Indian Journal of Agricultural Economics 79: 3 (2024):394-406
DOI:10.63040/25827510.2024.03.005

Forecasting of Retail Price Rise in Tomato, Onion and Potato Commodities in India

G.K. Vani, A. Verma, D. Singh and A. Shrivastav*

ABSTRACT

The study analyzed the price trends of tomato, onion, and potato commodities in India from January 2014 to April 2024, using Consumer Price Index (CPI) and Wholesale Price Index (WPI) data. It applied autoregressive models to forecast the CPI sub-indices for these three commodities. Results indicated that tomato had the highest price rise but the lowest instability, while potato exhibited the highest instability. Seasonality was vital, with onion prices peaking in November, tomato in July, and potato in November. The model achieved high forecasting accuracy, with Mean Absolute Percentage Error (MAPE) values of 4.58 per cent, 2.02 per cent, and 1.30 per cent for tomato, onion, and potato, respectively. The study also identified key determinants of price volatility, emphasizing the need for seasonal agricultural planning at macro, meso, and micro levels. Particular attention should be paid to high-price months like November and July. The study recommends strengthening market intelligence efforts to ensure stakeholders have equal access to current and forecasted market prices. Policymakers are urged to closely monitor CPI sub-indices and develop price stabilization strategies.

Keywords: CPI, WPI, autoregression, growth, instability

JEL codes: Q02, Q11, Q13

I

INTRODUCTION

The Consumer Price Index (CPI) is the most commonly used index to measure the cost of living worldwide. The CPI in India has increased from 5.7 percent in 2022 to 6.4 percent in 2023 (RBI Annual Report, 2024), marking a 0.7 percent gain with a growth of 0.12 percent per annum**.** Several researchers have studied the index to understand the factors affecting it. They have identified food & beverages, fuels & lights, and housing as the major contributors to this index (Ministry of Statistics and Programme Implementation, GOI). Among these factors, agriculture had a large share due to a considerable weightage assigned to it. Within agriculture, cereals, milk & milk products, and vegetables are major articles with weightage of 9.67 percent, 6.61 percent, and 6.04 percent in the CPI, respectively. Within the Wholesale Price Index (WPI), these items have weights of 2.83 percent, 4.43 percent, and 1.87 percent, respectively (Official website of Office of Economic Advisor, GOI). Research on the CPI and WPI has revealed that due to their high price volatility, tomato, onion, and potato have contributed significantly to changes in both CPI and WPI. These commodities occupy a sensitive portion of consumers' plates and are major vegetable staples in the Indian diet. In the past, Tiwari *et al.* (2021) noted that onion had higher

---

*Department of Agricultural Economics and Farm Management, JNKVV, Jabalpur.

price volatility among the TOP commodities than potato and tomato. These commodities contributed 51 percent of total horticultural production in India during 2020-21, with potato contributing the most, followed by onion and tomato. The bulk of the TOP output in India is contributed by Madhya Pradesh, Maharashtra, and Uttar Pradesh, respectively, during 2020-21 (Agri-exchange, APEDA). Researchers have tried to forecast the price of tomato, onion, and potato. Birthal *et al.* (2022) forecasted food inflation using the CPI of TOP crops, as these crops play a major role in contributing to the volatility of food inflation. The quantification of news articles and various communications, such as audio and video, related to TOP crops and their price fluctuations was done via natural language processing (NLP) to create sentiment scores. This provided information about future prices and thus helped forecast food inflation. Ajmal *et al.* (2024) analyzed the wholesale price index of TOP crops to understand the dynamics and volatility of their prices from 2005 to 2021. The results showed that West Bengal experienced significant price fluctuations in TOP crops, where tomato prices peaked from July to November, onion prices were the highest from August to January, and potato prices saw their peak from July to December. Additionally, states like Rajasthan, Karnataka, and Punjab experienced the highest monthly price instability for TOP crops.

The government has always been keen to keep the price of these commodities in check, given the political repercussions of price rises in these commodities. Hence, it is pertinent to study the dynamics of changes in the CPI sub-index of these commodities and forecast the same. This study aimed to explore the variation in the CPI sub-index of TOP commodities. The remaining sections of the paper are divided into three parts. Section II includes data and methodology, Section III deals with the analysis of data and the discussion of results, and Section IV discusses the conclusion and policy implications for managing TOP commodity price rises.

II

DATA AND METHODOLOGY

The study was solely based on secondary data from January 2014 to April 2024 on the sub-index on the TOP wholesale price index collected from the official website of the Office of Economic Advisor, GOI, and the combined CPI collected from the Ministry of Statistics and Program Implementation, GOI. These indices had a common base year of 2012. The changes in the economic environment post-2013 dictated the choice of the study period. By the end of 2014, India had entered a soft inflation period compared to the running inflation during 2011-14.

The data was divided into training data and testing data. The training data pertained to January 2014 to April 2023, and the testing data was from May 2023 to April 2024. Gretl software was used to forecast rises in tomato, onion, and potato prices. The autoregressive model was employed to forecast the consumer price sub-index of these three commodities using the OLS. Lagged values of wholesale and consumer price sub-index of three commodities were used as regressors in addition to several lags of dependent variables.

$$\ln Y = \alpha + \ln X \beta + e$$

Where Y is a vector of dependent variable values, α is the intercept, β is a vector of partial regression coefficients, and X is a matrix of values of the regressors. All variables in the regression model were taken in log form because log transformation of variables can better capture the non-linearity in the process. It is also appropriate since the geometric mean is more appropriate than the arithmetic mean for index number aggregation.

Since this is a time-series data, the problem of autocorrelation can arise. The Breush-Godfrey (BG) test was used to detect it. This test runs an auxiliary regression of estimated residuals on its lagged values and original regressors in the model. The test has the null hypothesis of no autocorrelation. Also, the residuals in a time-series regression must be white noise. To detect this, autocorrelation and partial autocorrelation functions (ACF & PACF) of residuals were plotted. If, for all chosen lags, neither autocorrelation nor partial autocorrelation is outside the confidence interval, then it is proven that estimated residuals are white noise. Time-series regression can suffer from heteroscedasticity, and White's test was used with the null hypothesis of homoscedasticity. White's test also uses auxiliary regression of squared estimated residuals upon original regressors in the model. The choice of this test was dictated by its non-reliance on the normality assumption for residuals and ease of implementation. An ARCH test was conducted to detect autoregressive conditional heteroscedasticity (ARCH) in residuals of time-series regression. The ARCH test runs a regression of squared estimated residuals upon its several lags. The null hypothesis of the ARCH test is that there is no ARCH effect. Time-series regression can suffer from parameter instability. The CUSUM test was used with the null hypothesis of no parameter instability to detect this.

The accuracy of the estimated model was assessed using the mean absolute percentage error (MAPE) value and Theil's U2. These measures are preferred over RMSE and mean percentage error due to their compatibility across different commodities and periods; a lower Theil's U2 and MAPE value implies greater accuracy (Gohain, 2021). Theil's U2 measures relative accuracy by comparing the anticipated results of the naïve model against the results of the given forecasting method (Jadhav *et al.*, 2017). The thumb rule for Theil's U2 is as follows:

- If Theils-U< 1: The forecast method is better than the naive method.
- If Theils-U> 1: The naive method is better than the forecast method.
- If Theils-U=1: Both the Naive and forecast methods are suitable for forecasting.

The Growth rate was calculated using the average annual growth rate (AAGR), the average yearly growth in the variable over its previous year. Instability was calculated using Sen's method and was measured as 100 times the standard deviation of the natural logarithm of the ratio of current value over its last value (in %).

III

RESULTS AND DISCUSSION

Table 1 provides summary statistics for the CPI and WPI sub-index for TOP. Among all indices, the CPI of onion and WPI of tomato had the highest mean value, respectively. The WPI for tomato had the highest coefficient of variation and range values between January 2014 and April 2024. Lower mean and coefficient of variation (CV) values of the CPI for potato and tomato also indicate that not all variation in wholesale prices transmits to retail prices of these two commodities. This contradicts the assumption that retail prices had higher inflation than wholesale prices. Among the CPI sub-index of three commodities, the potato had the lowest value for CV, which can be confirmed by Figure 1.

TABLE 1. SUMMARY STATISTICS OF CPI AND WPI SUB-INDEX FOR TOMATO, ONION AND POTATO

Variable (1)	Mean (2)	Median (3)	S.D. (4)	C.V. (%) (5)	Min (6)	Max (7)
CPI_Onion	208	187	81.7	39.28	116.0	617
CPI_Tomato	161	143	70.1	43.54	83.0	607
CPI_Potato	142	137	42.1	29.65	83.9	325
WPI_Potato	195	189	69.7	35.74	91.1	491
WPI_Onion	212	185	102.0	48.11	102.0	730
WPI_Tomato	223	177	140.0	62.78	79.5	1040

Note: S.D. stands for Standard Deviation and C.V. for coefficient of variation.

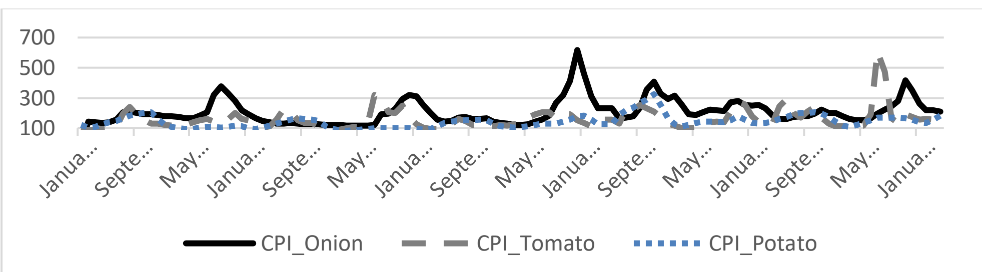

Figure 1. Time-series plot of CPI sub-index of Onion, Tomato and Potato

Table 2 provides a month-wise comparison of the CPI and WPI sub-index for three commodities. The onion CPI and WPI sub-index had the lowest value during May, while November had the highest value. For the retail index for onion, the lowest CV was in June, with December showing the highest variation. The tomato CPI and WPI sub-index had the lowest value during March, while July had the highest, with the highest variation. Similarly, the potato CPI and WPI sub-index had the lowest value in March and February, with the lowest variation among all months, while November had the highest value and the highest variation.

TABLE 2. MONTH-WISE COMPARISON OF CPI AND WPI SUB-INDEX FOR TOMATO, ONION AND POTATO

Month (1)	CPI_Onion (2)	CPI_Tomato (3)	CPI_Potato (4)	WPI_Onion (5)	WPI_Tomato (6)	WPI_Potato (7)
January	235	129	129	238	164	160
	(41.36)	(20.7)	(21.78)	(44.54)	(37.2)	(28.56)
February	207	116	115	202	145	143
	(33.57)	(18.36)	(18.52)	(39.8)	(37.38)	(22.73)
March	182	116	113	170	142	149
	(26.59)	(20.26)	(17.52)	(29.47)	(34.08)	(25.97)
April	165	121	122	149	161	171
	(22.12)	(19.01)	(21.31)	(22.95)	(27.52)	(29.53)
May	156	140	126	137	178	184
	(22.12)	(32.86)	(16.98)	(17.96)	(57.3)	(25.76)
June	158	167	142	157	225	200
	(16.08)	(35.39)	(20.63)	(22.61)	(57.78)	(27.1)
July	177	244	154	184	354	217
	(18.19)	(56.15)	(23.9)	(28.32)	(73.73)	(27.47)
August	201	221	160	217	314	223
	(24.53)	(46.15)	(26.13)	(40.88)	(63.06)	(29.28)
September	220	172	161	234	229	224
	(31.55)	(24.24)	(30.68)	(43.16)	(41.97)	(35.36)
October	247	178	168	266	270	233
	(31.42)	(24.21)	(32.68)	(36.73)	(43.33)	(37.25)
November	280	185	175	304	291	247
	(37.5)	(28.97)	(35.2)	(42.76)	(47.42)	(42.11)
December	275	156	154	291	231	202
	(52)	(29.36)	(28.7)	(60.82)	(52.38)	(36.63)

Note: Figures within parentheses are coefficient of variation in percentage.

Table 3 compares the CPI and WPI sub-index year-wise for three commodities. These have risen over the years, with some fluctuations.

TABLE 3. YEAR-WISE COMPARISON OF CPI AND WPI SUB-INDEX FOR TOMATO, ONION AND POTATO

Year	CPI_Onion (2)	CPI_Tomato (3)	CPI_Potato (4)	WPI_Onion (5)	WPI_Tomato (6)	WPI_Potato (7)
2014	173	135	156	181	150	235
	(16.59)	(37.41)	(22.63)	(22.76)	(51.53)	(27.96)
2015	233	145	107	263	176	126
	(32.36)	(17.45)	(7.48)	(39.54)	(28.01)	(12.38)
2016	140	143	133	127	133	183
	(14.64)	(29.58)	(18.72)	(16.46)	(25.56)	(24.43)
2017	172	173	94.7	178	224	113
	(42.97)	(48.09)	(6.75)	(52.36)	(60.71)	(11.95)
2018	183	119	130	173	138	188
	(27.65)	(17.23)	(19.54)	(37.46)	(29.13)	(25.74)
2019	228	164	128	253	246	170
	(67.54)	(20.91)	(17.42)	(75.49)	(29.96)	(19.94)
2020	276	181	208	260	259	307
	(34.89)	(25.47)	(32.74)	(46.92)	(44.4)	(28.4)
2021	243	153	143	251	257	189
	(17.45)	(39.8)	(12.8)	(22.87)	(59.53)	(15.87)
2022	201	187	171	196	300	243
	(16.32)	(28.61)	(17.02)	(21.84)	(40)	(20.91)
2023	227	210	149	232	336	190
	(36.87)	(76.19)	(15.03)	(49.14)	(85.42)	(15.47)
AAGR (%)	5.98	7.00	4.23	8.02	14.94	7.41
Instability (%	24.54	19.85	30.43	33.42	31.74	44.00

Note: Figures within parentheses are coefficient of variation in percentage

The most notable period for onion was in 2016 and 2021 when both sub-indices declined. For tomato, both sub-indices declined in 2018 and 2021. The CPI and WPI sub-index for potato had risen and fallen every alternate year. This starkly contrasts onion and tomato, where increases in the CPI did not consistently accompany increases in the WPI. It may also be said that only part of the changes in the WPI are transmitted to the respective CPI. WPI has been a leading indicator of changes in the CPI, but the reverse is not necessarily true. This reasoning has been utilized further in the regression analysis to help forecast the CPI sub-index for TOP commodities. This reasoning is further supported by the fact that the average annual growth rate and instability in the TOP commodity CPI sub-indices are lower than their respective WPI sub-indices.

Table 4 provides a correlation structure among TOP's CPI and WPI sub-indices. It can be observed that the CPI of each commodity was strongly correlated with the WPI of that same commodity. However, the correlation among the commodities was weak. This doesn’t necessarily imply that these commodities' price formation and consumption habits are not linked.

TABLE 4. CORRELATION AMONG CPI AND WPI SUB-INDEX OF TOP COMMODITIES

(1)	CPI_Onion (2)	CPI_Tomato (3)	CPI_Potato (4)	WPI_Potato (5)	WPI_Onion (6)	WPI_Tomato (7)
CPI_Onion	1.000					
CPI_Tomato	0.152	1.000				
CPI_Potato	0.321	0.349	1.000			
WPI_Potato	0.254	0.316	0.967	1.000		
WPI_Onion	0.974	0.154	0.273	0.198	1.000	
WPI_Tomato	0.213	0.949	0.394	0.331	0.204	1.000

Determinants of Tomato CPI and its Forecasting

Table 5 provides the determinants of the tomato CPI sub-index. It is observed that the tomato price sub-index of the CPI for the current month depended upon its wholesale price sub-index prevailing in the past two months, the second lag of the WPI sub-index for potato, and the third lag value of the CPI potato sub-index. The dependence of CPI tomato on the CPI potato sub-index also reveals that these two commodities complement each other, as indicated by the negative regression coefficient. All these regression coefficients were significant. All the numbers are numerically below the unit value, meaning that a percentage point change in the regressor causes less than a percentage point change in the tomato CPI sub-index. This means there is less than complete transmission of changes in the regressor to the dependent variable. The tomato CPI sub-index depended on its third lag, which indicated that the entire price formation process for tomato was three months. A percentage point increase in tomato CPI sub-index in the past three months had a negative impact on its current value by 0.526 per cent.

TABLE 5. DETERMINANTS OF CPI SUB-INDEX FOR TOMATO(ALL VARIABLES IN LOG-FORM)

Variable (1)	***Coefficient*** (2)	***Std. Error*** (3)	***t-ratio*** (4)	***p-value*** (5)	
const	3.64726	0.345551	10.55	<0.0001	***
CPI_Potato_3	−0.525619	0.122589	−4.288	<0.0001	***
WPI_Potato_2	0.331391	0.0972527	3.408	0.0009	***
WPI_Tomato_1	0.621314	0.0621015	10.00	<0.0001	***
WPI_Tomato_2	−0.200376	0.0626115	−3.200	0.0018	***

Mean dependent var	4.998541	S.D. dependent var	0.308448
Sum squared resid	3.520687	S.E. of regression	0.183991
R-squared	0.657359	Adjusted R-squared	0.644180
F(4, 104)	49.88115	P-value(F)	2.28e-23
Log-likelihood	32.41740	Akaike criterion	−54.83480
Schwarz criterion	−41.37806	Hannan-Quinn	−49.37760
Rho	0.199783	Durbin-Watson	1.592412

White's test for heteroskedasticity: $\chi^2(df = 14) = 16.4546$, p-value = 0.286407
LM test for autocorrelation up to order 12: F (12, 92)= 1.5683, p-value =0.114847
CUSUM test: Harvey-Collier t(103) = -0.854501, p-value=0.39481
ARCH test: $\chi^2(df = 12) = 4.85827$, p-value = 0.962546
Mean Absolute Percentage Error =4.5805, Theil's U2 = 0.81335

Significance Code: *** for p-value <.001, ** for p-value <0.01 and * for p-value <0.05.

The sum of all the coefficients of the regressors in the model is less than one, meaning that a percentage point increase in all regressors would lead to a less than one percent increase in the tomato CPI sub-index. This phenomenon was attributed to the wide variation in the planting dates of tomato in these states, leading to its staggered harvest. Additionally, year-round planting and the availability of varieties and hybrids suitable for different climatic conditions ease the pressure on the price of tomato through regular market arrivals. The increased shelf life of tomato is also responsible for reducing the time lag associated with the CPI sub-index. The regressors in the regression accounted for 64 percent of the variation in the tomato CPI sub-index. The remaining variation of 36 percent was accounted for by the factors not included in the model, reflecting the need to search for more powerful predictors of the tomato CPI sub-index.

The model didn’t suffer from autocorrelation, heteroscedasticity, parameter instability, and non-linearity, as indicated by the p-value for the relevant tests. Also, the ARCH process was not present in the residuals. All the residuals were white noise, as indicated by autocorrelation and partial-autocorrelation falling within the confidence intervals (Figure 2). The absence of parameter instability and white noise residuals allows the use of estimated regression for forecasting. Forecasting the tomato CPI sub-index revealed an inaccuracy of only 4.58 percent (Table 5 and Figure 3). Theil’s U2 value was less than one, indicating that the regression model was better at predicting the tomato CPI sub-index than the naïve forecast. The excellent performance of this simple autoregressive model suggests that higher-order models like vector autoregression, ARIMA, Neural Network, GARCH, etc., were not necessary. This simplified the task of predicting the tomato CPI sub-index. The forecast results from May 2023 to April 2024 are provided in Table 6.

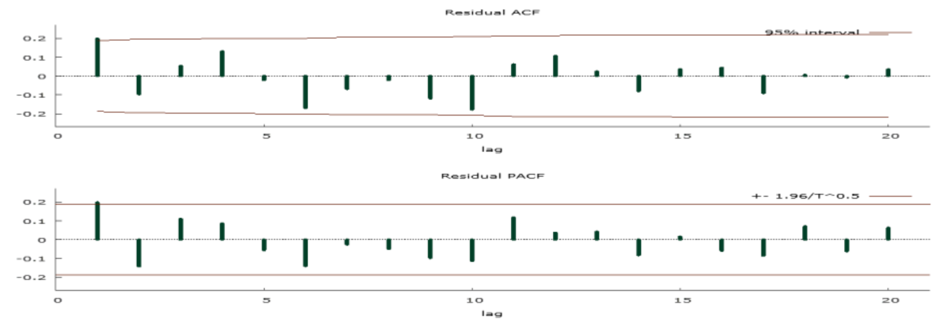

Figure 2. ACF and PACF plot of residuals from regression of Tomato CPI sub-index

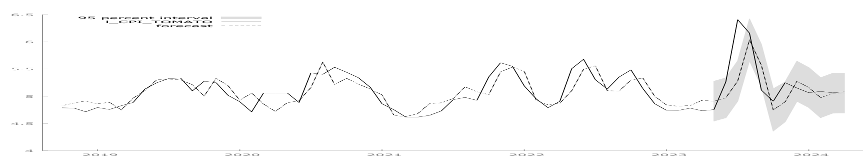

Figure 3. Time Series plot of Tomato CPI sub-index and its forecast with confidence interval (All values in log form)

TABLE 6. COMPARISON OF TOMATO CPI SUB-INDEX PREDICTION WITH ACTUAL OBSERVATIONS FOR TEST DATA SERIES (ALL VALUES IN LOG FORM)

Year-Month (1)	Actual (2)	Forecast (3)	std.error (4)	95% interval (5)
2023:05	4.76	4.92	0.20	4.522 - 5.319
2023:06	5.26	4.89	0.20	4.492 - 5.289
2023:07	6.41	5.23	0.20	4.826 - 5.626
2023:08	6.16	6.11	0.21	5.696 - 6.528
2023:09	5.11	5.53	0.21	5.113 - 5.943
2023:10	4.91	4.76	0.22	4.318 - 5.207
2023:11	5.25	5.05	0.22	4.611 - 5.484
2023:12	5.15	5.40	0.20	5.004 - 5.803
2024:01	5.07	5.21	0.20	4.809 - 5.609
2024:02	5.09	5.06	0.20	4.660 - 5.456
2024:03	5.06	5.17	0.20	4.776 - 5.571
2024:04	5.08	5.10	0.20	4.699 - 5.493

Determinants of Onion CPI and its Forecasting

Table 7 provides the regression analysis results for the determinants of the onion CPI sub-index. It reveals that the onion retail price sub-index of the CPI depended upon its first lag, its own two lags of the WPI sub-index, and the tomato CPI sub-index. The positive sign on the partial regression coefficient of the tomato CPI sub-index meant that the two commodities were substitutes. The negative sign on the partial regression coefficient of the second lag of the onion WPI sub-index meant that a one percent point increase in onion WPI leads to market adjustments and production pattern changes, which result in a decline in the onion CPI sub-index in the current period by 0.31 percent point. A one percent point increase in the onion CPI in the previous month transferred 66 percent of that increase to the next month. The sum of all estimated partial regression coefficients is nearly one, implying that a one percent point increase in all regressors could increase the current month's onion CPI by one percent point. All the partial regression coefficients were jointly significantly different from zero.

TABLE 7. DETERMINANTS OF ONION CPI SUB-INDEX (ALL VARIABLES IN LOG-FORM) HAC STANDARD ERRORS

	Coefficient	Std. Error	t-ratio	p-value	
CPI_Tomato_1	0.169008	0.0396918	4.258	<0.0001	***
WPI_Onion_1	0.479861	0.110650	4.337	<0.0001	***
WPI_Onion_2	−0.309393	0.0967574	−3.198	0.0018	***
CPI_Onion_1	0.669401	0.175804	3.808	0.0002	***

Mean dependent var	5.265686	S.D. dependent var	0.344306
Sum squared resid	1.369085	S.E. of regression	0.114736
R-squared	0.999545	Adjusted R-squared	0.999532
F(4, 104)	77067.59	P-value(F)	1.5e-179
Log-likelihood	82.62601	Akaike criterion	−157.2520
Schwarz criterion	−146.5235	Hannan-Quinn	−152.9020
rho	0.034545	Durbin's h	NA

LM test for autocorrelation up to order 12: F(12, 92) = 0.473458, p-value= 0.925474
White's test for heteroskedasticity: $\chi^2(df = 14) = 36.8513$, p-value = 0.000776966
CUSUM test: Harvey-Collier t(103) = 0.286121, p-value=0.77536
ARCH test: $\chi^2(df = 12) = 15.2779$, p-value = 0.22659
Mean Absolute Percentage Error (MAPE)=2.0226, Theil's U2 = 0.82653

The model didn't suffer from heteroscedasticity in residuals and parameter instability but suffered from autocorrelation. For this reason, the regression results were provided with Heteroscedasticity-Autocorrelation-Consistent (HAC) standard errors. Before utilizing the forecasting model, the ACF and PACF plot was checked for significant autocorrelation and partial autocorrelation, but none were significant (Figure 4). Hence, the given model was found suitable for forecasting.

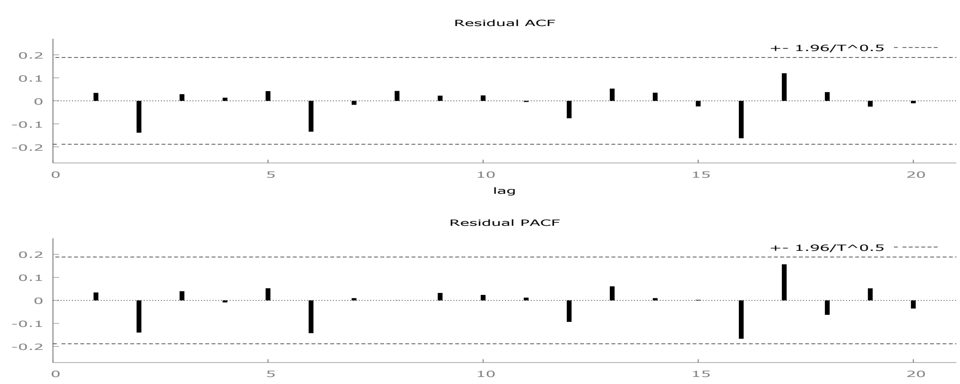

Figure 4. ACF And PACF Plot Of Residuals From Regression Of Onion CPI Sub-Index

The results of the onion CPI sub-index are provided in Table 8 and are graphically shown in Figure 5. The predicted values were very close to the actual onion CPI numbers. A MAPE value of 2.02 percent revealed the level of inaccuracy. The model was better than the naïve forecast method as Theil's U value was less than one.

TABLE 8. COMPARISON OF TOMATO CPI SUB-INDEX PREDICTION WITH ACTUAL OBSERVATIONS FOR TEST DATA SERIES (ALL VALUES IN LOG FORM)

Year-Month	Actual	Forecast	std. error	95% interval
2023:05	5.03	4.98	0.114	4.75-5.20
2023:06	5.11	4.97	0.138	4.69-5.24
2023:07	5.28	5.10	0.147	4.80-5.39
2023:08	5.40	5.46	0.151	5.16-5.76
2023:09	5.50	5.70	0.153	5.39-6.00
2023:10	5.64	5.68	0.154	5.38-5.99
2023:11	6.03	5.70	0.154	5.39-6.01
2023:12	5.86	5.87	0.154	5.56-6.17
2024:01	5.56	5.70	0.154	5.39-6.01
2024:02	5.38	5.50	0.154	5.19-5.81
2024:03	5.39	5.39	0.154	5.08-5.70
2024:04	5.35	5.40	0.154	5.09-5.70

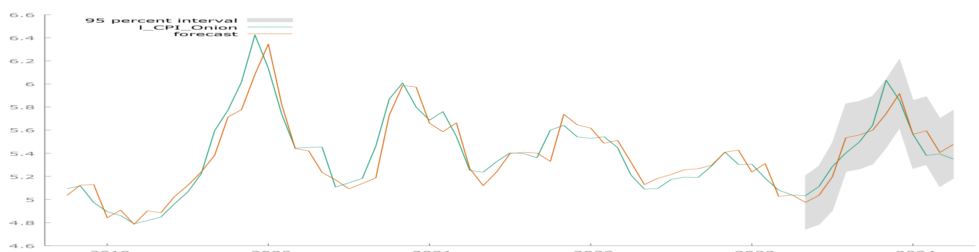

Figure 5. Time Series plot of Onion CPI sub-index and its forecast with confidence interval (All values in log form)

Determinants of Potato CPI Sub-Index

Table 9 presents the results of the regression analysis for the determinants of the potato CPI sub-index, revealing the dependence on the first two lags of onion and tomato CPI sub-indices, its own first, second, eleventh, and twelfth lag, and the WPI sub-index of onion, tomato, and potato at different lags. The intercept is not statistically significant. Hence, the regression model is valid for the long run. Its dependence on higher lags suggests that potato retail market pricing is subject to the supply of potato stored several months earlier and the decision of farmers made at least a year in advance. Although potato is produced at different times across India, their supply to various parts of India is subject to the prices prevailing in the destination local markets. The sum of all partial regression coefficients associated regressors is 1.96, implying that a one percent point change in all regressors simultaneously can lead to nearly a two percent point change in the potato CPI sub-index. Thus, compared to onion and tomato, the retail price of potato is more sensitive to changes in its influencing factors.

TABLE 9. DETERMINANTS OF POTATO CPI SUB-INDEX (ALL VARIABLES IN LOG-FORM)

(1)	Coefficient (2)	Std. Error (3)	t-ratio (4)	p-value (5)	
Const	0.163171	0.233752	0.6981	0.4870	
CPI_Onion_1	0.761523	0.182940	4.163	<0.0001	***
CPI_Onion_2	−0.205734	0.0649290	−3.169	0.0021	***
CPI_Tomato_1	0.123234	0.0353868	3.482	0.0008	***
CPI_Tomato_2	−0.109247	0.0411903	−2.652	0.0095	***
WPI_Potato_3	0.145569	0.0530850	2.742	0.0074	***
WPI_Onion_1	−0.479329	0.116552	−4.113	<0.0001	***
WPI_Tomato_3	0.0589490	0.0217442	2.711	0.0081	***
CPI_Potato_1	1.37191	0.0966047	14.20	<0.0001	***
CPI_Potato_2	−0.750927	0.123011	−6.105	<0.0001	***
CPI_Potato_11	0.261107	0.0563921	4.630	<0.0001	***
CPI_Potato_12	−0.231243	0.0565136	−4.092	<0.0001	***

Mean dependent var	4.891122	S.D. dependent var	0.281018
Sum squared resid	0.377906	S.E. of regression	0.065532
R-squared	0.951663	Adjusted R-squared	0.945621
F(11, 88)	250.0672	P-value(F)	2.26e-61
Log-likelihood	137.0201	Akaike criterion	−250.0402
Schwarz criterion	−218.7782	Hannan-Quinn	−237.3879
rho	−0.015083	Durbin's h	−0.583775

LM test for autocorrelation up to order 12: F(12, 76) = 0.539532, p-value= 0.882068
White's test for heteroskedasticity: $\chi^2(df = 77) = 87.5871$, p-value = 0.192133
CUSUM test: Harvey-Collier t(87) = -0.260412, p-value=0.795161
ARCH test: $\chi^2(df = 12) = 14.2093$, p-value = 0.287544
Mean Absolute Percentage Error (MAPE)=1.302, Theil's U2 = 0.92049

All the regressors in the regression model explain 95 per cent of the variation in the potato CPI sub-index and are jointly significant. The estimated model doesn't suffer from autocorrelation, heteroscedasticity, or parameter instability. All autocorrelations and partial-autocorrelations were within the confidence interval, meaning that residuals were white noise (Figure 6). There was no need for an ARCH model to be fitted to residuals, as the null hypothesis of no ARCH effect was accepted.

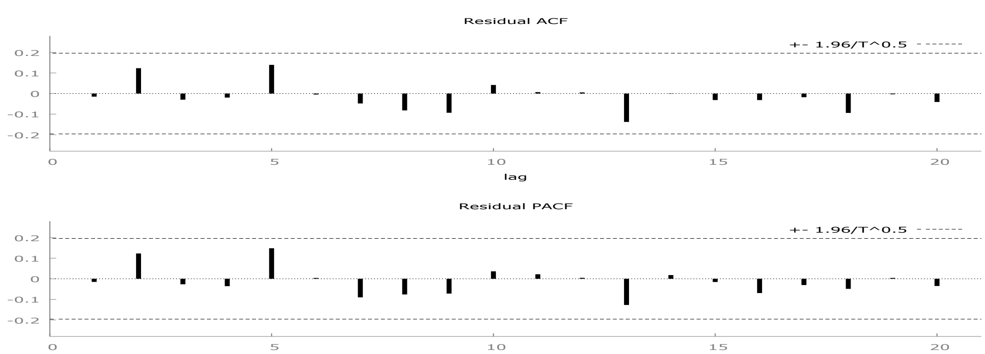

Figure 6. ACF and PACF plot of residuals from regression of Onion CPI sub-index

Table 10 presents the forecast results for the potato CPI sub-index, as shown in Figure 7. The forecasted values were a close approximation to actual values for the test period. The MAPE value was 1.302, a very low error (Table 9). Theil's U value is within one, meaning the regression model has better forecast ability than the naïve forecast model.

TABLE 10. COMPARISON OF POTATO CPI SUB-INDEX PREDICTION WITH ACTUAL OBSERVATIONS FOR TEST DATA SERIES (ALL VALUES IN LOG FORM)

Year-Month (1)	Actual (2)	Forecast (3)	std. error (4)	95% interval (5)
2023:05	4.92	4.94	0.068	4.80-5.07
2023:06	5.02	5.05	0.069	4.92-5.19
2023:07	5.12	5.17	0.070	5.03-5.30
2023:08	5.15	5.33	0.081	5.17-5.49
2023:09	5.13	5.11	0.088	4.94-5.29
2023:10	5.13	5.02	0.073	4.88-5.17
2023:11	5.12	5.07	0.074	4.92-5.22
2023:12	5.06	5.11	0.070	4.97-5.25
2024:01	4.95	4.91	0.069	4.77-5.05
2024:02	4.90	4.85	0.068	4.71-4.98
2024:03	5.04	4.91	0.067	4.78-5.05
2024:04	5.22	5.14	0.067	5.01-5.28

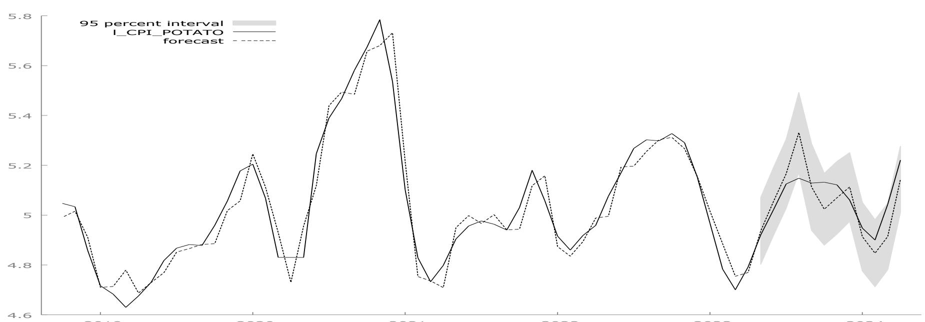

Figure 7. Time Series Plot of Onion CPI Sub-Index and Its Forecast With Confidence Interval (All values in log form)

IV

CONCLUSION

The study examined the CPI and the indices trends for tomato, onion, and potato over the past decade. The regression models provided accurate forecasts for the CPI sub-index of these commodities. Policymakers should closely monitor the factors influencing these sub-indices and incorporate seasonal agricultural planning at macro, meso, and micro levels, particularly in months with higher prices, like November and July. Potatoes showed the strongest linkages among the three due to their diverse determinants and exhibited the greatest price instability. While tomato experienced the highest price increases, they had the least instability. To ensure equal access to market information for all stakeholders, it is recommended that commodity-specific market intelligence and forecasting efforts be enhanced.

REFERENCES

Ajmal, S., S, Rohith., Unniravisankar, P., & Osman, N. (2024). Price Dynamics of Tomato, Onion and Potato (TOP) in India. *Asian Journal of Agricultural Extension, Economics & Sociology*, *42(3)*, 134-143. https://doi.org/10.9734/AJAEES/2024/v42i32387

Anonymous (2024). Economic Adviser, Office of the Economic Adviser. Department for Promotion of Industry and Internal Trade, Government of India. Retrieved June 23, 2024, from https://eaindustry.nic.in/

Anonymous (2024). Percentage share. Government of India. Retrieved [June 23, 2024], from https://mospi.gov.in/percentage-share

Anonymous. (2024). Annual report publications: Monetary policy. Reserve Bank of India. Retrieved [June 24, 2024], from https://m.rbi.org.in/Scripts/AnnualReportPublications.aspx?Id=1403#:~:text=Monetary%20Policy-,III.,in%20several%20food%20sub%2Dgroups

Anonymous. (2024). Consumer Price Index. Ministry of Statistics and Programme Implementation. Retrieved [May 17, 2024], from https://cpi.mospi.gov.in/Default1.aspx

Anonymous. (2024). India production: Vegetables. APEDA. Retrieved [June 24, 2024], from https://agriexchange.apeda.gov.in/India%20Production/India_Productions.aspx?cat=Vegetables&hscode=1088

Gohain, K. (2021). Evaluation of theils u: a naïve forecast application. *Quantum Journal of Engineering, Science and Technology*, *2(5)*, 26-31.

Jadhav, V., Chinnapa, R. B., & Gaddi, G. M. 2017. Application of ARIMA model for forecasting agricultural prices.

Birthal, P., Ranjan, A., Kishore, V., & Bhoi, B. B. (2022). Forecasting Food Inflation using News-based Sentiment Indicators. *Reserve Bank of India Occasional Papers*, *42(2)*.

Tiwari, A., Afroz, S. B., & Kumar, V. (2021). Market Vulnerabilities And Potential Of Horticulture Crops In India: With Special Reference To Top Crops. *Indian Society Of Agricultural Marketing*, *35*(3).