Forecasting Consumer Credit Card Balances for Large Banks: A Time Series Analysis

MOTHI GOWTHAM ASHOK KUMAR

Abstract

This research explores the quarterly trends of consumer credit card balances for large banks and aims to forecast future balances using advanced time series methods. Insights from this study can help banks in strategic decision-making, risk assessment, and resource allocation.

Background Information

Consumer credit card balances reflect the borrowing behavior of the population and provide insights into consumer confidence and financial health. For large banks, understanding these trends and forecasting future balances is crucial for optimizing lending strategies, managing risk, and anticipating market dynamics.

Initial Literature Review

From a report published by CNBC on August 10, 2023 [1], it was highlighted that the U.S. is witnessing a growing reliance on credit. The New York Federal Reserve reported that total credit card debt in the U.S. crossed the $1 trillion threshold for the first time. This indicates an almost 20% increase in credit card balances from the previous year, leading to the average balance per consumer reaching $5,947 – the highest in the past decade. Furthermore, a striking 60% of cardholders who carry balances have been in this debt situation for over a year. The report also shed light on the challenges of high interest rates on borrowing, with the average credit card rate now soaring beyond 20%. Additionally, with the U.S. personal savings rate standing at 4.3%, significantly below the long-standing average of 8.9%, concerns about inflation's adverse effects on households have been raised. The credit card market has also seen a surge in participation from the younger generation, specifically Gen Z. Approximately 19 million new credit accounts were opened recently, bringing the total credit cards in circulation to a record 530.6 million. Despite this, delinquencies, defined as payments overdue by 90 days or more, have been on the rise.

In another report released by the Consumer Financial Protection Bureau (CFPB) on October 25, 2023 [2], concerning insights into the consumer credit card market, it was found that in 2022 credit card companies charged consumers an astounding amount of over $105 billion in interest and more than $25 billion in fees. This has resulted in the unprecedented total credit card debt surpassing $1 trillion. Notably, the profits of credit card companies have exceeded pre-pandemic levels, suggesting a possible lack of competition in the industry. The annual percentage rates (APRs) have also been escalating, continuing to surpass major index rates. An alarming figure from the report was the combined charge of over $130 billion in interest and fees to cardholders in 2022. With the termination of pandemic relief programs, delinquency rates have been on the rise, exemplified by the substantial late fees amounting to $14.5 billion. The report also highlighted the growing challenge of "persistent debt," where cardholders are charged more in interest and fees annually than what they pay toward the principal. This pattern could pose severe challenges for consumers, especially if interest rates stay elevated. Lastly, a significant shift towards digital communication in the credit card industry was noted, with a majority of cardholders now favoring mobile apps and digital portals for their card management and payment activities.

Hypothesis/ Theories

Through this project we will be testing the following hypotheses,

H1: Seasonal trends influence consumer credit card balances, with certain quarters consistently showing higher balances.

Based on patterns observed in consumer spending behaviors, it is postulated that there are specific times of the year when consumers are more likely to accumulate higher credit card balances. These times might coincide with major holidays, festive seasons, or school starting dates, during which consumer expenditure naturally increases. The hypothesis suggests that certain quarters will consistently demonstrate higher balances due to these seasonal shopping trends, regardless of broader economic conditions.

H2: Economic downturns, as evidenced by events like the 2020 pandemic, have a significant effect on credit card balances.

The global economic slowdown caused by significant events, like the 2020 pandemic, is hypothesized to have a profound impact on how consumers utilize credit. With uncertainties surrounding employment and overall economic health, consumers might resort to credit cards as a primary source of funds, thus increasing their credit card balances. This theory suggests that during these economic downturns, there is a notable surge in credit card balances as consumers navigate financial challenges and unexpected expenses.

H3: Advanced forecasting models can accurately predict consumer credit card balances in the future.

By understanding underlying trends, seasonal patterns, and the impact of external factors, models like ARIMA, SARIMA, and Holt-Winters can provide insightful projections. This theory posits that, with the right data and methodology, future consumer behaviors in terms of credit card usage can be predicted, allowing financial institutions to better prepare and strategize.

Data Overview & Visualization

The dataset consists of quarterly data from 2012 to 2023 on consumer credit card balances for large banks in the U.S. The frequency of the data is quarterly.

The data is extracted from FRED which is sourced from Federal Reserve Bank of Philadelphia.

There is one series that shows the quarterly consumer credit card balances and forms the main part of data for our analysis and forecasting results. There are approximately 44 data points in the series.

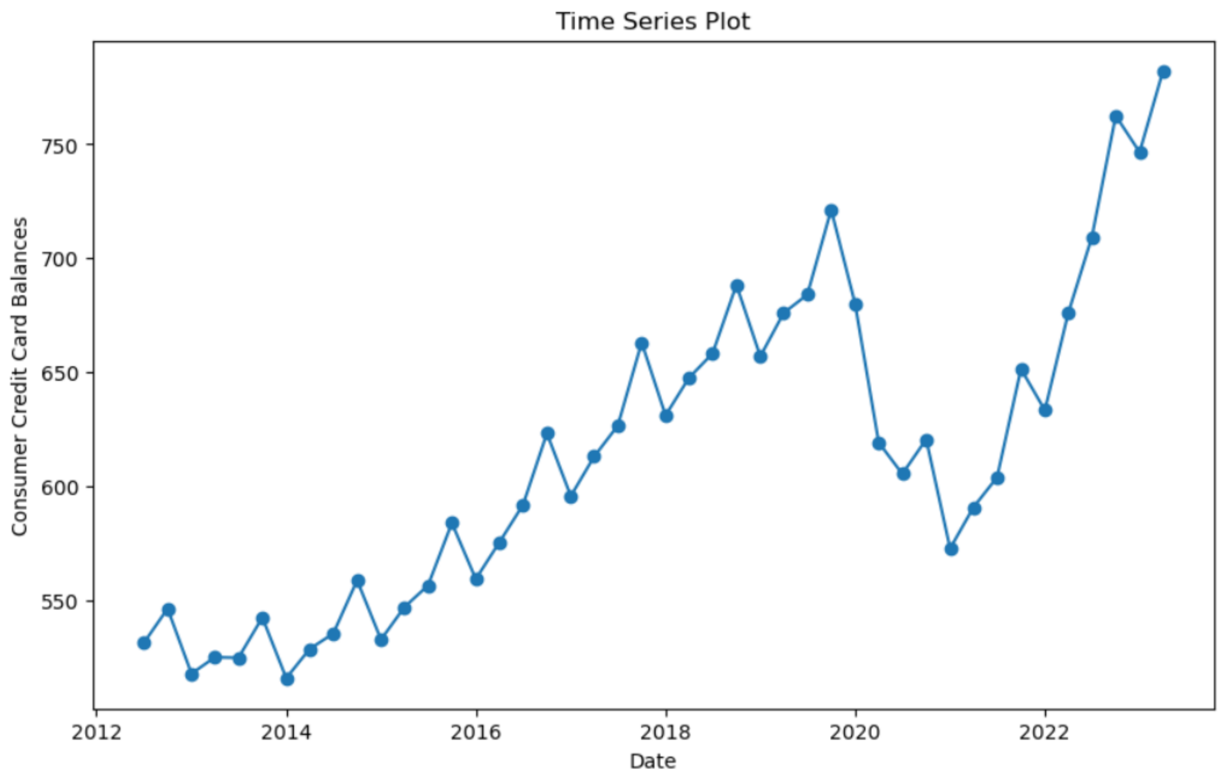

Figure 1. Consumer credit card balances

Figure 1 presents a time series plot of consumer credit card balances from 2012 to 2023. The plot reveals a clear upward trajectory, punctuated by regular fluctuations that suggest seasonal or cyclical effects. Notably, a pronounced dip aligns with the 2020 economic downturn, illustrating the pandemic's immediate impact on consumer credit usage. This is followed by a sharp rebound, indicating a resurgence in credit card balances as economic conditions began to improve. This visual evidence supports the hypothesis that external economic events significantly affect consumer credit behavior, while also hinting at underlying seasonal patterns.

Summary statistics and data cleaning procedure

The data has been verified for missing values and there are no missing values.

The summary statistics of the data series is as follows,

Frequency	Quarterly
data points	44
mean	613.82
median	609.15
std dev	70.2
Q1	549.21
Q3	609.14

Range	515 - 782

Table1. Summary Statistics. Values in Billions of U.S. Dollars

From the data series there seems to be a downtrend during the covid period. This will be verified while testing our hypotheses.

Seasonal Decomposition and Initial Assessment

The time series data from 2012 to 2023 shows fluctuations in consumer credit card balances. Notably, there appears to be a dip around 2020, which could be attributed to economic factors arising from the global pandemic. Further analysis will delve into seasonal trends, potential cyclic behavior, and the forecasting of future data points.

Additive decomposition was performed to analyze existence of a trend and seasonality,

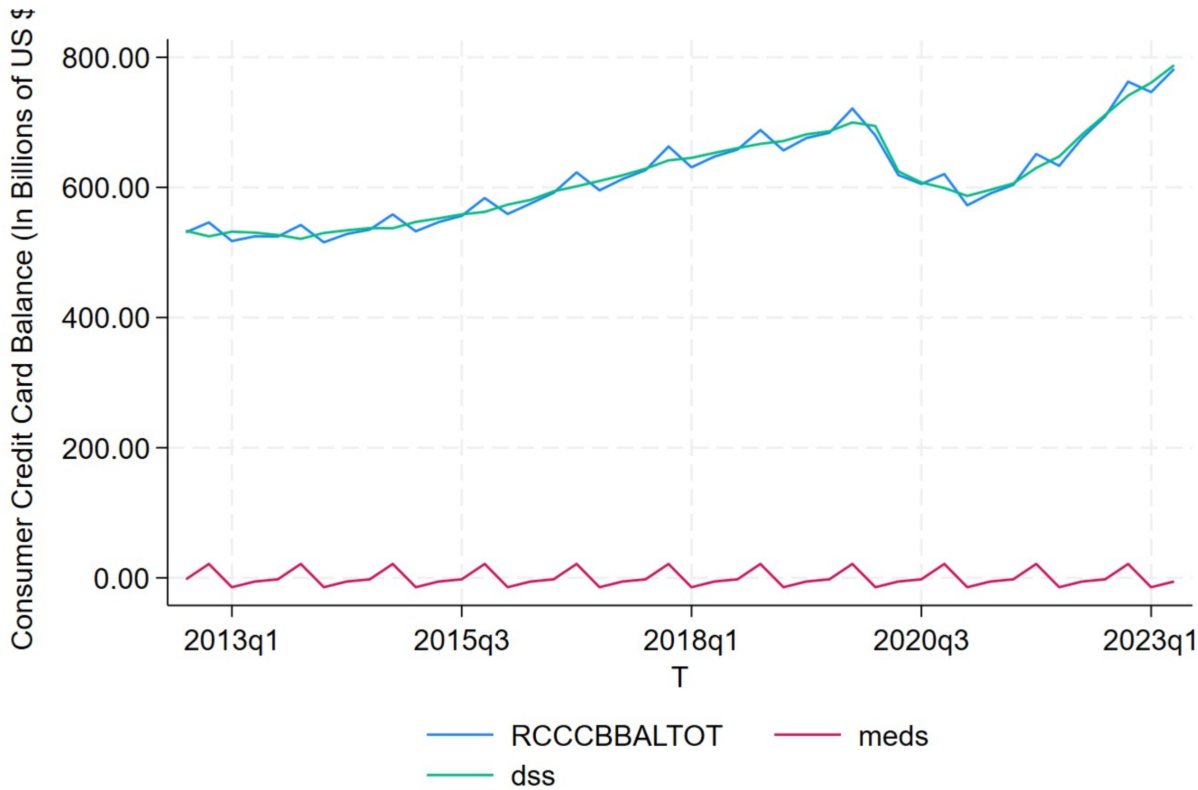

Figure 2. Seasonal decomposition produced using stata

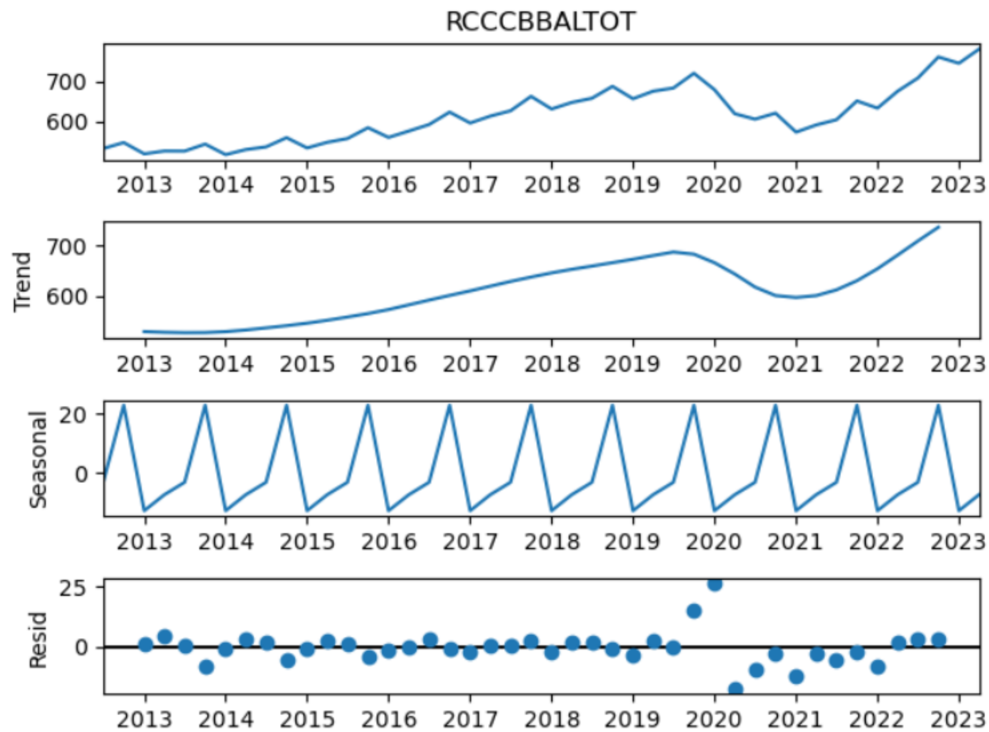

Figure 3. Seasonal decomposition produced using Python Statsmodel

From the plot there seems to be a clear seasonality in the data. We will need to handle the downtrend around years 2020 to 2021 and use seasonal models to forecast the future data. We will take part of 2022 and year 2023 as test data to evaluate our forecasts.

Preliminary Forecasting Techniques

In the initial stage of our analysis, we designated the period from July 1, 2022, to April 1, 2023, as our testing window, utilizing the historical data preceding this interval as the training dataset. Our approach began with the application of various rudimentary forecasting techniques, each with unique assumptions and mechanisms for predicting future values. The methods implemented include the Naïve, Average, Seasonal Naïve, Drift, and Moving Average forecasts. These methods were chosen for their foundational relevance in time series forecasting, serving as a benchmark against more sophisticated models.

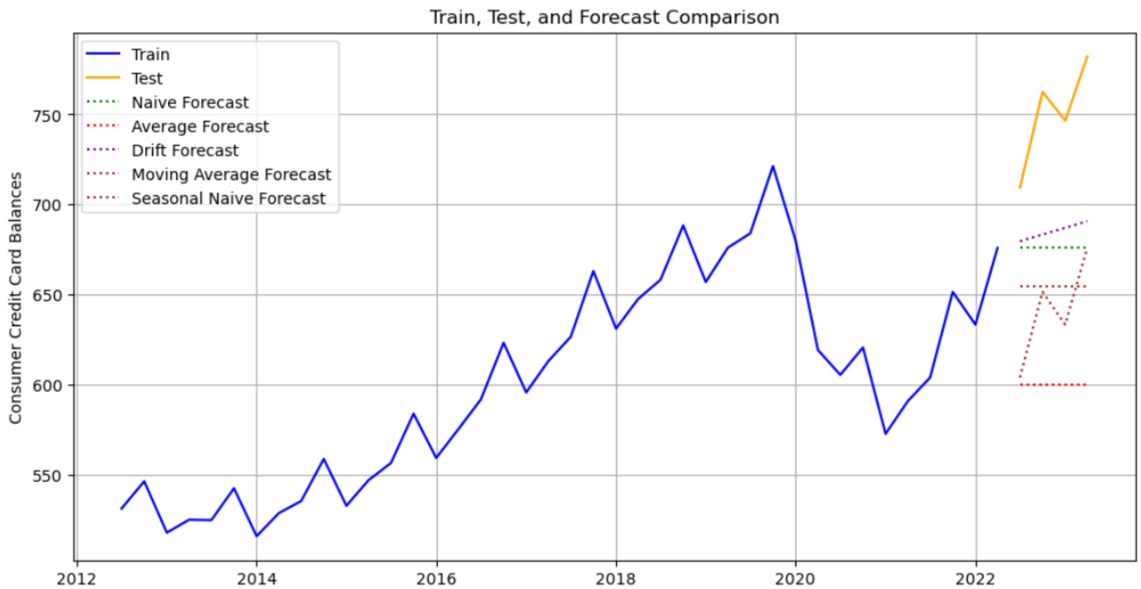

Figure 4. Preliminary forecasting methods prediction

Upon plotting the forecasts, we observed that none of the rudimentary methods could closely match the actual values. However, the drift forecast seemed to follow the observable trend to a certain extent, performing better than the other naive models applied. This relative performance suggests that while simplistic, the drift method can somewhat account for the trend component present in the data.

The forecasted and actual values for the test period are tabulated as follows:

Date	**Actual**	**Naive Forecast**	**Average Forecast**	**Drift Forecast**	**Moving Average Forecast**	**Seasonal Naive Forecast**
7/1/2022	709.3	675.9	600.2	679.6	654.6	603.8
10/1/2022	762.6	675.9	600.2	683.3	654.6	651.4
1/1/2023	746.6	675.9	600.2	687	654.6	633.3
4/1/2023	782.2	675.9	600.2	690.7	654.6	675.9

Table 2. Forecasted values

Subsequently, we calculated the forecast errors using Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), which are summarized as follows:

Forecast Method	**MAE**	**RMSE**
Drift	65	69.1
Naive	74.3	79
Moving Average	95.6	99.3
Seasonal Naive Forecast	109.1	109.1
Average	150	152.3

Table 3. Forecast error comparison

The error metrics underscore the superior performance of the drift method in capturing the series' trajectory, albeit still leaving room for improvement. The higher errors in other models point to the need for more advanced techniques that can better encapsulate the complex dynamics of the series.

Forecasting Techniques Continued

Building on our preliminary models, we progressed to more advanced techniques with the ability to capture different levels of data patterns such as trend and seasonality. We applied Exponential Smoothing, Double Exponential Smoothing, and Holt-Winters Forecasting models to our dataset.

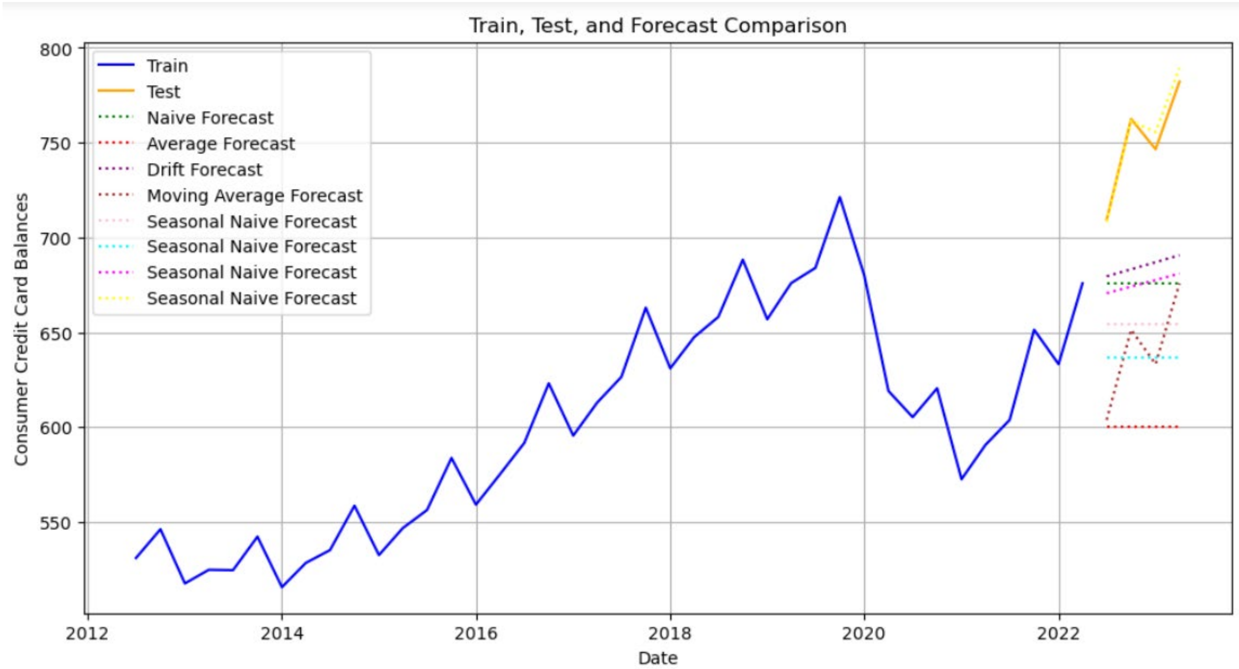

Figure 5. Expanding from basic forecasting methods. Holt Winters, Exp Smoothing, Double Exp Smoothing

The forecasts generated using these advanced techniques for the test period are provided below:

Data	Actual	Exponential Smoothing	Double Exponential Smoothing	Holt-Winters
7/1/2022	709.26	636.45	670.68	708.24
10/1/2022	762.56	636.45	674.13	761.61
1/1/2023	746.56	636.45	677.59	755.57
4/1/2023	782.22	636.45	681.05	789.42

Table 4. Forecasted values.

To evaluate the accuracy of these forecasts, we computed the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) for each method:

Forecast Method	MAE	RMSE
Holt-Winters	4.544965	5.809062
Double Exponential Smoothing	74.287207	77.942767
Exponential Smoothing	113.695672	116.805604

Table 5. Forecast error metrics.

The Holt-Winters method demonstrated the highest accuracy with the lowest MAE and RMSE, suggesting its effectiveness in capturing the underlying patterns in the data. This is consistent with the seasonal patterns observed in consumer credit card balances, as Holt-Winters is designed to account for such seasonality. Conversely, Exponential Smoothing showed the highest errors, likely due to its inability to accommodate the trend and seasonal variations present in our series.

Advanced Forecasting Techniques – SARIMA

Given the pronounced seasonal variations in our dataset, we elected to utilize the Seasonal Autoregressive Integrated Moving Average (SARIMA) model for forecasting. SARIMA is particularly adept at handling data with both non-stationarity and seasonality, making it well-suited to our series.

```
Dickey-Fuller Test:
ADF Statistic: -1.285181
p-value: 0.635925
Critical Values:
        1%: -3.639
        5%: -2.951
        10%: -2.614
```

Our initial step involved testing for stationarity with the Dickey-Fuller test, which indicated a non-stationary series. To address this, we applied first-order differencing, which brought the p-value closer to the critical threshold, rendering the series tentatively stationary.

Perform first order differencing and check again,

```
Dickey-Fuller Test:
ADF Statistic: -2.789006
p-value: 0.059861
Critical Values:
        1%: -3.679
        5%: -2.968
        10%: -2.623
```

After differencing we see the p-value to be near the critical value which is acceptable

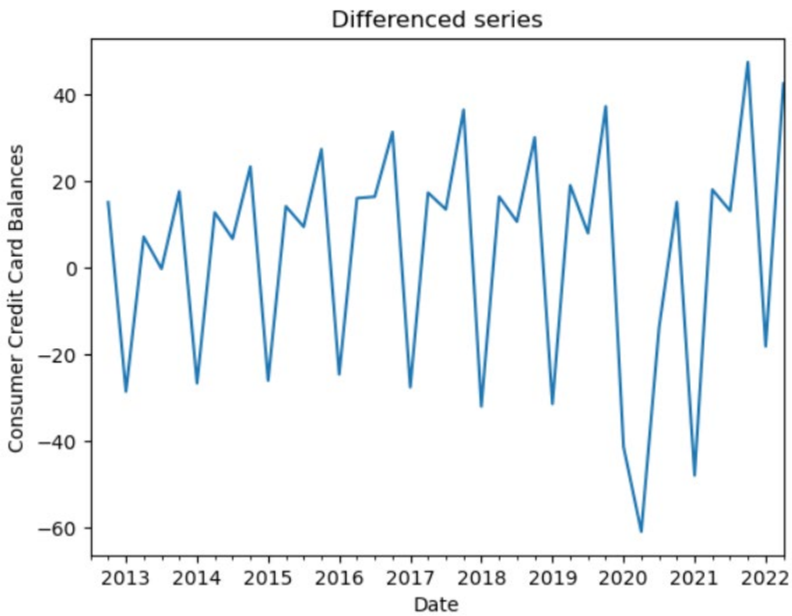

Figure 6. First order differenced series

Since we see a clear seasonality, we also perform a seasonal differencing on the data,

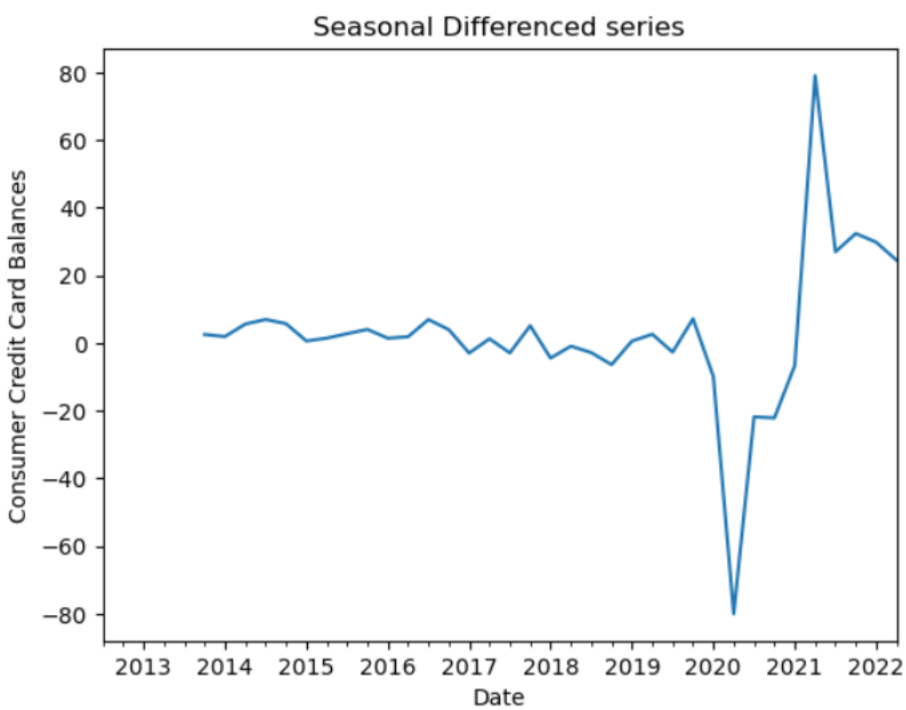

Figure 7. First order seasonally differenced series

Perform dickey fuller test on seasonally differenced data,

```
Dickey-Fuller Test:
ADF Statistic: -2.637512
p-value: 0.085490
Critical Values:
        1%: -3.700
        5%: -2.976
        10%: -2.628
```

After seasonal differencing we see the p-value to be at ~0.085. This is a little above the critical value of 0.05. The test might inflate the value because of the rapid change in the series during covid period.

Acknowledging the strong seasonal component, we further conducted seasonal differencing. Post-seasonal differencing, the Dickey-Fuller test yielded a p-value of approximately 0.085. While slightly above the conventional critical value of 0.05, this result may be attributed to the abrupt shifts during the COVID-19 period. We proceeded with additional testing, excluding data from the peak of the pandemic (prior to April 1, 2020), which provided a more stable basis for our models.

Perform Dickey fuller test on the series without the covid period < "2020-04-01"

```
Dickey-Fuller Test:
ADF Statistic: -4.238568
p-value: 0.000567
Critical Values:
        1%: -3.724
        5%: -2.986
        10%: -2.633
```

We then plot the ACF and PACF plots so that we can decide on the SARIMA model specification.

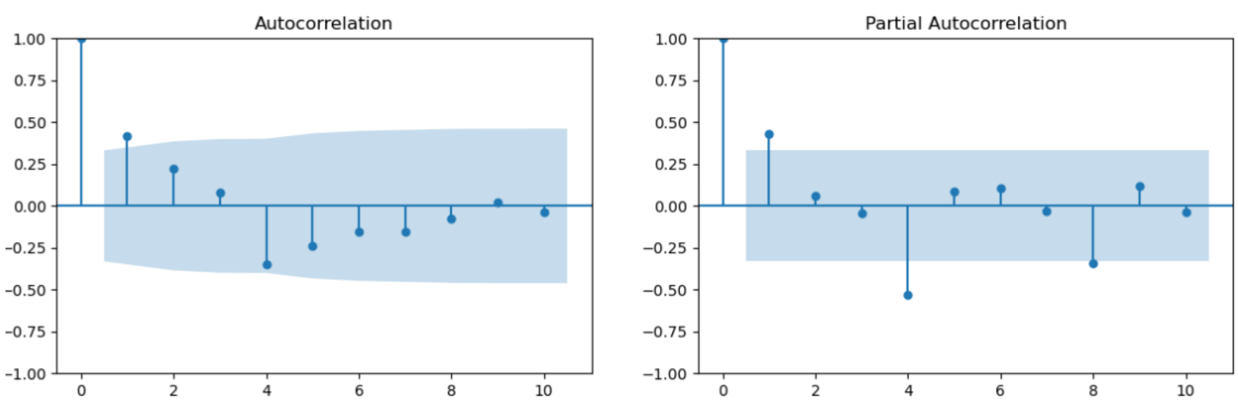

Figure 8. ACF and PACF Plots

The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots were instrumental in identifying significant lags for the AR and MA components. The first lag's significance in both plots suggested the inclusion of AR(1) and MA(1) components. We tested various model specifications to determine the optimal combination, guided by the lowest Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and error metrics.

We evaluated multiple SARIMA configurations, each yielding its own set of metrics. The following table presents a comparative analysis of these configurations:

pdq	seasonal_pdq	aic	bic	mae	rmse
(1, 1, 1)	(0, 1, 1, 4)	301.8	308.1	29.5	33.5
(1, 1, 1)	**(1, 1, 0, 4)**	**310.3**	**316.5**	**4.0**	**5.4**
(1, 1, 1)	(1, 1, 1, 4)	303.8	311.6	29.4	33.4
(2, 1, 2)	(2, 1, 0, 4)	307.3	318.1	24.7	35.8
(2, 1, 2)	(2, 1, 1, 4)	307.2	319.7	30.4	38.6
(1, 2, 1)	**(0, 1, 1, 4)**	**298.5**	**304.7**	**11.5**	**14.1**
(1, 2, 2)	(0, 1, 1, 4)	299.2	306.8	21.6	24.3
(1, 2, 2)	(1, 1, 1, 4)	301.2	310.4	22.6	25.3
(0, 2, 2)	**(0, 1, 1, 4)**	**298.5**	**304.6**	**10.3**	**12.6**
(2, 2, 2)	(0, 1, 1, 4)	300.9	310.1	30.4	34.3

Table 6. SARIMA forecast AIC, BIC, error metrics comparison.

While some models excelled in specific metrics, a holistic view of AIC, BIC, MAE, and RMSE was necessary to select the most robust model. Upon plotting the forecasts from these models, we observed varied performance across the different specifications. Plotting them we see,

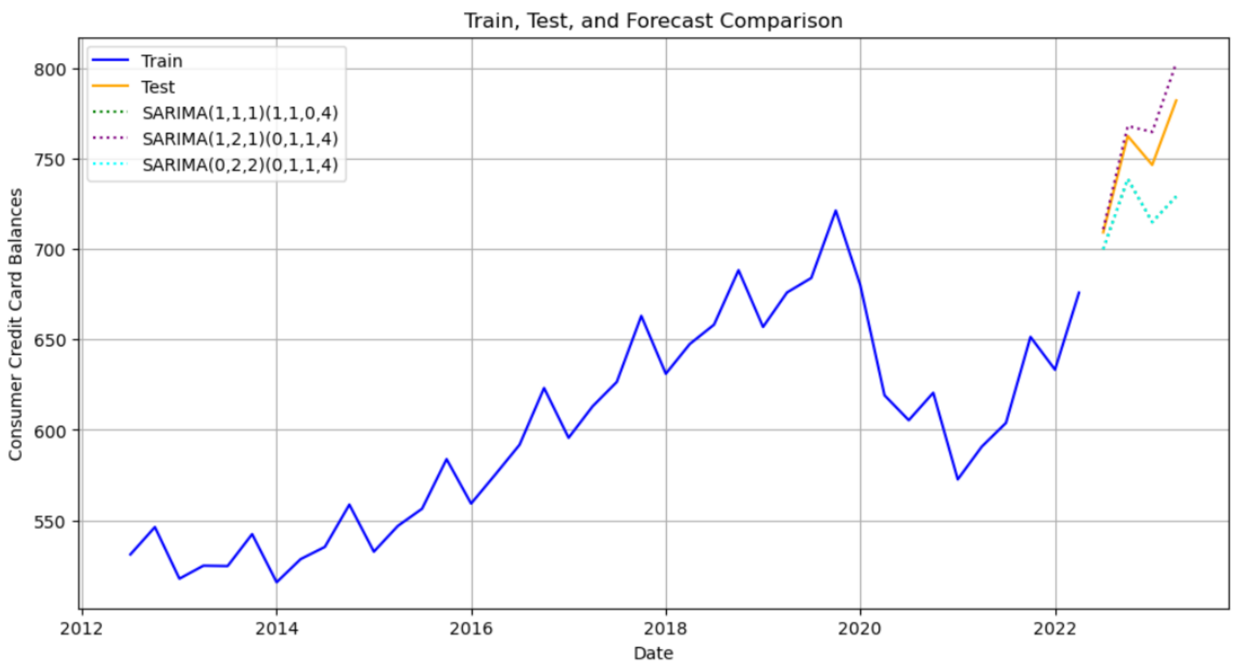

Figure 9. SARIMA Forecasts

The SARIMA model's efficacy is evidenced in its ability to adapt to the intricacies of the data, capturing the seasonal trends and adjusting for non-stationarity. The diversity in model performance underscores the necessity of a nuanced approach to selecting the appropriate model, balancing complexity against predictive accuracy.

Advanced Forecasting Techniques – SARIMAX

In our quest to refine the forecasting model, we introduced the Seasonal Autoregressive Integrated Moving Average with eXogenous variables (SARIMAX) model. This model incorporates external factors—in our case, the COVID-19 pandemic—identified as significant disruptors of consumer credit card spending patterns.

The model was augmented with a 'covid index' to represent the period between April 1, 2020, and January 1, 2021, encapsulating the timeframe when COVID-19 markedly altered consumer behavior. The inclusion of this variable aimed to quantify the pandemic's impact on the credit card balances. The specific range was chosen in line with the pattern disruption in the series.

We explored various SARIMAX configurations, each assessed by their respective Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE):

pdq	seasonal_pdq	aic	bic	mae	rmse
(1, 1, 1)	(0, 1, 1, 4)	242.2	250.0	9.5	12.3
(1, 1, 1)	**(1, 1, 0, 4)**	**240.8**	**248.5**	**7.2**	**10.0**
(1, 1, 1)	(1, 1, 1, 4)	242.1	251.4	8.9	12.6
(2, 1, 2)	(2, 1, 0, 4)	235.6	248.0	55.1	59.4
(2, 1, 2)	(2, 1, 1, 4)	237.6	251.6	54.4	58.7
(1, 2, 1)	(0, 1, 1, 4)	236.5	244.1	28.5	36.3
(1, 2, 2)	(0, 1, 1, 4)	233.7	242.8	23.4	30.6
(1, 2, 2)	**(1, 1, 1, 4)**	**232.2**	**242.9**	**21.7**	**29.5**
(0, 2, 2)	**(0, 1, 1, 4)**	**232.7**	**240.4**	**25.8**	**33.4**
(2, 2, 2)	(0, 1, 1, 4)	235.0	245.6	19.2	26.7

Table 7. SARIMA forecast AIC, BIC, error metrics comparison

The assessment of various SARIMAX models revealed that certain specifications outperformed others in specific metrics. For instance, some configurations demonstrated lower AIC and BIC values, while others excelled in minimizing forecast errors, MAE, and RMSE. These metrics informed the selection of the most appropriate model for our data.

The significance of the covid index was evaluated to ascertain its contribution to the model. The variable exhibited a p-value indicating strong statistical significance, confirming its role as a meaningful predictor within the specified timeframe.

```
                                  SARIMAX Results
==========================================================================================
Dep. Variable:                          RCCCBBALTOT   No. Observations:                 40
Model:             SARIMAX(2, 2, 2)x(0, 1, [1], 4)   Log Likelihood                -110.478
Date:                              Thu, 14 Dec 2023   AIC                            234.957
Time:                                      22:25:08   BIC                            245.641
Sample:                                  07-01-2012   HQIC                           238.600
                                       - 04-01-2022
Covariance Type:                                opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
covid_index  -65.7384      5.660    -11.615      0.000     -76.832     -54.645
ar.L1          1.0374      0.374      2.777      0.005       0.305       1.770
ar.L2         -0.7821      0.571     -1.369      0.171      -1.902       0.338
ma.L1         -0.8959      6.902     -0.130      0.897     -14.424      12.632
ma.L2          0.9956     15.110      0.066      0.947     -28.619      30.611
ma.S.L4       -0.1482      0.239     -0.620      0.536      -0.617       0.321
sigma2        34.7289    514.126      0.068      0.946    -972.939    1042.397
===================================================================================
Ljung-Box (L1) (Q):                   0.74   Jarque-Bera (JB):                 0.54
Prob(Q):                              0.39   Prob(JB):                         0.76
Heteroskedasticity (H):               5.13   Skew:                            -0.25
Prob(H) (two-sided):                  0.01   Kurtosis:                         3.36
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```

From the p-value of the covid index we can conclude the newly added variable is statistically significant and is a useful predictor.

Based on the metrics such as AIC, BIC, MAE and RMSE we see multiple SARIMAX specifications to perform best in at least one of the metrics. Plotting them we see,

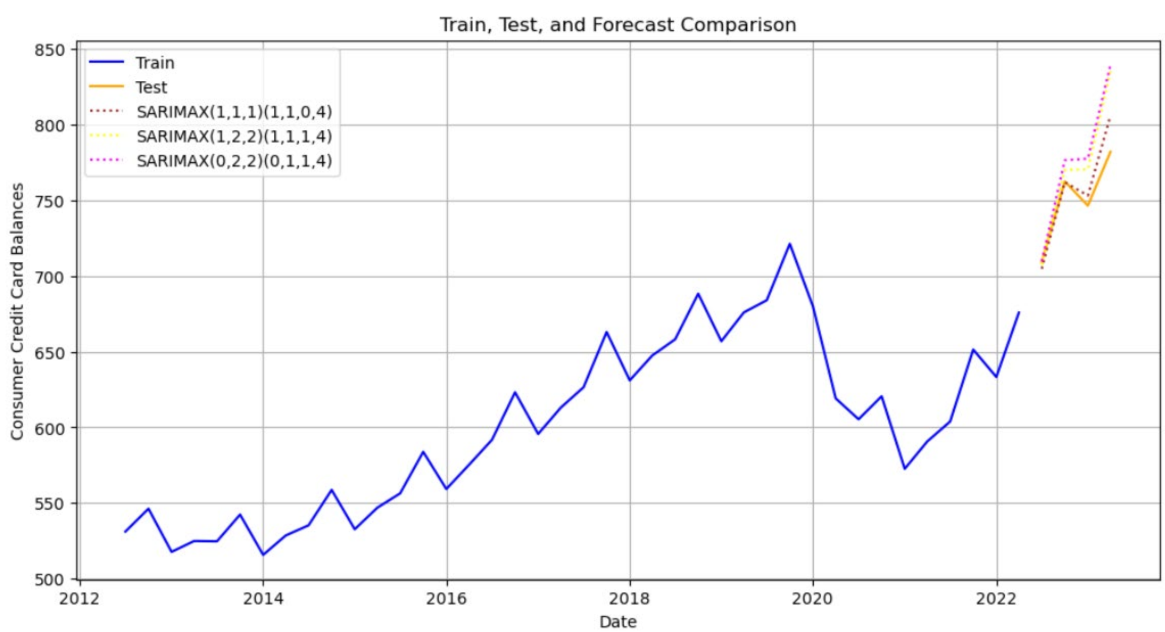

Figure 10. SARIMAX Forecasts

The integration of the covid index into the SARIMAX model provided valuable insights, allowing us to accommodate the anomalous effects of the pandemic on consumer spending. This advanced technique enhanced our forecasting capability, ensuring that the model remained robust and relevant in the face of unprecedented external shocks.

Conclusion

Through rigorous testing and refinement of forecasting methods, we have determined that the Holt-Winters method excels in predictive accuracy as evidenced by its superior performance across AIC, BIC, and error metrics during the specified testing period. However, the applicability of a forecasting method extends beyond numerical metrics; it encompasses the method's ability to adapt to the data's inherent patterns and external influences.

Our exploration of SARIMA and SARIMAX models has revealed that both exhibit commendable forecasting prowess. These methods, grounded in a thorough understanding of the data's seasonal behavior and non-stationarity, have shown a strong alignment with observed values, as illustrated by our error metrics, and corroborated by visual inspection of the forecasts.

In particular, the SARIMAX model, with its integration of the covid index, adeptly captures the nuances introduced by the pandemic, a testament to its capability to assimilate significant exogenous variables. This model's adaptability in tracing both trend and seasonality, while factoring in the extraordinary circumstances posed by the pandemic, demonstrates its robustness and potential utility in predictive analytics within the banking sector.

It is important to acknowledge the temporal boundary of our tests. Forecasting is inherently an extrapolation of past behavior into the future, and while our models have been validated against recent data, their long-term efficacy would benefit from continuous evaluation against emerging trends and economic conditions.

In conclusion, while the Holt-Winters method stands out based on selected metrics, the advanced SARIMA and SARIMAX models offer a nuanced understanding of complex behaviors and external disruptions, such as those encountered during the COVID-19 pandemic. The choice of model thus hinges on the specific forecasting objectives, the nature of the data, and the contextual relevance of external factors. As the financial industry continues to evolve amidst a landscape shaped by unforeseen events, our approach underscores the importance of flexible, responsive, and comprehensive modeling techniques in strategic decision-making and risk assessment.

References

1. https://www.cnbc.com/2023/08/10/average-consumer-carries-5947-in-credit-card-debt-a-10-year-high.html
2. https://www.consumerfinance.gov/about-us/newsroom/cfpb-report-finds-credit-card-companies-charged-consumers-record-high-130-billion-in-interest-and-fees-in-2022/