B184777 Financial Econometrics Project

Part A - Literature Review on Extensions of the Dickey-Fuller Framework

A major problem faced while analysing time series data is that of non-stationarity. Stationary data is that which has a constant mean, variance and auto-covariance. It's an important assumption typically made before modelling data and it holding is important, as it means that theorems such as the Law of Large Numbers and Central Limit theorem hold. Furthermore, the assumption not holding, and data being non-stationary can lead to poor or false inference from models, or the issue of spurious regression occurring where non-stationarity is interpreted as a causal relationship between variables (Ryan, Haslbeck and Waldorp, 2023).

Non–stationarity can exist in two ways, through a deterministic or stochastic trend. A deterministic trend violates the constant mean assumption, with a measurable trend over time, and is modelled as follows:

$$X_t = \alpha + \delta t + \phi X_{t-1} + \epsilon_t$$

Where δt represents the trending part of the data, and so long as the absolute value of ϕX_{t-1} is <1 the process is mean reverting. Data showing only this kind of trend can be thought of as trend-stationary in that once de-trended and the deterministic part of the data is removed, it will be stationary.

A stochastic trend, on the other hand, is one that is random and not mean reverting, and can be modelled as follows:

$$X_t = \alpha + 1X_{t-1} + \epsilon_t$$

This is known as having a unit root, as the coefficient of the lag of X is 1, meaning X is measured purely as a combination of its last value and other random, non-predictable forces captured by the epsilon term. This means that variance is non-constant, being a function of time as the impacts captured by epsilon never die out due to the coefficient of X being 1. Because of these properties, it is important to be able to identify and account for a unit root, which is where the Dickey Fuller test comes in.

Standard Dickey Fuller / Augmented Dickey Fuller Test

The original, most basic Dickey Fuller test (although not initially named this) was published in 1979, and works as follows (Dickey & Fuller, 1979):

$$\Delta y_t = \beta y_{t-1} + \nu_t \tag{1}$$

$$\Delta y_t = \alpha + \beta y_{t-1} + \nu_t \tag{2}$$

$$\Delta y_t = \alpha + \delta t + \beta y_{t-1} + \nu_t \tag{3}$$

The three above models account for: eq 1) no mean and no drift eq 2) mean (inclusion of the alpha term) eq 3) mean and drift(inclusion again of the alpha term for the mean, and of the delta t term to represent the deterministic aspect of the data).

It tests against the null hypothesis that β is $= 0$, as it works with differenced data and this corresponds to a coefficient of 1 for the undifferenced data, which is the definition of a unit root. The test statistic is calculated as follows:

$$t_\beta = \frac{\hat{\beta}}{\mathrm{se}(\hat{\beta})} \tag{4}$$

and then tested against the Dickey Fuller critical values, which are different to the standard t values (Dickey & Fuller, 1979). Testing for a unit root with a mean/drift uses the same test statistic, but instead makes the alternative hypothesis that the data is a stationary process, exhibiting a deterministic trend.

This test is simplistic, and fails to account for autocorrelation potentially present in the error term. This is particularly prevalent in many financial series, which are often correlated over time. As such, an improved version of this framework was explored in 1984 by Said and Dickey to account for this autocorrelation through the inclusion of additional lags of Y as $\sum_{i=1}^{p} \phi_i$. This is known as the Augmented Dickey Fuller Test(ADF):

$$\Delta y_t = \beta y_{t-1} + \sum_{i=1}^{p} \phi_i \Delta y_{t-i} + \nu_t \tag{5}$$

This model improves on the original by accounting for autocorrelation through the inclusion of lagged variables. Deciding on the number of lags to include can be effectively done through the use of information criteria, be it Akaike or Bayes, or an improved modified criterion proposed by Ng and Perron in 2001. By better specifying the model, this test accounts for serial correlation and improves the test power compared to the standard Dickey Fuller test (Said and Dickey, 1984). Most further improvements to the Dickey Fuller model build on this augmented model as a result.

Augmented Dickey Fuller with Structural Breaks

One possible extension to the standard ADF framework is to include the impact of structural breaks. Perron (1989) highlighted the issues around detecting stationarity around a broken linear trend using the standard ADF method. He discovered that regardless of whether or not the trend is stationary around a break, ADF will have limited power, and the null of a unit root will not be rejected nearly as much as it should. This has since been coined the 'Perron Phenomenon'. Building on this, the 'Converse Perron Phenomenon' has also been researched in detail, and it has been found that when a structural break occurs early in a time series the opposite can occur, with ADF often causing a spurious rejection of the null of a unit root (Leybourne, Mills & Newbold, 1998).

Clearly, the inability to detect stationarity around a break point could be an issue. To account for this, Perron, and as a result much of the subsequent literature, proposed the inclusion of dummies around 'break points' in the ADF model to account for them. The inclusion of these dummies can be done slightly differently depending on whether the break is believed to just change the intercept :

$$\Delta Y_t = \beta Y_{t-1} + \alpha + \delta t + \sum_{i=1}^{p} \phi_i \Delta Y_{t-i} + \gamma \text{SB}_t + \nu_t \tag{6}$$

Or add a new time trend to the data:

$$\Delta Y_t = \beta Y_{t-1} + \alpha + \delta t + \sum_{i=1}^{p} \phi_i \Delta Y_{t-i} + (\gamma_\alpha + \gamma_\delta t)\text{SB}_t + \nu_t \tag{7}$$

The first part of these models are just the ADF discussed above, but it is now modified to either add just a dummy variable at the time of the structural break: γSB_t, or a new additional time trend as well as an intercept change: $(\gamma_\alpha + \gamma_\delta t)$SB. As with the inclusion of lags, Perron(1989) found that the inclusion of these dummies is effective in increasing the power of ADF around a break in structure. It should however be noted that this test requires different critical values to the standard ADF.

Generalised Least Squares Detrending

Typically when testing for a unit root, the first step would be to de-trend the data by removing the deterministic element as follows: When the initial data is modelled as eq 8 and eq9:

$$Y_t = \alpha + \delta t + u_t \tag{8}$$

$$u_t = \phi u_{t-1} + \nu_t \tag{9}$$

And we want to test:

$$H_0 : \phi = 1 \quad \text{[Nonstationary] against } H_1 : \phi < 1 \quad \text{[Stationary]}$$

The trend and mean are estimated, and removed from the data to give us detrended data like such in eq10:

$$Y_t^* = Y_t - \hat{\alpha} - \hat{\delta} t. \tag{10}$$

The ADF can then be used on the detrended data. However, there are certain issues with this method of testing, as highlighted most notably by Elliot, Rothenberg and Stock (1996). Their paper argued that this way of testing has limited power, proving that detrending in this way led to a consistent failure to reject the null of non-stationarity when a series was close to unity. This distinction is very important, as even if a process is close to having a unit root, it will still have very different statistical properties to a true unit root process. A true unit root won't for instance be mean reverting, whereas over time a close to unity series will be (so long as the coefficient to the lag is <1). Eliot, Rothenberg and Stock (1996) proposed an alternative way of detrending data and testing for a unit root, which through Monte-Carlo simulations they proved to have a higher power than the standard ADF.

The detrending method works as follows: Say we have our data modelled as above. We first define a new coefficient in eq 11:

$$\phi^* = 1 + \frac{\tilde{c}}{T} \tag{11}$$

so that it is close to being one (or close to unity), where ($\tilde{c}$) is defined as -7 in the case of a mean but no deterministic trend of -13.5 in the case of both a mean and time series trend. These values were calculated through simulations so the test reaches the envelope of maximum power (Elliot, Rothenberg and Stock (1996). Once calculated, we are able to define new transformed variables as follows, partially differenced to account for near-unity behaviour in eq12:

$$Y_t^* = Y_t - \phi^* Y_{t-1}, \quad \alpha^* = 1 - \phi^*, \quad t^* = t - \phi^*(t-1). \tag{12}$$

We then run another regression on these newly defined variables, to estimate the trend and mean part of the series (eq 13) (as we would have in a standard ADF test):

$$y_t^* = \pi_0 \alpha^* + \pi_1 t^* + u_t^*, \tag{13}$$

and then de-mean and de-trend the residuals (eq 14) to now perform our standard ADF test (though with different critical values as a result of the transformation).

$$\hat{u}_t^* = y_t^* - \hat{\pi}_0 \alpha^* - \hat{\pi}_1 t^*. \tag{14}$$

This process is known as the ADF-GLS test. By transforming to account for the potential for a near–unity ϕ^* before detrending, the ADF-GLS method is able to account for ADF's shortfall in failing to reject non-stationarity in close to unity cases, greatly increasing the power of the test. There are however still some issues with this. For one, there is some disagreement on the best way to select, carry out and define critical values for the test. Ng and Perron (2001) suggest a modified Akaike Information Criterion to find optimal lag length, while Perron and Qu (2007) estimate critical values in a different way. Sephton (2001) demonstrates that especially in smaller sample sizes, these two methods gave different results, arguing that the presence of some subjectivity in how the test is carried out, and the ability of different methods to give slightly different results, is a weakness of the test.

Unit Root With a Null of Stationarity

Typically in an ADF test, as has been seen with all of the extensions here, we take the null hypothesis as there being a unit root present, and the alternative hypothesis of stationary, or stationary with a trend/mean. Kwiatkowski et al (1992) investigate an alternative method, setting the null to be stationary and the alternative to the presence of a unit root. They note that although it changes with the use of ADF-GLS testing, standard ADF models fail to reject a null of a unit root in most aggregate US economic time series. While the standard explanation for that was that they simply contained stochastic trends, they wanted to challenge this view by switching the null hypothesis to see if that made a difference.

The methodology proposed worked as follows:

$$y_t = \alpha + \delta t + w_t + u_t, \tag{15}$$

Where u_t is a disturbance term, δt models the deterministic trend, and w_t is a stochastic trend modelled by the following:

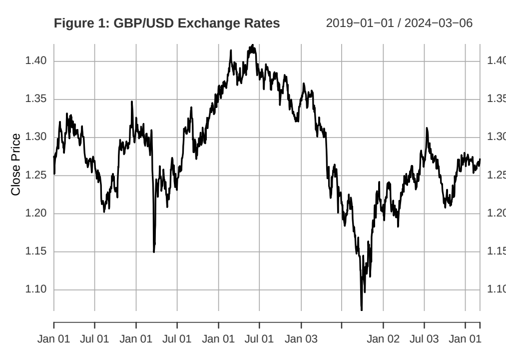
$$w_t = w_{t-1} + v_t, \quad v_t \sim \text{iid } N(0, \sigma_v^2). \tag{16}$$

The new null hypothesis is $H_0 : \sigma_v^2 = 0$. No variance in the stochastic element indicates it is constant, and as such makes the process stationary. Upon using this alternative method of testing, Kwiatkowski et al (1992) found that for 6 out of 14 commonly tested US economic time series aggregates (such as unemployment, GNP etc...) that were declared to have a unit root under ADF, the null of stationarity was not rejected. This led the paper to conclude that there was simply not enough information available in the data to declare either a unit-root or stationarity, and that there was evidence to suggest either one could be true. Although the other 8 matched up to the ADF prediction, 6/14 being inconclusive is an alarmingly high amount and does highlight a potential issue with the ADF test. This model has also been expanded since to work not just on a univariate, but also a multivariate allowing for multiple testing (Choi and Ahn 1999)

Part B - USD/GBP Analysis

B.1

The data being looked at here is the USD/GBP currency pairing from the start of 2019 to the present. Figure 1 shows this data and provides an initial place to look to get an idea of the data.



At first glance, the data seems to be non-stationary. Although it seems there could be a constant mean, variance does not seem constant, and there seems to be autocorrelation. Increases in price seem to be often followed by increases, and the same for decreases, and the size of movements in price also seems to be non constant, though more testing is needed to prove this.

Moments. Mean, Variance, Skewness and Kurtosis.

Figure 2 presents some summary facts about the exchange rate series. Although interesting, return data is what will be looked at to establish the presence of styalised facts about financial data. Typically in return series we see heavy tails (measured by kurtosis, the 4th moment), where there is a higher than normally distributed proportion of extreme values, as well as negative skewing (3rd moment).

Fig 2: Exchange Rate Moment Data

Mean	Variance	Skewness	Kurtosis
1.283043	0.004267651	-0.06711437	2.731465

Fig 3: Return Moment Data

Mean	Variance	Skewness	Kurtosis
1.66366e-05	3.470121e-05	-0.1914126	8.227198

Figure 3 represents data on the 4 moments on our return series. Here we see a mean and variance close to 0. This in itself isn't exceedingly useful, more investigation into autocorrelation of the first two moments (for instance) would be more interesting. Skewness is negative, at -0.19(2d.p), which is typical in financial returns data, implying frequent small gains, and infrequent larger losses. Returns also show kurtosis of 8.23(2d.p) indicating that the data is leptokurtotic (excess kurtosis of 8.23 - 3 = 5.23). This suggests heavier than normally distributed tails, which means more frequent returns at extremes than with data that is normally distributed. Again this is typical with financial data.

Departure From Normality

To test for departure from normality, a Jarque-Bera test was conducted. Its test statistic is found from the following formula: The Jarque-Bera test statistic is defined in eq 1 as:

$$JB = T\left(\frac{SK^2}{6} + \frac{(KT-3)^2}{24}\right) \sim \chi_2^2 \tag{1}$$

and is tested against the null of the data being normally distributed. Conducting this test we find a chi-squared test statistic of 1546.341, giving us a p-value of ~ 0 at 2.2e-16 and allowing us to reject normality at a 0.1% significance level.

Integration

Often in financial data, log prices can be an integrated process, meaning they are non stationary and contain a unit root. Although looking at the graph suggests non-stationarity, in order to test for this a Dickey-Fuller test was conducted. Eq 3 provides a reminder of how we set up the test:

$$y_t = \rho y_{t-1} + u_t \tag{3}$$

where y_t is log price. We test against the null hypothesis that $\rho = 1$ and hence a unit root is present. Conducting this test, we cannot reject the null that a unit root is present, finding a test statistic of -2.32(2d.p), with a p value of 0.44. The presence of a unit root can lead to spurious regressions and makes it very difficult to gain any kind of inference from models run on the data. One way of accounting for it is to difference the process, which works as follows to remove the unit root:

Given

$$y_t = \alpha + \beta y_{t-1} + \varepsilon_t, \tag{4}$$

lag both sides by 1 to obtain

$$y_{t-1} = \alpha + \beta y_{t-2} + \varepsilon_{t-1}. \tag{5}$$

Subtract eq 5 from eq 4 to get:

$$\Delta y_t = \beta \Delta y_{t-1} + \varepsilon_t - \varepsilon_{t-1}. \tag{6}$$

Upon differencing the data, and running an Augmented Dicky Fuller test on the differenced data we now obtain a DF test statistic of -25.27 (2d.p), allowing us to confidently reject the null of non-stationarity at the 1% significance level. This implies the initial process was integrated to the first degree, containing one unit root. We are left with the I(0) part of the series we can now analyse.

Variance Clustering

Finally, we test for variance clustering. This is another property of financial returns data, where periods are likely to have similar levels of variance to their surrounding period. As stated, at first glance this looks to be the case with this price series as areas of high variance seem to be followed by areas with similarly high variance, but to test for this, both the Generalised Auto Regressive Conditional Heteroskedacity (GARCH) and standard ARCH models were considered. Figure 4 shows that comparing information criteria found that both AIC and BIC were lower for the GARCH model for this dataset.

Figure 4: Return Moment Data

Model	AIC	BIC
ARCH(1)	-7.49	-7.47
GARCH(1,1)	-7.58	-7.57

GARCH is an extension of the ARCH model. While the ARCH model works by regressing conditional volatility on the lags of squared residuals, GARCH improves the model by adding in lagged conditional variance terms, helping it capture the more long-term volatility components by assuming variance can be modelled by ARMA. The model is as follows by eq 7:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \tag{7}$$

where ω represents the intercept, ε_{t-i}^2 represents the conditional variance term from the GARCH model and σ_{t-i}^2 the squared-residuals variance term from the ARCH model. Upon running the model, we get the following regression shown by eq 8:

$$\sigma_t^2 = \underset{(5.35e-07)}{1.5e-06^{**}} + \underset{(2.08e-2)}{0.01^{***}} \varepsilon_{t-1}^2 + \underset{(3.19e-2)}{0.085^{***}} \sigma_{t-1}^2 \tag{8}$$

Running the LM - ARCH test, we find the TR^2 test statistic as 15.2(2d.p), allowing us to reject the null of no volatility clustering at the 1% significance level.

Long-Range Dependence

One potential explanation for volatility clustering is the concept of long-range dependence, where strong serial correlation can be observed in the variance of returns. Yaya, Saka and Olawale (2014), looked at the Hurst coefficient to measure this relationship, calculating it on the squared log-returns as a proxy for long-range volatility dependence. They make reference to how a Hurst value > 0.5 implies long-range dependency, while a value $= 0.5$ implies a random walk process, and a value < 0.5 implies a mean-reverting process. There are a number of ways to calculate the Hurst coefficient, but here for simplicity, the R/S method (Hurst, 1956) was used even though it may not be the most robust (Cannon et al, 2011), calculating the coefficient H in accordance with the following relationship shown in eq 9:

$$\mathrm{RS} = (\mathrm{RS})_0 n^H \tag{9}$$

Where R is the range of the integral of the data, S is the standard deviation of the data, and n is the number of evenly split bins of data, giving us a value of 0.63 (2d.p). This leads us to believe there is some form of long-range dependency in the variance of returns, as $0.63 > 0.5$.

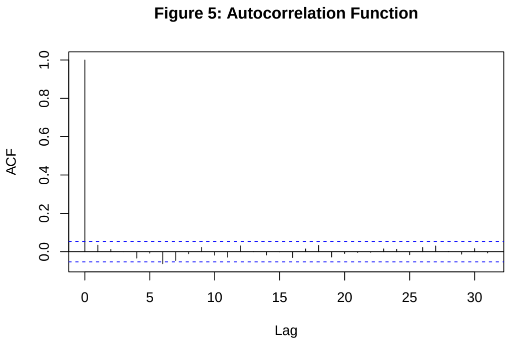
B.2

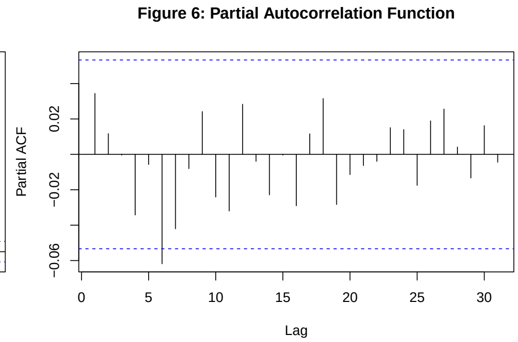
The Auto Regressive Integrated Moving Average (ARIMA) model has proven itself to be effective in forecasting exchange rates, even found to outperform more complex neural network models (Babu and Reddy, 2015). For this reason, ARIMA(p,d,q) was used to conduct the ex-ante forecast. The model works by first differencing the data d times to make it stationary, and then regressing the stationary dependent variable on both p lags of itself, and q lags of its forecast error, or white noise. The model looks as follows:

In order to find the optimal order of the ARIMA model, we first find out to what order the data is intigrated.

Conducting a Dickey-Fuller test on the close values, we get a test statistic of -2.25(2d.p), not allowing us to reject H0 of non-stationarity with a p value of 0.47(2d.p). To combat this, we difference once and repeat the test, now achieving a test statistic of -25.27(2d.p), allowing us to reject non-stationarity at a 1% significance level, proving our required order of differencing is 1.

To now find the optimal MA and AR lags, we can start by looking at the autocorrelation and partial autocorrelation functions of the differenced closing prices:





The lack of statistically significant lags in the ACF and PCF (figs 5&6), apart from potentially at the 6th lag, initially suggests that just a constant mean model without lags could be the most appropriate, but further testing is necessary.

Recent studies that use ARIMA to forecast exchange rates often take a more robust, machine-learning like approach to model specification, initially looping through all possible combinations between certain lags (we initially investigated between 0 and 2 lags) and computing their Bayes and Akaike Information Criterion, and then estimating their Root Mean Squared Error (RMSE) against test data. (Babu and Reddy, 2015. Ying and Hailun, 2007) For this project, a 'corrected' AIC was included, argued to be preferable in selecting ARIMA models by Hyndman and Athanasopoulos (2018) by accounting for a bias in AIC in models in small sample sizes. The values are computed as follows:

$$\text{AIC} = -2\log(L) + 2(p+q+k+1) \tag{10}$$

$$\text{BIC} = -2\log(L) + \log(T)(p+q+k+1) \tag{11}$$

$$\text{AICc} = \text{AIC} + \frac{2(p+q+k+1)(p+q+k+2)}{T-p-q-k-2}, \tag{12}$$

Where p,q,T and k represent the number of AR and MA lags, number of time periods and the intercept of the model, respectively. After checking the models, we get the following information criteria values for up to two lags:

Model	AIC	BIC	AICc
ARIMA(0,1,0)	-9441.108	-9435.899	-9441.105
ARIMA(0,1,1)	-9440.669	-9430.252	-9440.660
ARIMA(0,1,2)	-9438.904	-9423.278	-9438.886

ARIMA(1,1,0)	-9440.707	-9430.290	-9440.698
ARIMA(1,1,1)	-9438.799	-9423.173	-9438.781
ARIMA(1,1,2)	-9437.063	-9416.229	-9437.033
ARIMA(2,1,0)	-9438.892	-9423.267	-9438.875
ARIMA(2,1,1)	-9436.824	-9415.989	-9436.794
ARIMA(2,1,2)	-9437.206	-9411.163	-9437.162

Although the table isn't shown to save space, the maximum lags were increased to 8 to see if higher order models could improve performance, especially in the presence of the potential spikes in ACF and PACF at lag 6, which found ARIMA(8,1,2) to have a AIC, BIC and AICc values of 9441.654 -9384.360 -9441.457 respectively, marginally giving them the lowest AIC and AICc values. The coefficients are shown in eq 13:

$$\begin{aligned}\Delta y_t = &\underset{(0.0324)}{-1.0120^{***}}\Delta y_{t-1} + \underset{(0.0434)}{-0.9171^{***}}\Delta y_{t-2} + \underset{(0.0459)}{0.0430}\Delta y_{t-3} + \underset{(0.0459)}{-0.0235}\Delta y_{t-4} \\ &+ \underset{(0.0460)}{-0.0371}\Delta y_{t-5} + \underset{(0.0460)}{-0.0969^{**}}\Delta y_{t-6} + \underset{(0.0391)}{-0.0855^{***}}\Delta y_{t-7} + \underset{(0.0278)}{-0.0835^{***}}\Delta y_{t-8} \\ &+ \underset{(0.0183)}{1.0490^{***}}\varepsilon_{t-1} + \underset{(0.0209)}{0.9693^{***}}\varepsilon_{t-2} + \varepsilon_t,\end{aligned} \tag{13}$$

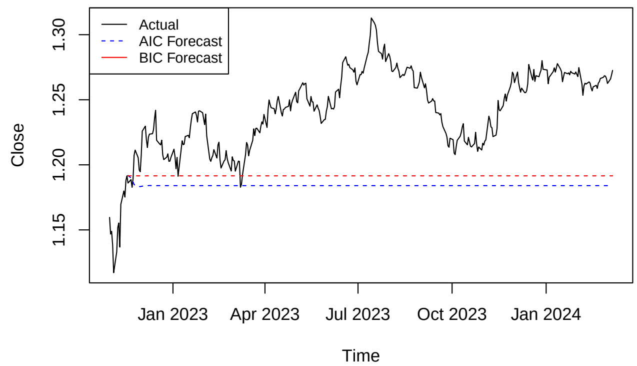
Although here overfitting is a concern due to the high complexity of the model, for comparative purposes we include it to compare its forecasting power to ARIMA(0,1,0), which still has the lowest BIC value. This is to be expected as BIC generally penalizes complexity more than AIC, with the penalty term being a product of log T rather than 2, where T is the sample size (i.e so long as the sample size is above 7, as it is here, it will be higher penalized).

To compare the two models selected by AIC and BIC, we split the exchange rate data into a test (1/4 of the data) and training sample (3/4 of the data), and then compute the RMSE of the test sample from the model found through the training data, based on each of the ARIMA models selected by the information criteria. The goal is to minimise RMSE, as this suggests a closer fit of the model to the data. RMSE is calculated as follows in eq 14:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{14}$$

where $\hat{y}_i$ is the forecast value, and y_i the true value.

Figure 7: Actual vs. Forecasted Values



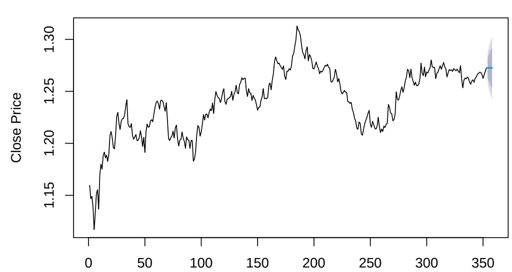
We find that ARIMA(0,1,0) has a lower RMSE than the higher dimension ARIMA(8,1,2), of 0.059(2d.p) compared to 0.066(2d.p), suggesting the ARIMA(0,1,0), a constant mean model on the differenced data, has the highest predictive power. ARIMA(0,1,0) is a random walk and implies that there is no information to be gained on future price from past information from AR or MA lags, in alignment with the efficient market hypothesis. Although from the figure 7 for this particular forecast the higher order model seems to fit the data best initially, testing this on other windows finds this is not the general case. Intuitively, a constant mean model does make some sense. It fits with ideas in the efficient market hypothesis, which states that it should not be possible to use past information to predict future price movements. It also fits with what the ACF and PACFs suggest, of no lags providing useful information.

The differenced ARIMA(0,1,0) model can be written as:

$$\Delta Y_t = \mu + \varepsilon_t \tag{15}$$

where μ is 0. in this case.

Figure 6: Forecast GBP/USD Exchange Rates



Using this model, as it just forecasts a constant mean, we predict the next 5 days will all have the same exchange rate of the series constant mean, 1.272(3d.p) USD/GBP, shown in figure 6, which represents the exchange rate from 2022-10-31 to 2024-03-01, with the blue line representing the 5 day forecast values. The shaded area represents the 95% and 99% confidence intervals for the forecast.

B.3

Value at risk (VaR) is a useful tool to help estimate how much, in an unlikely scenario, is at risk of being lost. So by calculating the 1% VaR, we calculate what we are at risk of losing in the worst 1% of the time for a given time period. Here we calculate the 1% VaR for the return series in 1 and 5 days. There are three ways to calculate it:

Historical Simulation

This method is the simplest, just calculating the first percentile of distributed returns. Doing this, we find a one day 1% VaR value of 0.015 (3d.p).

Variance method

This method works by assuming that returns are normally distributed (though we have proven this is not the case with this series, so this is not the most accurate way of calculating VaR) and then as we know that 1% of the value lies at the tail with the critical value of -2.33, we can calculate the following:

$$VaR(1\%) = \hat{\mu} - 2.33 \times \hat{\sigma} \tag{16}$$

, where μ is the estimated mean and sigma the estimated standard deviation.

Following this formula, we get a 1% one day VaR of 0.014(3d.p).

Monte-Carlo Simulation

This method works by first estimating an appropriate model for the historical data. Here we used the ARIMA(0,0,0) model as it best fits the data (found in the same way as the exchange forecast), storing the residuals. Then, you recursively forecast the model h times, randomly assigning a residual to each estimation and storing the new estimated value for Yt+h. As our model is ARIMA(0,0,0) with 0 mean, this looks like the following:

$$\begin{gathered}
\hat{r}_{T+1} = \mu + \tilde{V}_{T+1}, \\
\hat{r}_{T+2} = \mu + \tilde{V}_{T+2}, \\
\vdots \\
\hat{r}_{T+H} = \mu + \tilde{V}_{T+H},
\end{gathered} \tag{17}$$

where $\tilde{V}_{T+1}$ is a random draw from the collected residuals and μ is 0.

This process is repeated i times (in this case 1000), a distribution of the values is stored and the first percentile is used as the 1% value at risk.

This process found the one day 1% VaR to be 0.014(2d.p), and our five day 1% VaR to be 0.030(2d.p). All three methods find similar values of 1% VaR.

B.4: Covid Event Study

The impact of Covid 19 being declared a pandemic is the event being studied. This event was chosen to try to quantify the impact of the initial chaos around the start of the pandemic around a measureable point. In order to study the impact of this declaration, dummy variables were created for a week prior, the week of, the day of and the week after the declaration was made, as well as when the state of emergency was declared to be over. They were picked to attempt to model how the market responded, week by week, to the build up and declaration of a pandemic, and then to the pandemic being declared over. These were added to the ARIMA(0,0,0) model built on normal returns. The dummy variables, in turn, measure the abnormal returns from the event.

Running this regression, we get the following output:

$$\begin{aligned} \hat{r}_T = & -\underset{(0.0026)}{0.0048^{*}} \text{ prepandemic} - \underset{(0.0058)}{0.0127^{**}} \text{ pandemicday} - \underset{(0.0028)}{0.0127^{***}} \text{ pandemic} \\ & -\underset{(0.0026)}{0.0110^{***}} \text{ postpandemic} + \underset{(0.0058)}{0.0011} \text{ pandemicover} \end{aligned} \tag{17}$$

Another way of quantifying the impact of an event is to first calculate the expected returns through our ARIMA model in the time period we are measuring, find the difference between these excpeted returns and the actual returns, and then test for the significance of these abnormal returns. (Mestel & Gurgul, 2003) The abnormal return is calculated as follows in eq 18:

$$AR_t = R_t - E[R_t|X_t] \tag{18}$$

where AR_t is the abnormal return at time t, R_t is the actual return at time t, and $E[R_t|X_t]$ is the expected return given information up to time t from ARIMA(0,0,0).

Conducting this test, we find returns over this time period to be 0.012(2d.p) lower than expected. Running a students t-test gets a test statistic of -2.85, indicating a decrease in returns significant at the 5% significance level, allowing us to conclude that there was very likely a significant decrease in returns at the time because of impacts around the declaration of the pandemic.

Moving back to the regression in eq17, we see that all statistically significant regressors have had a negative impact on returns. It also suggests that the week leading up to the pandemic being declared, as well as the day of the pandemic being declared over were not in fact that significant in impacting returns. The downwards pressure on the return series indicates the pound weakening relative to the dollar, implying the declaration of the pandemic negatively impacted the value of the pound more than the dollar. The lack of significance of the pandemic being declared over is not entirely suprising, as by that point events like vaccines being introduced, lockdown being lifted and the world generally recovering had already occured. The declaration did not really add any new information to the markets that hadn't already priced in. This is was not necessarily the case with the day the pandemic was declared, as there was far more uncertainty around how serious the disease was, and the declaration of a pandemic could have certainly negatively impacted many peoples' outlooks explaining at least in part the impact in returns.

The lack of significance of the pre-pandemic regressor is however initially a little suprising, but there could be a number of reasons for this. Firstly, the return series being looked at is the return of dollars against the pound. Returns increasing or decreasing indicate a relative strengthening weakening pound compared to the dollar, which is also being heavily impacted by covid. As both currencies were being impacted. No initial change in returns could indicate both currencies being similarly impated in that time period, not that there was no impact, and so their relative prices to eachother were not impacted and as a result returns on the exchange rate didnt change. The time periods of the dummies being selected also may not truely capture when the start of the chaos around covid was. Although they represent the time periods around the official declaration, people knew and were concerned about it before then and that information could have been captured prior to the dummy, potentially explainig why the pre-pandemic dummy is insignificant at a 5% level.

Another thing to notice about this study as a whole is that there is likely to be a lot of ommited variable bias that could inflate the coefficients around the announcement. The factors that came with covid like uncertainty, impacted supply chains and tanked interest rates all had very strong impacts on currency values and as such returns, and its likely some of this impact was captured by the event dummies. With that being said, the change from insignificance to significance at the time of the pandemic being officially declared does certainly imply some kind of impact of the announcement. It could be that it made real the severity of COVID, came as a suprise to some people or caused people to panic.

Bibliography

Reference listCannon, M.J., Percival, D.B., Caccia, D.C., Raymond, G.M. and Bassingthwaighte, J.B. (1997). Evaluating scaled windowed variance methods for estimating the Hurst coefficient of time series. 241(3-4), pp.606–626.

Choi, I. and Chul Ahn, B. (1999). Testing the null of stationarity for multiple time series. Journal of Econometrics, 88(1), pp.41–77.

D.A. and Fuller, W.A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. Journal of the American Statistical Association, [online] 74(366), pp.427–431.

Elliott, G., Rothenberg, T.J. and Stock, J.H. (1996). Efficient Tests for an Autoregressive Unit Root. Econometrica, 64(4), p.813.

H.E. (1956). METHODS OF USING LONG-TERM STORAGE IN RESERVOIRS. Proceedings of the Institution of Civil Engineers, 5(5), pp.519–543.

R.J. and Athanasopoulos, G. (2018). Forecasting : Principles and Practice. 2nd ed. Heathmont, Vic.: Otexts.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. Journal of Econometrics, 54(1-3), pp.159–178.

S.J., C. Mills, T. and Newbold, P. (1998). Spurious rejections by Dickey–Fuller tests in the presence of a break under the null. Journal of Econometrics, 87(1), pp.191–203.

R. and Gurgul, H. (2003). ARIMA Modeling of Event Induced Stock Price Reactions in Austria. [online] Available at: https://web.p.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=0&sid=92464e2e-a475-45d0-b2c1-501e6e4c451d%40redis.

Ng, S. and Perron, P. (2001). LAG Length Selection and the Construction of Unit Root Tests with Good Size and Power. Econometrica, [online] 69(6), pp.1519–1554.

Perron, P. (1989). The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis. Econometrica, 57(6), p.1361.

Perron, P. and Qu, Z. (2007). A simple modification to improve the finite sample properties of Ng and Perron's unit root tests. Economics Letters, 94(1), pp.12–19.

Reddy SK, B.A. (2015). Exchange Rate Forecasting using ARIMA, Neural Network and Fuzzy Neuron. Journal of Stock & Forex Trading, 04(03).

Ryan, O., Haslbeck, J.M.B. and Waldorp, L.J. (2023). Non-Stationarity in Time-Series Analysis: Modeling Stochastic and Deterministic Trends. PsyArXiv (OSF Preprints).

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

Sephton, P.S. (2022). Finite Sample Lag Adjusted Critical Values of the ADF-GLS Test. Computational Economics, 59(1), pp.177–183.

Yaya, O.S., Saka, L. and Akanbi, O.B. (2019). Assessing Market Efficiency And Volatility Of Exchange Rates in South Africa and United Kingdom: Analysis Using Hurst Exponent. The Journal of Developing Areas, 53(1), pp.127–145.

Y. and Wu, H. (2006). A comparison of the prediction performances by the linear models and the ARIMA model: Take AUD/JPY as an example. [online] Available at: https://www.diva-portal.org/smash/get/diva2:140033/FULLTEXT01.pdf.