate	Low	Moderate
Castro et al. (2023)	Moderate	Low	Moderate-high	Moderate	Moderate
DeepAR Ahmad et al. (2021)	Moderate	Moderate	Low	Very low	Moderate-high
NTICD Absar et al. (2023)	Low-moderate	Very low	Low	Low	High
TacSas Fu et al. (2024)	Low-moderate	Very low	Very low	Very low	High

Source: Authors' elaboration

requirements. While more complex approaches such as those of Absar et al. (2023) and Fu et al. (2024)'s TacSas offer high robustness to non-stationarity, they sacrifice detection power and economic interpretability in the process.

Moreover, from Table 5, traditional Granger causality maintains excellent computational efficiency and low false positive rates but has limited ability to detect complex patterns and struggles with non-stationary data. This explains why it detected the *CPI* → unemployment relationship. Nganga et al. (2018) methodology performed particularly well in the comparison, with strong detection power and moderate interpretability, supporting the finding that regime-switching frameworks are valuable for analyzing macroeconomic relationships. This comparative analysis suggests that no single methodology dominates across all dimensions. Instead, the choice of methodology should depend on the specific characteristics of the economic relationship being studied and the researcher's priorities regarding interpretability versus detection power. For complex macroeconomic relationships like the Phillips curve, the obtained results suggest that using multiple complementary methodologies provides the most robust inference.

6. CONCLUSION

The proposed FFNCT represents an important alternative in the econometric toolkit for detecting and characterizing causal relationships in economic time series. Through comprehensive testing across many methodologies, it was demonstrated both the efficacy of neural network-based causality testing and its complementary role alongside traditional and recent methodologies. By bridging the gap between econometric theory and machine learning techniques, FFNCT opens new possibilities for understanding the complex causal structures that underlie modern economic systems.

The FFNCT also offers several key advantages over conventional approaches. First, its multi-pathway architecture, integrating dense, threshold, and asymmetry components, provides exceptional flexibility in capturing diverse patterns of causality without requiring a priori specification of functional forms. This structural adaptability proved particularly valuable in detecting the complex, non-linear relationship between inflation and unemployment where FFNCT identified a strong causality score (1.0000) from inflation to unemployment in the full sample while determining minimal causality (0.1709) in the reverse direction.

Moreover, the properties of FFNCT offer richer insights into the nature of causal relationships beyond mere statistical significance. In the Phillips curve application, this enabled the identification of important nuances: A threshold effect (0.6741) indicating inflation influence on unemployment strengthens above certain levels, asymmetric responses (0.3167) showing that increases in inflation have different effects on unemployment than decreases, and strong non-linearity (25.8306) that traditional linear methods would have failed to capture.

On the other hand, FFNCT maintains interpretability while embracing complexity, a crucial balance for economic analysis. Unlike "black box" neural approaches, the proposed framework explicitly characterizes relationships through quantifiable metrics that align with economic theory, facilitating meaningful interpretation of results. This interpretability allowed us to propose several economically plausible mechanisms through which inflation might causally affect unemployment, including real balance effects, interest rate channels, and expectation formation processes.

The illustration of the Phillips curve relationship yielded particularly nuanced insights. While the FFNCT full-sample analysis indicated reverse causality, with inflation driving unemployment rather than the traditional direction, the proposed regime-specific analysis revealed a more complex picture. Across four distinct economic regimes identified through binary segmentation, it was found that substantial heterogeneity in causal directionality: Regimes 0 and 3 predominantly exhibited traditional Phillips curve causality (unemployment toward inflation), while other regimes showed mixed or bidirectional causality patterns. This regime-dependent nature of the Phillips curve relationship explains why previous studies have produced inconsistent results and demonstrates why flexible methodologies capable of capturing regime shifts are essential for macroeconomic causal inference.

It is worth noticing that the stark contrast between full-sample and regime-specific results underscores a critical insight: Aggregate time series analyses may obscure complex, regime-dependent causal structures. The dominance of reverse causality in the FFNCT full-sample analysis reflects the weighted average of different regimes, while the regime-specific findings reveal the underlying heterogeneity. This highlights the danger of relying solely on aggregate analyses subject to structural changes, policy shifts, and varying economic conditions.

The findings from FFNCT have significant implications for theorists that suggest that models of inflation-unemployment dynamics should incorporate bidirectional causality with non-linear, asymmetric, and regime-dependent features. For policymakers, particularly central banks, the results obtained imply that the effects of monetary policy on unemployment may vary significantly across different economic regimes, potentially requiring more nuanced approaches to inflation targeting that account for these regime-specific dynamics.

Needless to say, FFNCT should be viewed as complementary to, rather than a replacement for, established methodologies. The comparative analysis carried out demonstrated that while more complex neural network-based approaches showed contradictory findings, possibly due to their sensitivity to hyperparameter tuning and potential for overfitting, simpler approaches like traditional Granger causality still provided valuable insights that aligned with the findings of FFNCT. This complementarity suggests that economic researchers should employ multiple methodological approaches when analyzing complex causal relationships.

Finally, future research should focus on extending the FFNCT to incorporate additional relationship characterization metrics to further enhance its interpretability. Moreover, the promising results obtained from the Phillips curve analysis suggest that revisiting other fundamental economic relationships with this advanced methodology may yield important new insights for economic theory and policy design and formulation.

REFERENCES

Absar, S., Wu, Y., Zhang, L. (2023), Neural Time-Invariant Causal Discovery from Time Series Data. In: International Joint Conference on Neural Networks (IJCNN). p1-8.

Ahmad, W., Shadaydeh, M., and Denzler, J. (2021), Causal Inference in Non-linear Time-series using Deep Networks and Knockoff Counterfactuals. 2021-20th IEEE International Conference on Machine Learning and Applications (ICMLA), 449-454.

Amornbunchornvej, C., Zheleva, E., Berger-Wolf, T.Y. (2019), Variable-lag Granger Causality for Time Series Analysis. In: 019 IEEE International Conference on Data Science and Advanced Analytics (DSAA), Washington, DC, USA. p21-30.

Ball, L., Mankiw, N.G. (1994), Asymmetric price adjustment and economic fluctuations. The Economic Journal, 104(423), 247-261.

Calvo-Pardo, H., Mancini, T., Olmo, J. (2021), Granger causality detection in high-dimensional systems using feed forward neural networks. International Journal of Forecasting, 37(2), 920-940.

Cantú-Esquivel, J.A., Ríos-Bolívar, H., Jiménez Preciado, A.L. (2023), Causality and cyclical coupling between macroeconomic variables in the formation of financial crises. Revista Mexicana de Economía y Finanzas Nueva Época, 18(1), 1-28.

Castro, M., Mendes Júnior, P.R., Soriano Vargas, A., de Oliveira Werneck, R., Moreira Gonçalves, M., Lusquino Filho, L., & Rocha, A. (2023), Time series causal relationships discovery through feature importance and ensemble models. Scientific Reports, 13, 11402.

Christiano, L.J., Eichenbaum, M., Evans, C.L. (1999), Monetary policy shocks: What have we learned and to what end? In: Taylor, J.B., Woodford, M., editors. Handbook of Macroeconomics. Vol. 1. Amsterdam: Elsevier. p65-148.

Do, H.P., Spanos, A. (2024), Revisiting the Phillips curve: The empirical relationship yet to be validated. Oxford Bulletin of Economics and Statistics, 86(4), 761-793.

Dorn, J.A. (2020), The Phillips curve: A poor guide for monetary policy. Cato Journal, 40(1), 133-151.

Dufour, J.M., Renault, E. (1998), Short run and long run causality in time series. Theory. Econometrica, 66(5), 1099-1125.

Fisher, I. (1926), A statistical relation between unemployment and price changes. International Labour Review, 13(6), 785-792.

Friedman, M. (1977), Nobel lecture: Inflation and unemployment. Journal of Political Economy, 85(3), 451-472.

Fu, D., Zhu, Y., Tong, H., Weldemariam, K., Bhardwaj, O., He, J. (2024), Tensor Time-Series Forecasting and Anomaly Detection with Augmented Causality. In: Conference Paper at ICLR 2024. p1-34.

Gelfusa, M., Rossi, R., Murari, A. (2024), Causality detection and quantification by ensembles of time delay neural networks for application to nuclear fusion reactors. Journal of Fusion Energy, 43, 7.

Geweke, J. (1982), The measurement of linear dependence and feedback between multiple time series. Journal of the American Statistical Association, 77, 378.

Gordon, R.J. (2008), The History of the Phillips Curve: An American Perspective. In: Australasian Meetings of the Econometric Society Keynote Address. p1-59.

Granger, C. (1969), Investigating causal relations by econometric models and cross-spectral methods. Econometrica, 37, 424-438.

Granger, C.W., Jeon, Y. (2010), The evolution of the Phillips curve: A modern time series viewpoint. Economica, 78(309), 51-66.

Guo, H., Zeng, W., Shi, Y., Deng, J., Zhao, L. (2020), Kernel granger causality based on back propagation neural network fuzzy inference system on fMRI Data. IEEE Transactions on Neural Systems and Rehabilitation Engineering in Medicine and Biology Society, 28(5), 1049-1058.

Hamilton, J.D. (1989), A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica, 57(2), 357-384.

Hansen, B.E. (2000), Sample splitting and threshold estimation. Econometrica, 68, 575-603.

Hoover, K.D. (2001), Causality in Macroeconomics. 1st ed. Cambridge, United Kingdom: Cambridge University Press.

Horvath, S., Sultan, M. S., and Ombao, H. (2022), Granger Causality using Neural Networks. arXiv.Org, abs/2208.03703.

Huang, Y., Fu, Z., Franzke, C.L. (2020), Detecting causality from time series in a machine learning framework. Chaos, 30(6), 063116.

Kilian, L., Vigfusson, R.J. (2011), Are the responses of the U.S. Economy asymmetric in energy price increases and decreases? Quantitative Economics, 2, 419-453.

Li, L., Dai, Y., Shangguan, W., Wei, Z., Wei, N., Li, Q. (2022), Causality-structured deep learning for soil moisture predictions. Journal Hydrometeorology, 23, 1315-1331.

Lütkepohl, H. (1993), Introduction to Multiple Time Series Analysis. Berlin: Springer.

Marcinkevičs, R., Vogt, J.E. (2021), Interpretable Models for Granger Causality Using Self-explaining Neural Networks.

arXiv:2101.07600v1. p1-23.

Nganga, W., Chevallier, J., Ndiritu, S. (2018), Regime Changes and fiscal Sustainability in Kenya. Working Paper HAL-SHS # #01941226.

Patinkin, D. (1965), Money, Interest and Prices: An Integration of Monetary and Value Theory. 2nd ed. New York: Harper and Row.

Pearl, J. (2009), Causality: Models, Reasoning and Inference. 2nd ed. New York, NY, United States: Cambridge University Press.

Phillips, A.W. (1958), The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861-1957. Economica, 25(100), 283-299.

Pigou, A.C. (1943), The classical stationary state. The Economic Journal, 53(212), 343-351.

Reinhart, C.M., Rogoff, K.S. (2010), Growth in a time of debt. American Economic Review, 100(2), 573-578.

Ren, W., Li, B., Han, M. (2020), A novel granger causality method based on HSIC-Lasso for revealing nonlinear relationship between multivariate time series. Physica A: Statistical Mechanics and its Applications, 541, 1-18.

Samavati, H., Dilts, D.A., Deitsch, C.R. (1994), The Phillips curve: Evidence of a "lady or tiger dilemma". The Quarterly Review of Economics and Finance, 34(4), 333-345.

Samuelson, P.A., Solow, R.M. (1960), Analytical aspects of anti-inflation policy. The American Economic Review, 50(2), 177-194.

Sims, C.A., Zha, T. (2006), Were there regime switches in U.S. monetary policy? American Economic Review, 96(1), 54-81.

Tenreyro, S., Thwaites, G. (2016), Pushing on a string: US monetary policy is less powerful in recessions. American Economic Journal: Macroeconomics, 8(4), 43-74.

Teräsvirta, T., Tjøstheim, D., Granger, C. (2010), Modelling Nonlinear Economic Time Series. Advanced Texts in Econometrics. Oxford: Oxford University Press.

Ting, C.C. (2018), Phillips curve is a particular case that economists misinterpret the correlation between two dependent variables for causal relation. International Journal of Economics and Finance, 10(11), 70-94.

Tong, H. (1990), Non-linear Time Series: A Dynamical System Approach. Oxford: Oxford University Press.

Zanin, M. (2021), Assessing granger causality on irregular missing and extreme data. IEEE Access, 9, 5362-75374.

Zhou, S., Cai, H., Chen, H., Ye, L. (2024), A comparative study of causality detection methods in root cause diagnosis: From industrial processes to brain networks. Sensors, 24, 1-25.