CHAPTER VI
FORECASTING OF INFLATION

In this section, an investigation of better forecast model of inflation in Nepal is examined. It is also discussed core inflation.

1. FORECASTING OF INFLATION IN NEPAL

Forecasting is generally used to predict the future values of the economic phenomenon. An important reason for formulating an econometric model is to generate forecasts of one or more economic variables. There are two forecasting methodologies; first is, econometric forecasting- based on a regression model that relates one or more dependent variables to a number of independent variables; and second is, time series forecasting- based on attempts to predict the values of a variable from past values of the same variable. In the present study, both methodologies have been applied to forecast inflation in Nepal. The validity of the forecasting models has also been examined. Moreover, within one methodology, the forecastability of the model is examined by considering quarterly data frequency and annual data frequency.

(A) Trend Method

Under the trend method, extrapolation and interpolation of historical data is attempted through estimation of alternative trend equations. A trend equation is one in which the variable under forecast is made simply as a function of time. A basic characteristic of time series is its long-run growth pattern. In addition, the autoregressive trend model is another variant of the trend model that expresses the forecast variable as a function of its own lagged values. If we believe that this upward trend exists and will continue, we can construct a simple model that describes the trend, and can be used to forecast the series. There are various extrapolation models that characterize the time series. Some important extrapolation trend models for consumer price index of Nepal for both the annual and quarterly data frequency are examined.

Table 6.1
Inflation Forecasting by Trend Methods
(1975-2002)

Model	Data Frequency		R2		DW	
	Quarterly	Annual	Qtr	Anm	Qtr	Anm
Linear trend model	$P_t = -7.61 + 1.34T$ (-3.87)* (43.62)*	$P_t = -5.69 + 5.351T$ (-1.45)** (21.48)*	0.94	0.94	0.06	0.10
Logarithmic linear trend model (exp. growth)	$\ln P_t = 2\ 66 + 0.02T$ (189.81)* (106.03)*	$\ln P_t = 2.69 + 0\ 09T$ (103.22)* (55.83)*	0.99	0.99	0.26	0 34
Autoregressive trend model	$P_t = 0.60 + 1.01P_{t-1}$ (1.33)* (176.05)*	$P_t = 2\ 27 + 1\ 04P_{t-1}$ (2.36)* (82.50)*	0 99	0.99	2.30	1.19
Logarithm autoregressive trend model	$\ln P_t = 0.03 + 1.00 \ln P_{t-1}$ (1.65)** (205.32)*	$\ln P_t = 0.13 + 0.99 \ln P_{t-1}$ (3 21)* (93.92)*	0.99	0.99	2.39	1.60

Figures given below the coefficients in the parenthesis are t values. Asterisks (*) signifies coefficients significant at 1% level, asterisks (**) signifies coefficient significant at 5% level, and asterisks (***) signifies coefficients significant at 10% level.

The estimated coefficients of inflation equations, as shown in Table 6.1, give in-sample forecast for data ranging from 1975:I to 2002:IV (in the case of quarterly data) and 1975 to 2002 (in the case of annual data). On the basis of in-sample estimation, ex-post forecasts for the year 2003 are estimated to examine the predictability of the in-sample forecast model in Table 6.2. The evaluation of forecasts is made by Mean Absolute Prediction Error (MAPE), which is considered as surrogative measure that combines all errors into one measure. Lesser is MAPE found, better is the forecastability of a model. An ex-post forecast of inflation and their corresponding MAPE for quarterly and annual data are presented in Table 6.2.

Table: 6.2

Ex-post Forecast of CPI for the Year 2003

	Price Indices		Model			
Quarterly	Actual	Actual (log)	Linear Trend	Exponential Growth	Auto-regressive	Log/Auto-regressive
(1)	(2)	(3)	(4)	(5)	(6)	(7)
2003 I	156 20	5.0511	141.94	5.2442	154.56	5.0446
2003 II	154 43	5 0394	143.28	5.2673	156.64	5.0617
2003 III	153 30	5.0324	144.61	5.2904	158.75	5.0788
2003 IV	155 17	5.0445	145.94	5.3135	160 87	5.0985
MAPE	-	-	**10.83**	**0.2369**	**2.9305**	**0.0282**
Annual						
2003	154.78	5.0421	144.02	5.2800	157 45	5.0732
MAPE	-	-	**10.78**	**0.2378**	**2.6466**	**0.0311**

The actual data in absolute and logarithmic forms are given in column (2) and (3) respectively. Similarly, the results of ex-post forecasts are presented in column (4) to (7). In both the quarterly and annual data frequency, the MAPEs of autoregressive model are found to be lesser than that of the linear trend model evaluating better forecasting ability. Similarly, the log/autoregressive model is seen to be possessing lesser MAPE in comparison to the exponential growth model, and hence found to be better forecasting performance. Autoregressive model is found better, in the case of annual data, whereas, log/augoregressive model is in the case of quarterly data.

(B) Regression Method

Under the regression method, a causal model is first formulated, and estimated using the historical data. It is verified on the basis of theoretical and statistical tests, and then used to derive ex-post forecasts. Because of the lack of availability of quarterly data series other than price indices and monetary aggregates, annual data are used to examine the forecastability of multivariate model, and quarterly data series are used to forecast bi-variate model. Therefore, forecastability of pure monetarists model is examined using quarterly data. The estimated results of the best fitted inflation function and its regression coefficients with test statistics are shown in Table 6.3.

Table 6.3
Inflation Forecasting by Regression Model
(1975:I-2002:IV)

Data Frequency		R2		DW		MAPE For the Year 2003	
Quarterly	Annual	Qtr	Anm	Qtr	Anm	Qtr	Anm
$\ln P_t = -1.63 + 0.59 \ln M1_t$ (-32.37)*(111.94)*	$\ln P_t = -1.67 + 0.60 \ln M1_t$ (-21.14)*(71.73)*	0.99	0.99	0.94	0.76	0 0968	0 1084
$\ln P_t = -1.61 + 0.29 M1_t + 0.31 M1_{t-1}$ (-33 74)* (3 36)* (3 57)*	$\ln P_t = -4.27 + 0.52 \ln M1_t +$ (-1 40)** (5 55)* $0\ 28 \ln GDPR_t - 0.01 RLR_t$ (0.86) (-2.20)**	0.90	0.99	0.69	0 50	0 1006	0 1082
$\ln P_t = -1.42 + 0.53 \ln M2_t$ (-33.83)* (129.32)*	$\ln P_t = -7\ 36 + 0\ 24 \ln M1 +$ (-2.74)* (2.35)** $0.63 \ln GDPR + 0.35 \ln IWPI + 0.11 EP$ (2.18)** (2.28)** (3.65)*	0 99	0.99	0 57	1 65	0 1075	0.0391

Note: M1 and M2 are narrow and broad monetary aggregates, GDPR is real gross domestic product, IWPI is Indian wholesale price index, EP is expected rate inflation.

If we compare the forecasting performance of the regression model estimated by using quarterly and annual data frequencies, the regression model of inflation on M1 using quarterly data estimation is found to be better than that use is made of annual data. It is confirmed by the lower value of MAPE found in using quarterly data as compared to annual data. Similarly, M1 has a better predictive capacity than M2, if it is compared within the quarterly data frequency. In case of the use of annual data, a regression model of inflation on M1, GDPR, IWPI and EP, as shown in third equation, is found to have the highest predictive power as compared to the first and second equations after a number of trials. This model is found to be the best in terms of other test statistics also. The histogram of residuals of the regression model of inflation on M1, GDPR, IWPI and EP is shown below. It is found to be approximately following normal distribution satisfying the residual term of classical Linear Regression Model (CLRM). Similarly, actual and forecasted value of the same equation is also shown in figure 2.

Figure1 : Histogram of Residuals and the Normal Density

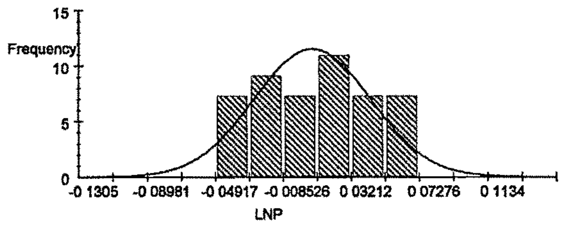
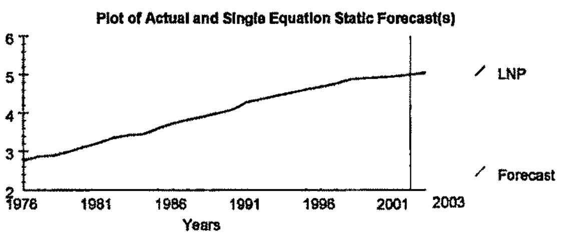


Figure 2: Plot of Actual and Forecast of Inflation



(C) Box Jenkins (B-J) Methodology

The Box-Jenkins method (also known as the Auto-Regressive Integrated Moving Average (ARIMA) method) of forecasting involves four steps: identification, estimation, verification of the model, and derivation of forecasts. First, the identification of the model can be chosen either AR(p) or MA(q) or combination of ARMA(p,q) with or without integrated of different orders for the specification of the model. The ARIMA models of CPI can be represented as follows:

$$CPI_t = \alpha_0 + \alpha_1 CPI_{t-1} + \alpha_2 CPI_{t-2} + ... + \alpha_p CPI_{t-p} + \varepsilon_t \qquad \text{AR}(p) \text{ process}$$

$$CPI_t = \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + ... + \phi_q \varepsilon_{t-q} \qquad \text{MA}(q) \text{ process}$$

$$CPI_t = \alpha_0 + \alpha_1 CPI_{t-1} + \alpha_2 CPI_{t-2} + ... + \alpha_p CPI_{t-p} + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + ... + \phi_q \varepsilon_{t-q} \qquad \text{ARMA}(p,q)$$

One criterion of selecting (p,d,q) terms in ARIMA is to estimate and examine the adequacy of the estimated model by the test of significance of the estimated parameters by t-tests and goodness of fit of the model representing coefficient of determination (R^2). Thus insignificant parameters would be dropped from the model.

The next step is to follow a diagnostic checking for the evaluation of the model by examining the properties of the residuals from the estimated model. In particular, if the model is correctly specified then, by definition, the ε_t must be random. Autocorrelation (AC) and partial autocorrelation (PAC) coefficients of various orders, and testing the significance of a group of ACs on the basis of the Box-Pierce Q statistic are the various statistical tools to examine the randomness of ε_t. If ε_t is random then we would expect all the autocorrelations coefficients in different lags to be insignificant. The coefficients of AC and PAC and Q statistics depend on the sample size. The appropriate size of the lags in calculating AC and PAC is to be about one fourth of total observations. The estimated value of Q can then be compared to the critical value of Q applying χ^2 distribution for (K-p-q) degree of freedom where 'K' is number of lags introduced in AC and PAC estimation.

If the estimated value of Q statistic is found to be unacceptably large (exceeding the table value of χ^2), then we would conclude that Q is significant, and hence the null hypothesis of no AC and PAC should be rejected. It implies that the model is found to be inadequate, and hence requires re-specification and re-estimation of ARIMA (p,d,q) terms. Indeed, the nature of the estimated AC of ε_t indicates the requirement of model re-specification (that is, whether additional AR or MA parameters should be introduced in the model). This process is continued unless and until a good identification of model is found. Once a Box-Jenkins model has been satisfactorily identified and estimated, it can then be used for the forecasting purpose.

The ARIMA (p,d,q) model specification and the application of the model for the forecasting of Consumer Price Index (CPI) of Nepal are examined in this study. Forecasts are based on both the annual and quarterly data from 1975 to 2003.

In the case of annual data, only AR(p) model is initially identified for the trial. Though AR(1) and AR(2) specifications were found to be statistically significant in terms of t-test and goodness of fit criteria, these models were found to be inadequate in the diagnostic checking stage on the basis of visualization of AC and PAC functions and corresponding Q statistic. The residuals (ε_t) show a non-random pattern. Therefore, the re-identification, re-estimation and re-checking of ARIMA (p,d,q) is made. Ultimately, ARIMA(1,0,1) data generating process of CPI of Nepal is found to be satisfactory. The AC and PAC function and corresponding Q statistics of the ε_t up to seven lags as a criteria of diagnostic checking of ARIMA(1,0,1) model is shown in Table 6.4.

Table: 6.4

AC and PAC of CPI on ARMA(1,0,1) Specification

(1975 – 2003)

<table>
<tr><th>AC</th><th>PAC</th><th>Lag</th><th>AC</th><th>PAC</th><th>Q-Stat</th><th>Prob</th></tr>
<tr><td>| |</td><td>| |</td><td>1</td><td>0 052</td><td>0 052</td><td>0 0850</td><td>0 771</td></tr>
<tr><td>|** |</td><td>|** |</td><td>2</td><td>0 247</td><td>0.245</td><td>2 0623</td><td>0 357</td></tr>
<tr><td>| |</td><td>| |</td><td>3</td><td>-0 024</td><td>-0 049</td><td>2 0812</td><td>0 556</td></tr>
<tr><td>|* |</td><td>| |</td><td>4</td><td>0 098</td><td>0 043</td><td>2 4157</td><td>0 660</td></tr>
<tr><td>| |</td><td>| . |</td><td>5</td><td>-0 001</td><td>0.011</td><td>2 4158</td><td>0 789</td></tr>
<tr><td>| |</td><td>. | |</td><td>6</td><td>0 002</td><td>-0 036</td><td>2 4159</td><td>0 878</td></tr>
<tr><td>|* . |</td><td>|* |</td><td>7</td><td>0 135</td><td>0 151</td><td>3 1483</td><td>0 871</td></tr>
</table>

The Q statistic of ε_t at 7th lag is 3.15, which is less than the critical value of χ^2 at 11.07 at (K-p-q), i.e. (7-1-1)=5 degree of freedom. The insignificant Q statistic signifies randomness of ε_t and hence adequate description of the data generating process CPI by ARIMA (1,0,1). The model explains as follows:

$$CPI_t = \alpha_0 CPI_{t-1} + \varepsilon_t + \phi_1 \varepsilon_{t-1} \qquad \text{ARMA(1,0,1)}$$

This model can be used to generate forecasts of CPI_t

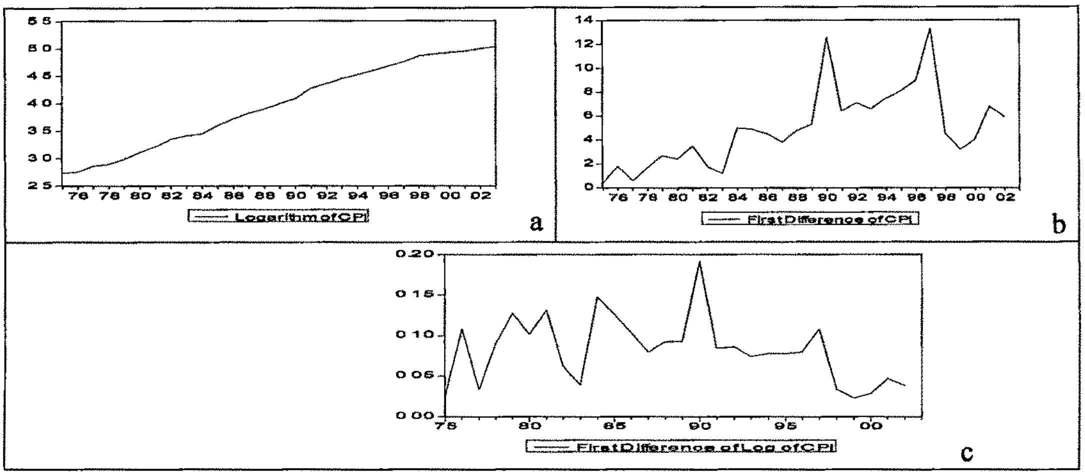
$$CPI_t = 1.062CPI_{t-1} + \varepsilon_t + 0.380\varepsilon_{t-1}$$

(117.74) (2.04) R2=0.99 DW=1.74

Both the coefficients of AR and MA are statistically significant at 5 percent level. R2 is quite high and DW statistic is not less so. The above model was estimated using a sample of 29 observations of CPI. Thus for the last sample observation the model could be written as: $CPI_{2003} = \alpha_0 PI_{2002} + \varepsilon_{2003} + \phi_1 \varepsilon_{2002}$. The observation of ε_t can be found from the estimated residual from the ARIMA(1,0,1) model. Ex-ante forecast can be derived by the recursive method, where the forecasted value of inflation is considered to be actual value for the derivation of next years forecast. The ex-ante forecast was found by applying the above model at 157.86 for the year 2004. However, the BJ methodology is a better representative for the forecasting of short period of one or two years. The Absolute Mean Prediction Error (MAPE) which is summarized measure of error of ex-post model is 3.51 for the year 2003.

If we introduce integrated term 'I' in the ARMA (p,q), the specification becomes ARIMA(p,d,q). Here, we are considering the difference data in order to derive the model before the identification and estimation of the model. If a series is stationary by first differencing, the resulting series is considered as I(1), where d=1 in 'd' term of ARIMA(p,d,q) model. A forecasting model identification based on first difference CPI series has been analyzed subsequently.

Figure 3: Plot of actual, logarithmic and first difference of CPI



The raw data of inflation shows a clear upward trend, as shown in panel 'a' of the above diagram. Thus, data transformation by first is will be required. The data could also be interpreted as being homoscedastic as shown in panel 'b'. The data exhibits so much variation over time that it requires data to be taken into logarithm. If we take the logarithmic data into first difference, it becomes stationary as shown in panel 'c'. We can now attempt to fit a B-J model to the data of logarithm of first difference which are represented by the rate of growth of original data. Re-identification and re-examination of different ARIMA(p,d,q) carried an adequacy of model of ARIMA(1,1,1) with their corresponding coefficients and statistical values as follows.

$$\Delta CPI_t = 0.978\Delta CPI_{t-1} + \varepsilon_t + (-0.942)\Delta\varepsilon_{t-1}$$

ARIMA*(1,1,1)*

(37.80) (-6.43) DW=1.59

An evaluation of diagnostic checking for ARIMA(1,1,1) model can be made by deriving AC and PAC functions of ε_t up to lags 7 and their corresponding Q statistics are given below:

Table: 6.5
AC and PAC of CPI on ARMA (1,1,1) Specification
(1975 to 2003)

AC	PAC	Lag	AC	PAC	Q-Stat	Prob
|* |	|* |	1	0 161	0 161	0 8101	0 368
*| |	*| |	2	-0 082	-0 111	1 0268	0 598
*| |	| |	3	-0 071	-0 040	1 1950	0 754
| |	| |	4	0 003	0 014	1 1953	0 879
|** |	|** |	5	0 258	0 255	3 6280	0 604
|** |	|** |	6	0 289	0 223	6 8127	0 339
*| |	**| |	7	-0 167	-0 231	7 9273	0 339

Considering the properties of ε_t from the specified ARIMA(1,1,1), the model of CPI seems adequate for the forecasting purpose.

We can then go on using the above equation for forecasting; where dependent variable is changes in the CPI over the previous year. What it is obtained by the above model is forecast of CPI changes in 2004 over the 2003, or forecast of 2005 over the 2004. In order to obtain the forecast of CPI level rather than its change form, first-difference transformation should be done. On the basis of reverse transformation, forecast can be made into log form. At last forecast of CPI index can be obtained by taking antilog.

$$\Delta \log CPI_t = 0.978 \Delta \log CPI_{t-1} + \varepsilon_t + (-0.942) \Delta \log \varepsilon_{t-1}$$

ARIMA(1,1,1)

$$\log CPI_{2004} = (1-0.978)\log CPI_{2003} - 0.978 \log CPI_{2002} - 0.942 \log \varepsilon_{2003} + 0.942 \log \varepsilon_{2002} + \log \varepsilon_{2004} - \log \varepsilon_{2003}$$

If we substitute the known values of CPI and estimated residual terms up to the last observation, the forecast value for 2004 is found to be log 5.07897. If we take antilog, the absolute value of forecast figures becomes 160.61. In the case of ARIMA(1,1,1) model, the AEMP for the year 2003 is 3.25 which is smaller than the AEMP of 3.51 of ARIMA(1,0,1) model as examined above. Therefore, ARIMA(1,1,1) gives better forecasting performance than ARIMA(1,0,1).

In the case of quarterly data, the ARIMA(4,0,5) data generating process is found to be adequate after a re-identification and re-estimation of ARIMA(p,d,q). The model includes four AR terms (lag 1,4,8 and 16) and five

MA terms (lag 1,4,8,12,and 16). The data generating process of ARIMA(4,0,5) may show a tendency of seasonality. The estimated coefficients and their different test statistics as per the above specification are presented as follows:

$$CPI_t = 1.036CPI_{t-1} - 0.263CPI_{t-4} + 0.151CPI_{t-8} + 0.114CPI_{t-16} + 0.113\varepsilon_{t-1} + 0.928\varepsilon_{t-4} +$$

(17.20) (-3.56) (3.08) (2.08) (1.88) (87.38)

$$1.250\varepsilon_{t-8} + 0.781\varepsilon_{t-12} + 0.336\varepsilon_{t-16} + e_t \qquad \text{ARMA(4,0,5)}$$

(22.34) (13.70) (29.52) Adj R2=0.99 DW=2.10

Above model seems adequate in terms of significant coefficient of t-statistic and value of R2. The model is considered as correctly specified if the ε_t is found to be randomness. The AC and PAC functions of ε_t up to lags of 29 quarters 29 show a random pattern in terms of visual expression. Hence the model specification of ARIMA(4,0,5) is considered adequate.

Table: 6.6

AC and PAC of CPI on ARMA(4,0,5) Specification (1975:I-2003:IV)

AC	PAC	Lag	AC	PAC	Q-Stat	Prob
|*	|*	1	0 100	0 100	1 1690	0 280
|*	|*	2	0 127	0 118	3 0897	0 213
|*	|*	3	0 137	0 116	5 3300	0 149
|	|	4	0 062	0 027	5 7901	0 215
|*	|*	5	0 125	0 093	7 7054	0 173
|	|	6	0 048	0 008	7 9913	0 239
|	|	7	0 054	0 018	8 3518	0 303
|*	|*	8	0 153	0 122	11 303	0 185
|	|	9	0 023	-0 016	11 373	0 251
|	*|	10	-0 039	-0 091	11 570	0 315
.|	|	11	-0 013	-0 043	11 590	0 395
|*	|*	12	0 168	0 182	15 297	0 226
|.	*|	13	-0 036	-0 074	15 465	0 279
|	|	14	0 043	0 019	15 714	0 331
|*	|*	15	0 170	0 167	19 618	0 187
|**	|**	16	0 217	0 205	26 005	0 054
|	*|	17	-0 044	-0 170	26 275	0 070
|	*|	18	-0 009	-0 061	26 287	0 093
|	*|	19	-0 048	-0 067	26 605	0 114
|*	|*	20	0 195	0 180	31 968	0 044
|	|	21	0 017	-0 042	32 008	0 058
*|.	**|.	22	-0 154	-0 207	35 446	0 035
|*	|*	23	0 103	0 085	36 986	0 033
|*	|*	24	0 128	0 171	39 409	0 025
|	|	25	0 006	0 042	39 414	0 033
.|*	|	26	0 083	0 035	40 451	0 035
|*	|*	27	0 105	0 095	42 141	0 032
|**	|*	28	0 232	0 159	50 446	0 006
|	|	29	0 055	0 007	50 915	0 007

We can then go on forecasting CPI, applying the above model. The variable of the model being level form data, ex-ante forecast value is found by directly applying the above model without considering difference form. Substituting all the related values from the actual data and estimated residual terms in the above ARIMA(4,0,5) model, the forecast of CPI one period ahead (that is, for the first quarter of 2004 is found 158.3). The AEMP of ARIMA(4,0,5) model for the last quarter of 2003 is found 2.56.

Table 6.7
Ranking AEMP of various ARIMA Specifications

Data Frequency	ARIMA Specification	AEMP	Rank
Annual (2003)	ARMA(1,1)	3.51	III
Annual (2003)	ARIMA(1,1,1)	3.25	II
Quarterly (2003 IV quarter)	ARIMA(4,0,5)	2.56	I

The ARIMA (4,0,5) has least AEMP. It is most efficient in forecasting quarterly inflation in Nepal. In terms of annual data, ARIMA (1,1,1) model is found to be better in comparison to ARMA(1,0,1) model.

The weakness of B-J methodology is that, the forecast values of the distance future that is obtained by using this methodology, are found to be converging or reverting back to the mean of the series. This phenomenon is dominant particularly in the case of pure MA process rather than AR process because of the MA depending fully on ε_t terms, that are considered zero in the future. However, in the case of ARMA model, the impact of the past values of AR terms are considered but not that of the MA terms, which are assumed equal to be zero after one period ahead of the last observation. Therefore, it can be stated that B-J methodology is useful for short-term forecasting. The manner in which B-J models are used in practice is that as new information on dependent variable is available, the reliability of the forecast values increases.

(D) Vector Autoregression (VAR) Methodology

If simultaneity is found among a set of variables, they should all be treated on an equal footing; there should not be any a priori restriction between endogenous and exogenous variables (Sims, 1980). Using this argument, Sims developed the Vector Autoregression (VAR) method for data analysis. The term 'Autoregessive' is due to the appearance of the lagged value of the dependent variable on the right-hand side, and the term 'Vector' is due to the fact that we are dealing with a vector of two or more variables. In VAR analysis, all the variables are considered endogenous. VAR specification between CPIO and M1 monetary aggregate in Nepal is as follow:

$$CPIO_{1t} = \alpha + \sum_{j=1}^{k} \beta_j CPIO_{t-j} + \sum_{j=1}^{k} \gamma_j M1_{t-j} + u_{1t} \qquad (1)$$

$$M1_{1t} = \delta + \sum_{j=1}^{k} \beta_j CPIO_{t-j} + \sum_{j=1}^{k} \gamma_j M1_{t-j} + u_{1t} \qquad (2)$$

Where, the u's are the stochastic error terms, called impulses or innovations or shocks. In order to select the appropriate lag length, Akaike or Schwarz criteria are used. In the present CPIO and M1 vector autoregression estimation, only two lags of each variable are considered parsimonious. The result is presented as follows:

Table 6.8

Vector Autoregression Estimation between CPIO and M1

After adjustment (1975:3 2002:4)

	LOG(CPIO)	LOG(M1)
C	-0.224972 (-2.24501)	1.055213 (5.93973)
LOG(CPIO(-1))	0.696239 (7.14037)	0.345576 (1.99913)
LOG(CPIO(-2))	0.140256 (1.31378)	0.245885 (1.29918)
LOG(M1(-1))	-0.036958 (-0.65902)	0.592638 (5 96092)
LOG(M1(-2))	0.133227 (2.61713)	0.048983 (0.54277)
R-squared	0.997731	0.997425
Adj. R-squared	0.997644	0.997327
Akaike AIC	-3.761025	-2.615888
Schwarz SC	-3.638276	-2.493138
F Statistic	11540.0	

t-statistics in parentheses

The above estimated value of VAR coefficients are used for forecasting inflation. Data for inflation forecasting by VAR method covers the period from 1975I to 2003IV, but we have not used the values for 2003 in estimating the VAR models. The reason for not including last four quarters of 2004 in estimation is to compare the forecastability of the model. In-sample forecast for 2003I can be specified as:

$$\log(CPIO)_{2003I} = c + \log(CPIO)_{2002IV} + \log(CPIO)_{2003III} + \log(M1)_{2002IV} + \log(M1)_{2003III}$$

If we substitute estimated coefficients in the above specification from Table 29, it is found as:

$$\log(CPIO)_{2003I} = -0.225 + 0.696\log(CPIO)_{2002IV} + 0.140\log(CPIO)_{2002III} - 0.037\log(M1)_{2002IV} + 0.133\log(M1)_{2002III}$$

In-sample forecast of inflation using the above specification is given in Table 6.9:

Table 6.9
In-Sample Forecast of Inflation
(1975:I-2003:IV)

	Log(CPIO)	Log(M1)	Log (CPIO) Forecasted	Forecasted Inflation Indices	Actual Indices	Difference
(1)	(2)	(3)	(4)	(5)	(6)	(7)
2002 III	2 17026	4 89363				
2002 IV	2 18336	4 90475				
2003 I	2 19368	4 90519	2 071238	117 8	156 2	39 2
2003 II	2 18874	4 90652	2 081724	121 1	154 4	33 3
2003 III	2 18554	4 92909	2 079741	120 2	153 3	33 1
2003 IV	2 19080	4 94669	2 076169	119 2	155 2	36 0

The forecast values of CPIO from 2003I to 2003IV are shown in 5th column of Table 6.9. In order to compare forecast values with actual values, actual indices are also presented in column 6. The differences between the actual and forecast values are shown in column 7. These are significantly large showing very poor forecastability of inflation adopting only two variable VAR model of inflation.

(E) Smoothing Method

Another class of deterministic models, which is often used for forecasting, consists of moving average models. There are two versions of the smoothing method: simple smoothing (averaging) and weighted smoothing. In the simple smoothing method, a simple average of the specific number of observations (called the order) is taken, where the higher the order selection, the more the resulting series is smoothed at the cost of redundancy of observations. The smoothing formulae are as follows:

Single smoothing Model, $\hat{Y} = \frac{1}{n}(Y_t + Y_{t-1} + ... + Y_{t-k+1})$

Double Smoothing Model, $\hat{\hat{Y}} = \frac{1}{n}(\hat{Y}_t + \hat{Y}_{t-1} + ... + \hat{Y}_{t-k+1})$

Where, $\hat{Y}$ is single smoothed series, $\hat{\hat{Y}}$ is double smoothed series, 'n' is number of order, Y_t number of observation in period t, and 'k' is number of lagged observation. The choice of 'n' depends on the time path of the time series. A large 'n' should be used when there is a lot of randomness in the data, i.e. when time series are relatively stable. The moving average model is useful if we believe that a likely value for our series next month is a simple average of its values over the past observations. In moving average, higher the order of moving average, the resulting estimated series is highly smoothed in comparison to less order moving average.

In weighted smoothing method, weights of observations are taken into consideration. It is often more reasonable to have more recent values of observation playing a greater role than earlier observations. In this case it is assumed to follow geometrical progression, i.e. like $\alpha, \alpha(1-\alpha), \alpha(1-\alpha)^2...$, where α is the weight attached to the most current observation, $\alpha(1-\alpha)$ to the one period back observation, $\alpha(1-\alpha)^2$ to the two period back observation and so on. The sum of all these weights equals unit, and the values of α lie between zero and unit. A recursive formula for the computation of weighted moving average is as follows:

$$\hat{Y}_t = \alpha Y_t + (1-\alpha)\hat{Y}_{t-1}$$

The equation shows that the estimated value of Y at time 't' i.e. $\hat{Y}_t$ is a sum function of current value of the series i.e. Y_t plus estimated value of Y at time t-1 i.e. $\hat{Y}_{t-1}$ with respective weights. The closer 'α'is to 1, the more heavily the current value of Y_t is weighted in generating $\hat{Y}_t$. Thus smaller values of 'α' imply a more heavily smoothed series. The first value of $\hat{Y}_{t-1}$ has to be computed through some approximate method in order to generate $\hat{Y}_t$. Generally, first $\hat{Y}_{t-1}$ value is chosen either first actual value of the series i.e. Y_t or the simple average of the fist two actual value of the series i.e. $\hat{Y}_{t-1} = \frac{1}{2}(Y_t + Y_{t-1})$. In exponential weighted moving average, the higher the estimated series is smoothed, the lesser the weighted value of current observation is selected, so that, past observations are given weight to generate the estimated series. The procedure to choose their values is to simulate the historical data set using alternative values for 'α'. The value of 'α' is chosen in such a way that the Mean-Square-Error (MSE) of the simulated series is minimized i.e.

$$MSE = \frac{1}{n}\sum_{t=1}^{n}[Y_t - \hat{Y}_t]^2$$

Where, Y_t is actual series and $\hat{Y}_t$ is estimated series. Two period ex-ante forecast using simple moving average and exponential moving average in case of quarterly data of price index are presented in Table 6.10. Considering the presence of seasonality in the price indices, both the actual series and deseasonalized series are used to forecast the Consumer Price Index (CPI) in Nepal.

Table: 6.10

Simple Moving Average and Exponential Moving Average Forecast of CPI (1975I -2003IV)

	Forecast from Actual Index						Forecast from Deseasonalized Index				
			Moving Average		Exponential Moving Average			Moving Average		Exponential Moving Average	
Quarters	Actual Indices		3 periods	7 periods	0 2 (heavy)	0 8 (Light)	Deseas onalized Index	3 period s MA	7 periods MA	Exp 0 2 (heavy)	Exp 0 8 (Light)
1	2	3	4	5	6	7	8	9	10	11	13
2003 I	156 2	t-3	152 2	147 7	147 3	155 3	159 3	151 7	147 5	147 3	157 1
2003 II	154 4	t-2	154 4	149 9	148 7	154 6	156 0	154 5	149 9	149 1	156 2
2003 III	153 3	t-1	154 6	151 3	149 6	153 6	152 8	156 0	151 9	149 8	153 5
2003 IV	155 2	T	154 3	152 3	150 7	154 8	150 7	153 2	151 8	150 0	151 2
2004 I	Forecast	t+1	154 3	153 1	150 7	154 8	Forecast	152 2	152 3	150 0	151 2
2004 II		T+2	154 6	153 9	150 7	154 8		152 0	153 0	150 0	151 2

The forecast value for the two period ex-ante forecast taking 3 and 7 period moving average and exponential moving average with 0.2 as a heavy smoothing and 0.8 as a light smoothing parameters are presented in this study. The forecast is made using smoothed series by adaptive forecast technique. By 'adaptive' we mean that they automatically adjust themselves to the most recently available data. The last estimated value of in-sample observation is considered actual observation while deriving next period estimated value. If we forecast for a number of periods ahead, the estimated series converges into mean value. However, additional information in the actual series modifies the forecast in the future. Therefore, the shorter the forecast period, the higher will be the reliability of the forecasted value. Only two periods ahead forecasts have been presented here.

As depicted by the results, the forecast values of both the moving average and exponential moving average using actual index show an overestimation as compared to deseasonalized index (comparing column 4-7 with 9-13). The forecast based on deseasonalized index can be considered as superior for the long-run forecast. The choice between exponential moving average and simple moving average depends on the variability and the magnitude of smoothing parameters, for both actual and deseasonalized index, which are somehow considered as discretionary.

In the case of failure to find heavily smoothed series with the small value of 'α' without giving much weight to the past data point, a double exponential smoothing will come into operation. In this case 'α' can be used, and the resulting series will be heavily smoothed. In this case double as well as triple exponential smoothing formula can be used to find heavily smoothed series applying the following formula:

Double Exponential Smoothing Formula: $\hat{\hat{Y}}_t = \alpha\hat{Y}_t + (1-\alpha)\hat{\hat{Y}}_{t-1}$

Triple Exponential Smoothing Formula: $\hat{\hat{\hat{Y}}}_t = \alpha\hat{\hat{Y}}_t + (1-\alpha)\hat{\hat{\hat{Y}}}_{t-1}$

The alternative specification of the double and triple exponential smoothing specification can be shown as follows:

$$F_{t+n}(DE) = (2SE_t - DE_t) + \frac{an}{1-a}(SE_t - DE_t) \qquad (1)$$

$$F_{t+n}(TE) = (3SE_t - 3DE_t + TE_t) + \frac{an}{2(1-a)^2}[(6-5a)SE_t - (10.8a)DE_t + (4-3a)TE_t] + \frac{a^2n^2}{2(1-a)^2}(SE_t - 2DE_t + TE_t) \qquad (2)$$

Where, *DE* and *TE* stand for double and triple exponential smoothing series. Using the above formula, four period ahead forecast value of CPI have been presented in table 6.11.

Table: 6.11

Ex-ante Forecast Using Single and Double Exponential Moving Average (1975I -2003IV)

Quarterly Forecasts	Simple Double Smoothing Method	Double Exponential Smoothing Method	Triple Exponential Smoothing Method
2004 I	158 2	156 6	155 7
2004 II	159 9	157 6	156 3
2004 III	161 6	158 7	156 8
2004 IV	163 3	159 8	157 4

Where, *DE* and *TE* stand for double and triple exponential smoothing series. Using above formula, four period ahead forecast value of CPI have been presented in Table 6.11. The forecast values are found to be decreasing when the sizes of the exponential parameters are increasing.

The smoothing method of forecast is considered as somehow arbitrary in the sense that the choice of smoothing and exponential parameters is considered as discretionary in nature. These parameters depend on the extent of variability of the data series and effect of seasonal variation in the data frequency. This method lacks theoretical justification and hence can not be considered as handy tool of forecasting. The forecast values, derived using this method, are found to be mean reverting as a result of an increase in the magnitude of smoothing and exponential parameters.

To sum up the forecasting models of inflation in Nepal, time series forecasting method is found to be better than econometric forecasting particularly for the short-term forecast. Time series forecasting method is based on an attempt to predict the values of a variable from past values of the same variable. Among the several methods of time series forecasting, B-J methodology is useful for short-term forecasting. This study found that the ARIMA(4,0,5) data generating process fits better forecasting of inflation in Nepal. However, for the long-term forecast, regression method (a variant of econometric forecasting) would be a better tool of forecast.

2. CORE INFLATION

In an inflation-targeting regime, future inflation is the final target, but current inflation is a determinant of future inflation, and thus can be used as one of a set of indicators to suggest whether policy should be tightened or loosened immediately in order to achieve the final target in the future (Hogan, 2000). Such considerations raise the question of what measure of inflation would be the best to use as this indicator. Most inflation-targeting countries employ a definition of 'core' or 'underlying' inflation that seeks to capture the underlying trend in inflation.

Core inflation is generally associated with expectations and demand pressure components of measured inflation and excludes supply shocks (Roger, 1998). Core inflation includes a persistent or steady element of inflation which will tend to be incorporated into expectation and, consequently, will be comparatively benign. Intermittent or transient inflation, however, will be much less benign, precisely because it will be less readily anticipated. Therefore, core inflation and trend inflation are essentially synonymous. Eckstein (1981) defined core inflation as ...the trend increase of the cost of the factors of production. He decomposed measured inflation into (a) core contributed by factor prices, (b) a portion attributable to aggregate demand and (c) a portion which could be attributed to supply shocks. Core Inflation is that component of measured inflation which is output neutral over the medium to long term, it must be the component of inflation that feeds or reflects inflation expectations (Quah and Vahey (1995). In summing up, the core inflation rate should exhibit more persistence or less variability than the aggregate measured inflation rate. Supply shocks are the most important source of relative price changes. Therefore, supply-driven relative price changes affecting the aggregate inflation rate should only have a transient impact on the aggregate inflation rate. Bryan and Cecchetti (1994), Cecchetti (1997) agreed on the concept of core inflation as one that it should capture just the component of price change that is common to all items and exclude changes in the relative prices of goods and services.

Core inflation should track the component of overall price change that is expected to persist for several years and therefore be useful for short-term and medium-term inflation forecasting (Blinder, 1997, Bryan and Cecchetti, 1994). Folkertsam and Hubrich (2001) have given the following reasons for not using CPI as an ideal measure of inflation for monetary policy purpose: (a) CPI is a noisy signal of the inflation pressures in an economy, such as seasonal influences, changes in the indirect tax rate, purely relative price changes, etc. (b) Monetary policy operates on inflation with a long and variable time lag. Therefore, from the perspective of policymakers, a measure of inflation (like core inflation) would be useful for a leading indicator of future CPI. (c) Credibility is crucial to central bank performance, an operational inflation concept which only reflects price level movements for which the monetary authority is accountable.

(A) Measurement of Core Inflation

A host of methods have been developed to estimate core inflation. The methods suggested are either based on cross-sectional information, i.e. the distribution of individual price changes with respect to some reference period, on univariate or multivariate time series, or on pooled cross-sectional and time-series data (Wynne, 1999). Cross-sectional methods attempt to refine the CPI, aiming to eliminate its transitory movements and to increase the signal-to-noise ratio. Three different types of inflation measures rely on purely cross-sectional information. The most well-known type encompasses price indices that simply exclude allegedly volatile components, such as food or energy prices. A more sophisticated class of measures contains the various trimmed mean estimators. These measures do not *a priori* exclude specific commodities from the price index once and for all, but remove those commodities of which the observed price change relative to the previous period is an 'outlier' (Bryan, Cecchetti and Wiggins, 1997). The weighted median belongs to this class of inflation measures.

Finally, there are price indices which use the full cross-sectional information but aggregate the price changes with weights which are inversely related to their volatility. Clearly, since the weights of these price indices are not derived from budget shares, the scope of these inflation measures is not restricted to consumer prices. Univariate time-series methods remove high-frequency noise from the CPI in- inflation series by smoothing or filtering with, for example, moving averages or Kalman filters. The smoothed series are estimates of the core inflation process. Methods combining cross-sectional and time-series information apply the dynamic factor model to price data. The common component in all price changes is interpreted as core inflation.

(i) Stochastic Measures of Core Inflation

Based on the premise that extreme price changes are not indicative of the persistent component of inflation Bryan and Cecchetti (1993), Roger (1998) studied the measure of core inflation by removing or reducing the weight of the components with extreme price changes. Under this assumption "trimmed mean" is used which excludes a proportion of each tail of the cross-section distribution of price changes (i.e. extremely low and extremely high change) and takes the weighted average of price changes of the rest of the commodities to estimate the central tendency or underlying core of distribution.

A robust measure of core inflation is devised through the statistical measures of trimmed mean or weighted median. The use of such a measure reflects the intuition that the type of shocks that may cause problems with price measurement is infrequent and agents do not instantly adjust prices to every change in circumstances as there are 'menu costs' of price adjustment (Ball, 1991). If the distribution of shocks is skewed the aggregate price level will temporarily deviate resulting in transitory movement of headline inflation from its long-run trend (Alvarez and Matea 1997).

The stochastic measure of core inflation presupposes the individual price changes that involve a common, generalized inflation component plus idiosyncratic relative price shocks. Therefore, the distribution of consumer price changes is almost always found to be highly kurtotic. In such circumstances, a variety of measures such as the median or trimmed-mean will be far more efficient as estimators of the general tendency of price changes than the CPI mean-based estimator.

(ii) Exclusion-Based Measures

By this measure, core inflation is calculated by excluding certain commodities/components from the basket. Such commodities are believed to be unrepresentative of market-induced inflation trend over a short horizon. Such excluded commodities are assumed to be either seasonal and relatively more volatile leading to quick reversal, or else as infrequent and sudden changes due to administrative control (Brayn and Cecchetti 1993, Kearns, 1998).

(B) Core Inflation as Generalized Price Movement

Under this concept, core inflation measurement is made through some specific adjustment on measured inflation for some supply shocks or re-weighting all individual components of measured inflation, according to their contributions to the common price trend with a view to eliminating or diminishing the effect of supply shocks. Some shocks like international trade prices (New Zealand), Changes in exchange rate (Sweden), indirect taxes etc., are having temporary effects inflationary effect.

(C) Estimation of Core Inflation in Nepal

This section attempts to find out the components of Consumer Price Index (CPI) that are to be excluded before calculating core inflation in Nepal. The necessary condition for the calculation of core inflation is to find out the noisy components of CPI that are independent of effect of any aggregate demand management policies adopted by the government. In this study, core inflation is not calculated because it is outside the scope of this study. However, one of the preliminary jobs for calculation of core inflation using exclusion method is discussed.

The standard approach to estimation of core inflation by exclusion is to remove the noisy elements from the headline rate of inflation. The noise, measured by the Coefficient of Variation (CV) reveals that four out of 22 components are found to be excluded.

Table 6.12

Measures of Dispersion of Major Commodity Groups of CPI in Nepal

S.N	Commodity Name	Weight	Mean	S.D.	Median	Coefficient of Variation (CV)	Proportion of CV of Commdt. out of total CV
1	**Cleaning Supplies**	1.26	2 05	10 75	-0 06	524 39	**157.76**
2	**Cloths**	2.28	1 90	9 03	0 64	475 26	**142.98**
3	**Vegetables**	7.89	3 82	17 73	0 77	464 14	**139.63**
4	**Milk and Milk Products**	4.05	2 65	11 86	0 85	447 55	**134.64**
5	Sugar	1.21	1 77	7 30	0 87	412 43	124 08
6	Oil and Clarified Butter	3.07	2 48	10 04	1 18	404 84	121 79
7	Rice	14.16	1 94	6 94	0 99	357 73	107 62
8	Transport and Communication	4.03	2 20	7 75	0 22	352 27	105 98
9	Spices	1.85	2 44	8 39	2 63	343 85	103 44
10	House Rent	4.19	2 77	9 51	0 00	343 32	103 29
11	Beverages	2.28	2 14	7 19	0 20	335 98	101 08
12	Grains and cereal product	3.8	1 90	6 26	0 63	329 47	99 12
13	Foot-ware	2.2	1 52	4 82	0 23	317 11	95 40
14	Pulse	2.73	2 50	7 63	0 66	305 2	91 82
15	Meat, Fish and Eggs	5.21	2 41	7 15	0 18	296 68	89 25
16	House Furnishing and H. goods Household goods	3.5	2 04	5 70	0 32	279 41	84 06
17	Restaurant Meals	6.91	2 63	6 94	0 66	263 88	79 39
18	Fuel, Light and Water	5.92	2 82	6 70	0 46	237 59	71 48
19	Cigarettes	1.66	1 65	3 72	0 54	225 45	67 82
20	Clothing and Sewing Services	6.64	1 82	4 08	0 46	224 18	67 44
21	Medical and Personal Care	8.03	1 80	3 74	0 80	207 78	62 51
22	Education, Reading Materials and Recreation	7.09	2 22	3 64	0 89	163 96	49 33

The commodities which have CV over the 125 percent are practiced to be excluded. On the basis of this criterion, cleaning supplies, clothes, vegetables and milk and milk products are the noisiest components of the CPI of Nepal. Measurement of inflation, after excluding these noisy components, gives the core inflation rate which is considered to be controlled aggregate demand management, such as controlling money supply can control the rate of inflation.

Summing up, the basic objective of this section is to introduce various methods applicable for the calculation of core inflation rather than applying it in the Nepalese context. Measurement of core inflation and examining its implication would be an important area of further research.