Available online at website: https://jurnal.iaii.or.id/index.php/RESTI

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 9 No. 6 (2025) 1417 - 1427 e-ISSN: 2580-0760

A New Approach for Dynamic Analysis of Indonesian Food Prices using the PC Algorithm and Vector Autoregression

Teny Handhayani[1], Yudistira Permana[2], Akmal Farouqi[3], Naufal Firdausyan[4], Raffy Sonata[5], Marcel Yusuf Rumlawang Arpipi[6], Irvan Lewenusa[7]
[1,6,7]Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara, Jakarta, Indonesia
[2]Department of Economics and Business, Vocational College, Universitas Gadjah Mada, Yogyakarta, Indonesia
[3,4]Center for Energy Studies, Universitas Gadjah Mada, Yogyakarta, Indonesia
[5]Department of Economics, University of California, Los Angeles, USA

[1]tenyh@fti.untar.ac.id, [2]yudistira_hp@ugm.ac.id, [3]muhammadakmalfarouqi@gmail.com, [4]naufalmohamad@mail.ugm.ac.id, [5]raffysonata@ucla.edu, [6]marchel.535210039@stu.untar.ac.id, [7]irvanl@fti.untar.ac.id

Abstract

Food prices are important global issue and their relationship with fuel prices has become a main concern in society. An increase in the subsidized fuel price on 3 September 2022 has allegedly caused a rise in food (grocery) prices. This paper conducts an empirical study to analyze the relationships between food prices in Indonesia: rice, chicken, beef, egg, red chili, cayenne, shallot, garlic, cooking oil, and sugar. The study uses time series data of food prices from 1 January 2018 to 31 December 2023, which consists of food prices from 87 traditional markets in Indonesia. The commodity prices are obtained from online public data provided by Bank Indonesia. It divides the analysis (pre- and post-3 September 2022) to see how the relationship between food prices changes due to the increase in the subsidized fuel price. It performs the Peter Clark (PC) algorithm to generate causal graphs from real datasets where the true graphs are unknown. It complements the analysis by performing Vector Autoregression (VAR) to investigate the dynamic relationship between food prices, especially how the subsidized fuel price increase changes its dynamic relationship. The causal graphs from pre- and post-increasing fuel prices show the changes in the role of variable relationships, e.g., sugar and beef. The VAR results also show an interesting change in the IRF pattern. The results from both the PC algorithm and VAR show that there is a structural change in the relationship between food prices and that there is a different effect of price shock due to the subsidized fuel price increase. It might have been an indication of a change in the consumption pattern in society as a response to a food price increase. This must be a huge task to do in maintaining food prices when there is an adjustment in the subsidized fuel prices

Keywords: bayesian network; causal graph; food prices; PC algorithm; VAR

How to Cite: T. Handhayani, "A New Approach for Dynamic Analysis of Indonesian Food Prices using the PC Algorithm and Vector Autoregression", *J. RESTI (Rekayasa Sist. Teknol. Inf.)* , vol. 9, no. 6, pp. 1417 - 1427, Dec. 2025.
Permalink/DOI: https://doi.org/10.29207/resti.v9i6.6601

Received: April 28, 2025
Accepted: December 16, 2025
Available Online: December 30, 2025

This is an open-access article under the CC BY 4.0 License
Published by Ikatan Ahli Informatika Indonesia

1. Introduction

Machine learning has been widely adopted across various fields, bringing about significant advancements. Machine learning has been instrumental in enhancing the monitoring and optimizing processes in sectors like cement production, showcasing its effectiveness in addressing industrial challenges [1]. Machine learning has also emerged as a powerful tool in the field of econometrics, offering advancements over traditional econometric techniques in data processing, prediction, and regression analysis [2] . The application of machine learning in econometric research has gained significant popularity in recent years, driven by the need for more sophisticated analytical tools that can handle complex datasets and uncover intricate relationships among variables. Machine learning algorithms like random forests, support vector machines, and neural networks have been introduced into the econometric toolbox, expanding the range of analytical tools available for economic analysis and policymaking [3], [4]. Machine learning offers numerous benefits across various fields, including improved prediction accuracy, enhanced decision-making processes, and the ability to handle large and complex datasets efficiently[5]. Machine learning techniques, such as deep learning, have shown advancements over traditional econometric models in

tasks like forecasting, classification, and regression analysis [6].

Machine learning is not without limitations with one significant challenge being the potential bias present in machine learning models, which can lead to unfair outcomes and discriminatory decisions [7]. Biases can arise from various sources, including biased training data, algorithmic design, and human input, impacting the accuracy and reliability of machine learning predictions. Addressing bias in machine learning models is crucial to ensure fairness and equity in decision-making processes. While machine learning models excel in prediction accuracy and flexibility, they may lack interpretability and transparency compared to traditional econometric models [8]. Explainability remains a critical issue in machine learning, especially in complex models like neural networks, where understanding the reasoning behind predictions is challenging. This lack of transparency can hinder the establishment of causal inferences and limit the trustworthiness of machine learning outcomes.

Fuel subsidies have become one of the government's policies to assist poor households. A Reduction in fuel subsidies may affect economic sectors, i.e., micro-industrial and food prices. The Indonesian government also applies a fuel subsidy policy. Some research has been conducted to study the policy of fuel subsidies in this country [9], [10]. In Indonesia, the reduction in fuel subsidies is an interesting issue. Social media analysis found that a reduction in fuel subsidies brings negative sentiment from society [11], [12]. One of the issues is the concern for the affordability of the increase in food prices. An interesting problem in this research is how to investigate the effect of the reduction in fuel subsidies on the increase in food prices. This paper aims to analyze the dependence relationships among the variables of food prices pre- and post-reduction in fuel subsidies. This paper proposes to use a hybrid approach of machine learning and economic model. This paper implements the Peter-Clark (PC) algorithm and Vector Autoregression (VAR). The PC algorithm, known for its ability to infer causal relationships from observational data, can be utilized to identify potential causal links between variables in economic systems [13]. Meanwhile, VAR is a stochastic process model and is traditionally used for capturing linear interdependencies among multiple time series [14], [15]. One of the primary advantages of the VAR model is its flexibility in modeling multiple time series without requiring a priori assumptions about the relationships among the variables. This hybrid approach is addressed to improve the understanding and predictive capabilities of complex systems; and the limitations of traditional econometric approaches, such as endogeneity and multicollinearity. However, it is essential to consider the potential biases that may arise in machine learning models when integrating them with econometric frameworks. Bias in machine learning algorithms can impact the reliability and fairness of the results, potentially leading to erroneous conclusions in economic analysis [7].

It combines these two techniques (PC algorithm and VAR) to investigate the impact of a fuel price increase on food prices. Both techniques are complementary, with the PC algorithm providing us with a causal relationship of combined food prices, while VAR provides the (dynamic) long-run relationship between food prices. Like many countries, the increase in fuel prices is crucial in affecting food prices through increased production and transportation costs. The hypotheses lie in two concerns: first, there is a significant change in the long-run relationship between food prices due to the fuel price increase; second, the fuel price increase changes the variation effect of food prices. It takes a case of the subsidized fuel price increase in Indonesia (in September 2022) and its effect on the basic food prices. The contribution of this paper is the hybrid model of PC algorithm and VAR for analyzing food prices pre-and post-increasing fuel prices, a case study in Indonesia.

2. Methods

A research workflow is displayed in Figure 1. The main steps are data collection, data pre-processing, and analysis. The analysis consists of causal analysis and VAR analysis. The causal analysis consists of generating causal graphs using the PC algorithm and then analyzing the graph.



Figure 1. Research workflow

The PC algorithm is one of the examples of algorithms for structure learning Bayesian Networks. The PC algorithm is a causal learning method with two main steps: generating skeletons and orienting the edges [16] [17]. The first step is generating a completed undirected graph from a dataset. The PC algorithms run conditional independence tests to find the independence relationships among variables [18] [19]. The conditional independence test follows equation (1), where $n, S, \alpha, \hat{\rho}$ and Φ represent number of samples, separation set, significant level, partial correlation and cumulative distribution function of normal distribution respectively [18]. Equation 1 tests the dependence relationship between variables Z_u and Z_v given variable Z_S. Figure 2 shows the first step of the PC algorithm to generate a graph's skeleton.

$$Z_u \perp Z_v | Z_S \Leftrightarrow \sqrt{n - |S| - 3} \left\| \frac{1}{2} log\left(\frac{1+\hat{\rho}_{uv|S}}{1-\hat{\rho}_{uv|S}}\right) \right\| \leq \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (1)$$

The information from the conditional independent test will be used in the second step to orient the edges. Figure 3 illustrates how to orient the edges. The output of the PC algorithm is a graph represented by a completed partially directed acyclic graph (CPDAG). A graph generated by the PC algorithm represents causal

relationships among variables. This graph can be called a causal graph. A simple causal graph $A \rightarrow B$ implies that A is a cause of B and B is an effect of A. Variance of the PC algorithms has been developed to generate causal graphs from mixed data [16].

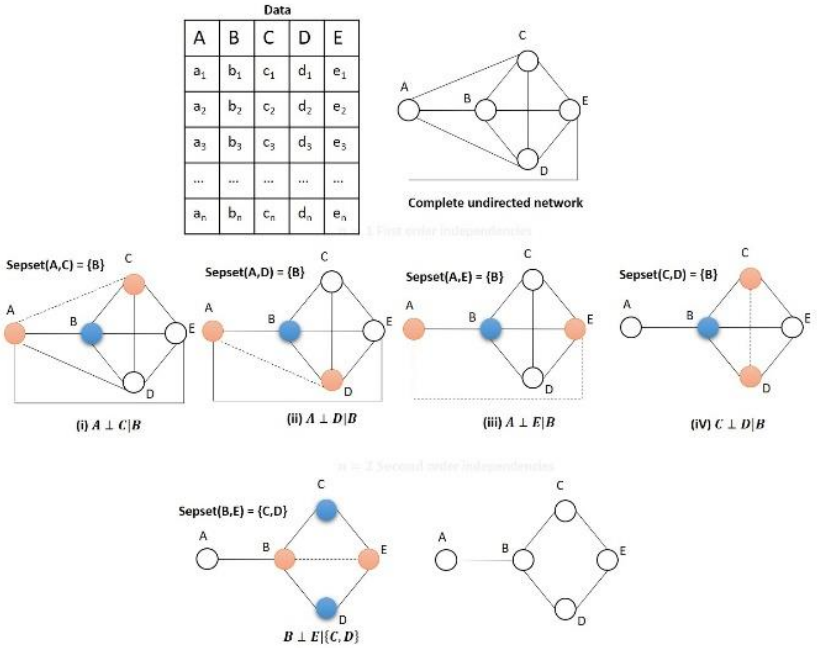


Figure 2. The first step of the PC algorithm: generating a skeleton

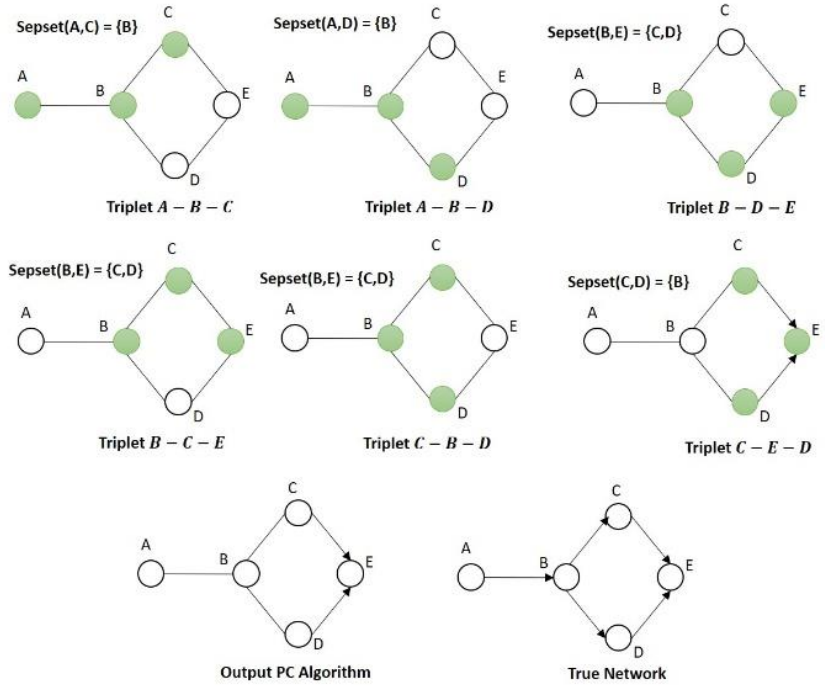


Figure 3. The second step of the PC algorithm: orienting the edges

In economics, the implementation of the PC algorithm has started to grow recently as this method is widely used to discover causal relationships among variables. For example, research done in the US found that corn cash prices in Iowa dominate crop pricing over the year [20] or find causal patterns on nearby futures, spot, and first-distant futures [21]. The modified PC algorithm has been used for learning the causal nexus between monetary policy and inflation [22].

It complements the analysis by performing the Vector Autoregression (VAR) which is popular for analyzing the dynamic relationship and long-run, equilibrium between time series variables by generalizing the univariate autoregression models [23]. VAR modeling requires several procedures as follows [24]: i) performing unit root test to check variable stationery at I(0), ii) specifying lags for model specification, iii) estimating the VAR model including impulse response

and variance decomposition, and iv) diagnostic check for the estimation results and error measurement.

Statistically check variables' stationarity, ii) investigate cointegration between non-stationary variables, iii) specify lag in VECM, and iv) check the goodness-of-fit of VECM. The Augmented Dickey-Fuller (ADF) [25], Phillips-Perron (PP) [26], and Zivot-Andrew (ZA) [27] tests are used to check variables' stationarity. The ADF test uses an autoregressive process which is given by Equation 2, where α is a constant, β is the coefficient on a time trend t, $i \in J$is the lag order of the autoregressive process, and ε is the innovation process following a zero-mean value. This ADF test is carried out under the null hypothesis of $\lambda = 0$ indicating the presence of unit root process in the variable tested; against the alternative hypothesis of $\lambda < 0$ given the asymmetrical setup in the critical value setup of ADF test. The PP test utilizes a non-parametric method of controlling for serial correlation in the variable tested. It gives a better unit root test when there exists a structural break in the process; in our case, the Coronavirus disease 2019 (COVID-19) pandemic has been present since early 2020. The PP test specifies an autoregressive process by Equation 3, where α is the drift coefficient, δ is the deterministic trend coefficient, and ε is the mean zero innovation process. This PP test takes a null hypothesis of unit root presence $\phi = 1$ against the alternative hypothesis of $\phi < 1$. The ZA test accommodates the existence of structural break in the dataset which may lead to potentially spurious results when using standard unit root tests. The ZA test is specified by Equation 4, where μ is the change in the intercept before and after the break, β is the trend slope before the break, α is the slope coefficients and are assumed to be constant, $DT_t(T_b)$ is a one-time break, b is the estimated trend and $\{\theta, c\}$are the estimated break parameters. The VAR for stationary variables the k dimensional VAR (p) process can be computed using Equation 5, where $\{\beta, \gamma\}$ are the estimated coefficients of lagged variables and u are the stochastic error terms, or called impulses or shocks in the VAR model. It takes a strategy by separating the analysis of pre- and post-fuel price increases in September 2022. This is to investigate if there is a significant change in the relationship between observed food prices due to fuel price increases.

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{j=1}^{J} \lambda_j \Delta y_{t-j} + \varepsilon_t \quad (2)$$

$$y_t = \alpha + \delta t + \phi y_{t-1} + \varepsilon_t \quad (3)$$

$$y_t = \mu + \beta t + \gamma DT_t(T_b) + \alpha y_{t-1} + \sum_{i=1}^{k} c_i \Delta y_{t-i} + \varepsilon_t \quad (4)$$

$$y_t = \alpha + \sum_{j=1}^{J} \beta_j y_{t-j} + \sum_{k=1}^{K} \sum_{j=1}^{J} x_{k,t-j} + u_{kt} \quad (5)$$

3. Results and Discussions

3.1 Dataset

This paper utilizes the PC algorithm to determine food commodity price causality in Indonesia namely rice, chicken, beef, egg, shallot, garlic, red chili, cayenne, cooking oil, and sugar. These commodities are collected in traditional markets across 87 major cities in each province in Indonesia. The commodity prices are obtained from online public data provided by Bank Indonesia. These commodities were sourced from https://www.bi.go.id/hargapangan/TabelHarga/PasarTradisionalDaerah. For context, commodity prices are the average offered price of a given traditional market, so it is not the transaction price because frequently there were negotiations between buyer and seller. The observation spanned from 1 January 2018 to 31 December 2023. Moreover, since we wanted to capture the different behaviors before and after the subsidized fuel price increase, we separated the period into two groups, before the subsidized fuel increase ranging from January 2018 to August 2022, and after the subsidized fuel increase, which occurred from September 2022 onwards.

3.2. Machine Learning Approach

The preprocessing steps consist of missing values handling and generating monthly average prices. The dataset is a monthly time series of food prices from January 2018 - December 2023. The dataset is divided into two sections based on the rise of fuel prices on 3 September 2023. The period before fuel price increases was from January 2018 to August 2022 and from September 2022 to December 2023. Based on the data pattern, it found that food prices in Indonesia are possibly different each month. However, the price pattern changes after the fuel price rises as represented in Figure 4. Some commodities that have major changes after the fuel price increases are rice, egg, beef, and cooking oil. On the other hand, commodities such as red chili, cayenne, shallot, and garlic–volatile price commodities, do not appear to have changed significantly. It might correlate with the harvesting period since those commodities are classified as seasonal agriculture. Thus, the increase in fuel prices likely does not affect the prices of these volatile commodities. Besides, chicken and sugar prices show a slightly different move during changes in fuel prices. It utilizes correlation analysis to observe the relationship for each commodity variable in the dataset. It applies the Pearson correlation to compute the correlation coefficient among variables. The correlation coefficients show that they perform all positive correlations. It means that the prices move together in the same direction. It takes a correlation coefficient value $\rho \geq 0.4$ that is strong enough for detailed analysis. Before the fuel price increases, eggs had many significant correlations with other commodities, such as chicken, shallot, red chili, cayenne, and sugar. However, after the fuel price increases, the price of rice and cooking oil are two variables that experience an increase in correlation coefficient with some variables raising fuel prices. The strongest correlation coefficient is between red chili and cayenne which is around 0.7. It is not surprising because red chili and cayenne are the main ingredients in most Indonesian cuisines and they can substitute each other as a spicy flavoring.

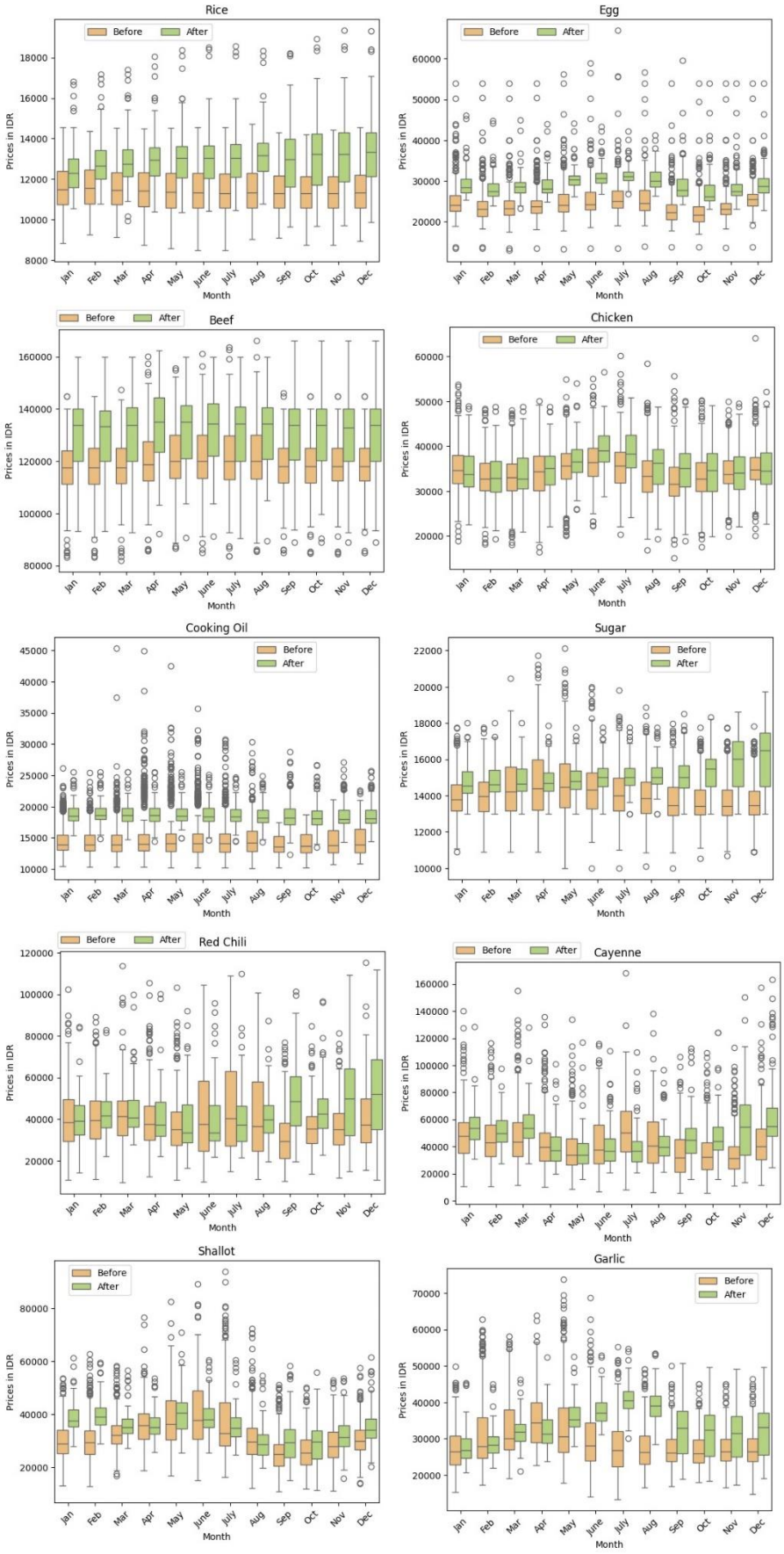


Figure 4. Monthly trend of food prices before and after increasing fuel prices.

Before increasing fuel prices, strong positive relationships happened between the prices of chicken meat - egg, egg - shallot, egg - red chili, egg - cayenne, egg - sugar, shallot - garlic, shallot - red chili, shallot - sugar, garlic - sugar, and red chili - cayenne. The strong positive relationships after increasing fuel prices consist of the prices of rice - garlic, rice - red chili, rice - sugar, chicken meat - egg, chicken meat - shallot, chicken meat - garlic, chicken meat - cooking oil, shallot - cooking oil, garlic - cooking oil, garlic - sugar, red chili - cayenne, red chili - sugar, and cayenne - sugar. The price of rice and cooking oil are two variables that experience increasing in correlation coefficient with some variables after rising fuel prices. The strongest correlation coefficient is between red chili and cayenne which is around 0.7. It is not surprising because red chili and cayenne are the main ingredients in most Indonesian cuisines and they can substitute each other as spicy flavoring. In this case, it cannot be claimed that this correlation is a causation. Thus, the data shown above does not mean that one commodity causes the others.

Causal learning inference is a machine learning approach to investigate the possible dependent relationships among variables. This paper implements the PC algorithm from the R package *pcalg*. The graphs are generated from dataset using the PC algorithm at the significance level $\alpha = 0.01$. Figure 5 shows a graph generated from a dataset before increasing the fuel prices using the PC algorithm at $\alpha = 0.01$.

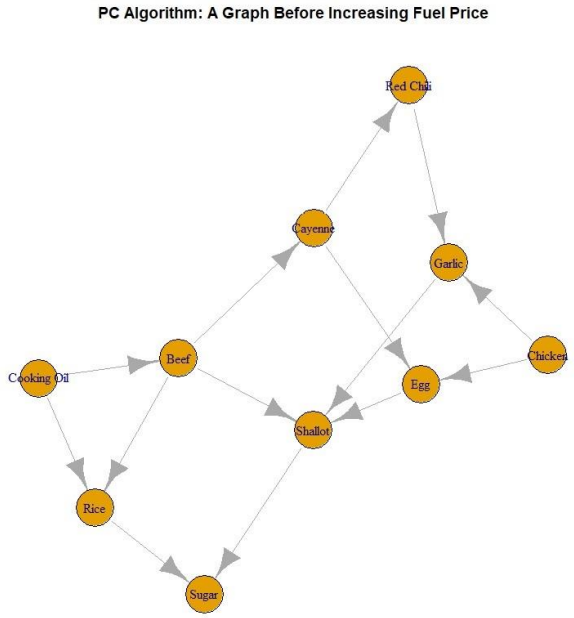


Figure 5. A causal graph was generated from a dataset pre- the rise of fuel prices

A learned graph generated using the PC algorithms from a dataset before rising fuel prices says that red chili prices influence garlic prices, garlic prices affect shallot prices, and shallot prices work on sugar prices. It also shows that cayenne price is a cause variable of red chili and egg price. Egg price is a cause variable of shallot price and rice price is a cause variable of sugar price. Beef prices affect shallot, cayenne, and rice prices. Cooking oil prices influence beef and rice prices.

Figure 6 displays a graph generated from the data after the increase of fuel prices using the PC algorithm. A learned graph shows red chili prices influence sugar and cayenne prices. Cayenne's price is also affected by sugar prices and beef prices. Cayenne's price is found to affect the egg price. Beef prices predispose cayenne and chicken prices. Chicken and shallot prices affect egg prices. Rice prices sway sugar prices. Garlic and cooking oil prices influence rice prices.

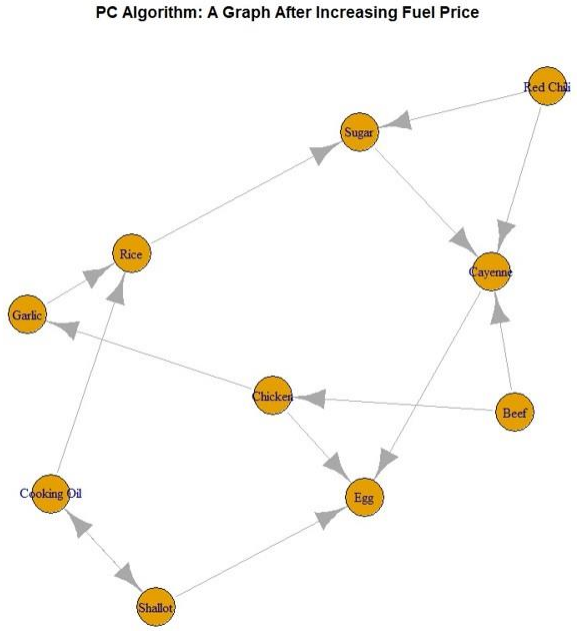


Figure 6. A causal graph was generated from a dataset post the rise of fuel prices

Comparing two causal graphs pre- and post-increasing fuel prices, sugar and beef have changed their influence. Before increasing fuel prices, sugar was an effect variable of rice and shallot but then changed into a cause variable of cayenne. Before increasing fuel prices, there were no direct dependent relationships between beef and chicken. However, after increasing fuel prices, beef becomes a cause variable for chicken. It shows that the reduction the fuel subsidies influences some food prices and possibly leads to changes the household consumption.

3.3 Vector Autoregression (VAR) Approach

The results using VAR can be explained as follows. Tables 1 and 2 show ADF, PP, and ZA tests before and after increasing fuel prices. In this analysis, it found that the optimal lag selection for the pre-period is 6 and the post-period is 2. Determining optimal lag selection would lead the VAR model to be parsimonious so the models are simple with great explanatory predictive power. We also conducted a stability test to check whether the models were stable. The IRF and FEVD results would not have a clear interpretation if the models are not stable because IRF and FEVD do not

reach equilibrium. After testing, both are stable using AR roots of Characteristic polynomials.

Table 1. ADF, PP and ZA tests pre increasing fuel price

Variables	ADF		PP		ZA
	Level I(0)	First-difference I(1)	Level I(0)	First-difference I(1)	Level
Beef	0.56	0.01	0.07	0.01	-4.51
Cayenne	0.01	0.01	0.06	0.01	-5.75
Chicken	0.01	0.01	0.01	0.01	-5.12
Cooking oil	0.01	0.01	0.92	0.01	-3.41
Egg	0.58	0.01	0.13	0.0	-5.45
Garlic	0.01	0.01	0.59	0.01	-4.15
Red chili	0.23	0.01	0.01	0.01	-4.63
Rice	0.03	0.01	0.95	0.01	-5
Shallot	0.03	0.01	0.68	0.01	-4.84
Sugar	0.01	0.01	0.92	0.01	-3.91

In Table 2, note that the data is stationary if ADF & PP is significant and ZA is not significant. Orthogonal impulse response and variance decomposition can be found in Appendix 1 and Appendix 2 (Supplementary File).

The first model estimate was presented using the Impulse Response Function (IRF) for each commodity in the data. The IRF is attached to the appendices where it is divided into two periods of time, which are pre- and post-subsidized fuel price increases in Indonesia. The pre-period showed that shock in beef and red chili has overall short-period effects on other commodities. In beef's IRF graph, cayenne and red chili have the longest duration shock effect compared to other commodities. It indicates that cayenne and red chili, in this regard, have a complementary relationship with each other where people usually use beef, red chili, and cayenne as combined food. The same goes for red chili and cayenne graphs. Yet, these two commodities have no huge effect on each other. Cayenne's graph exhibited an effect on chicken, while it is not shown in red chili. Thus, red chili and cayenne have a mutual relationship with beef, and cayenne has the same pattern as chicken price. On the other hand, rice price has the longest effect on other commodity prices. Rice as a primary food for most Indonesians showed as a driver for other commodity prices. It is also supported by Statistics Indonesia that rice contributed 5 percent (on average) to total expenditure per capita in Indonesia.

Table 2. ADF, PP and ZA tests post increasing fuel price

Variables	ADF		PP		ZA
	Level I(0)	First-difference I(1)	Level I(0)	First-difference I(1)	Level
Beef	0.43	0.01	0.29	0.01	-5.4
Cayenne	0.58	0.01	0.83	0.01	-4.22
Chicken	0.47	0.01	0.8	0.01	-3.19
Cooking oil	0.5	0.01	0.95	0.01	-5.19
Egg	0.33	0.01	0.88	0.01	-3.9
Garlic	0.75	0.01	0.94	0.01	-3.17
Red chili	0.46	0.01	0.72	0.01	-3.51
Rice	0.52	0.01	0.95	0.01	-3.05
Shallot	0.34	0.27	0.96	0.01	-2.97
Sugar	0.74	0.02	0.99	0.01	-3.5

However, post-period time showed an interesting change in the IRF pattern. The rise of subsidized fuel prices has smoothened the IRF pattern in almost all commodities in the data. The shock effect has been shortened since the subsidized fuel price rose. This might indicate that the consumption patterns of households also changed. In brief, fuel shock on food commodities has a huge effect on their prices and leads to changes in household consumption patterns in Indonesia.

The results of variance decomposition for each commodity observed using the VAR approach are used to determine the relationship between these food commodities and others. We divided the analysis into two periods of time, which are pre and post-period. This period is based on the time when the subsidized fuel price increases during September 2022. The variance decomposition would be interpreted as a change in other commodities' influences on the commodity observed.

Table 3 and 4 show the forecast error of Variance Decomposition (FEVD) for commodities during pre- and post-fuel price increases. In the pre-period, almost all commodities were affected by the condition of their market, meaning that other commodities had a lower influence on the commodity's price. However, an interesting result was shown in the price of red chili. The red chili market only contributed less than 50% of its prices, while other commodities had, at least, more than 70%. In this regard, red chili's prices were influenced by beef (around 30%-32%) and cayenne (around 18%-19%).

(1) = d(log(beef)), (2) = d(log(cayenne)), (3) = d(log(chicken)), (4) = d(log(cooking oil)), (5) = d(log(egg)) , (6) = d(log(garlic)), (7) = d(log(red chili)), (8) = d(log(rice)), (9) = d(log(shallot)), (10) = d(log(sugar))

Table 3. Forecast Error of Variance Decomposition (FEVD) for Commodities during Pre-Fuel Price

					d(log(beef))					
Horizon	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	100%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
10	91.99%	0.87%	0.66%	2.1%	0.53%	0.37%	2.47%	0.41%	0.19%	0.32%
20	91.89%	0.88%	0.70%	2.18%	0.58%	0.39%	2.47%	0.41%	0.19%	0.32%
30	91.89%	0.88%	0.70%	2.18%	0.58%	0.39%	2.47%	0.41%	0.19%	0.32%
					d(log(cayenne))					
Horizon	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	5.27%	94.73%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
10	8.84%	87.51%	0.86%	1.14%	0.50%	0.65%	0.42%	0.26%	0.59%	0.23%
20	8.58%	86.68%	1.15%	1.18%	0.71%	0.98%	0.44%	0.27%	0.73%	0.28%
30	8.56%	86.63%	1.16%	1.18%	0.71%	1.01%	0.45%	0.27%	0.73%	0.30%
					d(log(chicken))					
Horizon	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	13.8%	1.50%	84.64%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
10	13.60%	1.9%	81.11%	0.31%	0.63%	0.23%	0.59%	0.42%	0.52%	0.63%
20	13.43%	2.17%	79.85%	0.40%	0.65%	0.23%	0.60%	0.79%	1.02%	0.86%
30	13.37%	2.21%	79.47%	0.41%	0.72%	0.23%	0.60%	0.89%	1.22%	0.87%
					d(log(cooking oil))					
Horizon	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	0.05%	0.14%	0.05%	99.76%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
10	0.21%	0.34%	0.20%	94.33%	0.33%	1.28%	0.19%	1.41%	0.45%	1.27%
20	0.23%	0.35%	0.21%	93.16%	0.37%	2.01%	0.21%	1.66%	0.55%	1.25%
30	0.23%	0.35%	0.21%	93.06%	0.37%	2.04%	0.22%	1.71%	0.55%	1.26%
					d(log(egg))					
Horizon	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	1.79%	1.98%	1.54%	0.0%	94.68%	0.0%	0.0%	0.0%	0.0%	0.00%
10	1.31%	3.57%	1.72%	0.38%	89.86%	1.12%	0.13%	0.94%	0.9%	0.07%
20	1.32%	4.17%	1.69%	0.54%	86.41%	1.22%	0.19%	1.89%	2.4%	0.12%
30	1.31%	4.28%	1.67%	0.59%	85.40%	1.22%	0.19%	2.17%	3.07%	0.12%
					d(log(garlic))					
Horizon	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	0.66%	0.06%	0.21%	0.37%	0.74%	97.97%	0.00%	0.00%	0.00%	0.00%
10	1.00%	0.49%	0.55%	0.56%	0.83%	95.02%	0.39%	0.15%	0.7%	0.2%
20	0.99%	0.60%	0.69%	0.57%	1.34%	93.26%	0.59%	0.22%	1.47%	0.28%
30	0.99%	0.62%	0.69%	0.57%	1.44%	93.02%	0.59%	0.22%	1.58%	0.28%
					d(log(red chili))					
Horizon	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	32.11%	18.85%	0.31%	0.05%	0.04%	0.49%	48.14%	0.00%	0.00%	0.00%
10	30.77%	19.41%	0.78%	0.66%	0.72%	1.61%	44.57%	0.48%	0.32%	0.68%
20	30.25%	19.58%	0.83%	0.77%	1.04%	2.22%	43.66%	0.48%	0.33%	0.83%
30	30.21%	19.59%	0.83%	0.79%	1.05%	2.25%	43.62%	0.49%	0.34%	0.84%
					d(log(rice))					
Horizon	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	0.00%	0.05%	0.77%	3.98%	0.26%	0.31%	0.80%	93.82%	0.00%	0.00%
10	0.39%	0.42%	0.85%	3.89%	0.85%	0.59%	0.88%	90.64%	1.01%	0.47%
20	0.46%	0.47%	0.82%	4.51%	1.20%	0.60%	0.85%	88.98%	1.63%	0.48%
30	0.48%	0.46%	0.81%	4.68%	1.35%	0.60%	0.85%	88.45%	1.86%	0.47%
					d(log(shallot))					
Horizon	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	2.78%	0.69%	0.29%	0.66%	0.89%	1.54%	1.71%	2.41%	89.03%	0.00%
10	2.27%	2.70%	0.52%	0.58%	2.89%	1.38%	1.78%	1.96%	85.45%	0.46%
20	2.10%	3.87%	0.47%	0.55%	5.19%	1.60%	1.58%	2.05%	81.87%	0.72%
30	2.08%	4.14%	0.46%	0.61%	5.70%	1.62%	1.56%	2.15%	80.71%	0.97%
					d(log(sugar))					
Horizon	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	0.01%	1.66%	1.52%	6.83%	0.05%	0.26%	0.47%	7.82%	3.88%	77.49%
10	0.37%	1.42%	1.32%	6.42%	0.13%	1.16%	1.04%	6.28%	3.01%	78.86%
20	0.47%	1.34%	1.23%	6.72%	0.16%	1.83%	1.05%	5.86%	2.86%	78.49%
30	0.49%	1.35%	1.22%	6.83%	0.17%	2.02%	1.03%	5.78%	2.87%	78.24%

This result demonstrated that both commodities markets had a higher influence on red chili's price compared to other commodities markets. This would indicate that red chili, beef, and cayenne had a complementary good relationship. On the other hand, the result from the pre-period has no huge difference from the post-period. Red chili stayed as the commodity and the market only drove out half of its price (around 50%). Yet, its market control increased to 50%-51% compared to the pre-period (43%-48%). This could be interpreted as red chili's market dependencies on beef and cayenne decreased in the period observed, meaning that the increase in subsidized fuel price gradually

phased out its hinges on other commodities although it was not significant.

(1) = d(log(beef)), (2) = d(log(cayenne)), (3) = d(log(chicken)), (4) = d(log(cooking oil)), (5) = d(log(egg)) , (6) = d(log(garlic)), (7) = d(log(red chili)), (8) = d(log(rice)), (9) = d(log(shallot)), (10) = d(log(sugar))

Table 4. Forecast Error of Variance Decomposition (FEVD) for Commodities during Post Fuel

					d(log(beef))					
Horizon	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	100%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
10	90.10%	0.80%	0.31%	3.60%	0.45%	0.34%	3.37%	0.84%	0.09%	0.12%
20	90.05%	0.80%	0.31%	3.59%	0.47%	0.35%	3.37%	0.86%	0.09%	0.12%
30	90.05%	0.80%	0.31%	3.59%	0.47%	0.35%	3.37%	0.86%	0.09%	0.12%
					d(log(cayenne))					
Horizon	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	3.21%	96.79%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
10	4.70%	89.99%	0.72%	1.25%	0.30%	0.48%	0.26%	0.16%	1.07%	1.06%
20	4.69%	89.89%	0.73%	1.25%	0.33%	0.49%	0.27%	0.17%	1.09%	110%
30	4.69%	89.89%	0.73%	1.25%	0.33%	0.49%	0.27%	0.17%	1.09%	1.10%
					d(log(chicken))					
Horizon	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	10.61%	3.52%	85.87%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
10	10.46%	3.24%	80.98%	0.31%	0.27%	0.28%	1.80%	0.60%	1.95%	0.12%
20	10.45%	3.24%	80.92%	0.31%	0.27%	0.29%	1.81%	0.61%	1.98%	0.12%
30	10.45%	3.24%	80.92%	0.31%	0.27%	0.29%	1.81%	0.61%	1.98%	0.12%
					d(log(cooking oil))					
Horizon	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	0.09%	0.82%	0.02%	99.08%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
10	0.34%	5.14%	0.09%	79.88%	4.73%	3.36%	0.32%	3.44%	0.32%	2.38%
20	0.34%	5.12%	0.09%	79.53%	4.94%	3.39%	0.32%	3.57%	0.32%	2.38%
30	0.34%	5.12%	0.09%	79.53%	4.94%	3.39%	0.32%	3.57%	0.32%	2.38%
					d(log(egg))					
Horizon	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	1.74%	2.71%	2.24%	0.83%	92.48%	0.00%	0.00%	0.00%	0.00%	0.00%
10	1.61%	2.52%	3.81%	0.84%	84.96%	1.91%	0.76%	2.24%	1.06%	0.28%
20	1.61%	2.51%	3.81%	0.85%	84.78%	1.93%	0.75%	2.34%	1.10%	0.30%
30	1.61%	2.51%	3.81%	0.85%	84.78%	1.93%	0.75%	2.34%	1.10%	0.32%
					d(log(garlic))					
Horizon	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	0.27%	0.01%	1.12%	0.32%	0.00%	98.27%	0.00%	0.00%	0.00%	0.00%
10	0.28%	1.72%	1.75%	0.63%	0.05%	93.85%	0.40%	0.29%	0.38%	0.65%
20	0.28%	1.72%	1.76%	0.63%	0.06%	93.72%	0.40%	0.29%	0.38%	0.65%
30	0.28%	1.72%	1.76%	0.63%	0.06%	93.72%	0.40%	0.33%	0.38%	0.70%
					d(log(red chili))					
Horizon	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	16.02%	31.71%	0.05%	0.00%	0.28%	0.02%	51.92%	0.00%	0.00%	0.00%
10	15.09%	31.26%	0.79%	0.98%	0.42%	0.31%	50.83%	0.81%	0.28%	0.24%
20	15.09%	31.24%	0.80%	0.98%	0.43%	0.31%	50.81%	0.82%	0.29%	0.25%
30	15.09%	31.24%	0.80%	0.98%	0.43%	0.31%	50.81%	0.82%	0.29%	0.25%
					d(log(rice))					
Horizon	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	0.01%	1.16%	0.35%	0.20%	0.38%	0.12%	0.06%	97.72%	0.00%	0.00%
10	0.28%	1.60%	0.64%	1.10%	1.79%	0.30%	1.33%	92.64%	0.12%	0.19%
20	0.28%	1.61%	0.65%	1.11%	1.80%	0.30%	1.33%	92.60%	0.13%	0.19%
30	0.28%	1.61%	0.65%	1.11%	1.80%	0.30%	1.33%	92.60%	0.13%	0.19%
					d(log(shallot))					
Horizon	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	0.79%	0.76%	8.84%	1.31%	0.24%	7.23%	1.59%	0.57%	78.67%	0.00%
10	0.73%	0.56%	7.89%	1.32%	0.88%	5.58%	2.74%	3.18%	76.14%	0.98%
20	0.73%	0.56%	7.86%	1.31%	0.90%	5.55%	2.74%	3.31%	75.92%	1.12%
30	0.73%	0.56%	7.86%	1.31%	0.90%	5.55%	2.74%	3.31%	75.92%	1.12%
					d(log(sugar))					
Horizon	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	0.15%	2.23%	0.06%	0.11%	0.13%	0.76%	0.02%	0.36%	0.03%	96.15%
10	0.14%	1.77%	0.09%	0.16%	0.30%	0.56%	1.19%	0.43%	0.05%	95.31%
20	0.14%	1.77%	0.09%	0.16%	0.30%	0.56%	1.23%	0.43%	0.06%	95.25%
30	0.14%	1.77%	0.09%	0.16%	0.30%	0.56%	1.23%	0.43%	0.06%	95.25%

This result demonstrated that both commodities markets had a higher influence on red chili's price compared to other commodities markets. This would indicate that red chili, beef, and cayenne had a complementary good relationship. On the other hand, the result from the pre-period has no huge difference from the post-period. Red chili stayed as the commodity and the market only drove out half of its price (around 50%). Yet, its market control increased to 50%-51% compared to the pre-period (43%-48%). This could be

interpreted as red chili's market dependencies on beef and cayenne decreased in the period observed, meaning that the increase in subsidized fuel price gradually phased out its hinges on other commodities although it was not significant.

4. Conclusions

This study investigates the change in the long-run relationship between food prices in Indonesia and how the subsidized fuel price increase (in September 2022) may change it. It combines the PC algorithm and the VAR model for estimation purposes, which helps depict the causal relationship and dynamic relationship between food prices in pre- and post-fuel price increases. The results show that there is a structural change in the relationship between food prices, indicating a change in the market structure for basic groceries. In addition, this may indicate a change in the long-run household consumption as shown by IRF and variance decomposition analysis in VAR estimation. This method, by combining the PC algorithm and VAR, also provides a comprehensive approach to give a signal to the government and market regarding policy intervention needed in the future, especially when transportation costs are substantial, to determine the food prices.

The mixture method, in this context, should widen the realm of methodology used to determine the relationship between macroeconomic variables, particularly when it comes to studying shocks in the market. However, one should note the possible extended model and practicalities using different contexts and variables. *Firstly*, the use of Bayesian VAR (BVAR) with a non-linear setup may accommodate complex dynamic interactions, interdependent uncertainty and limited data; or make the estimation in the panel data setup at the provincial level. *Secondly*, one can use the consumption quantity of essential groceries to delve deeper into the change in consumption patterns and nutritional status following the fuel price increases. Overall, this typical study opens room for the development of dynamic relationship modelling and could benefit the government in providing a mitigation plan for policymakers to address the complex side effects of a policy.

Acknowledgements

This paper is supported by a research grant from Universitas Tarumanagara No: 014-Int-BGRA-PTNBH-KLPPM/UNTAR/XI/2023.

References

[1] H. Zermane and A. Drardja, "Development of an efficient cement production monitoring system based on the improved random forest algorithm," *The International Journal of Advanced Manufacturing Technology*, vol. 120, no. 3–4, pp. 1853–1866, May 2022, doi: 10.1007/s00170-022-08884-z.

[2] I. Shah, N. Gul, S. Ali, and H. Houmani, "Short-Term Hourly Ozone Concentration Forecasting Using Functional Data Approach," *Econometrics*, vol. 12, no. 2, p. 12, May 2024, doi: 10.3390/econometrics12020012.

[3] Y. Zhao, W. Liu, Y. Yan, and F. Li, "Navigating the confluence of econometrics and data science: Implications for economic analysis and policy," *Theoretical and Natural Science*, vol. 38, no. 1, pp. 26–31, Jun. 2024, doi: 10.54254/2753-8818/38/20240551.

[4] T. Handhayani, I. Lewenusa, and M. Y. R. Arpipi, "Forecasting Volatile Fresh Chili Prices In Indonesia Using Support Vector Regression," in *2024 International Conference on Information Technology Research and Innovation (ICITRI)*, IEEE, Sep. 2024, pp. 281–286. doi: 10.1109/ICITRI62858.2024.10699118.

[5] M. Aghdaee *et al.*, "An examination of machine learning to map non-preference based patient reported outcome measures to health state utility values," *Health Econ*, vol. 31, no. 8, pp. 1525–1557, Aug. 2022, doi: 10.1002/hec.4503.

[6] S. F. Lehrer, T. Xie, and G. Yi, "Do the Hype of the Benefits from Using New Data Science Tools Extend to Forecasting Extremely Volatile Assets?," in *Data Science for Economics and Finance*, Cham: Springer International Publishing, 2021, pp. 287–330. doi: 10.1007/978-3-030-66891-4_13.

[7] M. Yang, E. McFowland, G. Burtch, and G. Adomavicius, "Achieving Reliable Causal Inference with Data-Mined Variables: A Random Forest Approach to the Measurement Error Problem," *INFORMS Journal on Data Science*, vol. 1, no. 2, pp. 138–155, Oct. 2022, doi: 10.1287/ijds.2022.0019.

[8] A. G. F. Hoepner, D. McMillan, A. Vivian, and C. Wese Simen, "Significance, relevance and explainability in the machine learning age: an econometrics and financial data science perspective," *The European Journal of Finance*, vol. 27, no. 1–2, pp. 1–7, Jan. 2021, doi: 10.1080/1351847X.2020.1847725.

[9] M. R. Jazuli, I. Steenmans, and Y. Mulugetta, "Navigating policy dilemmas in fuel-subsidy reductions: learning from Indonesia's experiences," *Sustainability: Science, Practice and Policy*, vol. 17, no. 1, pp. 391–403, Jan. 2021, doi: 10.1080/15487733.2021.2002024.

[10] M. Ichsan, M. Lockwood, and M. Ramadhani, "National oil companies and fossil fuel subsidy regimes in transition: The case of Indonesia," *Extr Ind Soc*, vol. 11, p. 101104, Sep. 2022, doi: 10.1016/j.exis.2022.101104.

[11] N. Patria, B. Irawanto, and A. N. Abrar, "'(Don't) Stop the Rising Oil Price': Mediatization, Digital Discourse, and Fuel Price Controversies in Indonesian Online Media," *Journalism and Media*, vol. 6, no. 3, p. 124, Aug. 2025, doi: 10.3390/journalmedia6030124.

[12] T. Handhayani, M. Y. R. Arpipi, and J. Hendryli, "An Evaluation of the Performance Deep Learning Methods for Sentiment Analysis from Imbalanced Data," in *2025 IEEE International Conference on Artificial Intelligence and Mechatronics Systems (AIMS)*, IEEE, May 2025, pp. 1–5. doi: 10.1109/AIMS66189.2025.11229639.

[13] M. Huber, "An introduction to causal discovery," *Swiss J Econ Stat*, vol. 160, no. 1, p. 14, Oct. 2024, doi: 10.1186/s41937-024-00131-4.

[14] P. S. Hou, L. M. Fadzil, S. Manickam, and M. A. Al-Shareeda, "Vector Autoregression Model-Based Forecasting of Reference Evapotranspiration in Malaysia," *Sustainability*, vol. 15, no. 4, p. 3675, Feb. 2023, doi: 10.3390/su15043675.

[15] R. Nsor-Ambala and E. Bugri Anarfo, "A vector autoregression (VAR) analysis of corruption, economic growth, and foreign direct investment in Ghana," *Cogent Economics & Finance*, vol. 10, no. 1, Dec. 2022, doi: 10.1080/23322039.2022.2146631.

[16] T. Handhayani and J. Cussens, "Kernel-based Approach for Learning Causal Graphs from Mixed Data," in *Proceedings of the 10th International Conference on Probabilistic Graphical Models*, Skørping: PMLR, Sep. 2020, pp. 221–232.

[17] T. Handhayani and D. K. Caglar, "An Analysis of Dependence Relationships Between Covid-19 Pandemic

and Environmental Conditions Cases Study in Indonesia, England, and Türkiye," in *2024 7th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, IEEE, Dec. 2024, pp. 1–6. doi: 10.1109/ISRITI64779.2024.10963452.

[18] T. Handhayani and J. Pragantha, "Learning Dependence Relationships of Air Pollution and Meteorological Conditions in Yogyakarta Using Bayesian Network," in *2024 International Conference on Electrical Engineering and Computer Science (ICECOS)*, IEEE, Sep. 2024, pp. 262–267. doi: 10.1109/ICECOS63900.2024.10791241.

[19] N. N. Qomariyah, T. Handhayani, S. D. A. Asri, and D. Kazakov, "Exploring COVID-19 Vital Signs using Bayesian Network Causal Graph," in *2024 2nd International Conference on Software Engineering and Information Technology (ICoSEIT)*, IEEE, Feb. 2024, pp. 37–41. doi: 10.1109/ICoSEIT60086.2024.10497468.

[20] X. Xu, "Contemporaneous causal orderings of US corn cash prices through directed acyclic graphs," *Empir Econ*, vol. 52, no. 2, pp. 731–758, Mar. 2017, doi: 10.1007/s00181-016-1094-4.

[21] X. Xu, "Contemporaneous Causal Orderings of CSI300 and Futures Prices through Directed Acyclic Graphs," *Economics Bulletin*, vol. 39, no. 3, pp. 2052–2077, 2019.

[22] Rizwan Fazal, Md. Shabbir Alam, Umar Hayat, and Naushad Alam, "Effectiveness of monetary policy: Application of modified Peter and Clark (PC) algorithms under Graph-Theoretic Approach," *Scientific Annals of Economics and Business*, vol. 68, no. 3, pp. 333–344, Sep. 2021, doi: 10.47743/saeb-2021-0019.

[23] E. Zivot and J. Wang, "Vector Autoregressive Models for Multivariate Time Series," in *Modeling Financial Time Series with S-Plus®*, New York, NY: Springer New York, 2003, pp. 369–413. doi: 10.1007/978-0-387-21763-5_11.

[24] D. N. Gujarati, D. C. Porter, and M. Pal, *Basic Econometrics*, 6th ed. MC GRAW HILL India, 2019.

[25] R. N. Putri, M. Usman, Warsono, Widiarti, and E. Virginia, "Modeling Autoregressive Integrated Moving Average (ARIMA) and Forecasting of PT Unilever Indonesia Tbk Share Prices During the COVID-19 Pandemic Period," *J Phys Conf Ser*, vol. 1751, no. 1, p. 012027, Jan. 2021, doi: 10.1088/1742-6596/1751/1/012027.

[26] N. Haldrup and M. Jansson, "Improving Size and Power in the Unit Root Testing," *SSRN Electronic Journal*, 2005, doi: 10.2139/ssrn.1147594.

[27] O. Osabuohien-Irabor, "Unit root tests in the presence of structural breaks: Evidence from African stock markets," *Economic Journal of Emerging Markets*, vol. 12, no. 2, pp. 119–137, Oct. 2020, doi: 10.20885/ejem.vol12.iss2.art1.