Student Economic Review, Vol. 18, 2004

ANALYSIS OF UK INFLATION DYNAMICS USING ARCH AND ALLOWING FOR SEASONAL EFFECTS

BY DAVID MORRISSEY

Senior Sophister

Inflation is the one of the economic phenomena that is in the spotlight of forecasting. In this paper David Morrissey evaluates the dependency of UK inflation on its previous values, extending the analysis to include seasonality later in the text. The econometric analysis is conducted using an ARCH model with various extensions that appears to be superior to its OLS counterpart. Finally, the author concludes that his model has sound explanatory and forecasting powers and suggests further extensions for analysis.

Introduction

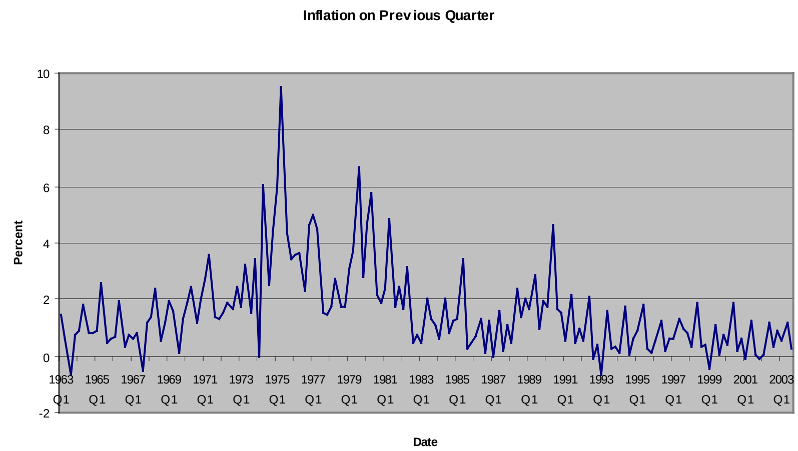
This paper intends to explain how the evolution of inflation over time depends on its previous values. Inflation is known to exhibit time varying volatility, hence traditional regression models, which assume homoscedastistic variance are inadequate. I will attempt to overcome this problem by using an ARCH model. It is instructive to begin by looking at a plot of inflation against time, Figure 1. I used UK data, a series of 168 quarterly CPI observations from 1962 to 2003, found on the UK's National Statistics website.

It is quite apparent from inspection that there are periods in which the variance of inflation is high (the 1970s) and others when it is comparatively low (the late 1990s). Hopefully an ARCH model will help to explain these dynamics.

The Model

I am going to run an autoregression; regressing inflation on its lagged values. Seasonal effects are often encountered in quarterly data, to allow for this I will also include seasonal dummy variables. In this project, I am using dummies to allow only for differential intercepts in each quarter, not for differential slope coefficients

Figure 1: UK Inflation 1963 - 2003



I will commence my analysis by regressing Inflation (Inf) on 5 lagged values of itself and 4 quarterly dummies. Note I do not include an intercept, so as to avoid the dummy variable trap.[1]

$$\pi_t = \alpha_1\pi_{t-1} + \alpha_2\pi_{t-2} + \alpha_3\pi_{t-3} + \alpha_4\pi_{t-4} + \alpha_5\pi_{t-5} + \beta_1 Q_1 + \beta_2 Q_2 + \beta_3 Q_3 + \beta_4 Q_4 + \mu_t$$

I will first use the traditional Ordinary Least Squares (OLS) method of estimation. The results are given in the table below.

Table 1: UK inflation regression

Regressor	Coefficient	Standard Error	T-Ratio [Prob]
INF(-1)	.46734	.079784	5.8576[.000]
INF(-2)	.28542	.087220	3.2725[.001]
INF(3)	.019862	.090287	.21998[.826]
INF(-4)	.13337	.087221	1.5291[.128]
INF(-5)	-.079349	.079622	-.99657[.321]
Q1	.25114	.21435	1.1716[.243]
Q2	1.3392	.21438	6.2468[.000]
Q3	-.56275	.23954	-2.3493[.020]
Q4	.054633	.21159	.25820[.797]

[1] If I used 4 dummies and an intercept there would be perfect multicollinearity, the data matrix would be non-singular and estimation impossible.

Table 2: Diagnostic Tests

Test Statistics	LM Version	F Version
A:Serial Correlation	CHSQ(4)= 17.2499[.002]	F(4, 150)= 4.4382[.002]
B:Functional Form	CHSQ(1)= .40350[.525]	F(1, 153)= .37969[.539]
C:Normality	CHSQ(2)= 281.6049[.000]	Not applicable
D:Heteroscedasticity	CHSQ(1)= 19.1078[.000]	F(1, 161)= 21.3796[.000]

A:Lagrange multiplier test of residual serial correlation
B:Ramsey's RESET test using the square of the fitted values
C:Based on a test of skewness and kurtosis of residuals
D:Based on the regression of squared residuals on squared fitted values

The model explains the data quite well with an $R^2 = 0.64813$. The F-Test for the joint significance of all the explanatory variables is also quite encouraging an F value of 35.4582 clearly rejecting the null that they have no explanatory power.

Non White Noise Errors

There are clearly quite severe problems with the residuals in this model with the diagnostics pointing towards serial correlation, non-normality and heteroscedasticity. Heteroscedasticity is however an encouraging sign for the presence of ARCH effects as we obviously expect non-constant variance.

With reference to the problems discussed above we must treat the t-statistics quoted with caution. However it is interesting to note that only the first 2 lags of inflation appear to have explanatory power. For the next 3, we fail to reject the null that their coefficients are in fact 0.

Testing Linear Restrictions using Wald Tests

To investigate further the lack of explanatory power of the 3rd, 4th and 5th lags of inflation, I conducted a Wald test with the null hypothesis that, $\alpha_3 = \alpha_4 = \alpha_5 = 0$. The Wald test is especially useful in this case as even in the absence of normally distributed residuals, it is asymptotically Chi Squared distributed. The Wald test returned a value W = 2.7806 with an associated p value of 0.427. Hence, we fail to reject H_0, therefore in subsequent versions of the model I only consider the 1st and 2nd lags.

I conducted another Wald test to check the validity of including the seasonal dummies. In this the null hypothesis was that $\beta_1 = \beta_2 = \beta_3 = \beta_4$. This would indicate the absence of an intercept differential between any of the quarters. The statistic returned was W = 43.1079 with an associated p value of 0.000. Hence we reject H_0 and conclude that we were right to include the seasonal dummies.

Testing for ARCH

The arch model was first proposed by Engle in 1982. He suggested modelling conditional variance as a function of past squared residuals.

$$\text{ARCH(1) model: } h_t^2 = E(\sigma_t^2 | \Omega_{t-1}) = \gamma_0 + \gamma_1 u_{t-1}^2$$

in general an ARCH(p) model includes p lags of squared residuals.

The testing procedure for ARCH is hence to run OLS, as we have above then save the residuals and run an AR(p) regression on their squares. A Lagrange Multiplier (LM) test is then conducted with null hypothesis that the coefficients on each of the p lagged squared residuals is 0. It is shown by Engel 1982 that this test boils down to obtaining R^2 from the AR(p) regression and then testing TR^2 as χ_p^2.

Microfit runs this test automatically see the output below,

Table 3: Autoregressive Conditional Heteroscedasticity Test of Residuals (OLS Case)

```
*******************************************************************************
 Dependent variable is INF
 List of the variables in the regression:
 INF(-1)         INF(-2)         INF(-3)         INF(-4)         INF(-5)
 Q1              Q2              Q3              Q4
 163 observations used for estimation from 1963Q2 to 2003Q4
```

P = 1

```
*******************************************************************************
 Lagrange Multiplier Statistic    CHSQ( 1)=   9.4400[.002]
 F Statistic                   F(  1, 153)=   9.4055[.003]
```

P = 12

```
*******************************************************************************
Lagrange Multiplier Statistic    CHSQ(12)=  22.0917[.037]
 F Statistic                   F( 12, 142)=   1.8552[.045]
*******************************************************************************
```

As evident above we reject the null hypothesis (at the 5% level of significance) of no arch effects for lag lengths of u_t^2, 1 through 12.

However, I was still unsure what lag length to use in my conditional variance function. So, I carried out a regression of u_t^2 on 8 lags of itself to investigate the t-statistics.

Table 4: Ordinary Least Squares Estimation

Dependent variable is USQ
155 observations used for estimation from 1965Q2 to 2003Q4

Regressor	Coefficient	Standard Error	T-Ratio[Prob]
C	.40073	.21735	1.8437[.067]
USQ(-1)	.20324	.082675	2.4583[.015]
USQ(-2)	.028079	.084211	.33343[.739]
USQ(-3)	.052507	.084203	.62358[.534]
USQ(-4)	.23428	.084148	2.7841[.006]
USQ(-5)	-.054013	.084181	-.64162[.522]
USQ(-6)	.036358	.084196	.43183[.666]
USQ(-7)	.063798	.084197	.75771[.450]
USQ(-8)	-.043619	.082383	-.52946[.597]

With an $R^2 = 0.12852$

The only significant t-statistics are the 1st and 4th lags of u_t^2 hence it appears that only they should be included in the conditional variance function.

ARCH Estimation

We are now trying to estimate the following model,

$$\pi_t = \alpha_1\pi_{t-1} + \alpha_2\pi_{t-2} + \beta_1 Q_1 + \beta_2 Q_2 + \beta_3 Q_3 + \beta_4 Q_4 + \mu_t$$

and we are going to assume the variance of μ_t follows the ARCH process

$$\text{var}(\mu_t) = h_t^2 = \gamma_0 + \gamma_1 u_{t-1}^2 + \gamma_2 u_{t-4}^2$$

When variances are estimated using an arch process, they will not be independent. This implies that Maximum Likelihood is a more efficient estimation procedure than Least Squares. For a proof of this see Engel 1982.

For MLE it is necessary to make assumptions about the distributions of the stochastic disturbances μ_t. In this case, it is assumed they are normally distributed with mean 0 and variance h_t^2.

Results from Arch estimation

Table 5: GARCH(0,4) assuming a normal distribution converged after 91 iterations

Dependent variable is INF
166 observations used for estimation from 1962Q3 to 2003Q4

Regressor	Coefficient	Standard Error	T-Ratio[Prob]
INF(-1)	.60157	.16043	3.7498[.000]
INF(-2)	.27662	.10116	2.7346[.007]
Q1	.087743	.19189	.45725[.648]
Q2	1.3519	.15607	8.6622[.000]
Q3	-.85778	.26458	-3.2420[.001]
Q4	.13304	.21543	.61756[.538]

The estimated model is hence:

$$\pi_t = (0.602)\pi_{t-1} + (0.277)\pi_{t-2} + (0.088)Q_1 + (1.352)Q_2 - (0.858)Q_3 + (0.133)Q_4$$

with conditional variance function:

$$h_t^2 = 0.45 + (0.283)u_{t-1}^2 + (0.110)u_{t-4}^2$$

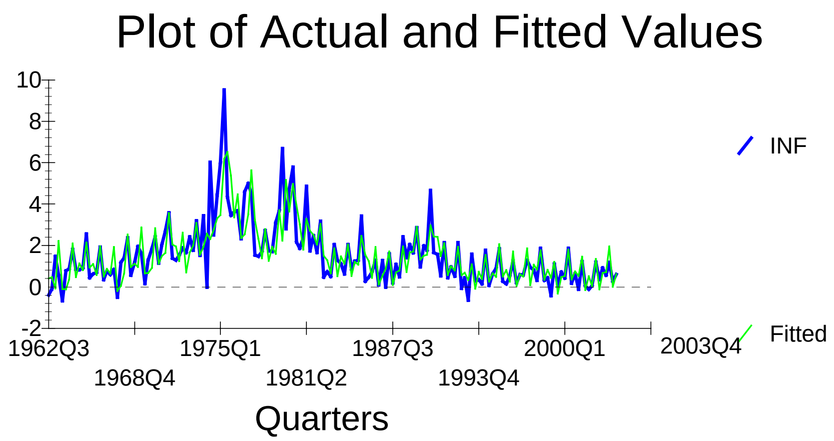
All coefficients are significant at the 1% level, except for those on Q1 and Q4. The 1st and 2nd lags of inflation have a positive impact on a given period's inflation, the first lag having a larger influence. It is also possible to isolate the equation for a given quarter with the dummy variables. For example the estimated equation for quarter 2 is:

$$\pi_t = 1.352 + (0.602)\pi_{t-1} + (0.277)\pi_{t-2}$$

The model has similar explanatory power to the one we estimated under OLS, the value of $R^2 = 0.62917$ see plot below. However, we can have more faith in the estimates as we have addressed the severe heteroscedasticity problem evident in the earlier model.

I now want to find some way to compare the OLS and the MLE ARCH models. I was thinking of comparing the dispersion of residuals, perhaps looking at outliers as in the Engel paper. However, I do not think this would be very instructive as we have a pre-defined the functional form for the variance of the ARCH residuals. Hence their distribution is endogenous to the model and does not give insight to the model's adequacy. So I am going to compare the models in terms of their forecasting ability.

Figure 2: Plot of actual and fitted values



Forecast Comparison

In comparing the models, I am using the version with 2 lags of inflation and the 4 dummies. I then estimate it using the observations from 1962–1999 and use this estimated model to forecast inflation for 2000-2003.

OLS Forecast

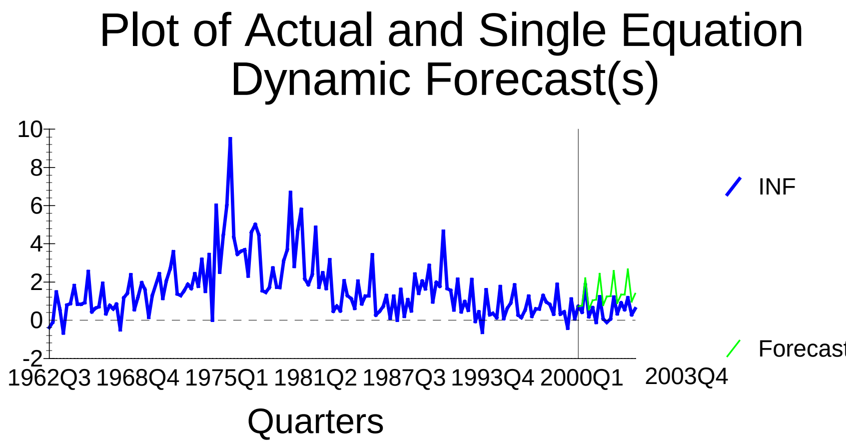
Table 6: Dependent variable is INF. 150 observations used for estimation from 1962Q3 to 1999Q4

Regressor	Coefficient	Standard Error	T-Ratio[Prob]
INF(-1)	.47034	.078577	5.9858[.000]
INF(-2)	.33125	.078439	4.2230[.000]
Q1	.40141	.17928	2.2391[.027]
Q2	1.5898	.18210	8.7305[.000]
Q3	-.72121	.22320	-3.2313[.002]
Q4	.058426	.23008	.25394[.800]
R-Squared	.62818	R-Bar-Squared	.61527
S.E. of Regression	.96904	F-stat. F(5, 144)	48.6578[.000]

Summary statistics for single equation dynamic forecasts

Based on 16 observations from 2000Q1 to 2003Q4			
Mean Prediction Errors	-.82998	Mean Sum Abs Pred Errors	.82998
Sum Squares Pred Errors	.84766	Root Mean Sumsq Pred Errors	.92069

Figure 3: Plot of actual and single equation dynamic forecast(s)



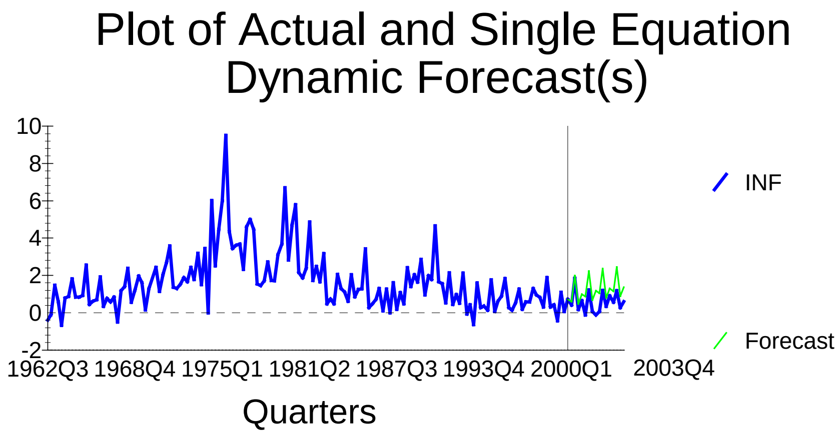
ARCH forecast

Table 7: GARCH(0,4) assuming a normal distribution converged after 146 iterations

```
*******************************************************************************
 Dependent variable is INF
 150 observations used for estimation from 1962Q3 to 1999Q4
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 INF(-1)                    .60962             .13884            4.3907[.000]
 INF(-2)                    .25576             .10418            2.4550[.015]
 Q1                         .14987             .20087            .74610[.457]
 Q2                         1.4290             .17158            8.3286[.000]
 Q3                        -.89312             .25442           -3.5104[.001]
 Q4                         .19358             .24292            .79690[.427]
*******************************************************************************
 R-Squared                     .61586    R-Bar-Squared                   .59692

           Summary statistics for single equation dynamic forecasts
*******************************************************************************
 Based on 16 observations from 2000Q1 to 2003Q4
 Mean Prediction Errors        -.69473    Mean Sum Abs Pred Errors       .69473
 Sum Squares Pred Errors        .62009    Root Mean Sumsq Pred Errors    .78746
*******************************************************************************
```

Figure 4: Plot of actual and single equation dynamic forecast(s)



It can be seen from the summary statistics that the ARCH model is definitely more accurate in forecasting inflation. Whilst both models tend to overpredict inflation, every one of the four summary statistics is in displays smaller prediction error with the ARCH model.

Conclusion

This paper has shown that past values of inflation have a considerable capacity to explain its current value. This paper also outlines the procedure involved in using arch estimation, and shows that it improves the forecasting ability of the model. It has also been seen that seasonal dummies are appropriate for modelling inflation and can be combined to good effect with ARCH. It might have been interesting to try including seasonal dummies in the conditional variance equation to allow for seasonal patterns in volatility. Another interesting expansion of the model would be to use the conditional variances of inflation computed here in a macroeconomic growth model. This would allow investigation of the impact of inflation uncertainty on growth.

Bibliography

Gujarati, D.N. (2003), *Basic Econometrics.* 5th ed. Singapore: McGraw-Hill

Johnston, J & J. DiNardo. (1997), *Econometric Method,* Singapore: McGraw-Hill

Engle, R. (1982), 'Autoregressive Conditional Heteroscedasticity with Estimates of United Kingdom Inflation' *Econometrica,* 987-1007.

National Statistics Online (2004), Available from: <*http://www.statistics.gov.uk*>