Pakistan Journal of Applied Economics, Vol.32 No.2, (169-190), Winter 2022

FORECASTING GROUP-WISE IMPORTS AND EXPORTS OF PAKISTAN

Ateeb Akhter Shah SYED*, Humaira KIRAN and Sumbal QURESHI*****

Abstract

This study forecasts imports and exports of Pakistan at a disaggregated level. Both in and out-of-sample forecasts are produced using conventional econometric time-series models and the Artificial Neural Network (ANN), a machine learning approach. The forecast performance is reported using the Root Mean Squared Error (RMSE). Given improved forecasts by the 'iterative optimization' nature of the long and short-term method, ANN outperforms other model in-sample. For the out-of-sample period, the autoregressive (AR) and ANN model outperforms other models for import groups, while all univariate approaches outperform each other for two out of six subgroups in out-of-sample forecasts. Hence, performing equivalently well for export groups.

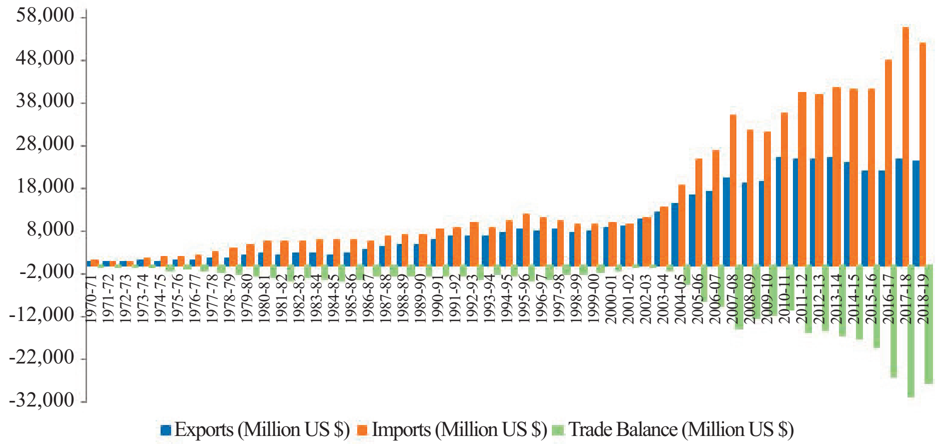
Keywords: Pakistan, Imports, Exports, Forecast, ANN.
JEL Classification: B22, B23, F17.

I. Introduction

Trade is a very important component of the economy of any country as it contributes in the national growth and stabilization of the economy to which, and Pakistan is no exception. International trade is one of the most important sources of foreign exchange income which helps in the stabilization of the balance of payments and increases employment opportunities in a country. Since independence, a trade deficit is one of the main problems of Pakistan's economy. Although Pakistan has achieved commendable progress in improving its trade situation over time; however, the imports are still more in comparison with exports, which is why Pakistan is facing a trade deficit and the impact of an increasing trade deficit on the economy has become more severe with the passage of time [Ali (2009)].

To better understand the importance of trade for the economy of Pakistan, we plot the exports, imports and trade balance for the past 49 years in Figure 1.

* Deputy Director, Research Dept., **Assistant Director, Statistics and Data Warehouse Dept., State Bank of Pakistan, Karachi ***M.Phil. Scholar, Dept. of Computer Science and Information, University of Punjab, Lahore, Pakistan.



Source: State Bank of Pakistan (Economic Data – Balance of Trade Archive).

FIGURE 1

Exports, Imports and the Trade Balance of Pakistan

We observe that throughout these 49 years, the balance of trade has been negative. It was US $-504,000,000 in FY70-71 and has further deteriorated to US $-2,493,000,000 in FY90-91, which is about a 5-fold increase. The trade deficit increased to a record high of US $-30,903,000,000 in FY17-18. The exports increased merely by 12.0 per cent in FY18, whereas the imports grew by 15.7 per cent. Some additional measures were taken to slow down the imports of the country at the end of 2018 by the new government in power. As a result, In FY19, the imports and exports fell by 6.8 and 2.1 per cent, respectively. A huge trade deficit threatens a default for a country like Pakistan, which already has a very low amount of foreign exchange reserves. Thus the descriptive, inferential analysis and forecasting of the elements of trade deficit and policy recommendations on the basis of such analysis become important.

This study contributes in the existing literature on forecasting imports and exports by considering all the groups of imports, overall imports, all groups of exports and overall exports and using different time series models. Furthermore, this paper strives to find the best forecasting model, both in-sample and out-of-sample, for all groups of imports and exports and for overall imports and exports. In this study, we exercise to forecast group-wise exports and imports of Pakistan using a monthly dataset.

There have been several studies that have forecasted exports and imports of Pakistan. Ghauri, et al., (2020) used a monthly dataset from July 2002 to June 2019 to forecast exports and imports of Pakistan using two univariate models. Box and Jenkins (1976) Autoregressive Integrated Moving Average (ARIMA) model and the AR model with seasonal dummies are used. The study reports the RMSE and Mean Absolute

Error (MAE). According to the MAE, the exports are best forecasted by the ARIMA model, while imports are forecasted best by AR model with seasonal dummies. According to the RMSE, they find that the exports are best forecasted by AR model with seasonal dummies; however, for imports, both the ARIMA and AR model with seasonal dummies depict similar forecast performance.

Farooqi (2014) also forecasted the total imports and exports series for the economy of Pakistan. He used the ARIMA model by applying Box and Jenkins (1976) approach. The sample period of his study extends on an annual basis from 1947 to 2013. The forecasts are obtained for the five years, 2014 to 2018. The author has applied the likelihood ratio test and minimized Akaike's Information Criteria (AIC) to find the models that fit best to the import and export series. He finds that the appropriate model to predict annual export and import of Pakistan are ARIMA (2, 2, 2) and ARIMA (1, 2, 2) respectively.

Mehmood (2012) forecasted the exports from Pakistan to South Asian Association for Regional Cooperation (SAARC) countries. The data used in his study is annual data from 1975 to 2009. The study uses the Box and Jenkins (1976) methodology to estimate the models and reports the results in terms of Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE), RMSE and MAE. According to all these criterion, the study finds that ARIMA (1, 1, 4) is the model that best forecasts Pakistan's exports to the SAARC countries.

Numerous other studies have forecasted the overall imports and exports of Pakistan. However, all of the studies either use annual or monthly overall exports and imports datasets of Pakistan. To the best of our knowledge, none of these studies uses a monthly dataset for predicting the group-wise exports and imports of Pakistan. Furthermore, in the literature, no study has used the univariate and the multivariate econometric models; compared their forecast performance against the machine learning model/models for exports and imports of Pakistan. This study fills these gaps by providing forecasts of overall and group-wise exports and imports using the traditional econometric models and a machine learning model for the economy of Pakistan.

The rest of the study is organized as follows. Section II presents the data of the study. Section III discusses the methodology employed in the study, i.e., Box-Jenkins Methodology for Univariate, Univariate Time-Series and Multivariate Time-Series Models. Results and analysis are presented in Section IV. Finally, Section V concludes the study and provides relevant policy recommendations.

II. Data

Monthly data (in USD million) of all groups and overall imports and exports from July 2006 to June 2019 is taken from the monthly statistical bulletin published by State Bank of Pakistan (SBP). All nine groups of imports and five groups of exports are considered for this research. As the primary source of this data set is the Pakistan

Bureau of Statistics (PBS), the classification of groups is performed by PBS. The information about the classification is not publicly available on the website of either PBS or SBP. However, the details are in Table A-1 and A-2 (Appendix A).

The details of import groups and commodities that belong to each of the specific clusters are in Table A-1 (Appendix A). Imports of Pakistan are classified into a total of nine groups; eight of these groups contain 53 commodities. The remaining commodities are part of all other items grouped - the details of export groups and commodities that belong to each sub-group are in Table A-2 (Appendix A). Exports of Pakistan are classified into five groups that contain 53 commodities and the remaining commodities are contained in the sub-group called all other items group.

The determinants of imports are listed in Table A-3 in Appendix A. These variables are the Quantum Index of Large-Scale Manufacturing (LSM), Real Effective Exchange Rate (REER) and Remittances (R). They are used in multivariate analysis. The source of this data is SBP. These determinants of imports are chosen following the work by Aker (2008). Table A-4 (Appendix A) contains the determinants of export used in multivariate analysis, and the data on these variables are obtained from SBP. These variables are LSM, REER and Indirect Taxes (IT). The determinants of exports are chosen following the work by Majeed, et al., (2006). All the variables (exports, imports and determinants) in our analysis are used in the logged form.

III. Methodology

The forecasting approaches/models used in this study are briefly discussed below.

1. Box-Jenkins Methodology for Univariate Models

Box and Jenkins (1976) suggested five steps: testing the data for stationary, identification of model parameters, estimation of the model parameters, model diagnostics and forecasting.

In the first step, stationary is checked. Many tests can be applied to determine stationarity. If the series is non-stationary, we have to make it stationary. The second step is model identification, in which visualization of data is performed, and the parameters are identified by exploratory analysis. The most commonly used tools for identifying the model are autocorrelation and partial autocorrelation function. In the next step, the parameters are estimated by using Ordinary Least Square (OLS), Maximum Likelihood Estimation (MLE) or Yule-Walker's estimation. In the diagnostic section, the plot of residuals is tested to see if they are white noise. Bayesian Information Criteria (BIC), AIC and quality measures are used for model diagnostics too. In the last step, a suitable model is selected for forecasting. Box-Jenkins developed the above procedure for applying ARIMA models and we followed these steps to choose our ARIMA/ARIMA models.

2. *Univariate Time-Series Models*

One of the distinct features of time series data is that a variable observed over time may contain useful information to predict future values. Therefore, we first check basic useful models for analyzing time-series data like AR and ARIMA. As mentioned earlier, we will use the Box-Jenkins methodology for the application of these models for each observed export group, overall exports, import groups and overall imports. The models are:

a) *Autoregressive Model*

AR model is given by regressing the variable of interest on its previous lags. For example, *AR(p)* or, equivalently, *ARIMA (p,0,0)* is given by the Equation (1):

$$Y_t = c + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \ldots + \Phi_p Y_{t-p} + \varepsilon_t \quad (1)$$

where c is the intercept, $\Phi_1, \Phi_2, \ldots, \Phi_p$ are the parameters and εt is the error term. In *AR* modeling, lags are used to predict the value of the current time period. If only one lag is used in the functional form of the AR model, it is referred to as AR (1). If the last two past values are used, then it is AR (2). Therefore, in general, a '*p*' order autoregressive model contains '*p'* lags of the dependent variable on the right-hand side in the form of linear regression and is denoted by *AR(p)*.

b) *Autoregressive Integrated Moving-Average*

ARIMA model considers the AR and Moving Average (MA) components to model a time series. ARIMA model involves the integrated factor and is a generalization of an ARMA model. The ARIMA model is also an excellent forecasting model compared to the AR model as it considers both AR and MA components.

As discussed above, ARIMA models comprise two parts: lagged values of the variable of interest (the AR component) and lagged values of the error term (the MA component). It is suggested by Crawford and Fratantoni (2003). ARIMA technique is utilized to design a huge category of stationary time series as long as a suitable order of '*p*' representing the number of AR terms and '*q*' representing the number of MA terms are identified properly [Makridakis and Hibon (1997)]. The ARIMA (*p,q*) can be presented in Equation (2):

$$Y_t = c + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \cdots + \Phi_p Y_{t-p} + \theta_1 \mu_{t-1} + \theta_2 \mu_{t-2} + \cdots + \theta_q \mu_{t-q} + \mu_t \quad (2)$$

with $E(\mu_t) = 0$; $E(\mu_t^2) = \sigma^2$ and $E(\mu_t, \mu_s) = 0$, such that $t \neq s$. Φ_t and θ_t are the coefficients, p and q are the orders of *AR* and *MA* polynomials, respectively, 'μ_t' is the pattern of irregular shocks that are supposed to be identically and independently distributed with zero mean and constant variance, which is uncorrelated.

c) Artificial Neural Network

Long and Short-Term Memory (LSTM) is a type of Recurrent Neural Network (RNN) architecture. It was developed to model temporal sequences more accurately with their long-range dependencies than the traditional RNNs. A simple RNN has a huge problem of vanishing gradients. The further the RNN model is traced back, the lower the gradient becomes. As a result, it becomes hard to train the weights throughout the network. The problem can be tackled by initializing the weights. Echo state networks can also be used to solve this gradient vanishing problem. We have opted for LSTM, which was specifically developed for this purpose. LSTM forgets all irrelevant data and remembers past knowledge that has been passed through the network.

LSTM contains special 'memory blocks' units in the hidden recurrent layer. These memory blocks consist of 'memory cells' with self-connections which store the network's temporal state. In addition to the temporal state, special multiplicative units called gates are also stored. A typical network consists of four different gates: input gate, output gate, cell state, and forget gate. These gates filter out the useless data and only keep what is required. After identifying the data, it can pass information down the chain of sequences in order to make predictions. The difference in the working of a simple RNN and LSTM network is that along with the hidden state; the internal cell state is also passed. Given the benefits of the LSTM over the simple RNN, we used the LSTM to model the neural network models used in this paper.

Finally, the appropriateness of each of the time-series models is ensured by testing the residuals from equations of each of these time-series models for each group (exports and imports) using the Ljung-Box (1978) test with the null hypothesis of no autocorrelation. The Q-statistics show that the residuals from each equation of these time-series models for each group (exports and imports) are white noise. Furthermore, the AR, ARIMA and VAR models selected for each group are not reported to conserve space. However, this information is available upon request.

3. Multivariate Time-Series Model

The multivariate forecasting time-series model used in our analysis is the Vector Autoregressive Model (VAR).

a) Vector Autoregressive Model

Using their past values, the VAR model permits feedback among the dependent and independent regressors. Introduced by Sims (1980), the VAR models have been the workhorse of macroeconomic analysis for more than 30 years. Generally, for this model, it is assumed that all the variables in the model are endogenous. Appendix B contains details about the VAR models. The model can be written as Equation (3):

$$Y_t = c + \theta_i \sum_{i=1}^{n} Y_{t-i} + \varepsilon_t \tag{3}$$

where Y_t is a vector of endogenous variables at time t, θ_i*'s* are the parameters and ε_t are the uncorrelated white noise disturbance terms. If the dataset of interest contains some or all non-stationary variables, co-integration is tested among the variables of the model. If there is no co-integration, Enders (2009) suggest that the VAR model in first differences is preferable. The VAR model in first differences can be written as Equation (4):

$$\Delta Y_t = c + \theta_i \sum_{i=1}^{n} \Delta Y_{t-i} + \varepsilon_t \tag{4}$$

where ΔY_t is a vector of first difference of the endogenous variables at time t, θ_i*'s* are the parameters and ε_t are the uncorrelated white noise disturbance terms.

IV. Results and Analysis

1. Seasonality Test

Upon analyzing the graphs of the series used in this study, some seemed to have seasonality. However, we did not rely on the graphs and tested each series for seasonality. Therefore, in this section, an OLS regression containing constant and 11 dummies for January to November is estimated for each export and import series. A joint significance test on the coefficients of these dummies is performed. The regression can be written as Equation (5):

$$y_t = \alpha_k + \sum_{k=1}^{11} \gamma_k D_{kt} + \varepsilon_t \tag{5}$$

Where α_k and γ_k are the constant and coefficients on '*kth*' seasonal dummy, respectively, D_{kt} is the value of '*kth*' seasonal dummy, ε_t is an error term for the current period, $\sum_{k=1}^{11} \gamma_k D_{kt}$ are seasonal components.

Table 1 contains the results of the Wald test (joint significance test) that is used for checking seasonality in each import series. For all import groups, the test does not reject the null hypothesis. Therefore, we conclude that there is no seasonality in any subcategory of imports and overall imports. Table 2 contains the results of the seasonality test for all groups of exports. For the food group, textile group and overall exports the null hypothesis is rejected, and the test finds at least one of the coefficients on the seasonal dummies to be significant. Hence, these series contain seasonality. However, there is no seasonality found in the petroleum and coal group, other manufacturers group and all other items group.

TABLE 1

Results of Wald Test for Seasonality (Imports)

	Statistic Value	P-Value
Food Group	0.5036	0.8977
Machinery Group	0.4895	0.9067
Transport Group	0.6388	0.7924
Petroleum Group	1.1575	0.3241
Textile Group	1.9024	0.0500
Agricultural and Other Chemical Group	0.9142	0.5292
Metal Group	1.2829	0.2425
Miscellaneous Group	1.0132	0.4391
All Other Items	0.9489	0.4968
Overall Imports	0.9897	0.4598

Source: Authors' estimation.

2. *Stationarity Test*

Testing the data for stationarity is the second step in our analysis. For this purpose, literature provides different unit root tests; however, Dickey and Fuller (1981) is the most popular and widely used unit root test. The null hypothesis of the Augmented Dickey-Fuller (ADF) test is that series is non-stationary against the alternative that the series is stationary or trend stationary.

TABLE 2

Results of Wald Test for Seasonality (Exports)

	Statistic Value	P-Value
Food Group	4.1003	0.0000
Textile Group	2.0139	0.0311
Petroleum Group and Coal	1.6399	0.0936
Other Manufactures Group	1.3982	0.1825
All Other Items	0.8120	0.6281
Overall Exports	2.3760	0.0099

Source: Authors' estimation.

a) ***Augmented Dickey-Fuller Test***

Said and Dickey (1984) augmented the basic autoregressive unit root test to accommodate general ARIMA (*p, q*) models whose orders are unknown. Their test is referred to as the ADF test. The three possible types of ADF tests are as follows in Equations (6), (7) and (8):

$$\Delta y_t = \rho y_{t-1} + \sum_{i=1}^{p} \phi_i \Delta y_{t-i} + \varepsilon_t \tag{6}$$

$$\Delta y_t = \alpha_0 + \rho y_{t-1} + \sum_{i=1}^{p} \phi_i \Delta y_{t-i} + \varepsilon_t \tag{7}$$

$$\Delta y_t = \alpha_0 + \beta t + \rho y_{t-1} + \sum_{i=1}^{p} \phi_i \Delta y_{t-i} + \varepsilon_t \tag{8}$$

where, y_t is the time series that is tested for unit root, α_0 is the intercept, βt is the coefficient with time trend and ε_t is the error term. The following tables contain the results of the stationarity tests.

Table 3 contains the ADF test results for the import groups. The results show that food, transport, textile, agricultural and chemical and all other items groups are stationary in log levels. The test statistic for these series is greater than the critical values in absolute terms. Machinery and petroleum group are also stationary when drift is considered. However, they become non-stationary in the presence of both trend and drift, so it is concluded that the machinery and petroleum group are stationary in log levels with a drift.

TABLE 3

Results of Augmented Dickey-Fuller Test (Imports by Commodity Groups)

	In Log-Levels (Trend and Drift)		In Log-Levels (Drift)	
	Statistic	Critical-Value	Statistic	Critical-Value
Food	-5.51	-3.45	-4.44	-2.89
Machinery	-2.37	-3.45	-3.24	-2.89
Transport	-6.91	-3.45	-3.39	-2.89
Petroleum and Coal	-3.41	-3.45	-3.37	-2.89
Textile	-5.24	-3.45	-3.79	-2.89
Agricultural and Chemical	-5.21	-3.45	-3.04	-2.89
Metal	-5.82	-3.45	-1.90	-2.89
Miscellaneous	-3.75	-3.45	-2.61	-2.89
All Other Items	-4.75	-3.45	-3.38	-2.89
Overall Imports	-4.10	-3.45	-2.62	-2.89

Source: Authors' estimation.

The metal group, miscellaneous group and overall imports are non-stationary in log levels in the ADF test with drift while stationary when both trend and drift are added. In such a case, Enders (2009) suggests performing the joint significance test on the coefficient of the lagged dependent variable and the trend term. Therefore, the joint significance test is performed on the coefficients of the lagged dependent variable and the trend term, that is, the null hypothesis '$\beta = \rho = 0$' in the third form of the ADF test. In Table 4, the test statistic for all import groups and overall import is greater than the critical values in absolute terms. Therefore, it is concluded that all import groups and the overall import are stationary in 1st difference.

Table 5 shows that the test statistic for all exports except the textile group, petroleum and coal group and overall export is greater than the critical value in absolute terms. Therefore, all export groups except these three are stationary in log levels. Table 6 presents the results of the ADF test for exports in the first differences. The results reveal that all the export groups and overall exports are stationary in the first differences. Table 7 shows the result of the ADF test for the selected determinants of exports and imports used in the VAR models.

The results show that the test statistic is greater than the critical values in absolute terms. Therefore, it fails to reject the null hypothesis and concludes that all determinants are non-stationary in log levels.

Table 8 shows the results of ADF tests for the determinants of exports and imports in first differences. The test statistic for each variable is greater than the critical value in absolute terms. Therefore, all these variables are stationary in first differences.

TABLE 4

Results of Augmented Dickey-Fuller Test (Imports by Commodity Groups)

	In First Differences (Trend and Drift)		In First Differences (Drift)	
	Statistic	Critical-Value	Statistic	Critical-Value
Food	-10.34	-3.45	-10.33	-2.89
Machinery	-13.26	-3.45	-18.53	-2.89
Transport	-14.43	-3.45	-14.48	-2.89
Petroleum and Coal	-10.41	-3.45	-10.45	-2.89
Textile	-8.22	-3.45	-8.25	-2.89
Agricultural and Chemical	-10.77	-3.45	-10.81	-2.89
Metal	-9.48	-3.45	-9.50	-2.89
Miscellaneous	-11.60	-3.45	-11.60	-2.89
All Other Items	-10.91	-3.45	-10.93	-2.89
Overall Imports	-11.02	-3.45	-11.05	-2.89

Source: Authors' estimation.

TABLE 5

Results of Augmented Dickey-Fuller Test (Exports)

	In Log-Levels (Trend and Drift)		In Log-Levels (Drift)	
	Statistic	Critical-Value	Statistic	Critical-Value
Food	-5.83	-3.45	-4.39	-2.89
Textile	-2.80	-3.45	-2.21	-2.89
Petroleum and Coal	-2.23	-3.45	-2.22	-2.89
Other Manufactures	-3.44	-3.45	-3.14	-2.89
All Other Items	-3.10	-3.45	-3.02	-2.89
Overall Exports	-2.84	-3.45	-2.82	-2.89

Source: Authors' estimation.

TABLE 6

Results of Augmented Dickey-Fuller Test

	In First Differences (Trend and Drift)		In First Differences (Drift)	
	Statistic	Critical-Value	Statistic	Critical-Value
Food	-8.04	-3.45	-8.21	-2.89
Textile	-12.17	-3.45	12.21	-2.89
Petroleum and Coal	-9.55	-3.45	-9.58	-2.89
Other Manufactures	-11.86	-3.45	-11.87	-2.89
All Other Items	-10.54	-3.45	-10.56	-2.89
Overall Export	-11.33	-3.45	-19.90	-2.89

Source: Authors' estimation.

TABLE 7

Results of Augmented Dickey-Fuller Test
(Determinants of Exports and Imports)

	In Log-Levels (Trend and Drift)		In Log-Levels (Drift)	
	Statistic	Critical-Value	Statistic	Critical-Value
LREER	-0.78	-3.45	-0.94	-2.89
LLSM	-2.58	-3.45	-0.30	-2.89
LIT	-3.53	-3.45	-0.91	-2.89
LR	-2.07	-3.45	-0.69	-2.89

Source: Authors' estimation.

TABLE 8

Results of Augmented Dickey-Fuller Test
(Determinants of Exports and Imports)

	In First Differences (Trend and Drift)		In First Differences (Drift)	
	Statistic	Critical-Value	Statistic	Critical-Value
DREER	-4.89	-3.45	-4.67	-2.89
DLSM	-15.32	-3.45	-15.30	-2.89
DIT	-7.97	-3.45	-8.02	-2.89
DR	-4.12	-3.45	-4.07	-2.89

Source: Authors' estimation.

The mean of the petroleum import group changed suddenly at several instants and remained in the new range for some time period; making it susceptible to structural breaks. Therefore, to check its stationarity following Zivot and Andrews (2002), a unit root test is applied with structural breaks on first difference of the petroleum import group. The test shows the series is stationary, with breaks in January 2009 and January 2015. The results are presented in Table 9.

TABLE 9

Results of Stationarity with Structural Breaks

Import Group	Test Type	Test Statistic	Critical Value	Break Point
Petroleum Group	Break in Trend	-9.67	-4.42	Jan-09
	Break in Drift	-9.77	-4.80	Jan-15

Source: Authors' estimation.

3. *Forecast Performance*

a) *In-sample Forecast Performance*

In this section, we report the results of Forecasting models. The forecast is performed recursively using the estimated models. It involves step-by-step forecasting; one period forecast is obtained; using this, a two periods forecast is performed and the process continues till all the forecasts have been obtained. The forecasts are computed by applying models to each series. The training data is from August 2007 to June 2018 and forecasts are computed for both the in-sample (August 2007 to June 2018) and the out-of-sample that consists of 12 months from July 2018 to June 2019.

Table 10 contains the in-sample forecasting performance of the competing models for export groups and overall exports. The numbers in the table represent RMSE. When read across rows, the number in bold shows the model that performed best in forecasting the corresponding export group. We observe that the ANN model outperforms other competing forecast approaches for the food group. Therefore, it is the best model to forecast food export groups. The VAR model outperforms the ANN, AR and ARIMA models for the textile group. Therefore, it is the best forecast model for the textile group. We observe that the ANN model outperforms the competing approaches for the petroleum and coal groups. Hence, it can be said that the ANN model obtains the best forecasts of this series. The ANN model outperforms the competing forecast models for other manufacturers' export groups. Therefore, it is the best model to forecast this group. The ANN model again outperforms other competing models for all other export group items. For the overall exports group, the VAR model outperforms the competing models. Therefore, it is the best model to forecast this group.

Measures such as RMSE provide a better idea of how well the models perform against other models in forecasting the series of interest. Results show that for the in-sample forecast exercise, the ANN model outperforms other models for four out of six export groups. Therefore, it is the best model to forecast export groups in-sample for Pakistan.

TABLE 10

In-Sample Forecast Performance Results for Export Groups

	ANN	AR	ARIMA	VAR
Food Group	0.1465	0.2539	0.2277	0.1476
Textile Group	0.0879	0.9524	0.0925	0.0849
Petroleum Group and Coal	0.1011	1.4125	1.3588	1.4161
Other Manufacture Group	0.0989	0.2013	0.1925	0.1501
All Other Items	0.0830	0.2882	0.2339	0.1859
Overall Exports	0.0930	.0.0954	0.0910	0.0876

Source: Authors' estimation.

Note: In-sample forecast accuracy for h = 1 for exports. Each entry shows RMSE for the forecasting models used. VAR model is used for the series of interest and the determinants of exports (LSM, REER and Indirect Taxes). One machine learning, two univariate and one multivariate econometric model are also considered; the models are ANN, AR model, ARIMA model and the VAR model. Bold entries indicate the lowest error measure achieved by the competing approach for the variable of interest across each row.

Table 11 contains in-sample forecast performance results for all the models for groups of imports. The ANN outperforms all the other models for all the import groups. The performance of ANN using the LSTM is better than other models for the in-sample data. The good performance observed in these cases is due to the 'iterative optimization' nature of LSTM. Because of these optimizing iterations that minimize RMSE, an under-

TABLE 11

In-Sample Forecast Performance Results for Import Groups

	ANN	AR	ARIMA	VAR
Food Group	0.1357	0.2429	0.2430	0.1828
Machinery Group	0.0904	0.3224	0.3072	0.1635
Transport Group	0.0563	0.4909	0.4923	0.4919
Petroleum Group	0.1106	0.3161	0.3176	0.2402
Textile Group	0.0764	0.2705	0.3149	0.1719
Agricultural and Other Chemical Group	0.0875	0.1607	0.1912	0.1452
Metal Group	0.0989	0.3014	0.2910	0.1887
Miscellaneous Group	0.0947	0.2549	0.2528	0.1594
All Other Items	0.1043	0.2629	0.2620	0.1866
Overall Imports	0.0903	0.2066	0.1978	0.1210

Source: Authors' estimation.

Note: In-sample forecast accuracy for h = 1 for imports. Each entry shows RMSE for the forecasting models used. VAR model is used for the series of interest and the determinants of imports (LSM, REER and Remittances). One machine learning, two univariate and one multivariate model are considered; the models are the ANN model, AR model, ARIMA model and the VAR model. Bold entries indicate the lowest error measure achieved by the competing approach for the variable of interest across each row.

fitted model becomes optimally fit. The tuning of parameters with respect to the gradient helps find the best results [Siami-Namini, et al., (2018)]. Another reason is the ability of LSTM to capture non-linear trends in the training data. By design, ANN can account for the non-linear trends in datasets [Kaushik, et al., (2020)]. Hence, the ANN model performs best as an in-sample forecasting model for exports and imports groups.

b) Out-of-Sample Forecast Performance

Table 12 contains out-of-sample forecasting performance of the competing models for export groups and the overall exports. As before, the numbers in the table represent RMSE. Read across rows, the number in bold shows the model that performed best in forecasting the corresponding export group. We observe that the ARIMA model outperforms other competing forecast approaches for the food group. Therefore, it is the best model to forecast food export groups. For the textile group, the ANN model outperforms the AR, ARIMA and the VAR model. Therefore, it is the best forecast model for the textile group. We observe that ANN outperforms the competing approaches for the petroleum and coal groups. Hence, it can be said that the ANN model obtains the best forecasts of this series. For other manufacturers' export groups, the ARIMA model outperforms the competing forecast models; hence it is the best model to forecast this group. The AR model outperforms the ANN, ARIMA and the VAR for all other items

TABLE 12

Out-of-Sample Forecast Performance Results for Export Groups

	ANN	AR	ARIMA	VAR
Food Group	0.2253	0.1765	0.1523	0.1783
Textile Group	0.1064	0.1291	0.1327	0.1283
Petroleum Group and Coal	0.1065	0.7560	0.7425	0.6612
Other Manufactures Group	0.1187	0.1087	0.1062	0.1522
All Other Items	0.2269	0.1727	0.2141	0.2632
Total	0.1058	0.1053	0.1338	0.1358

Source: Authors' estimation.
Note: Out-of-sample forecast accuracy for h = 1 for exports. Each entry shows RMSE. VAR model is used for the series of interest and the determinants of exports (LSM, REER and Indirect Taxes). One machine learning, two univariate and one multivariate model are considered. The models are the ANN model, AR model, ARIMA model and VAR model. Bold entries indicate the lowest error measure achieved by the competing approach for the variable of interest across each row.

export groups. For the overall exports group, the AR model stands out again and outperforms the competing models; therefore, it is the best model to forecast this group.

Table 13 contains the out-of-sample forecasting performance of the competing models for import groups and overall imports. We observe that the AR model outper-

TABLE 13

Out-of-Sample Forecast Performance Results for Import Groups

	ANN	AR	ARIMA	VAR
Food Group	0.2621	0.1340	0.1823	0.1457
Machinery Group	0.1136	0.1517	0.2041	0.1419
Transport Group	0.1568	0.3641	0.3778	0.4848
Petroleum Group	0.1285	0.1846	0.2169	0.1763
Textile Group	0.4767	0.1904	0.2137	0.2101
Agricultural and Other Chemical Group	0.1787	0.0626	0.1595	0.1095
Metal Group	0.1105	0.1073	0.1056	0.1477
Miscellaneous Group	0.3273	0.1375	0.1192	0.1007
All Other Items	0.1337	0.1242	0.1260	0.1661
Total Import	0.0786	0.0869	0.0845	0.1397

Source: Authors' estimation.
Note: Out-of-sample forecast accuracy for h = 1 for imports. Each entry shows RMSE. VAR model is used for the series of interest and the determinants of imports (LSM, REER and Remittances). One machine learning, two univariate and one multivariate model are considered; the models are the ANN model, AR model, ARIMA model and the VAR model. Bold entries indicate the lowest error measure achieved by the competing approach for the variable of interest across each row.

forms the competing approaches for the food group. Therefore, it is the best forecast model for the food group. When we analyze the forecast accuracy of the models for the machinery group, transport and petroleum groups, we see that the ANN model outperforms the other models. Hence, the ANN is the best model to forecast these import groups. For the textile group, among the competing models, AR outperforms the ANN, ARIMA and VAR; thus, it is the best forecast model to forecast the textile group.

The AR outperforms the competing approaches for agricultural and other chemical groups and is the best forecast model to forecast this group. The ARIMA model outperforms the ANN, AR and VAR models for the metal group. For the miscellaneous group, the VAR model provides the minimum RMSE among the models used and is the best model to forecast this group. For all other items grouped, the AR model outperforms the competing models. The ANN model outperforms the other competing methods for the imports group. Therefore, it is the best model to forecast this group. The ANN and the AR models outperform other models for four of ten import groups. Hence, it is concluded that both these models perform equally well for out-of-sample forecasts for the import groups.

V. Conclusion and Policy Recommendations

Recently, due to the advancements in the fields of econometrics in general and the use of machine learning in economic forecasting, the use of disaggregated data series and forecasts of these data series have come to the forefront. This is motivated by this fact and the absence of a study that forecasts the group-wise exports and imports in Pakistan. This paper attempts to forecast these group-wise indicators. We use econometric models and a machine learning model to forecast these groups. Although machine learning algorithms have been stronger regarding forecasting/predicting, we find that conventional models also perform well out-of-sample for our variables of interest.

As Ghauri et al., (2020) mentioned, the extension of forecast on imports and exports may be performed using the ANN and other sophisticated models. To the best of our knowledge, our study is the first one that provides forecasts for Pakistan's overall and group-wise imports and exports. From the policy perspective, our paper gives the policymakers a new dimension; that is, the policymakers using the best forecasting models found in this study can forecast each group and form group-specific policies. It will allow them to do two things; first, they will be able to see which groups will perform better or worse, and second, it will allow them to direct specific policies towards declining export groups and increasing import groups. It ultimately will lead to a betterment of the trade balance. For example, suppose the policymakers find that the petroleum import group will soar. In that case, they may investigate if the effect comes from the quantity of petroleum imported or a rise in international oil prices and plan accordingly.

In the future, we intend to use multivariate machine learning methods and the most commonly used dynamic factor econometric model to forecast these groups. However, this paper simultaneously opens up recommendations to use machine learning methods using multiple variables, which will surely help policymakers at the central bank and the central/provincial governments.

References

Aker, S. L., 2008, Major determinants of imports in Turkey, Turkish Studies, 9(1): 131-145, https://doi.org/10.1080/14683840701814059.

Ali, M. S., 2009, Effects of major exports and imports on the balance of foreign trade in Pakistan, Masters' Thesis, Lund University: Sweden, 1-36.

Box, G., and G. Jenkins, 1976, Time series analysis, Forecasting and Control, San Francisco: Holden-Day, 21: 575.

Crawford, G. W., and M. C. Fratantoni, 2003, Real estate: Assessing the forecasting performance of regime-switching, ARIMA and GARCH, Real Estate Economics, 31(2): 223-243.

Dickey, D. A., and W. A. Fuller, 1981, Likelihood ratio statistics for autoregressive time series with a unit root, Econometrica, 49(4): 1057-1072, https://doi.org/10.2307/1912517.

Enders, W., 2009, Applied Econometric Times Series, 3rd Edition, Enders Walter, John Wiley and Sons, University of Alabama: United States.

Farooqi, A., 2014, ARIMA model building and forecasting on imports and exports of Pakistan, Pakistan Journal of Statistics and Operation Research, 10(2): 157-168, https://doi.org/10.18187/pjsor.v10i2.732.

Ghauri, S. P., R. R. Ahmed, D. Streimikiene, and J. Streimikis, 2020, Forecasting exports and imports by using autoregressive (AR) with seasonal dummies and Box-Jenkins approaches: A case of Pakistan, Engineering Economics, 31(3): 291-301, https://doi.org/10.5755/j01.ee.31.3.25323.

Kaushik, S., A. Choudhury, P. K. Sheron, N. Dasgupta, S. Natarajan, L. A. Pickett, and V. Dutt, 2020, AI in healthcare: Time series forecasting using statistical, neural, and ensemble architectures, Frontiers in Big Data, 3: 1-16.

Ljung, A. G. M., and G. E. P. Box, 1978, Biometrika trust on a measure of lack of fit in time series models: Oxford University Press on behalf of Biometrika Trust Stable, 65(2): 297-303, http://www.jstor.org/stable/2335207.

Lutkepohl, H., 2005, New introduction to multiple time series analysis (1st edition), Springer-Verlag, European University Institute: Firenze, Italy.

Majeed, Muhammad Tariq, E. Ahmad, and M. I. Khawaja, 2006, Determinants of exports in developing countries, 45(4): 1265-1276.

Makridakis, S., and M. Hibon, 1997, ARIMA models and the Box-Jenkins methodology, Journal of Forecasting, 16(3): 147-163, https://doi.org/10.1002/(SICI)1099-131X(199705)16:3<147:AID-FOR652>3.0.CO;2-X.

Mehmood, S., 2012, Forecasting Pakistan's exports to SAARC: An application of univiriate ARIMA model, Lahore: University of Central Punjab.

Said, S. E., and D. A. Dickey, 1984, Testing for unit roots in autoregressive moving average models of unknown order, Biometrika, 71(3): 599-607.

Siami-Namini, S., N. Tavakoli, and A. S. Namin, 2018, A comparison of ARIMA and LSTM in forecasting time series: Proceedings-17th IEEE International Conference on Machine Learning and Applications (ICMLA), 2018: Orlando, United States, 1394-1401.

Sims, C., 1980, Macroeconomics and reality, Econometrica, 48(1): 1-48, https://doi.org/10.2307/1912017.

Zivot, E., and D. W. K. Andrews, 2002, Further evidence on the great crash: The oil price shock, and the unit root hypothesis, Journal of Business and Economic Statistics, 10(3): 251-270, https://doi.org/10.1080/07350015.1992.10509904.

APPENDIX A

TABLE A-1

Name of the Groups of Imports of Pakistan

Serial Number	Name of the Groups	Commodities in Each Group
1.	Food Group	Milk, Cream and Milk Food for Infants, Wheat un-milled, Dry Fruits and Nuts, Tea, Spices, Soya bean Oil, Palm Oil, Sugar, Pulses, and all other Food Items.
2.	Machinery Group	Power Generating Machinery, Office Mach. Incl. Data Processing, Equipment, Textile Machinery, Construction and Mining Machinery, Electrical Machinery and Apparatus, Telecom, Agricultural Machinery and Implements, and Other Machinery.
3.	Transport Group	Road Motor Vehicles (Build Unit, Ckd/Skd),Air-crafts, Ships and Boats, Others Transport Equip-ments
4.	Petroleum Group and Coal	Petroleum Products, Petroleum Crude, Natural Gas, Liquefied Petroleum Gas, Others
5.	Textile Group	Raw Cotton, Synthetic Fiber, Synthetic and artificial Silk Yarn, Worn Clothing, and Other Textile Items
6.	Agricultural and Chemical Group	Fertilizer Manufactured, Insecticides, Plastic Mate-rial, Medicinal Products, Others
7.	Metal Group	Gold, Iron and Steel Scrap, Iron and Steel, Alu-minum Wrought and Worked, All other Metals and Articles
8.	Miscellaneous Group	Gold, Iron and Steel Scrap, Iron and Steel, Alu-minum Wrought and Worked, All other Metals and Articles
9.	All Other Items	All Other Items or Commodities are included in this group

Source: Pakistan Bureau of Statistics.

TABLE A-2

Name of the Groups of Exports of Pakistan

Serial Number	Name of the Groups	Commodities in Each Group
1.	Food Group	Rice, Basmati, Others, Fish and Fish Preparations, Fruits, Vegetables/Leguminous vegetables, Tobacco, Wheat, Spices, Oil Seeds, Nuts and Kernels, Sugar, Meat and Meat Preparations, All Other Food
2.	Textile Group	Raw Cotton, Cotton Yarn, Cotton Cloth, Cotton Carded or Combed, Yarn Other than Cotton Yarn, Knitwear, Bed Wear, Towels, Tent, Canvas and Tarpaulin, Readymade, Made up, Articles (Ex. Towels and bed) Garments, Art, Silk and Synthetic Textile, Other Textile Materials
3.	Petroleum Group and Coal	Petroleum, Crude Petroleum Products (Excl. Naphtha), Petroleum Top Naphtha, Solid Fuels (Coal)
4.	Other Manufacturer Group	Carpets, Rugs and Mats, Sports Goods, Leather Tanned, Leather Manufactures, Footwear, Surgical Goods and Medical Instr., Cutlery, Onyx Manufactured, Chemical and Pharmaceutical Products, Engineering, Goods, Gems, Jewelry, Furniture, Molasses, Handicrafts, Cement, Guar and Guar Products
5.	All Other Items	All Other Items or commodities are included in this group

Source: Pakistan Bureau of Statistics.

TABLE A-3

Determinants of Imports of Pakistan

Serial Number	Name of the Variable
1.	Workers' Remittances (USD million)
2.	Real Effective Exchange Rate (Index)
3.	Quantum Index of Large-Scale Manufacturing Industries (Index)

Source: Authors' estimation.

TABLE A-4

Determinants of Exports of Pakistan

Serial Number	Name of the Variable
1.	Indirect Taxes (Million Rupees)
2.	Real Effective Exchange Rate (Index)
3.	Quantum Index of Large-Scale Manufacturing Industries (Index)

Source: Authors' estimation.

APPENDIX B

Details of the VAR Models

Before applying the VAR models, the data has to be tested for co-integration. Therefore, co-integration analysis is performed using the import and export series and its determinants for each import and export series. If co-integration is found, a VAR model in either levels or first differences will be miss-specified. The vector error correction model must be applied to forecast each series of interest. Lutkepohl (2005) states that if a stationary variable is included in the co-integration analysis, there is an additional cointegrating relationship for this stationary variable. The cointegrating rank will be at least as big as the number of stationary variables in the co-integration analysis.

Some of our analysis's imports and exports series are stationary in log levels. When we perform the co-integration analysis, it is found that whenever one of the stationary imports or exports series is present for the test along with its determinants in the co-integration analysis. Either a single cointegrating vector or no co-integration is found in the results. It signifies that the cointegrating relationship is present due to the addition of the stationary variable in the analysis. However, there is no co-integration present among the variables otherwise. In such a case, Enders (2009) suggest that the VAR model in first differences is preferable. Therefore, the VAR is used in the first difference in such a case. For the import and export series that are non-stationary in levels. No evidence is found for any co-integration among the variables in each model tested for co-integration. Therefore, the VAR model is estimated in the first differences.