

Modeling Investment Grade Corporate Bond ETFs

Charles Dotson

December 9, 2022

Contents

1	Introduction	2
2	Data	3
3	Evaluation	3
3.1	Stationarity	3
3.1.1	Augmented Dicky Fuller	3
3.1.2	Phillips-Perron	4
3.1.3	Autocorrelation Function	4
3.2	Conclusions	4
4	Transformations	5
5	Re-Evaluations	5
5.1	Timeseries	5
5.2	Augmented Dicky Fuller	6
5.3	Phillips-Perron	6
5.3.1	Autocorrelation Function	6
5.4	Conclusions	6
A	Figures	7
B	Tables	7
C	Code	7

1 Introduction

Exchange traded funds, ETF for short, are growing to be one of the largest new asset classes in the public markets today. In 2002 the total ETF assets under management (AUM) was just \$102 bb and in 2021, it had grown to \$7,191 bb. This number is still growing. There are ETFs which just follow an index all the way to those which are essentially multi strategy hedge funds whose AUM grows through the *create and redeem* process. This process is not the focus of this paper but to quickly discuss this. Imagine an actively managed hedge fund has a portfolio of assets, they then are able to have “*buckets*” for creating or redeeming a share of said ETF. For the create side, an investor would approach a broker dealer to with the said “*create bucket*”, just a portfolio of assets deemed by said ETF to be able to exchange for a share of the ETF. Then the broker dealer will initiate the exchange. The ETF will get the bucket of assets and the investor will get a share of the ETF. When the bucket in question is the “*redeem bucket*”, this exchange of assets is just in reverse. This new ability for people to have access to either indexes, money managers, sectors, leverage by investing in a levered ETF, etc., along with the added liquidity these products provide is the main reason for their exponential growth.

Since the ETF space is growing exponentially and allowing investors easier access to whole sectors, product types, strategies, funds, you name it, as well as the fact that these products are also tradeable in the market (regular buying and selling on an exchange), there comes a new opportunity. With a strong model for forecasting these products, a smart investor and/or trader now can take advantage of movements in areas once not able to take advantage of with the risks that come with trading derivatives. This is the goal of this paper.

We will be conducting research into formualting an ARIMA model for Investment Grade (IG) Corporate Bond ETFs. The bond is not as liquid as the equity/derivative market but through ETFs, this liquidity issue is dampened. Thus, forming a model to be able to forecast price movements of an IG corporate bond ETF would allow a fixed income portfolio manager (PM) to not suffer from the lack of liquidity that trading singular bonds entails as well as avoiding the risk of derivatives that come with trading the accompanying index/s. For our purposes we will choose to model the BlackRock ETF “*iShares iBoxx \$ Investment Grade Corporate Bond ETF*”, which has the ticker symbol **LQD**. LQD’s investment objective on [LQD’s homepage](#):

The iShares iBoxx \$ Investment Grade Corporate Bond ETF seeks to track the investment results of an index composed of U.S. dollar-denominated, investment grade corporate bonds.

LQD is also one of if not the most popular IG corporate bond ETFs. We can measure this “*popularity*” by KPIs such as average volume and net assets. These stats can be found by going to [yahoofinance.com](#) and searching “*LQD*” in the search bar (or by clicking [here](#)).

Our data, adjusted daily closing prices, was collected using the Python API **yfinance** ([documentation here](#)). This package is a webscraping package used to get historical data on any number of given names freely avaiable on [yahoofinance.com](#). The reasoning behind using adjusted closing prices instead of closing prices is the dividend adjustment. If we were to model the actual price, forecasts will be off due to the LQD paying dividends. Since we are delaing with a publicly traded asset, the market is closed on weekends and some holidays. Thus, instead of using 365 days for a year, our data is actually 252 days per year on average.

2 Data

We apply the general “10-year” rule to the amount of data we bring in (if the holding period of a strategy is a day or less, a backtest should be no less than 10 years). Our start and end dates for our data are 1/3/2012 and 1/3/2022 respectively. These are the first trading days of 2012 and 2022. This gives us a total of 2518 data points.



Figure 1: time-series plot of LQD’s price from 01/03/2012 - 01/03/2022

3 Evaluation

3.1 Stationarity

To begin formualting our ARIMA model, we must first check for stationarity. We will use the Augmented-Dickey Fuller test, and Phillips-Perron test, and the plotting of the Autocorrelation Function.

3.1.1 Augmented Dicky Fuller

Test Statistic	-0.164
P-value	0.943
Lags	21

Table 1: Augmented Dickey-Fuller Results

Trend: Constant

Critical Values: -3.43 (1%), -2.86 (5%), -2.57 (10%)

Null Hypothesis: The process contains a unit root.

Alternative Hypothesis: The process is weakly stationary.

3.1.2 Phillips-Perron

Test Statistic	-0.504
P-value	0.891
Lags	27

Table 2: Phillips-Perron Test (Z-tau)

Trend: Constant

Critical Values: -3.43 (1%), -2.86 (5%), -2.57 (10%)

Null Hypothesis: The process contains a unit root.

Alternative Hypothesis: The process is weakly stationary.

3.1.3 Autocorrelation Function

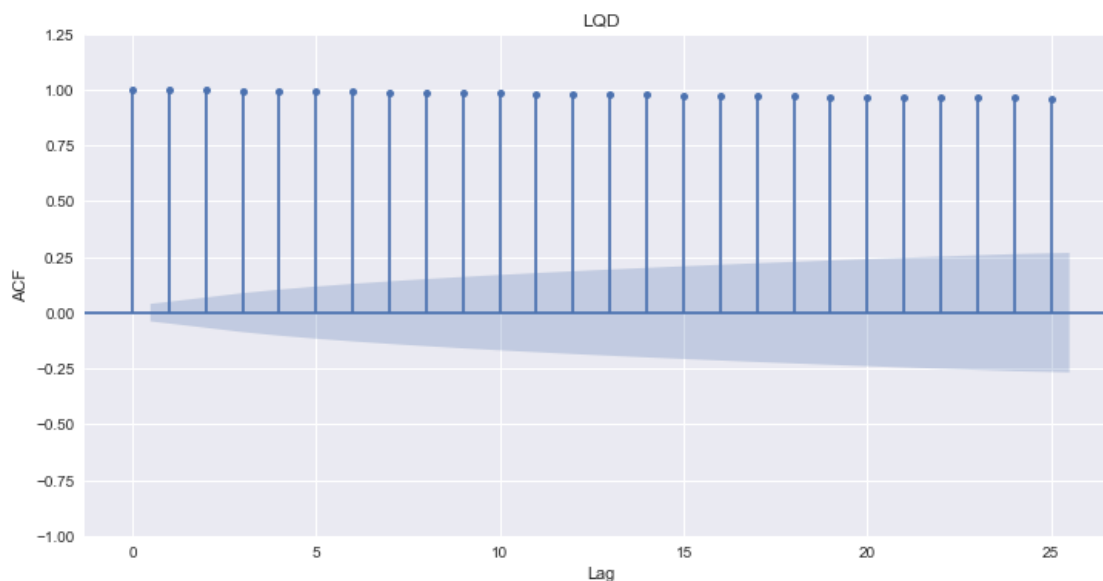


Figure 2: the Autocorrelation function of LQD's price series

3.2 Conclusions

Let us first start with the ADF test in 1. We can clearly see that with a P-value of 0.891, that the process clearly contains a unit-root. This is also corroborated by the PP-test in 2. Referring back to the ACF plot in ?? we can clearly see basically no decay and a clearly trending process. With all three of our methods returning non-stationarity in both mean and variance we must look to perform appropriate transformation/s on our and re-evaluate out our new process.

4 Transformations

It is common/best practice in the financial world to work with a log-returns when modeling price data. This is equivalent to a log-differencing transformation. Where our transform is,

$$y_{new}(t) = \log\left(\frac{y_t}{y_{t-1}}\right)$$

This transformation is advantageous due to their symmetry amongst other things. It should allow us to create a stationary mean and variance [1].

5 Re-Evaluations

5.1 Timeseries

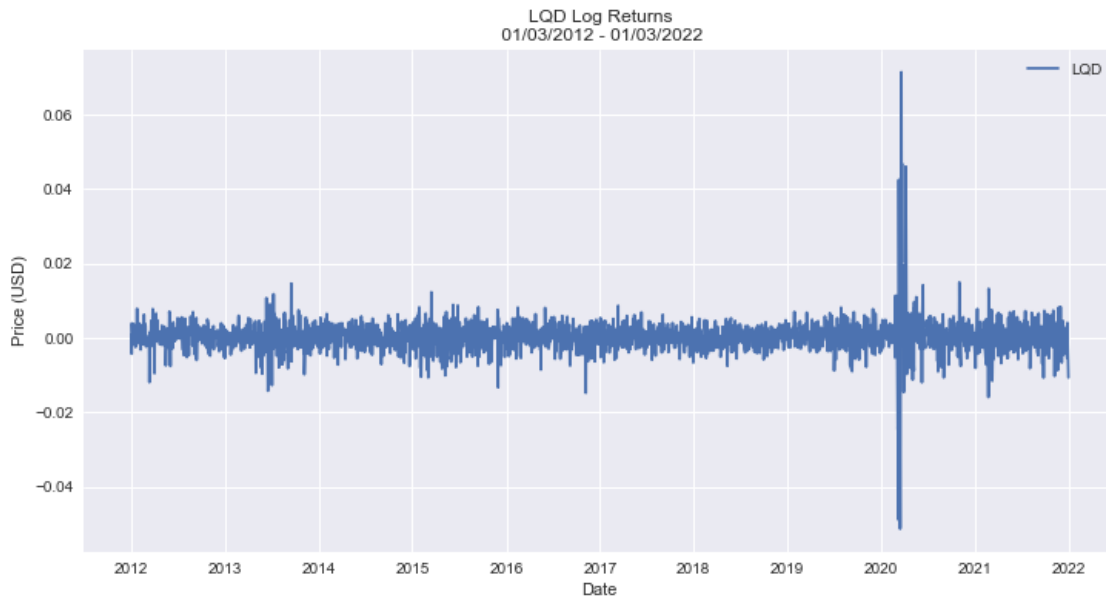


Figure 3: time-series plot of LQD's log-difference price-series 01/03/2012 - 01/03/2022

5.2 Augmented Dicky Fuller

Test Statistic	-13.630
P-value	0.000
Lags	20

Table 3: Augmented Dickey-Fuller Results of LQD log-returns

Trend: Constant

Critical Values: -3.43 (1%), -2.86 (5%), -2.57 (10%)

Null Hypothesis: The process contains a unit root.

Alternative Hypothesis: The process is weakly stationary.

5.3 Phillips-Perron

Test Statistic	-42.288
P-value	0.000
Lags	27

Table 4: Phillips-Perron Test (Z-tau) log-returns

Trend: Constant

Critical Values: -3.43 (1%), -2.86 (5%), -2.57 (10%)

Null Hypothesis: The process contains a unit root.

Alternative Hypothesis: The process is weakly stationary.

5.3.1 Autocorrelation Function

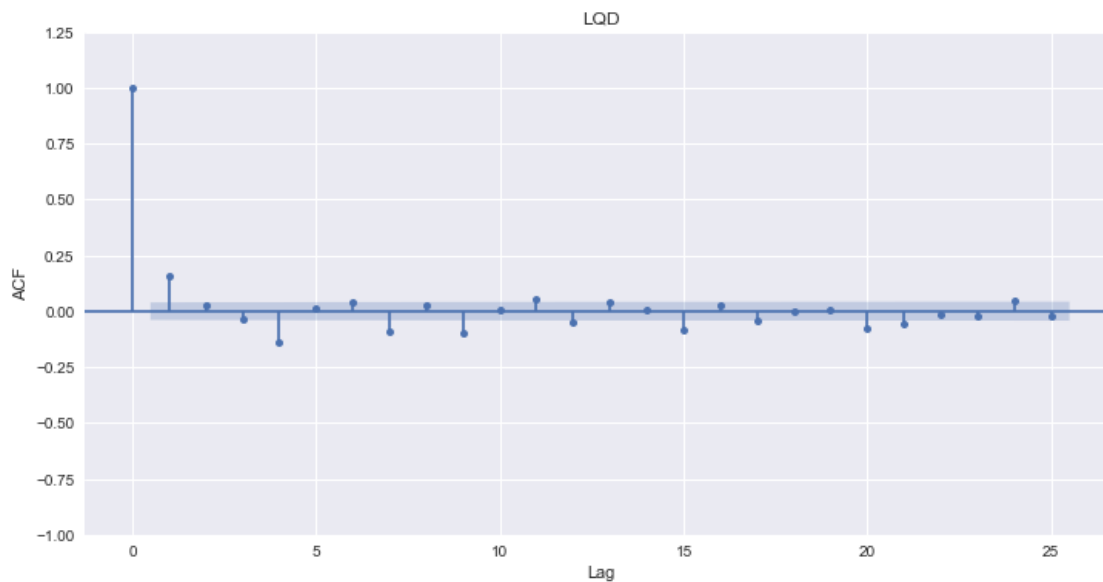


Figure 4: the Autocorrelation function of LQD's log-returns series

5.4 Conclusions

Appendix A Figures

List of Figures

1	time-series plot of LQD's price from 01/03/2012 - 01/03/2022	3
2	the Autocorrelation function of LQD's price series	4
3	time-series plot of LQD's log-difference price-series 01/03/2012 - 01/03/2022	5
4	the Autocorrelation function of LQD's log-returns series	6

Appendix B Tables

List of Tables

1	Augmented Dickey-Fuller Results	3
2	Phillips-Perron Test (Z-tau)	4
3	Augmented Dickey-Fuller Results of LQD log-returns	6
4	Phillips-Perron Test (Z-tau) log-returns	6

Appendix C Code

References

- [1] Robert Hudson and Andros Gregoriou. Calculating and comparing security returns is harder than you think: A comparison between logarithmic and simple returns, Feb 2010.

[76]: `convert('Modeling Investment Grade Corporate Bond ETFs.ipynb')`