

Analysis and Prediction on the Top Retailers in the UK

Quan Zhou

School of Mathematics, University of Bristol, Bristol, United Kingdom

wendy.zhou.2021@bristol.ac.uk

Abstract. The inflation in the UK has been increasingly serious, reaching approximately 10%. This directly causes the rise in the price of daily-used goods, including fresh and chilled, food cupboard and bakery, and household items. The growth in living costs has cut the shopping desire for the customers, pushing them to altering their original consumption habits. It is interesting to research in the behavior of the retailers during the inflation and the downturn of the UK economy. This paper has chosen four well-known retailers in the UK, Tesco, Sainsbury's, Marks & Spencer's and NEXT to represent the general development of the retail industry throughout the country. To evaluate the performance of the company before and after the influence, the stock price is used to indicate the current value of the four retailers. By collecting the data from Yahoo Finance and forecasting by applying Arima model, there will be no significant upwards or downwards tendency in the near future for Sainsbury's and NEXT, while slight fluctuations are predicted to be shown for Tesco and Marks & Spencer's stock price.

Keywords: Retail industry; UK economy; ARIMA model.

1. Introduction

With relevant indicators, the economic situation for the UK is not optimistic. The inflation rate reached 10.1% and the Consumer Price Index, which includes the cost of owner-occupied housing (CPIH), rose by 8.8% in the 12 months to September 2022, up from 8.6% in August and back to its recent peak in July [1]. The energy shortage, the devaluation of the British pounds and the unemployment crisis has increased the living costs and aggravated the financial pressure. The retailing industry has been straightly hit. The rising in transportation and the packaging costs encounter the food price inflation, while the heavy living costs pressure decreased the shopping desire. For instance, Tesco meal deal, sat at £3 the Clubcard and £3.50 for regular over a decade, has increased 40p to £3.40 and £3.90 in respective since the 21st of October [2].

To measure the future development in this industry, this paper has taken four well-known retailers, Tesco, Sainsbury's, Marks & Spencer and NEXT, to represent the industry in the UK. Tesco is the market leader of groceries, mainly focusing on the food grocery. Since there is a wide variety of products and services offered by Tesco, it is very hard to divide the customers by demographic and psychographic factors. But in general, Tesco is targeting for individuals who are aware of the costs and interested in special offers, discounts, and comparison shopping. Sainsbury's is the second largest groceries in the UK [3]. Sainsbury's supermarkets have implemented an omnichannel marketing focus so that their customers to buy from one place and of the best quality. This marketing strategy, as well as the outstanding service and quality, enabled them to become the first choice of all-aged customers. Sainsbury's also accelerates its growth in non-food sector so as to reach more customers [4]. Marks & Spencer (M&S) operates in the retail industry, but its food revenue has continuously taken over half of the annual revenue (6,639.6 of 9,166.9 million in GBP in 2021) [5]. M&S presently operates 959 locations around the UK, 615 of which are devoted solely to the sale of food items. Through its television advertising, M&S promotes the exclusivity and elegance of its eats and drinks. The major consuming group is mid-age, but with the collaboration to social media and live streaming platform, Marks & Spencer is trying hard to attracting more younger consumers and widening the variety of the customers [6]. NEXT, comparing the previous three, puts all its weight on the clothes retailing and pay more attention to the design, unlike Marks & Spencer. Their goal is to exceed the expectations of their customers by providing interesting, elegantly designed, high-quality apparel, footwear, accessories, and homeware [7]. Their present strategy in light of the decline in retail sales

is to maximise chances for online growth while minimising negative consequences by incorporating as much flexibility as possible into operations and cost structures.

The four retailers can be segmented into two pairs, Tesco and Sainsbury's aiming for food groceries, Marks & Spencer and NEXT operating in clothes retailing. Marks & Spencer can be seen as a transition and a combination of food industry and clothes industry due to its distinctive operation results. Therefore, they are taken to be the representative of the whole relevant industry. The share price of a company reflects the investor projection of its ability to earn and grow its profits in the future. Normally, the extent of how a company is operating can be measured by studying its stock price value and trend. Thus, this paper uses the stock price to be the database for the forecasting and indicate the business status.

As a collection of data points that have been indexed, listed, or graphically displayed in time order, the stock price can be thought of as a time series. A time series is most frequently a sequence captured at a series of equally spaced moments in time. As a result, it is a collection of discrete-time data. Forecasting is performed using the ARIMA model. This group of statistical models is used to examine and predict time series data. It specifically addresses a number of common time series data types and, as a result, offers a straightforward but effective technique for producing accurate time series forecasts. All coding is done using R studio, an integrated development environment for R, which is a popular coding language for statistics and data analysis.

2. Methodology

2.1 Data Selection

The four companies' stock price is collected from the website Yahoo Finance. With observing the stock prices, the past trends are shown clearly. The interval taken is from January 1st 2010 to September 15th 2022. A 12-year-long range is taken for a more precise forecasting. With a reasonably good length of data, the forecast is unlikely to be affected by anomalies. The results are preferred to follow the general trend and the development history of the company. The much earlier data are ignored since it is out of date. The excessive dataset will also alter the results. As it is selecting data from the website, data need to be cleaned and adjusted into R language format. This is done at the same time. A line graph of the stock price for each company is plotted.

Then the data is modeled into time series format, taking monthly frequency. Using monthly stock price, it is easier to operate for fewer numbers, and as a long interval is taken the property of the dataset is preserved. Since ARIMA model is univariate, only the prices are taken. The line graph of the stock price in time series has the same general shape as the raw data but in a smoother shape. This is because the raw data from the website is the daily price while for the time series plot is monthly. It is necessary to identify any unusual changes in the time plot and determine whether any data revisions seem to be required. If necessary, the variance can be stabilised using the Box-Cox transformation [8].

2.2 Modelling

Each time series' ACF and PACF are likewise plotted along with the data. The autocorrelation function (ACF) is so named. The ACF is a bar graph representing the correlation coefficients between a time series and its own lags. The linear relationship between a time series' lagged values is examined using autocorrelation, just as correlation establishes whether or not two random variables or bivariate data are causally related. There will be several autocorrelation coefficients assigned to each panel in the lag plot. For instance, r_1 assesses the link between y_t and y_{t-1} , r_2 assesses y_t and y_{t-2} , and so forth. As a result, the equation (1), where T is the length of the time series, can be used to express the value of r_k . The partial correlation coefficients between the series and its own lags are presented in the PACF plot.

$$r_k = \frac{\sum_{t=k+1}^T (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2} \quad (1)$$

Thus, the equation (2), where y'_t is the differenced series, can be used to express the value of r_k (it might have undergone multiple changes.). Both the lagged y_t values and the lagged errors are represented in the "predictors" on the right-hand side.

An ARIMA (p, d, q) model is what we are referring to here, where p denotes the order of the autoregressive part, d denotes the amount of initial differencing that is involved, and q is the order of the moving average part. The `auto.arima()` function is used in this situation. The `auto.arima()` function in R uses a modified Hyndman-Khandakar technique that incorporates unit root testing, AICc minimization, and MLE to produce an ARIMA model. With the help of numerous KPSS tests, the amount of differences $0 \leq d \leq 2$ is calculated. After differencing the data d times, the values of p and q are then determined by computing the AICc.

Instead of considering every potential pairing of p and q, the algorithm navigates the model space using a stepwise search. The "current model" is chosen to be the best model with the lowest fitted AICc score. Change p and/or q from the existing model by ± 1 ; include or exclude c from the present model are all variations that are taken into consideration. The current model or one of these alternatives, whichever is the best model up to this point, becomes the new current model. Up till the smallest AICc is founded, this process is repeated.

$$y'_t = c + \phi_1 y'_{t-1} + \dots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (2)$$

The ARIMA equation is enlarged and rearranged to have only y_t on the left side and all other terms on the right in order to forecast. The equation is then recast by substituting $T + h$ for t . Future observations are substituted with their forecasts on the right-hand side of the equation, future errors with zero, and previous errors with the corresponding residuals. These processes are then repeated starting with $h = 1$ for $h = 2, h = 3$, and so on until all [8].

3. Results

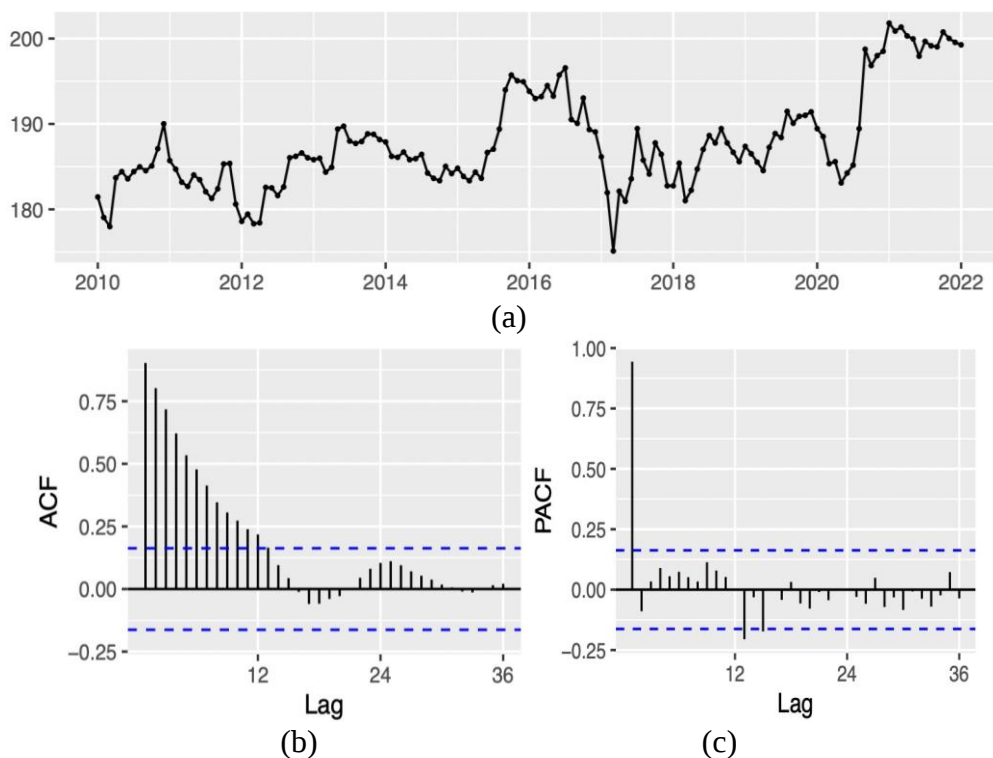


Fig. 1 ggtsdisplay(tsc0) (Photo credit: Original)

Figure 1 shows the time series model, ACF and PACF plot of the stock price of Tesco. The time plot does not show any sudden peak or drop. There is a drop at the beginning of 2020, which is due to the pandemic. A lot of restrictions are put on the outdoor activity. The time plot is otherwise normal, and it doesn't seem that any data modifications are necessary. As there is no evidence of changing variance, the Box-Cox transformation is unnecessary. From the ACF and the PACF plot, it is concluded that there is no seasonality, and it is not a stationary data. By checking the residuals, the function is white noise. Hence the auto.arima function is applied, where the order of the arima model is ARIMA (0, 1, 0) (1, 0, 0). As shown in Figure 2, It can be found that the forecasted Tesco stock price is mainly consistent, while a slight fluctuation in presented at the beginning.

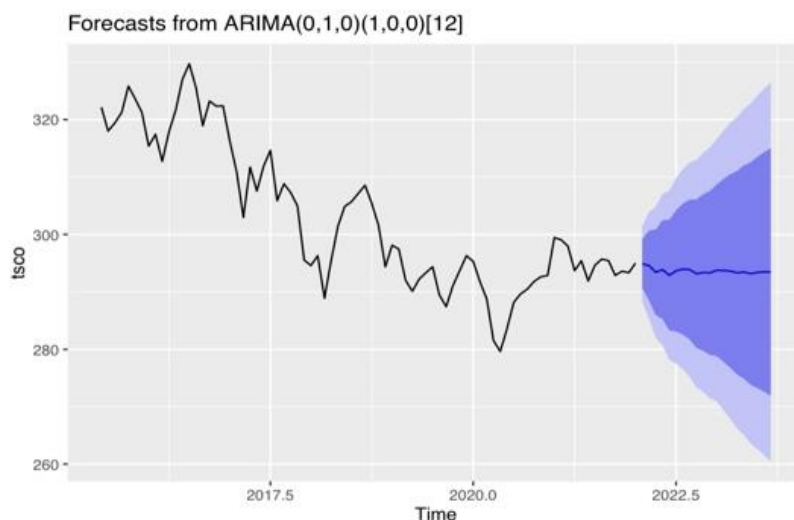


Fig. 2 Forecasts for Tesco (Photo credit: Original)

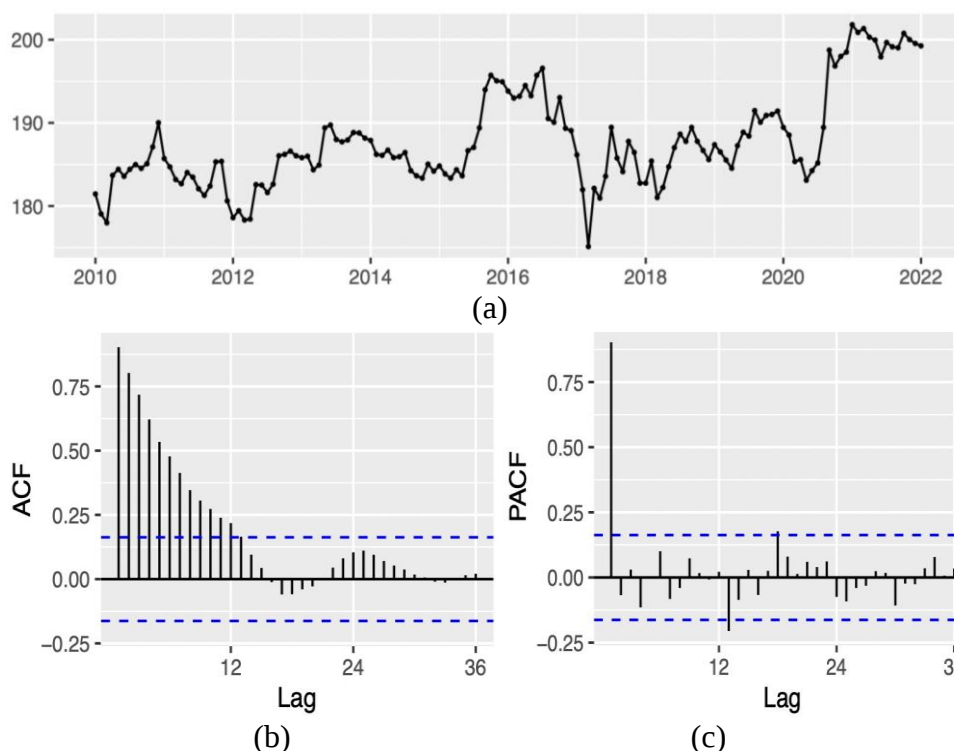


Fig. 3 ggtsdisplay (sbry) (Photo credit: Original)

The Sainsbury's stock price's time series model, ACF, and PACF plot are displayed in Figure 3. The time map reveals some abrupt shifts, especially the significant decline in 2017–2018. BREXIT might be the main reason for these modifications. The time plot is normal in all other aspects, and it

doesn't seem that any data modifications are necessary. Since there is no indication that variance is changing, Box-Cox transformation is not necessary. From the ACF and the PACF plot, it is concluded that there is no seasonality, and it is not a stationary data. By checking the residuals, the function is white noise. Hence the auto.arima function is applied, where the order of the arima model is ARIMA (0, 1, 0). According to Figure 4 of the forecasting, it can be easily seen that the forecasting line is horizontal. This is because the ARIMA (0, 1, 0) model is a random walk. That means the model, when differentiated once, becomes an ARMA (0, 0), which is random, uncorrelated, white noise. Additionally, the random work model is also called naive forecast. It totally sets all forecasts to be the value of the last observation (a reason why the following forecasts are the same value). The horizontal predicted line is marking at the historically high point. Thus, it can be concluded that Sainsbury's future business activity will be doing well.

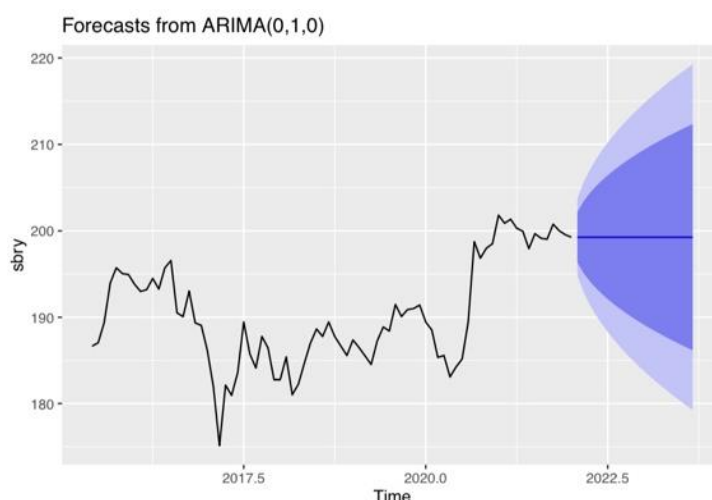


Fig. 4 Forecasts for Sainsbury's (Photo credit: Original)

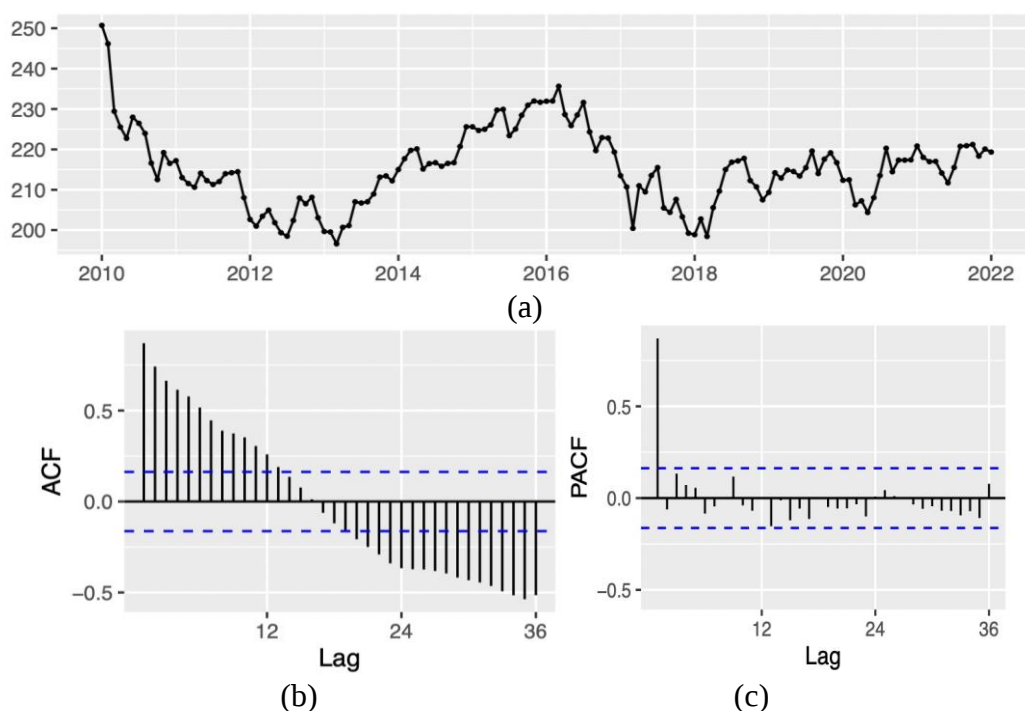


Fig. 5 ggsdisplay (mks) (Photo credit: Original)

Figure 5 shows the time series model, ACF and PACF plot of the stock price of Marks and Spencer. Not many abrupt shifts are depicted in the time plot. There doesn't seem to be a need to make any

data edits based on the time plot. Since there is no indication that variance is changing, Box-Cox transformation is not necessary. From the ACF and the PACF plot, it is concluded that there is no seasonality, and it is not a stationary data. By checking the residuals, the function is white noise. Hence the auto.arima function is applied, where the order of the arima model is ARIMA (1, 0, 0) (0, 0, 1). As shown in Figure 6, It can be found that the forecasted Marks and Spencer stock price is mainly consistent, while a gentle drop at the beginning. Overall, the stock price will sit at the middle point of the historical stock price range.

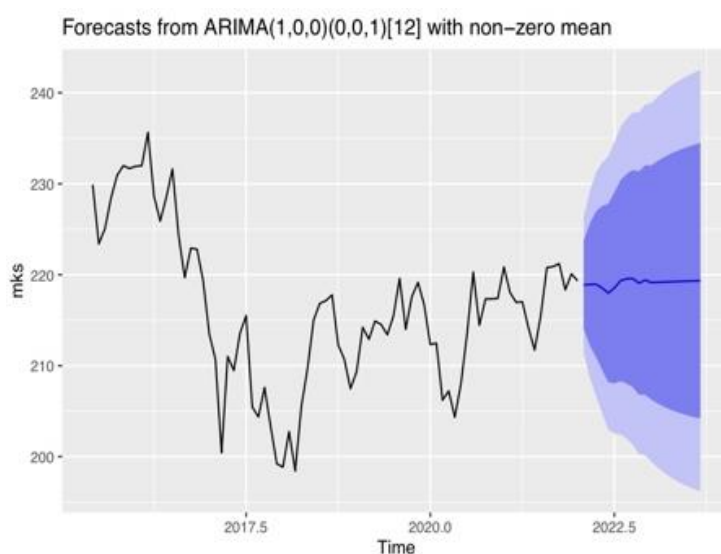


Fig. 6 Forecasts for Marks and Spencer (Photo credit: Original)

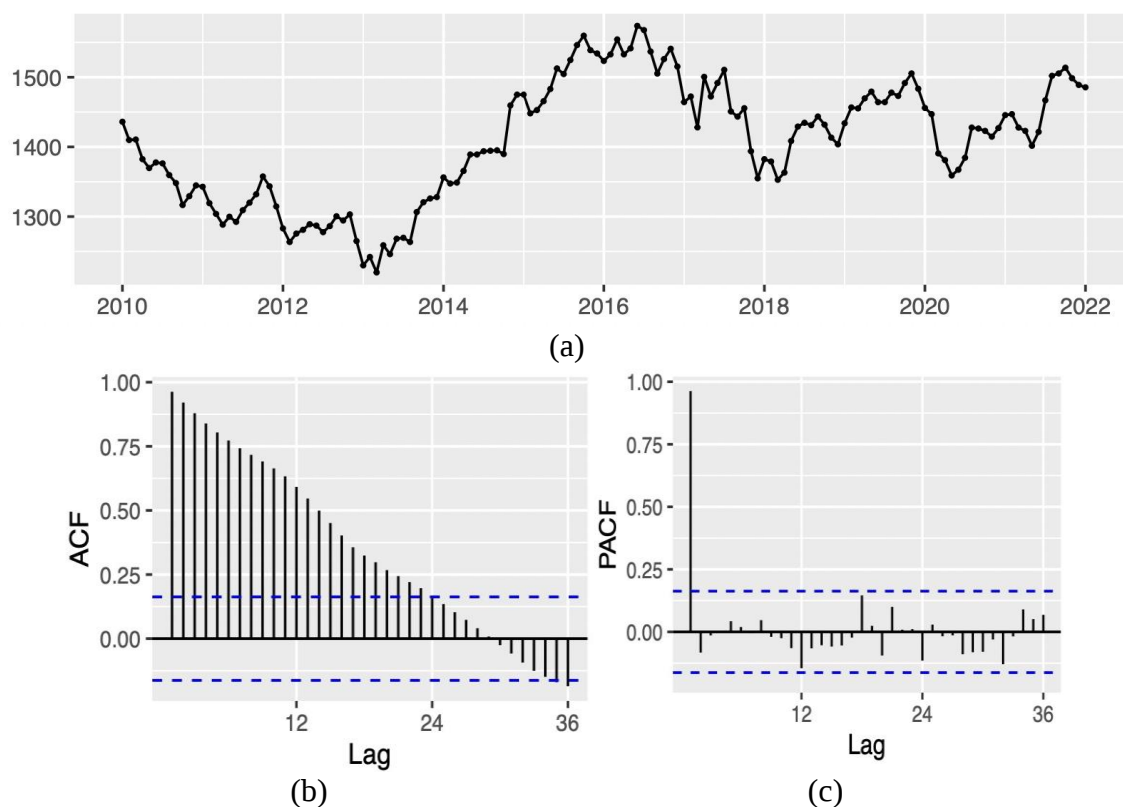


Fig. 7 ggtsdisplay (nxt) (Photo credit: Original)

The stock price of NEXT is plotted using a time series model, ACF, and PACF in Figure 7. The time plot does not reveal many abrupt shifts, and no data revisions appear to be required. A Box-Cox

transformation is not necessary here because there is no evidence of changing variance. From the ACF and the PACF plot, it is concluded that there is no seasonality, and it is not a stationary data. By checking the residuals, the function is white noise. Hence the `auto.arima` function is applied, where the order of the arima model is ARIMA (0, 1, 0). So, in this case the naïve method is used which is presented in the Arima model. Because when data follow a random walk, a naïve forecast is optimal. Figure 8 also complies with the fact that sometimes one of these basic forecasting techniques will be the greatest one that is available, but that in many instances, these techniques will be used as benchmarks rather than as the preferred technique.

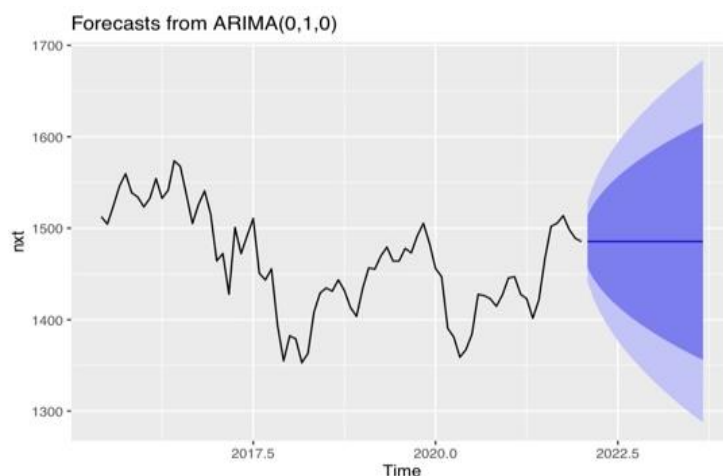


Fig. 8 Forecasts for NEXT (Photo credit: Original)

4. Discussion

Strong consumer demand and supply-chain constraints supported the rising cost of consumer products, which contributed to the increase in inflation. The percentage of income that low-income households spend on food and energy is higher than the national average. The conflict between Russia and Ukraine was the main cause of this cost-of-living problem, but many other causes will affect UK household budgets in 2022.

Even though the nation has recovered from the COVID-19, the long-term effects should not be understated. Aside from the effects on health, there have been social and economic upheavals, including millions of pounds in holiday expenditures and disruptions to the hospitality, travel, entertainment, and leisure sectors. Because of the nation's economic weakness, the cost-of-living dilemma is not something for which it is prepared. Although the government has issued a number of services, the needy receive little help. In the post-covid era, the industrial and distribution sectors are slowly coming back to life. This sparked a fresh spike in energy consumption, increasing pressure on prices but also boosting the nation's economic recovery.

Since the beginning of last year, more than 30 UK energy companies have filed for bankruptcy. Due to the growth of the green agenda and the rise in environmental concerns, the coal-fired power stations were shut down. However, renewable energy sources are not yet ready to deliver enough power. The main source of energy at that point was gas, which was maintained at an exorbitant cost. The cost of electricity not only squeezed household budgets but also put pressure on supplier networks. Both drivers and gasoline, which are in low supply, are essential for lorries. The price of products on the shelves rises along with the cost of distribution. Shipping a container from Asia to the UK used to cost £3,000, but it now costs £15,000.

Not only is the supply chain struggling, but suppliers are also on strike. Agriculture and food production take a significant amount of energy, which is currently one factor driving up costs. It also depends heavily on fertilisers, which are made using a lot of energy. It needs an efficient and cost-effective distribution system, which has been missing recently. Due to Covid restrictions, there are

fewer European workers available to perform picking and packing tasks. As a result, growers are paying higher rates, increasing labour costs. All of this translates to farm-gate inflation and higher retail pricing. The expense of transportation and distribution drives up the cost of raw materials, which inevitably drives up prices overall. Worryingly, a sizable percentage of the world's supply of wheat and other staple goods like vegetable oil is produced in Russia and Ukraine; hence, the longer the war lasts, the more the rest of the world will be impacted [9].

Historically, Sainsbury's was viewed as a more upscale store than Tesco. Although Sainsbury's was often more expensive than Tesco, the food was thought to be of higher quality. Early in the twenty-first century, Tesco was experiencing major quality issues and had out-of-date store infrastructure. Tesco turned its attention to quality and advancement as a result. This gap has slightly closed since then. Tesco has also made more of an effort to dominate the grocery business than Sainsbury's because Tesco Extra is a format that Sainsbury's does not provide. In addition, Sainsbury's has never ventured into a foreign market other than a brief stint in the now-defunct port city of Calais. Tesco once ran substantial grocery stores in Korea and other Asian countries in addition to having significant operations in Ireland and Eastern Europe. With the Fresh and Easy brand's attempts to break into the US market being regarded as a failure, the retailer lost emphasis on its primary UK operation [10].

5. Conclusion

According to the best fit model used in the model selection part, both Tesco and Marks and Spencer stock price have a small decrease but will be covered very soon. It might be difficult to recover to the past-covid level. For the other two retails, if there is an upward trend in the near future, it is believed that they will eventually recover to their previous level. However, the general economic environment is not optimistic, and it is doubtful whether this industry can recover to the previous level. Companies is impossible to operate well under a depressed market. This research has also contained limitation. Only big companies are selected. However, medium and small retailers are damaged worse as they tend to be private and there is no shareholder to share the responsibility of the loss. Countlessly closure is happened. Also, the stock price might not be precise enough as it is determined by the investors' opinions rather than based on the company realistic performance. This paper research in the further development in the retail and groceries industry. Apart from being an indicator, more concerns should be arisen on how to recover and fix the market.

References

- [1] Census 2021 Inflation and price indices. <https://www.ons.gov.uk/economy/inflationandpriceindices#:~:text=Consumer%20price%20inflation%2C%20UK%3A%20September%202022&text=The%20Consumer%20Prices%20Index%20including,returning%20to%20July%27s%20recent%20high>. Last accessed 2022/11/23.
- [2] Cost of living: Tesco raises meal deal price, with even Clubcard customers paying more. <https://news.sky.com/story/cost-of-living-tesco-raises-meal-deal-price-with-even-clubcard-customers-paying-more-12726307>. Last accessed 2022/11/23.
- [3] Tesco's Digital Marketing Strategies: A Detailed Case Study. <https://digitalscholar.in/tesco-digital-marketing-strategies/>. Last accessed 2022/11/23.
- [4] Sainsbury's Marketing Strategy: Becoming the Second-Largest Supermarket Chain in the UK. <https://blog.contactpigeon.com/sainsburys-marketing-strategy/>. Last accessed 2022/11/23.
- [5] Marks & Spencer food revenue in the United Kingdom (UK) from financial year 2009/10 to 2021/22. <https://www.statista.com/statistics/413627/food-sales-marks-and-spencer-mands-united-kingdom-uk/>. Last accessed 2022/11/23.
- [6] [https://issuu.com/feifeifeiwu/docs/next_marketing_audit_and_plan_print#:~:text=NEXT%20targets%20its%20consumers%20through,sensitive%20\(HOOVERS%2C%202015\)](https://issuu.com/feifeifeiwu/docs/next_marketing_audit_and_plan_print#:~:text=NEXT%20targets%20its%20consumers%20through,sensitive%20(HOOVERS%2C%202015)). Last accessed 2022/11/23.

- [7] ACF: (Partial) Autocorrelation and Cross-Correlation Function Estimation. <https://www.rdocumentation.org/packages/forecast/versions/8.18/topics/Acf>. Last accessed 2022/11/23.
- [8] Author R. J. Hyndman, Author G. Athanasopoulos Forecasting: Principles and Practice (2nd ed). <https://otexts.com/fpp2/>. Last accessed 2022/11/23.
- [9] Cost Of Living Crisis: Boosted Winter Fuel Payments Roll Out To 11.6 Million Pensioners. <https://www.forbes.com/uk/advisor/energy/cost-of-living-crisis/>. Last accessed 2022/11/23.
- [10] Why Do Companies Care About Their Stock Prices? <https://www.investopedia.com/investing/why-do-companies-care-about-their-stock-prices/#:~:text=A%20company%27s%20stock%20price%20reflects,and%20receive%20increases%20in%20compensation>. Last accessed 2022/11/23.