John Lambros
Regression Analysis
Spring 20102

Analysis of Inflation (Part 1)

I wanted to look at the value of money under a few different scenarios. In particular there are two things that are discussed regarding the US government and the US Dollar: whether there should be a federal reserve bank (The Fed), and whether we should be on a gold standard. To check the value of these, I want to test the predictability of inflation before and after The Fed was established, and before and after the US went off the Gold standard.

When considering inflation, the real question is whether next year's costs can be anticipated. To that end, let's assume that we know ahead of time a few other pieces of information and see if regression models work. Namely, I will test the Consumer Price Index against: the price of gold, the gold reserves of the US Treasury, and the Population of the US. Of course the first is likely to be meaningless (and certainly is before 1933 since the price was constant at $20.67/oz.), but it's included as a tractable piece of information. We might expect that the more gold the US has the more dollars it can put out, and so the more inflation that would occur. On the other hand, the larger the population, the fewer dollars there are per person, which might cause deflation (reduced CPI). Note that interest rates are known to have an impact on inflation, but I was unable to find data on interest rates going back far enough. The data I used is pulled from several sources (noted below in "Sources") and includes data between 1870 and 2013. There is interpolation between missing data (e.g. Treasury gold was only available every 5 years until 1950, and population data is from the census every 10 years). I selected 1870 as the start which would hopefully avoid the impact of both the CA Gold Rush and the Civil War.

The pre-Fed data then runs 1870-1913. The Fed data runs 1914-1972 (until the US came off the gold standard in 73). And the Modern data runs 1973-2012.

Regression Pre-Fed (Consumer Price Index between 1870 and 1913)

Regressing first on gold reserves: CPI = USTreasury * X_1 + X_0.

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	29.327	0.709	41.378	<2e-16
USTreasury	-0.002	0.001	-1.754	0.087

Residual standard error: 3.236 on 42 degrees of freedom
Multiple R-squared: 0.06823
Adjusted R-squared: 0.04605
F-statistic: 3.076 on 1 and 42 DF

This model has very little explanatory power with $R^2 = 0.068$, and there is an 8.7% chance of showing this coefficient when $H_0 (X_1 = 0)$ is true, so it would be difficult to not reject the null hypothesis.

A similar regression on population (in millions): CPI = USPop * X_1 + X_0

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	36.028	1.564	23.042	<2E-16
USPop	-0.116	0.023	-5.029	9.680E-06

Residual standard error: 2.648 on 42 degrees of freedom
Multiple R-squared: 0.3759
Adjusted R-squared: 0.361
F-statistic: 25.29 on 1 and 42 DF

This model also lacks explanatory power with $R^2 = 0.3759$, but there is only a 0.00097% chance of showing this coefficient when $H_0 (X_1 = 0)$ is true, so we would likely reject the null hypothesis. It does get interesting when you add the treasury gold reserves back in now:

Regressing on treasury reserves and population (in millions): CPI = USPop * X_2 + USTreasury * X_1 + X_0

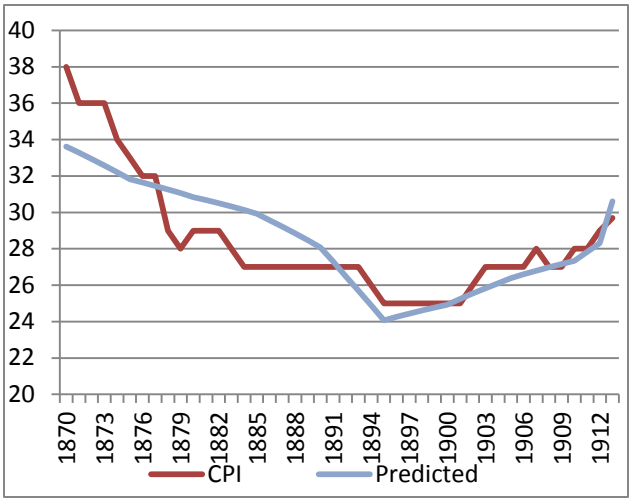
Coefficients:

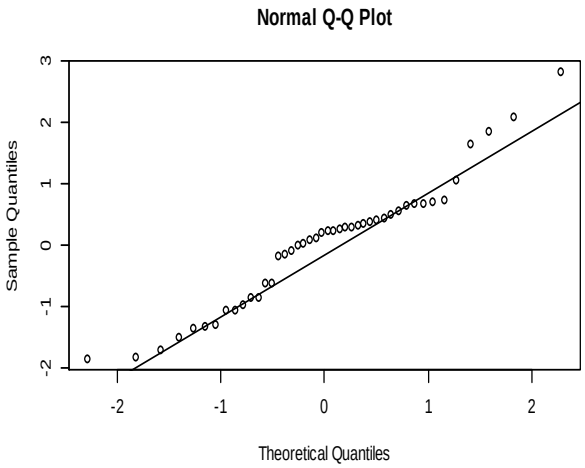
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	45.410	1.667	27.245	2.000E-16
USTreasury	0.007	0.001	7.240	7.560E-09
USPop	-0.326	0.033	-9.923	1.840E-12

Residual standard error: 1.776 on 41 degrees of freedom
Multiple R-squared: 0.7261
Adjusted R-squared: 0.7127
F-statistic: 54.34 on 2 and 41 DF

Note that now both X1 and X2 are significantly different from zero since they are more extreme due to treasury reserves and population having opposite effects on inflation. But more importantly, we see that R^2 is now .726 which means a significant portion of the CPI volatility is explained by these factors. One more statistic we will want for comparing this to other models will be the residual standard error (since the point is how close to reality we are predicting the CPI): $s_E = 3.153$. However, since CPI is, in general, increasing between the considered time periods, perhaps the ratio of standard error over the average CPI preferred: σ/μ=1.776/28.425=0.0625.

I also record here the QQ plot for this model which has a slightly positive skew suggesting an underestimation of the CPI in the extremes. On the bright side there is some (limited) responsiveness of the model when the CPI is changing.





Regression Fed (Consumer Price Index between 1914 and 1972)

Regressing first on the price gold: CPI = GoldPrice * X_1 + X_0.

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	4.547	12.462	0.365	0.717
GoldPrice	2.011	0.399	5.040	5.030E-06

Residual standard error: 20.8 on 57 degrees of freedom
Multiple R-squared: 0.3082
Adjusted R-squared: 0.2961
F-statistic: 25.4 on 1 and 57 DF

This model is interesting, and seems to suggest that a non-zero intercept in uncertain, but that CPI followed the price of Gold after the Fed was created. However, the R2 is only .308, so there is a significant amount of residual variance. Let's move on and see what else there is:

Regressing next on gold reserves: CPI = USTreasury * X_1 + X_0.

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	50.180	6.816	7.363	7.910E-10
USTreasury	0.001	0.001	2.577	0.013

Residual standard error: 23.67 on 57 degrees of freedom
Multiple R-squared: 0.1043
Adjusted R-squared: 0.08862
F-statistic: 6.64 on 1 and 57 DF

Across the board, this model suggests that the gold reserves don't explain much by themselves post Fed.

Moving on to regressing on population (in millions): CPI = USPop * X_1 + X_0.

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-38.242	5.202	-7.352	8.240E-10
USPop	0.717	0.035	20.497	2.000E-16

Residual standard error: 8.644 on 57 degrees of freedom
Multiple R-squared: 0.8805
Adjusted R-squared: 0.8784
F-statistic: 420.1 on 1 and 57 DF

Again, we see that the population appears to be significant, and further, the R2 is .88. However, now population is positively correlated with CPI. This is counter intuitive, and suggests significant influence by the Fed, likely in the speed of money and the ease of loan activity; this would lead to an inflationary environment, which is likely to mimic the population growth as another growing statistic rather than there being any true underlying correlation. Indeed, if I were to test H_0: $X_1 = 0$ against H_1: $X_1 < 0$ (which would correspond to the original expectation of why population was included), I would NOT reject the null hypothesis at any significance since the range of X_1 always intersects $X_1 \leq 0$.

Regressing now on all three: CPI = USPop * X_3 + USTreasury * X_2 + GoldPrice * X_1 + X_0

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-28.240	4.798	-5.886	2.460E-07
GoldPrice	-1.292	0.295	-4.387	5.240E-05
USTreasury	0.000	0.000	0.464	6.440E-01
USPop	0.910	0.043	21.166	2.000E-16

Residual standard error: 6.803 on 55 degrees of freedom
Multiple R-squared: 0.9286
Adjusted R-squared: 0.9247
F-statistic: 238.5 on 3 and 55 DF

Now we see that both the gold price and the population show as significant, which is consistent with the individual regressions. However, the next question is whether anything is added here. We can remove treasury reserves from the model, look at the F-test against the null hypothesis of the model: CPI = USPop * X_1 + X_0.

Coefficients:

	Estimate	Std.	Error	t
(Intercept)	-29.174	4.327	-6.742	9.240E-09
GoldPrice	-1.190	0.195	-6.110	1.000E-07
USPop	0.904	0.041	22.011	2.000E-16

Residual standard error: 6.755 on 56 degrees of freedom
Multiple R-squared: 0.9283
Adjusted R-squared: 0.9258
F-statistic: 362.7 on 2 and 56 DF

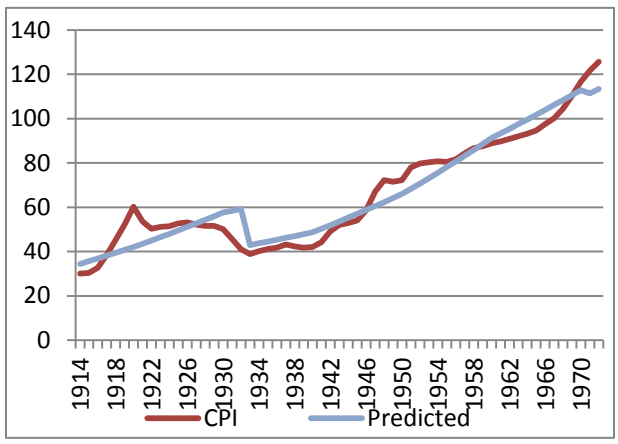
ANOVA

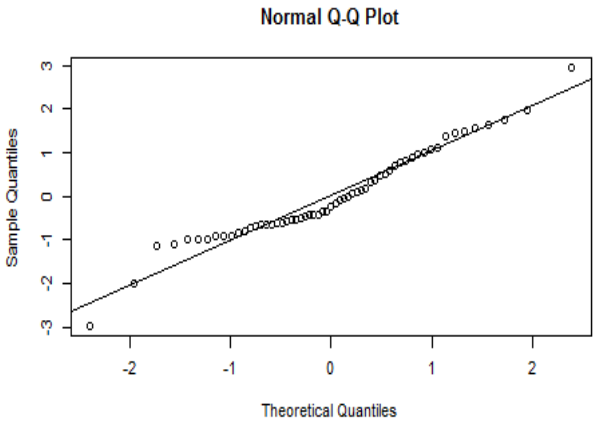
Model 1: CPI ~ USPop
Model 2: CPI ~ GoldPrice + USPop

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	57	4259.3				
2	56	2555.4	1	1703.8	37.338	1.00E-07

We see both an increase in R^2 and the F test, suggesting that the model with both variables is superior. Again we should note the residual error is: $s_E = 6.755$, which implies this is actually a worse model than before the Fed, despite explaining 92% of the variance, we are left with more residual variance. This is likely due to a large amount of the variance being "explained" by correlating to the population, which, as explained above may not really explain much. However, we should again consider the ratio of standard error over the average CPI preferred: $\sigma/\mu=6.755/65.85=0.1026$. This reaffirms that the model is weaker than the prior in actually predicting what the CPI is.

A quick look at the QQ plot shows no significant skew. But the predicted vs. actual CPI shows there is little prediction of the actual changes in the year to year CPI, instead it appears like a reactive "smoothing" of what happened to CPI.





Regression Modern (Consumer Price Index between 1973 and 2012)

For the modern era, I will skip reporting the details, as they are the same tests as above, and see the best linear fit is given by regressing: CPI = USPop * X_2 + USTreasury * X_1 + X_0. Though I note that the comparative F test, comparing this model vs. adding the price of gold gives 1.99 on 36 and 1 degrees of freedom. Hence, we do not reject the null hypothesis at a 90% confidence (since p = 16.7%).

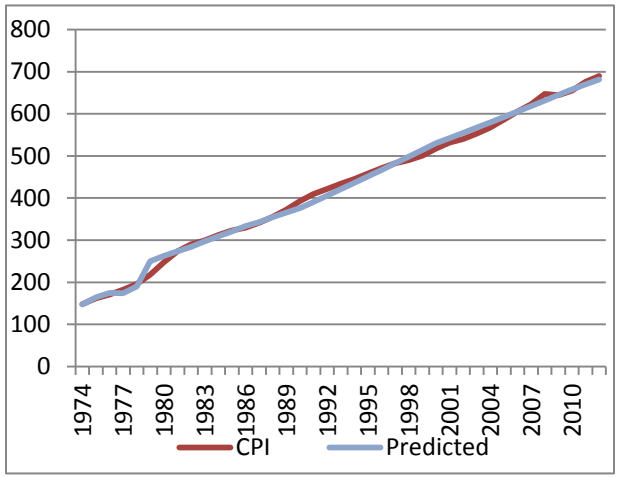
Coefficients:

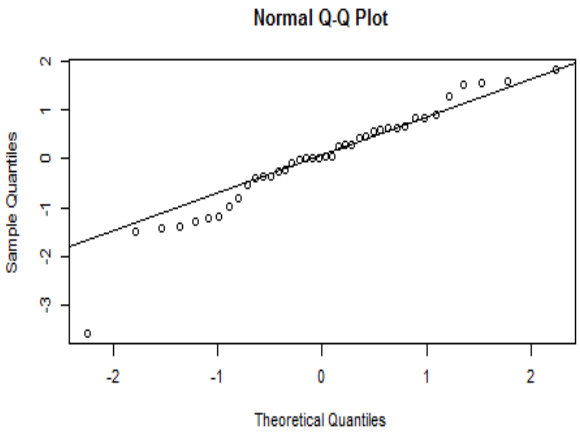
	Estimate	Std.	Error	Pr(>|t|)
(Intercept)	299.331	133.906	2.235	0.032
USTreasury	-0.133	0.015	-9.071	6.110E-11
USPop	4.671	0.072	64.672	2.000E-16

Residual standard error: 10.7 on 37 degrees of freedom
Multiple R-squared: 0.996
Adjusted R-squared: 0.9958
F-statistic: 4577 on 2 and 37 DF

Now, this looks like a great fit, but note that our mean squared error is now: $s_E = 10.7$. It is also interesting to note, that though the price of gold was significant by itself (with a p value of 3.04e-06 and an R^2 of .44), it was dominated by the introduction of the other variables… one of which (Treasury reserves) should have nothing to do with the CPI since the dollar is no longer tied to an underlying gold guarantee. Again, the ratio of standard error over the average CPI is: $\sigma/\mu=10.7/424.76=0.025$.

Here, the predicted and actual graphs look similar, but a simple line fit to time might have given similar results. Though, if we really wanted to, we might claim that it the reduction in the US treasury reserves in 1979 did anticipate the increase to CPI in 1980.





Again, in both the Fed and Modern periods, we see a positive correlation between population and CPI, suggesting that it is just a proxy variable represented the directed stable inflation. If we accept this as an OK way to use population, then the modern era does show the most stability, though (even with the exception of the 1979 outlier), the distribution of errors is still wider than the normal expectation.

Conclusion

Our ability to fit CPI to the variables given was always suspect at best, however, the fits are disturbingly good. The R^2 suggest improving quality of the models going from Pre Fed to Fed to Modern. The fact that the standard error is increasing somewhat dispels this, though one might expect that given the regression variables that the Fed environment (before going off the gold standard) should have been the most predictable. By looking at the graphs of actual vs. expected, I would argue that only the Pre-Fed model really predicted the changes in CPI. Overall, I don't really see much being gained from this modeling: none of the models really predicted the shocks to the CPI (though, I noted that the Modern model may have anticipated the CPI increase after the reserve sold some gold in 1979, the same model doesn't seem to have the sensitivity to detect the over inflation of 2008 or its decline in 2009); instead one could have simply looked at the graphs and figured out that the pre-Fed era had more volatility than the Fed era, which in turn had more volatility than the Modern era.

Sources

Data was compiled from information found in the following websites:

CPI data from Federal Reserve Bank:
http://www.minneapolisfed.org/community_education/teacher/calc/hist1800.cfm

Gold Prices:

http://www.onlygold.com/TutorialPages/PicesSince1972FS.htm

http://www.nma.org/pdf/gold/his_gold_prices.pdf

http://useconomy.about.com/od/monetarypolicy/p/gold_history.htm

Population:

http://en.wikipedia.org/wiki/Demographics_of_the_United_States