Peter Tom-Wolverton
Summer 2008
Regression Analysis: Student Project
October 26, 2012

Introduction

While the aims of the US Federal Reserve System are many, perhaps the most discussed under normal economic conditions is that of implementing the national monetary policy. In particular, two of the goals of said policy, maximum employment and stable prices, are sometimes referred to as the Federal Reserve's "dual mandate".[1] This project explores the relation between unemployment, inflation, and the federal funds target rate, which is perhaps the Fed's best-known tool for implementing monetary policy.

Data

The data for the dependent variable (the fed funds rate) is the data from October 1982 to November 2008. October 1982 is the earliest date for which the St. Louis Federal Reserve has data available, while November 2008 represents the point just before the Fed switched to reporting the target rate as a range and set said range to 0 – 0.25%; additionally, since then, the Fed has kept this range constant and increasingly resorted to other methods (*viz.* the "Quantitative Easing" programs) to carry out its monetary policy mandate.

Said data, as well as the data for real gross domestic product, were taken from the St. Louis Federal Reserve at http://research.stlouisfed.org. The data for unemployment and Consumer Price Index (the latter of which was then differenced to find inflation rates) are from the US Bureau of Labor Statistics at http://www.bls.gov. More specifically, the exact URL for each data set was recorded on the **Sources** worksheet in the accompanying Excel workbook.

All data is in monthly format. All regressions were performed with the independent variables on a one-month lag, to reflect the most recent data the Federal Open Market Committee would have available when making interest rate decisions each month. The unemployment rate data is seasonally adjusted.

Lastly, it must be noted that while the data extracts from the St. Louis Federal Reserve give the monthly dates as if the values are for the 1st of each month, they actually represent end-of-month values: for example, the fed funds target rate for 1982-11-01 is listed as 9.0%, but the 9.0% rate was not actually decided until November 22.[2] Therefore, the timing of the data is consistent with the data on CPI and unemployment.

Hypothesis

To reflect the Fed's dual mandate, we expect to see a statistically significant positive correlation between the fed funds target rate and inflation (raising rates when inflation is high to restrict the money supply and encourage saving) and a similar negative correlation between the fed funds target rate and unemployment (lowering rates when unemployment is high to encourage lending and stimulate demand). We also expect "core" inflation (that is, ignoring food and energy costs) to have a more

1 http://web.archive.org/web/20110512041758/http://www.federalreserve.gov/faqs/money_12848.htm
2 http://www.newyorkfed.org/markets/statistics/dlyrates/fedrate.html

significant relation to the fed funds rate than inflation across all items, as many economists consider food and energy costs to be too volatile to inform proper monetary policy.

First Regression

We attempt to regress the fed funds rate directly on annual inflation for all items (measured as the percentage change between the current month's CPI and the twelve-month-prior value) and unemployment rate:

Y = Federal funds target rate
X_2 = Annual inflation rate (all items)
X_3 = Unemployment rate

The summary is shown below, with beta1 = -0.1432, beta2 = 98.8782, and beta3 = 0.3984.

Regression Statistics	
Multiple R	0.519704
R Square	0.270092
Adjusted R Square	0.265398
Standard Error	2.119532
Observations	314

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	516.9933	258.49664	57.54067	5.47E-22
Residual	311	1397.141	4.4924162		
Total	313	1914.135			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-0.14315	0.593428	-0.2412211	0.809543
Ann inflation (all items)	98.87817	11.16749	8.8541052	6.55E-17
Unemployment rate (%)	0.398414	0.088868	4.4831878	1.04E-05

The F-statistic is highly significant (as are both t-statistics), so we can reject the null hypothesis of no relationship between the variables. However, the R^2 of 0.2701 could be better. Additionally, the Durbin-Watson statistic is 0.04, indicating severe positive serial correlation. Lastly, the positive coefficient for the unemployment variable is unusual; we expected this relation to be the opposite.

Second Regression

The only change between this and the first regression is the substitution of "core" inflation (i.e. less food and energy) for inflation across all items in the X2 spot. The summary is shown below, with beta1 = 1.2338, beta2 = 185.4682, and beta3 = -0.3121.

Regression Statistics	
Multiple R	0.753929
R Square	0.568409
Adjusted R Square	0.565634

Standard Error	1.62983
Observations	314

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	1088.012	544.00578	204.7949	1.79E-57
Residual	311	826.1232	2.6563446		
Total	313	1914.135			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	1.233839	0.411169	3.0008097	0.002911
Ann inflation (less food, energy)	185.4682	9.948628	18.642587	1.34E-52
Unemployment rate (%)	-0.31212	0.081126	-3.847376	0.000145

The R^2 of 0.5684 is much improved from the first attempt, while the F-statistic is even more significant than before. This supports our hypothesis that the FOMC pays more attention to core inflation than to general inflation when setting rates. Additionally, the coefficient corresponding to unemployment is now negative, as we would expect. However, there is still severe positive serial correlation indicated, with a Durbin-Watson statistic of 0.06.

In an attempt to correct for this serial correlation, we use the Hildreth-Lu procedure, performing an autoregressive transformation with first-order serial correlation coefficient rho and iterating the regression for various values of rho. The summary of these runs follows.

Rho	Sum of squared residuals	R^2	DWS
0	826.1231645	0.564608	0.057495
0.1	674.1472701	0.561891	0.069158
0.2	538.8305409	0.557839	0.086465
0.3	419.8804698	0.551599	0.112614
0.4	317.2483512	0.541592	0.15292
0.5	230.83071	0.524688	0.21613
0.6	160.3937344	0.494216	0.31553
0.7	105.3784547	0.434973	0.465454
0.8	64.44186568	0.312901	0.661226
0.9	34.93902683	0.103646	0.902658
1	20.97147674	0.049941	1.441004

The procedure dictates that we should select the equation with the lowest sum of squared residuals, in this case at rho = 1. However, this transformation makes the regression's fit all but disappear, as the R^2 drops to 0.0499. Additionally, the DW statistic of 1.44 is still significantly below 2, suggesting that higher-order serial correlation may be present.
The other details of the regression with rho = 1 are shown below.

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	1.10238009	0.55119	8.147682	0.000355937

Residual	310	20.97147674	0.0676499
Total	312	22.07385683	

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-0.031039148	0.014793355	-2.0981818	0.036698
X2*	3.771398228	9.202871127	0.4098067	0.682231
X3*	-0.376051703	0.095888264	-3.9217698	0.000108

One important takeaway here is that the F-statistic is still highly significant, so despite the pitiful R^2, we can reject the null hypothesis of no relation between the changes in the fed funds rate and the changes in inflation and unemployment. Another notable result is the difference in the t-statistics of the two dependent variables: X3*'s is highly significant but X2*'s is not. This indicates that most of the regression's explanatory power is due to changes in unemployment rate rather than changes in inflation rate, which gives an idea for the next attempt.

Third Regression

Substituting in the year-over-year percentage change in the unemployment rate for the unemployment rate itself, the variables are now:

Y = Federal funds target rate
X_2 = Annual inflation rate (less food and energy) in percent
X_3 = Unemployment rate year-over-year percentage change

The summary of this regression is given below:

Regression Statistics	
Multiple R	0.834985152
R Square	0.697200203
Adjusted R Square	0.695252938
Standard Error	1.365160913
Observations	314

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	1334.535113	667.2675565	358.0406354	2.08782E-81
Residual	311	579.5996028	1.863664318		
Total	313	1914.134716			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-0.1074962	0.235483558	-0.45649132	0.64835535
Ann inflation (%) (less food, energy)	1.675770815	0.069250589	24.19865084	1.75397E-73
Unemployment rate 1Y % chg	-7.164575816	0.578508986	-12.38455408	6.58895E-29

The R^2 has improved again, to 0.6972, and the F- and t-statistics remain highly significant, indicating that the change in unemployment rate carries more explanatory power than the unemployment rate itself. Also, both coefficients' signs are still as expected. Unfortunately, the DWS of 0.10 shows that taking annual differences of the unemployment rate has not helped with the serial correlation problem, and so we cannot trust that the true fit is as good as the R^2 would indicate.

Applying the Hildreth-Lu procedure to this regression gave the following results:

Rho	Sum of squared residuals	R^2	DWS
0	574.9727349	0.696874	0.099867
0.1	471.0502751	0.693878	0.124649
0.2	378.6626995	0.689272	0.1607
0.3	297.7619192	0.682013	0.213575
0.4	228.2559288	0.670182	0.291434
0.5	169.9615917	0.650026	0.40506
0.6	122.5003228	0.613709	0.564282
0.7	85.07199086	0.543854	0.763354
0.8	56.0282505	0.40261	0.950444
0.9	32.85553205	0.157098	1.072975
1	21.52252078	0.024977	1.344356

Again, the lowest ESS occurs with rho = 1, but this destroys the fit of the regression. As before, the F-statistic and the t-statistic for the unemployment variable are still significant, while the t-statistic for the inflation variable is not ($p = 0.32$).

Fourth Regression

A correlation coefficient of 0.56 was measured between the data sets for core inflation and unemployment, leading us to question whether multicollinearity might be affecting the regression. While the significance of the t-statistics leads us to believe that multicollinearity is probably not a problem, for completeness's sake we nevertheless investigate whether dropping either independent variable gives a better result.

For our fourth regression, we drop the inflation term, leaving only the year-over-year change in unemployment rate as our sole independent variable:

Regression Statistics	
Multiple R	0.356462444
R Square	0.127065474
Adjusted R Square	0.124267607
Standard Error	2.314194134
Observations	314

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	243.2204345	243.2204345	45.41512179	7.70208E-11
Residual	312	1670.914281	5.355494491		
Total	313	1914.134716			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	5.275329259	0.130996176	40.2708646	1.2659E-125
Unemployment rate 1Y % chg	-6.603546752	0.9798893	-6.739074253	7.70208E-11

This model can be rejected due to the much lower R^2 as compared to the three-variable regression.

Fifth Regression

For this attempt, we instead drop the unemployment term, leaving only the core inflation rate as our sole independent variable:

Regression Statistics	
Multiple R	0.740180486
R Square	0.547867152
Adjusted R Square	0.546418009
Standard Error	1.665489821
Observations	314

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	1048.691536	1048.691536	378.0626701	1.01286E-55
Residual	312	865.4431796	2.773856345		
Total	313	1914.134716			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	0.07745317	0.286710525	0.270144145	0.787227961
Ann inflation (%) (less food, energy)	1.641400149	0.084417516	19.44383373	1.01286E-55

This model does a better job keeping the fit, with $R^2 = 0.5479$. However, the DWS of 0.09 is still far too low, indicating that the fit may be overestimated due to serial correlation. It is interesting to note that the inflation variable alone does a better job of explaining the fed funds rate than the YOY unemployment rate change variable alone does, even though the inflation term is the one that becomes insignificant as rho approaches 1 in the Hildreth-Lu procedure.

Sixth Regression

Our last attempt at fitting a model takes into account the possibility that the strong positive serial correlation plaguing our results may indicate that we are unknowingly omitting a relevant independent variable. As the Fed's mission also includes "fostering a healthy economy"[3], we will see if adding a variable for year-over-year growth in gross domestic product improves the regression:

Y = Federal funds target rate
X_2 = Annual inflation rate (less food and energy) in percent
X_3 = Unemployment rate year-over-year percentage change
X_4 = Real GDP year-over-year percentage change

The results are summarized below.

Regression Statistics	
Multiple R	0.84600318
R Square	0.715721381

3 http://www.stlouisfed.org/inplainenglish/PDF/PlainEnglish.pdf, p. 5

Adjusted R Square	0.712970297
Standard Error	1.324882921
Observations	314

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	1369.987142	456.6623806	260.1598256	2.55237E-84
Residual	310	544.1475741	1.755314755		
Total	313	1914.134716			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-0.984561809	0.300525752	-3.276131256	0.001171693
Ann inflation (%) (less food, energy)	1.661305906	0.067284433	24.69079149	3.50853E-75
Unemployment rate 1Y % chg	-4.031632693	0.895095567	-4.504136586	9.44984E-06
GDP % chg YOY	30.3466274	6.75254346	4.494103234	9.87777E-06

This regression's R^2 is the best yet, at 0.7157 (though the improvement from the third regression is not that large), and the F- and t-statistics remain highly significant as well. However, the DWS of 0.10 indicates that we still have not solved the serial correlation problem. Trying the first-differences transformation gave a similar result to those of the second and third regressions: R^2 dropped to almost nothing, DWS remained substantially below 2, and the only independent variable that remained significant at anything even close to the 95% level was the unemployment rate change variable.

Conclusions

The principle of parsimony dictates that, while the sixth regression has the highest R^2, it does not give a sufficiently better fit than the third regression to justify the additional independent variable. Therefore, the best model out of those we have explored is the third regression:

$$Y = -0.1075 + 1.6758\ X_2 - 7.1646\ X_3$$

where:
Y = Federal funds target rate
X_2 = Annual inflation rate (less food and energy) in percent
X_3 = Unemployment rate year-over-year percentage change

Unfortunately, the high degree of serial correlation present casts doubt on the precision of the estimated parameters, and attempting to correct for this serial correlation invariably destroys the explanatory power of the model. It is likely that higher-order correlation exists among the residuals and that more complex correctional measures (outside the scope of this course) are needed. However, there are still some conclusions we can draw from these tests:

- Core inflation (that is, ignoring food and energy) affects the fed funds rate more than general inflation.
- Both inflation and unemployment have nonzero effects on the fed funds rate (concluded since all t-statistics were highly significant), even if we cannot estimate said effects with confidence.