Time Series Replication Assignment

22876634

2023-05-10

Introduction

In 1996, Crowder and Hoffman published a paper in which they found evidence for the existence of a cointegrating relationship between nominal interest rates and inflation. This would be known as the real interest rate, as per the tax-adjusted Fisher relation. The way that they estimated this relationship was via a Johansen method for Vector Error-Correction (VECM). In the paper, they postulate that though inflation and nominal rates are driven by non-stationaries, they can be predictably modeled within this framework.

The results of the paper yield stationary and economically sound impulse response functions and variance decompositions. In this way, the paper succeeds at the goal of estimating the real interest rate as the cointegrating relationship between the two variables discussed. The purpose of this replication is to firstly, prove whether the results are consistent, then whether they are founded in a good statistical framework, and whether they remain robust when used on data that the authors had not seen before. In this case, this replication study used data of the subsequent years to the paper's release, to test whether the real interest rate remained similar to what was postulated by the original paper's authors.

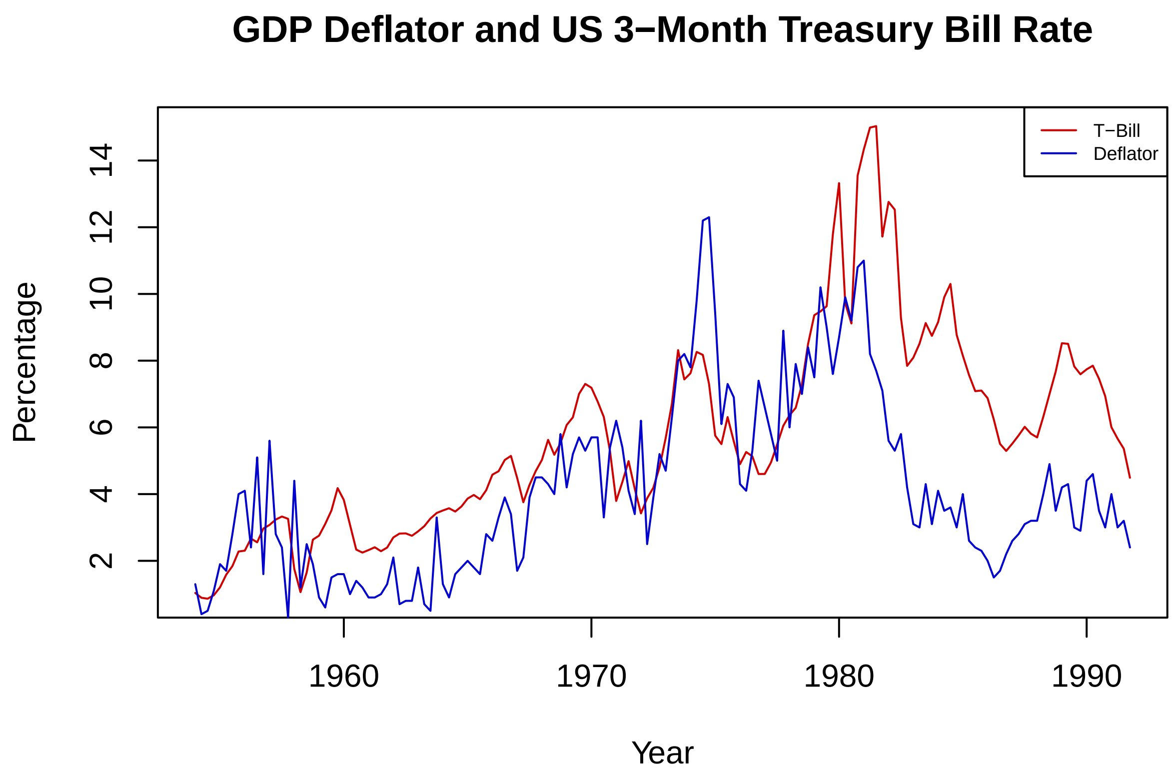
In order to analyse the robustness of this paper's results, this study will also pursue a Johansen method for A VECM. Additional checks to those in the original paper will be pursued, as well as discussions pertaining to why the results may differ despite similar procedures. The format of this study will be to first determine whether there is a unit root in these series, then test whether the residuals of an unrestricted Vector Autoregression (VAR) including these variables are "white noise". The paper then determines the number of cointegrating relationships, estimates a restricted VECM, and pursues forecasting in the form of impulse response functions (IRF) and variance decompositions within the framework of this model.

The paper by Crowder and Hoffman achieves what it aims to do. There cannot necessarily be many criticisms, other than the residuals (which go unmentioned in the original paper) are not strictly "white noise", the model is not incredibly parsimonious, and the application to the modern era's data produces results that economic theory would not predict.

Description of Data Used

In the original paper, the authors make use of the three-month United States (US) Treasury Bill rate as the nominal interest variable throughout the analysis (Crowder & Hoffman, 1996:107). For the expected inflation variable, they use Implicit Price Deflator for Total Consumption Expenditure. What is important to note is that this exact data is no longer available in that original format, if at all. As such, in this study I opt to further test the robustness of their results by using a simple GDP Deflator price series in its place. If their conclusions truly show evidence for the Fisher relation, then a replication utilizing a different widely accepted price series should not significantly impact their results. At the midpoint of the paper, the authors begin to use tax-adjusted data (Crowder & Hoffman, 1996:110). In neither their paper, nor that of the authors they cite as the inspiration for this, could I understand exactly what transformation was applied, even though they explain where they derived their tax estimates from. As such, this replication delves into the general methodology of their VECM. I will mention where our conclusions differ and whether these differences are due to the absence of tax-adjusted data.

It is important to note that the GDP Deflator series I found for this research was seasonally adjusted, and the original authors' probably is not. They do not mention that this is the case. Because of this, I have also seasonally adjusted the Treasury Bill rate. One does not wish to compare data in different formats, having been treated differently. In this case, I do not actually think nominal interest rates would be a seasonal variable, and the difference is negligible, so this adjustment was more a precaution than a necessity. The graphic of the time-series variables used by me can be seen below.



In the case of the above graphic, one can see the gradual climb of both inflation and nominal interest rates before the stagflation episode of the 1970s (Federal Reserve, 2023). There is a large peak in inflation in the middle of that decade, unmatched by nominal interest rates, until eventually Paul Volcker's Federal Reserve moved to increase nominal rates in a larger than one-to-one ratio with inflation. The pattern of nominal interest rates exceeding inflation seems to be present before the oil shock as well. I believe that this is why the authors are investigating whether they can detect a coefficient of larger than one (when inflation is normalised to one) in the cointegrating relationship between these variables.

When viewing Crowder & Hoffman's (1996:108) graphic (seen below), I argue that we observe these same patterns, as would be expected. Similar peaks, troughs, and eventual action by the federal reserve. It is important to note that in the two series used for this replication, interest seems significantly greater than inflation when compared to the graphic below. One still observes the interest hike in 1979, starting in 1976, but the nominal interest rate before and after the episode seem to be gravitating towards the inflation rate.

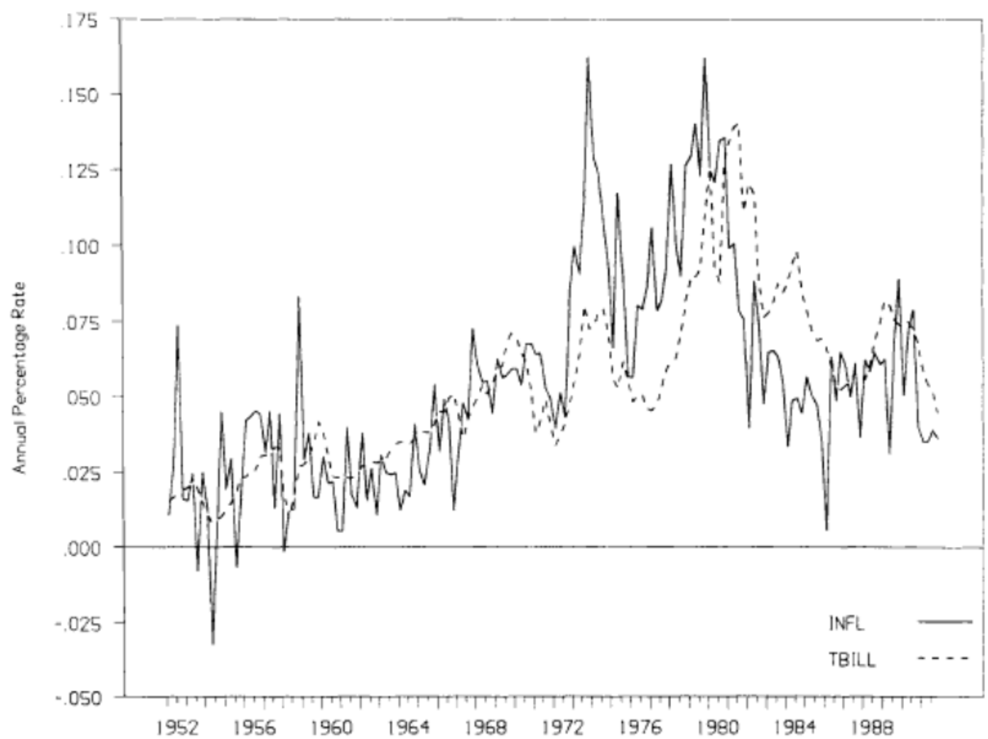


FIG. 1. U.S. Inflation and Three-Month T-Bill Rate, 1952:1 to 1991:4

Formal Dickey-Fuller tests for a Unit Root

In the case of the GDP Deflator, I tested for the presence of a unit root in the price series against the alternative that it is a stationary process about a time trend. This is a one-tailed test. The value of the test statistic is -2.3007, which is closer to zero than even the critical value for this test at the 10% level of significance. In studies desiring stationarity, this result is usually described as meaning one cannot reject the possibility of the presence of a unit root in the series. In the case of VECM analysis, one needs instead to be sure of a unit root. The p-value obtained for the aforementioned test statistic is 0.5384. This is reasonably high, and as such, I conclude that the inclusion of inflation in the cointegration analysis as a non-stationary process is not dubious.

What is interesting is that in my tests, these are not necessarily results which can be deemed on the brink of stationarity. In the original paper, Crowder & Hoffman (1996:107) claim that they could not argue for the presence of a unit root in the consumption expenditure price series without including moving average terms. Later in the paper, they state that it is modeled as an Auto-Regressive Integrated Moving Average (ARIMA) model, in the form ARIMA(0,1,1) (Crowder & Hoffman, 1996:111). This includes no lags for previous levels of inflation, which I find difficult to argue because inflation is a highly persistent economic phenomenon, where momentum and previous levels may play a large role.

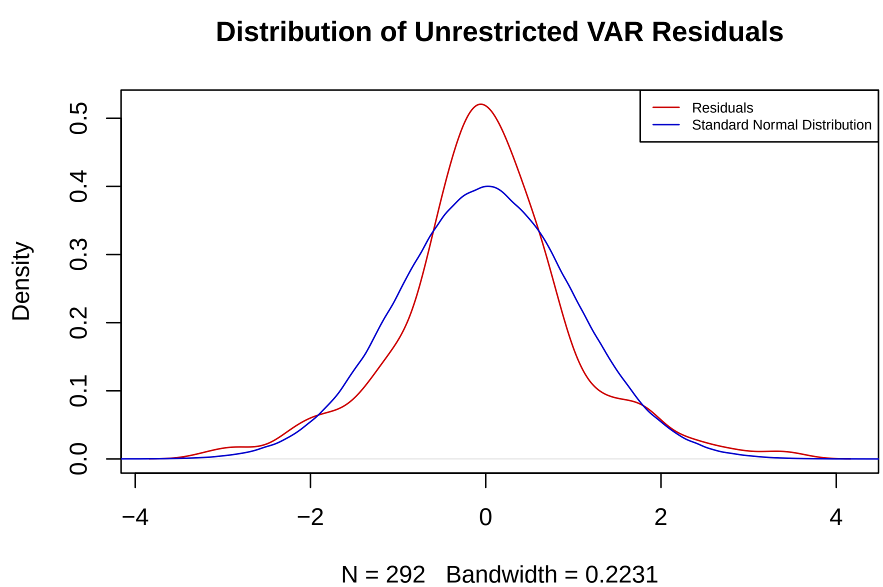
Crowder & Hoffman (1996:111) argue it is much easier to justify that the nominal interest series should be modeled as an ARIMA(1,1,0) process. When performing an augmented Dickey-Fuller test on this series, the test-statistic obtained is -2.6581, which most certainly is strictly larger than the critical value at all levels of significance, but the p-value of this statistic is only 0.1468. It is important to mention that this should be an identical series to the one used in the original paper, whereas the inflation series selected by me differs from theirs. They claim that they cannot reject the presence of a unit root at any level of significance for the nominal interest time series, which is true, but it is interesting that they could not argue similarly for the inflation series with the consumption deflator.

At this point in the paper, the authors begin performing Johansen tests for cointegration, such as the trace and maximum eigenvalue test. Before this venture is pursued in this study, it is important to first estimate an unrestricted Vector Autoregression (VAR) model in the levels of these variables, and to test for “white noise” residuals. More specifically, one would require that the mean of the residuals is insignificantly different from zero, that there is no serial autocorrelation in these residuals, and that there is no substantial clustering of errors which all exceed the chosen level of significance (Enders, 2015:65). The authors do not do this, perhaps it is taken as granted, but it is an important step in obtaining valid results.

Vector Autoregression and Tests for White Noise Errors

The focus of this section is not to hone in on the specifics of the VAR, because these objects are non-stationary and the VAR is simply used to verify whether the residuals of this relationship are white noise. This is important because all time-series properties of the relationship should be contained in the specified model, otherwise it is inadequate. If there are properties left outside the scope of the selected model, then it is not congruent and the results might be being biased by some unobserved variable that should have been included (Enders, 2015:66).

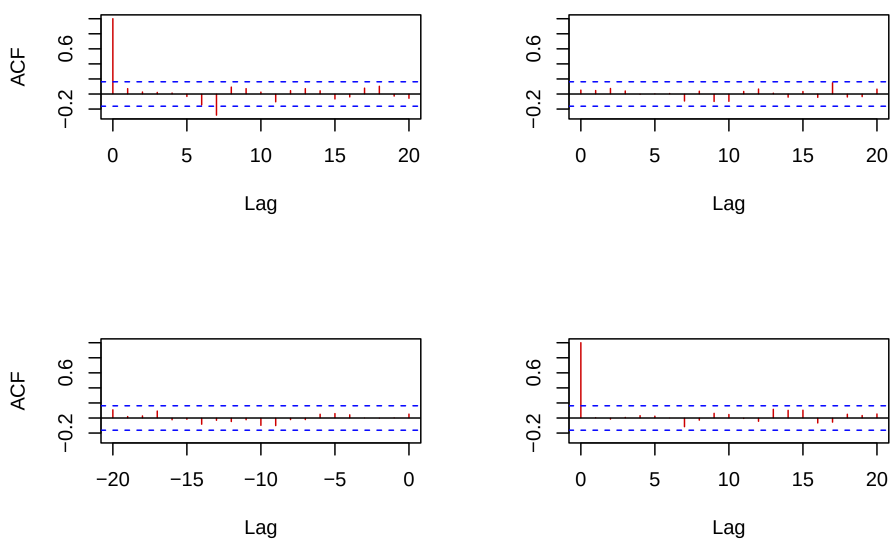
The first of these tests would be to see whether there is significant autocorrelation in the residuals up to some number of lags. For this, I performed a Portmanteau Test on the residuals of the VAR and obtained a p-value of 0.3271. This outcome indicates that one cannot reject the null hypothesis of no autocorrelation in the residuals at any commonly used significance level. This was never necessarily a fear, especially in the original paper, because they specifically accounted for a large number of lags and moving average terms in their final VECM formulation (Crowder & Hoffman, 1996:110). Though it must be made clear that serial autocorrelation is not the only aspect of white noise errors that is vital to valid time series analysis.



Further important considerations would be that the average of the residuals should be zero, and that they should be normally distributed. If the residuals exhibit a higher kurtosis than normal, Enders (2015:158) claims that Maximum Likelihood estimation (such as the Johannsen method for Vector Error Correction) might be inappropriate.This is because maximum likelihood is dependent on the normality of residuals for its limiting distributions. He recommends a GARCH specification. In the plot above, one can observe that the errors are centered around zero, and that the distribution is symmetrical. The mean of these residuals is isignificantly different from zero. The kurtosis of these residuals is not 3. The residuals of the inflation equation exhibit a kurtosis of 3.2, which is relatively normal, but the residuals of the nominal interest equation have a kurtosis of 5.67. This means that they do exhibit notable extreme residual values, which could distort and bias the maximum likelihood estimations. Enders (2015:158) refers to this as a “fat tailed” distribution, which is most often found in the analysis of financial variables.

The autocorrelation functions of these residuals can also be viewed below. There does not seem to be notable clustering where many consecutive residuals exceed the 95% confidence bands. In fact, very few residuals exceed these bands at all. This allows researchers to be more confident that the postulated model need not be one which has time-varying parameters or adapt to structural breaks (Enders, 2015:73). Though the residuals are not perfectly normally distributed, because of their status as mean zero, symmetrical, without any notable autocorrelation or clustering, it should not be considered entirely implausible that there is a valid relationship between the nominal interest rate and inflation, and that this model is congruent.

Autocorrelation Functions for Residuals of the Unrestricted VAR



Vector Error Correction Mechanism and Eigenvalue Tests

Trace and Maximum Eigenvalue Tests

The authors specify three separate formulations of the Johannsen tests for cointegration (Crowder & Hoffman, 1996:109). They mainly use the trace test for the diagonal matrix, but in this replication study both this and the maximum eigenvalue test are used. The first formulation they test is with the inclusion of a deterministic trend in the data. In their paper, they find that the computed trace statistic of 15.88 can be compared to the 5% critical value of 12.53, meaning that they can easily reject the null at the chosen level of significance (Crowder & Hoffman, 1996:109). This is also true in this replication, but the test statistic of 26.52 instead exceeds the critical value of 25.32 at the 5% level. In both cases, the computed test statistic for $H_1 : r < 1$ returns a very low level of evidence. This makes sense, because in a bivariate VECM, there cannot be more than one cointegrating relationship. The maximum eigenvalue test confirms these findings.

They do not explicitly outline the results of the other specifications; first being that there is no constant or trend, with the final postulation being that there is a constant in the relationship. All tests return the result that there is at least one cointegrating relationship. In the other formulations, the test statistics do not exceed the critical values as greatly at the 5% level, but using economic theory, and not just statistical results, it seems more difficult to argue that there is some time trend in the nominal interest rate or inflation. Especially considering that at any given time, more explicitly since the 1980s and 1990s, inflation is targeted at some (constant, 2% in the case of the US) value (Federal Reserve, 2023).

Specifying that there is instead a constant in the data yields a test statistic for $H_0 : r = 0$ of 15.99, which marginally exceeds the 5% critical value of 15.67. I argue that it is a better economic interpretation and as such I model it in this way instead. As will be observed later, it actually does not alter the results too greatly when analysing impulse response functions and variance decompositions. With this specification, I also find that four is the minimum number of lags which removes serial correlation concerns from the residuals. As such, it is the model specification to proceed with. The authors of the paper being replicated also found that four lags was the optimal decision, though they further found that increasing lags to five or decreasing them to three did not make a notable impact for the cointegration tests.

Imposing Restrictions and Estimating the Model

When estimating a Vector Error Correction Mechanism, accounting for a constant, with four lags, this study finds a β vector containing a normalised cointegrating relationship of [1, -0.404]. This is when the nominal interest rate has been normalised to one. The original paper finds that if inflation is normalised to one, the β vector would be [1, -1.35] (Crowder & Hoffman, 1996:110). When inflation is normalised to one in this model, the β vector becomes [1, -2.47], which shows the same relationship in terms of direction, but not in size. The error adjustment coefficients in α are [-0.0655**, -0.0634], showing that only the short run error correction of nominal interest is statistically significant, now proceeding in the case that nominal interest is normalised to one.

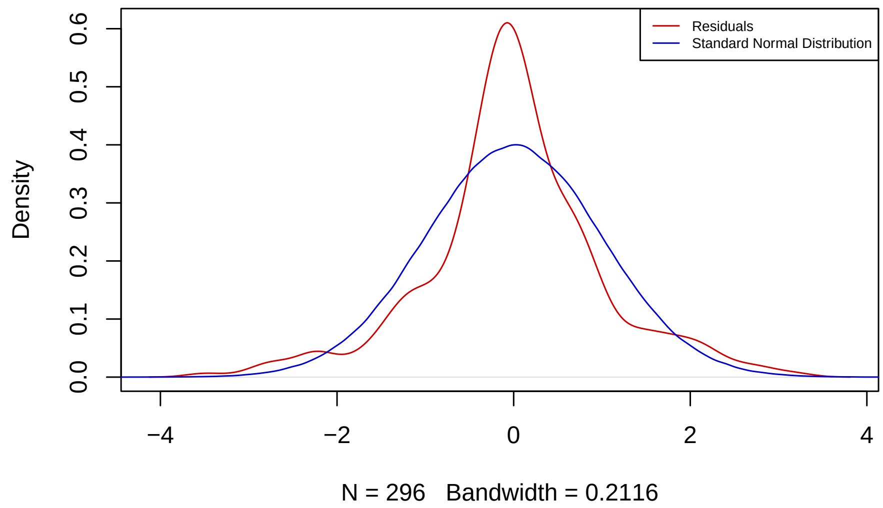
Though these numbers differ slightly from the original paper (likely due to different data), the general findings are similar, especially in terms of direction of change, and statistical significance. The model, as has been estimated in this replication, has insignificant estimators that could perhaps be tested for exclusion. The original papers' authors also find the error correction term for interest to be the only one of significance, leading the researchers to testing whether inflation is weakly exogenous to the system (Crowder & Hoffman, 1996:109). Because this study has found similar results, I have performed the same tests for weak exogeneity.

I performed two tests. The first test was derived from an Enders (2015:306) example, a Granger-Cauality test (similar to the original paper), where the null hypothesis is that nominal interest does not "granger-cause" inflation. The test returned a p-value of 0.013, indicating that at the 5% level, there is evidence for one to reject the null hypothesis and interest likely does "granger-cause" inflation over the long run. However, a separate test for instantaneous causality, resulted in a p-value of 0.5365. This means that there is not enough evidence to reject the null hypothesis that there is no contemporaneous relationship between the two variables. Therefore, one cannot conclude that there is a direct causal relationship between nominal interest and inflation in the short run, and inflation could be weakly exogenous. In the original paper they argue that this means that the α matrix can then be written as [0,1] and all long run error adjustments occur through the nominal interest rate (Crowder & Hoffman, 1996:113). At this point, they go on to display impulse response functions of the model under these restrictions, but, once again, this study will first assess model adequacy before viewing forecasting potential.

Model Adequacy

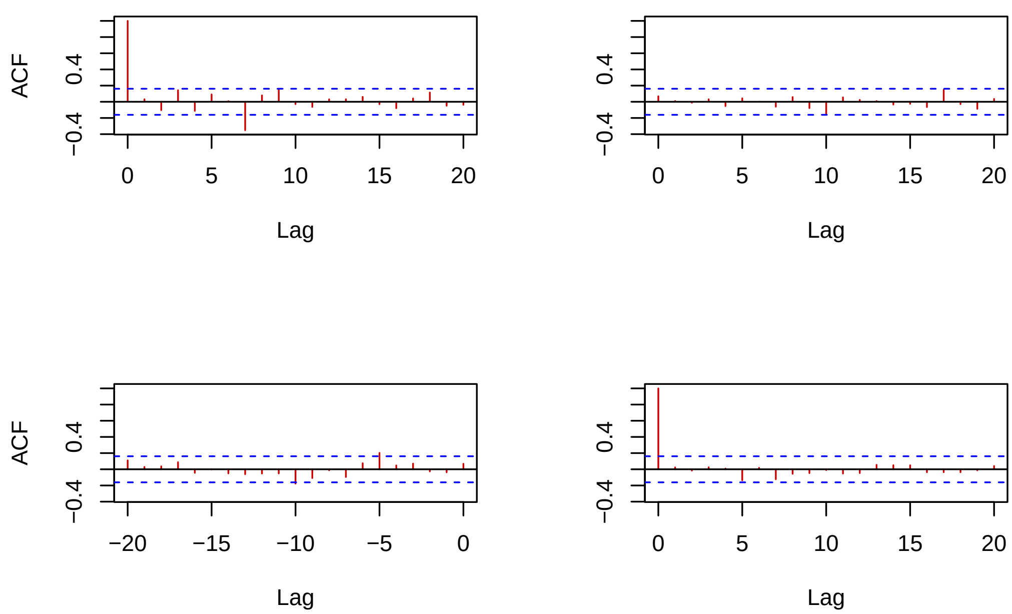
As mentioned before, the model specification includes the number of lags that removes serial autocorrelation based on a Ljung-Box test of the residuals. So to assess serial autocorrelation at this stage would be circular, as this was explicitly done in order to specify the VECM already. Interestingly, though the distribution of these residuals is still strictly mean zero and symmetrical, the kurtosis has become larger, as can be seen below. These residuals are not normally distributed, even less than before. At this point one might argue that there are notable, extreme values in the residuals and biased estimates should be a real concern.

Distribution of Restriced VECM Residuals

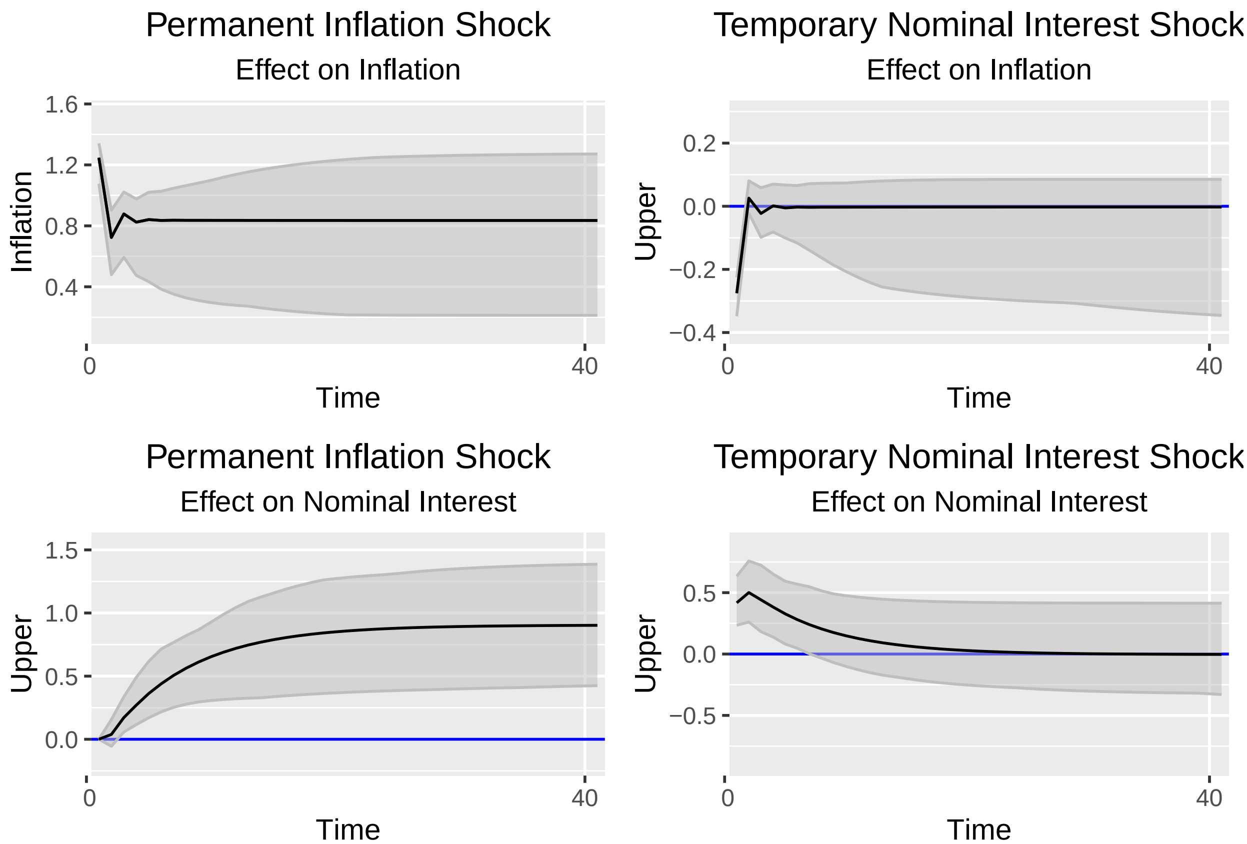


It must be said that a higher than normal kurtosis remains the only real potential obstacle to this model obtaining “white noise” residual status. There is no serial autocorrelation in the residuals, which are also mean zero. The ACF of the residuals displayed below further demonstrates that there is no clustering of residuals, nor is there a concerning number of lags exceeding the 95% confidence bands. What remains now is simply to shock the model and see whether the cointegrating relationship exhibits stationarity, which is the goal of cointegration analysis (Enders, 2015:343). If the variables behave in an explosive way, or do not return to steady state, then this entire process may have not yielded a stationary real interest rate.

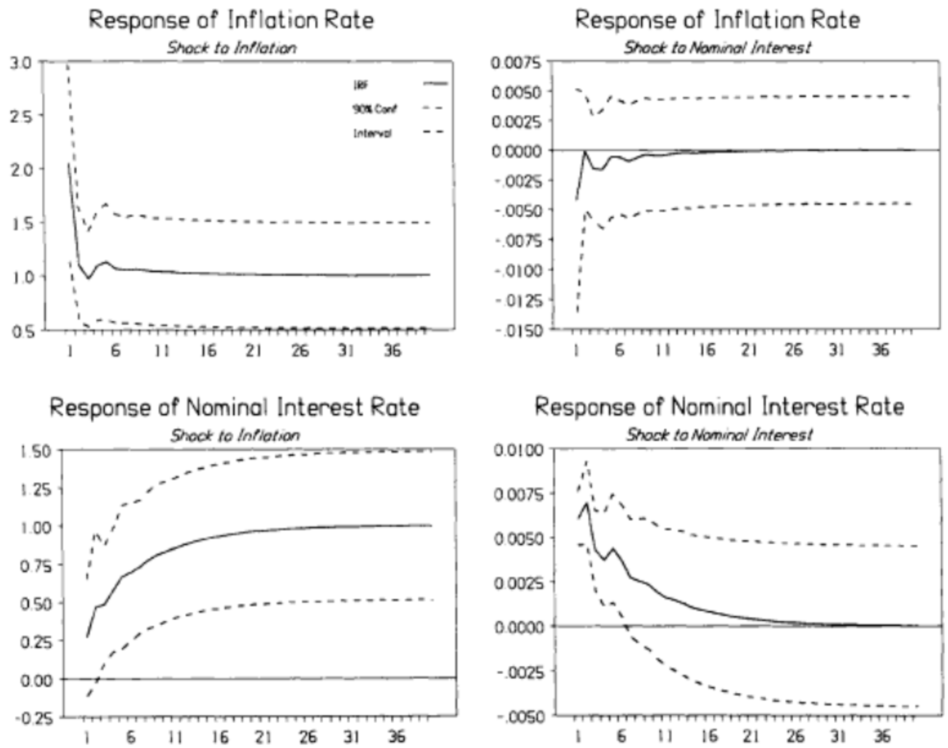
Autocorrelation Functions for the Restricted VECM



Impulse Response Functions and Variance Decomposition



These IRFs are estimated alongside 95% confidence bands, which are the result of 100 Monte Carlo estimations. In the paper, they use 1000 but I did not see an observable difference when increasing the number of runs by an order of magnitude. Similar to the results in the original paper, inflation exhibits a rapid decline to a new steady state value in response to a permanent shock, whereas nominal interest takes much longer, perhaps close to 20 quarters to reach a new steady state (Crowder & Hoffman, 1996:113). The impulse response functions for temporary shocks to the system continue to be similar to the original paper, where the authors also found a minor response from inflation in response to a one unit shock to interest. The interest rate also returns to its original steady state in response to a temporary shock. These results indicate that the cointegrating relationship, the real interest rate in this case, is stationary. This is because Enders (2015:295) explains that temporary shocks will not have permanent effects on stationary processes. The IRFs do not exhibit explosive behaviour, and all eventually stabilise. This adds further validity to their results, and to the results of this replication study. The original paper's IRFs are shown below.



What is notable is that the original authors used tax-adjusted data, which one would think is a significant change, but the general patterns are similar. Nominal interest stabilises close to one in both models, in response to a permanent shock. Inflation rapidly declines in both models, and all variables remain non-explosive. I do not know whether this adds robustness to the original findings as much as it questions the inclusion of tax-adjusted data, without showing the IRFs juxtaposed with one another. This may have been an interesting discussion, as then one could truly analyse which differences can be attributed to tax-adjustment, and which might have been found regardless. I do not dispute their finding that when using tax-adjusted data, the cointegrating coefficients become one-to-one, because I did not test this. I simply argue that in terms of responses to shocks, it cannot be argued that too much has changed.

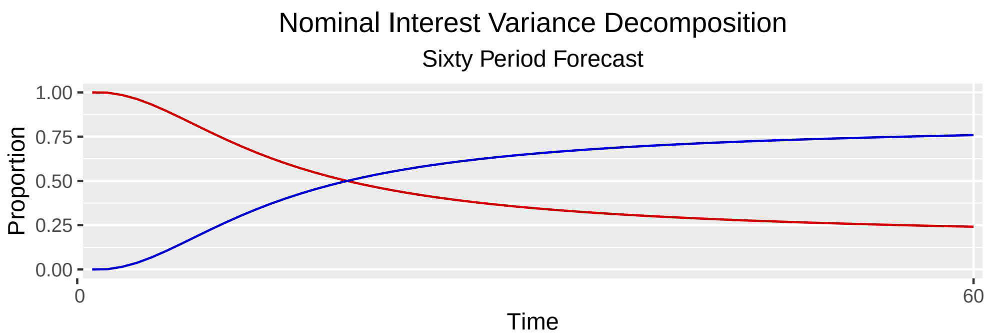
TABLE 3

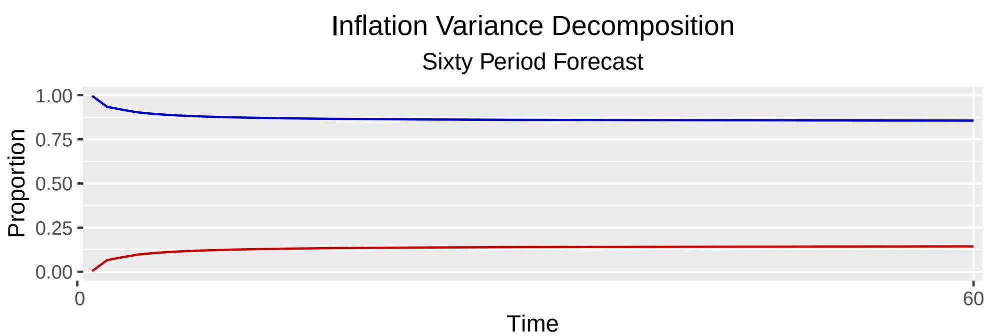
Variance Decompositions: Innovation to Common Trend

Forecast Horizon	Nominal Interest Rate	Inflation Rate
1	16.05 [58.20, 0.19]	93.85 [99.98, 59.21]
2	23.50 [66.54, 1.87]	94.89 [99.33, 63.12]
3	30.67 [72.16, 4.98]	94.91 [99.43, 62.94]
4	38.15 [76.67, 10.33]	94.85 [99.51, 62.59]
8	62.30 [87.13, 31.14]	95.89 [99.48, 66.05]
12	75.87 [89.93, 46.70]	96.24 [99.42, 68.50]
16	82.80 [91.70, 56.70]	96.53 [99.36, 70.41]
20	86.71 [92.85, 63.32]	96.65 [99.33, 71.89]
40	93.02 [96.05, 73.52]	97.14 [99.60, 74.03]
60	94.74 [97.29, 76.39]	97.27 [99.47, 75.41]

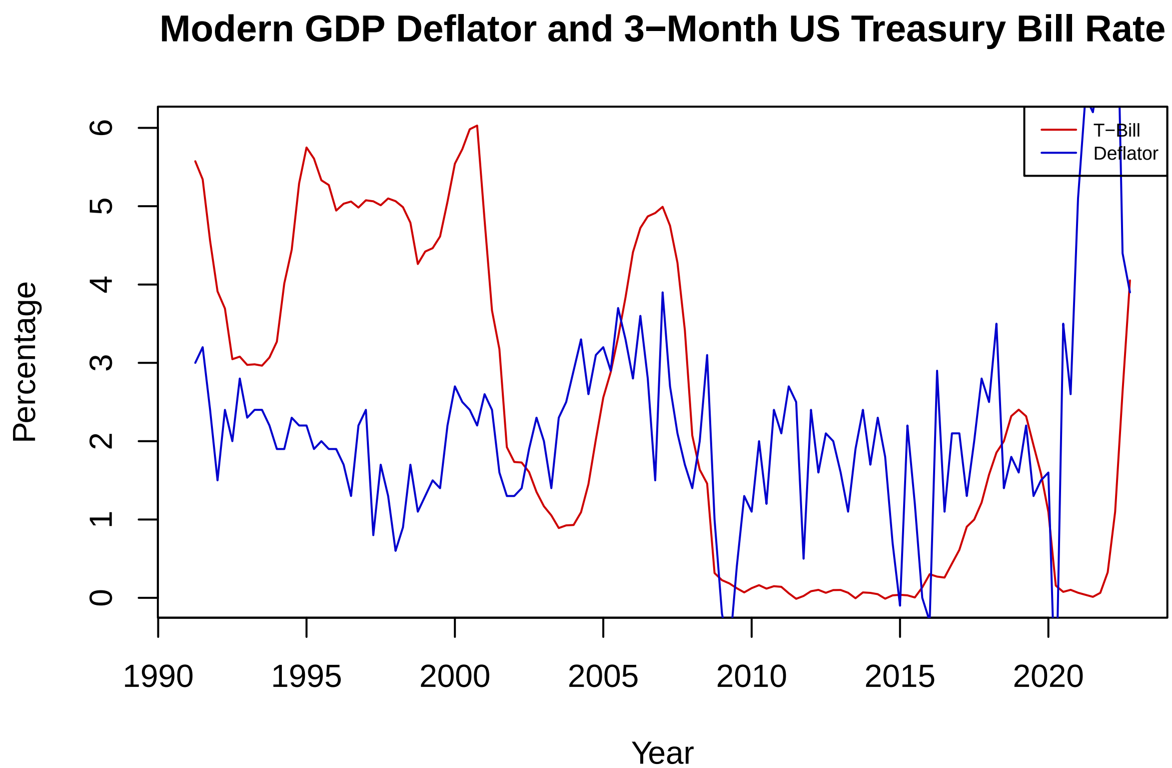
Notes: Numbers in column one represent the forecast horizon. Numbers in columns two and three represent percentages of forecast error variance explained by the common trend. Numbers in brackets represent the 95 percent confidence interval around the forecast error variance.

The paper also includes a variance composition (Crowder & Hoffman, 1996:114). From what can be seen above, the authors indicate that the permanent innovations in the system should be associated with inflation, due to its strong contribution to the forecast error at all time horizons. Below are the variance decompositions for the individual variables as estimated by the model in this replication study. This allows us to view the proportion of forecast error in one variable that can be attributed to another, and vice-versa (Enders, 2015:302). Clearly, inflation (blue) is responsible for a large proportion of variation in the system as a whole. In terms of its own variation, it very quickly stabilises to over 80%, and remains there for the entire forecast horizon. Forecast error in nominal interest (red) is also eventually dominated by inflation, with the cross-over occurring at about 20 time periods into the future. Clearly, from these results, inflation is largely the driving force behind variation in the system. This allows me to conclude that in general, the model the original papers' authors have described is one that is generally able to be replicated, and a valid representation of the real interest rate postualted as the cointegrating relationship between nominal interest rates and inflation.





Modernisation of the Model



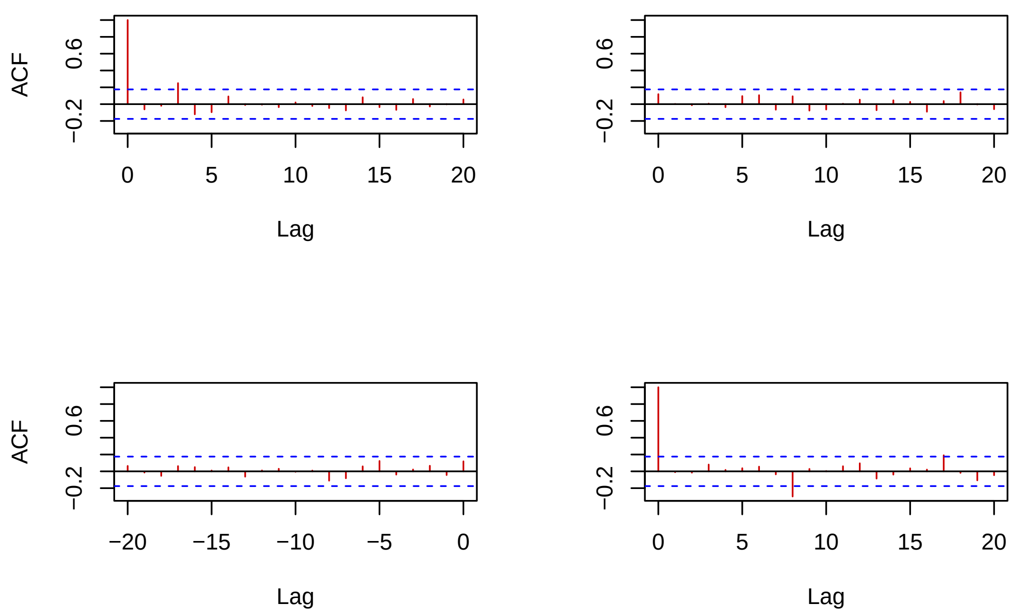
At first glance, the relationship between the nominal interest rate and inflation might seem very different in the subsequent years to the release of the original paper. This made me interested in seeing whether the Fisher relation the authors previously found evidence for would still hold. In the 1990s we began to see many nations formally begin inflation targeting (Federal Reserve, 2023). In the 2010s the world witnessed record low nominal interest rates, until the delayed response to the COVID-19 stimulus measures in 2021. In the 1990s we still observe the same pattern as before; nominal interest rates substantially larger than inflation (when the GDP Deflator is used). In the 2010s this no longer seems to be the aim. One could argue that the desire to keep nominal rates low could explain the hesitancy to raise rates in accordance with inflation following the pandemic, but we actually see that there was a similar failure to promptly raise rates in 2004 already, though both inflation and the required raise in rates were less substantial.

Preliminaries

Obviously any application of the former model in the subsequent time period would still require the same conditions to hold. It is very interesting that one can observe an almost reversal of the p-values obtained from the augmented Dickey-Fuller tests for these variables. The test for a unit root inflation now only returns a p-value of 0.09773, which would actually be enough to reject the presence of a unit root at the 10% level. In economics, we often strive for the 5% confidence level, so one may be tempted to simply claim that there is not enough evidence to reject the presence of a unit root in the price series. For the purposes of this replication, it will be argued that there could be a unit root in the inflation series, but one may have rejected this outcome if the 10% level is adequate for a given model. In terms of nominal interest, the test statistic of -2.6118 is similar to the previous analysis, but the p-value for the test is 0.6045. This means that one can argue for a unit root much more easily than in the analysis of the original paper.

For the unrestricted VAR, much has stayed the same. What is promising is that the VAR could be estimated with close to half the number of lags before autocorrelation was removed from the residuals. This bodes well for the hopes of a parsimonious VECM estimation later on in this modernisation. The mean of the residuals is still insignificantly different from zero, with the autocorrelation exhibited below still remaining within the bands without any concerning clustering. Previously, this was only achieved via four lags in the unrestricted VAR. Now, it can theoretically be done with only two - though this leads to less normally distributed errors, so there is a compromise at this marginal level if statistically significant estimators are desired.

Autocorrelation Functions of the Residuals for the Unrestricted VAR

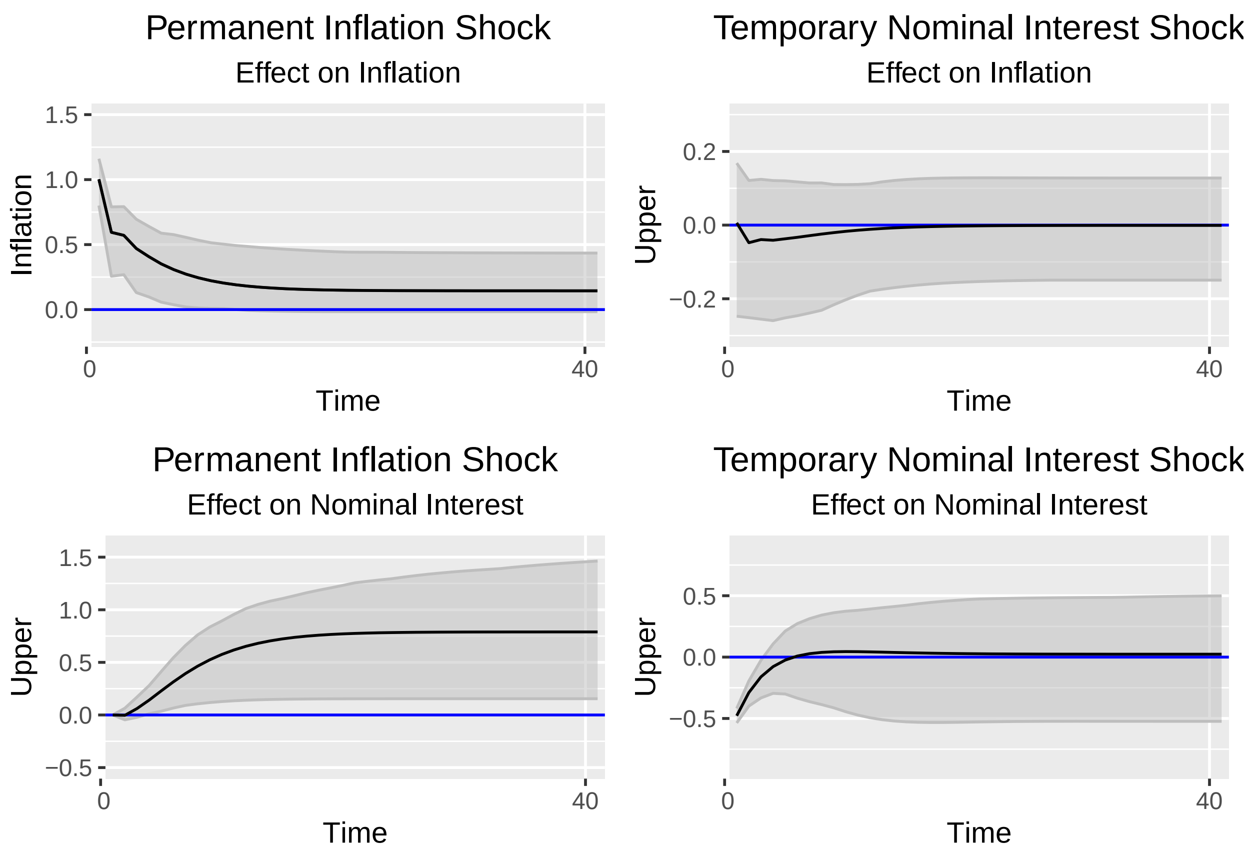


The only concern remains that the kurtosis of the residuals of the unrestricted VAR is still substantially higher than what a normal distribution would have. For nominal interest, the kurtosis is even higher at 6.06, with inflation also now deviating from normality at 5.2. This indicates that there have been extreme deviations from the estimated inflation and nominal interest rates more often in the modern era. This is not only empirically true, but one should also think so considering that both the Global Financial Crisis and the COVID-19 pandemic occurred within 15 years of one other. Regardless, the purpose of this modernisation of the previous model is to test whether the hypothesis of the authors has survived the test of time.

Modern Vector Error Correction Mechanism

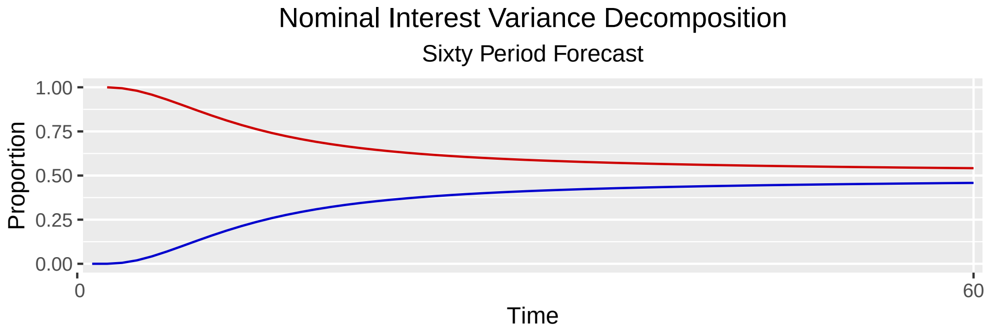
The modernisation of the model continues in a marginally different formulation from the original paper. In terms of the trace and maximum eigenvalue test, one can still find at least one cointegrating relationship when specifying that there is a constant in the data generating process. The maximum eigenvalue test statistic returns a value of 19.22, which exceeds the 5% critical value, 15.67. The trace statistic also exceeds the 5% critical value. If one estimates a restricted VECM with one cointegrating relationship, it can now be successfully achieved with only two lags while removing serial autocorrelation in the residuals. This results in more significant estimators, and a more parsimonious model than previously. The second lag for nominal interest should still be tested for exclusion, but overall, the following forecasts might show more accuracy due to the increased parsimony of the model. The error correction term of both variables is now statistically significant at the 90% level, and as such there is not as much evidence to estimate the model within the framework that inflation is weakly exogenous. Finally, the residuals of this model are mean zero, exhibit no autocorrelation, and no concerining clustering.

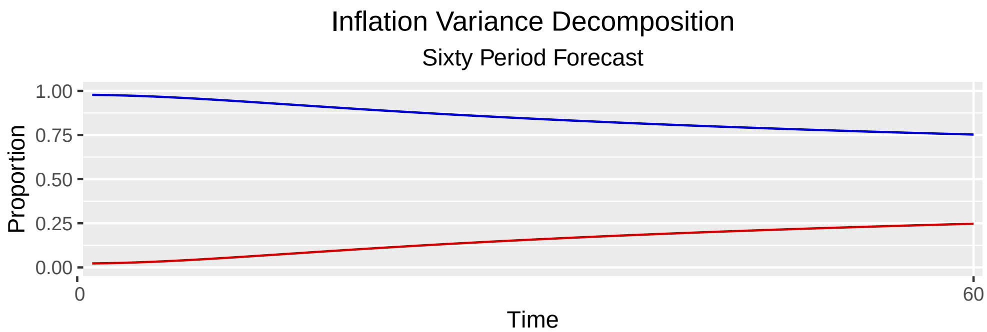
Modern Impulse Response Functions and Variance Decompositions



Most IRFs have actually remained consistent to their previous counterparts. One can still observe a slow move to a new steady state in response to a permanent inflation shock for nominal interest. It can also be said that inflation still resets to its mean rather quickly, both in terms of permanent and temporary shocks. There is a problem in that nominal interest seems to respond negatively to a one unit (positive) shock to itself, while still returning to its initial steady state value. This may indicate that the model is inadequately specified and might require further analysis, should one wish to formally give evidence for the Fisher relation in the modern day. The reason this is concerning is because a variable should increase in response to a positive shock to itself (Enders, 2015:297). Thus, this is an adequate outcome. Perhaps with different specifications this outcome would be more accurate to real world expectations.

There are minor differences in the variance decompositions of these variables, which can be seen below. Inflation continues to be the driving force in this cointegrating relationship, with the error correction term being statistically significant at the 10% level for both variables. It can be said that in the modern era, inflation never completely dominates the proportion of forecast error in the nominal interest rate, which is different to previously, where it would eventually explain most of the forecast error in interest.





Conclusion

The goal of the original paper by Crowder and Hoffman was to find evidence for the tax-adjusted Fisher relation, which the authors proposed to be the cointegrating relationship between nominal interest rates and inflation. The authors find economically sound results, and stationary forecasts. The authors do not convince the reader of “white noise” residuals, but upon testing the adequacy of the model in this study, it is reasonably congruent and the results should be considered valid, with the knowledge that they might include bias. It must be mentioned that the omission of certain details, such as white noise residuals, need not warrant suspicion. The specific criticism in this replication study is that the residuals cannot be considered strictly normally distributed, and the absence of this fact in the original paper is what is criticized. Despite this, the model adapts to the modern era with new data, which bodes well for supporters of Fisher relation as an economic truth. The most important aspect is that the model can be replicated when using a similar approach, and even a different price series. This was the main pursuit of this replication study, and the 1996 paper performed adequately.

Bibliography

Crowder, W.J & Hoffman, D.L. 1996. The Long-Run Relationship between Nominal Interest Rates and Inflation: The Fisher Equation Revisited. *Journal of Money, Credit and Banking*, 28(1):102-118.

Enders, W. 2015. Applied Econometric Time Series. 4th edition. New Jersey: John Wiley & Sons, Inc.

Gross Domestic Product: Implicit Price Deflator [Online]. [n.d.]. Available: https://fred.stlouisfed.org/series/A191RI1Q225SBEA [2023, May 10].

Monetary Policy Principles and Practice [Online]. [n.d.]. Available: https://www.federalreserve.gov/monetarypolicy/historical-approaches-to-monetary-policy.htm [2023, May 10].

3-Month Treasury Bill Secondary Market Rate, Discount Basis [Online]. [n.d.] Available: https://fred.stlouisfed.org/series/DTB3 [2023, May 10].