Recession Forecasting: The Probit Approach

VEE Regression Analysis Course (Students Project)

Fall 2011 Session

Nov 14th – Jan 8th 2012

Name: **Saurabh Khurana**

Email: pixelsallover.s@gmail.com

INDEX

INTRODUCTION

The project "**Recession Forecasting: The Probit Approach**" aims to analyze the Recessions occurred in the US economy since 1953 with the help of yield curve as the sole predictor variable. In the United States the beginning and ending dates of national recessions are determined by the **National Bureau of Economic Research (NBER)**. The NBER defines a recession as a recession as a general slowdown in economic activity over a long period of time, or a business cycle contraction visible in Gross Domestic Product (GDP), employment, investment spending, capacity utilization, household incomes, business profits and fall in inflation and rise in bankruptcies and the unemployment rates.

Many economist give the definition of recession as –"a true economic recession can only be confirmed if GDP (Gross Domestic Product) growth is **negative for a period of two or more consecutive quarters**"

Recession (or contraction) is a natural result of the economic cycle and will adjust for changes in consumer spending and consumption or increasing and decreasing prices of goods and labor.

Some of the **Recession periods** as given by the NEBR are:

NAME	DATE	DURATION
1953	July 1953 –May 1954	10 months
1958	Aug 1957 –April 1958	8 months
1960-61	Apr 1960 –Feb 1961	10 months
1969-70	Dec 1969 –Nov 1970	11 months
1973-75	Nov 1973 –Mar 1975	1 year 4 months
1980	Jan–July 1980	6 months
Early 1980s	July 1981 – Nov 1982	1 year 4 months
Early 1990s	July 1990 – Mar 1991	8 months
Early 2000	Mar – Nov 2001	8 months
Late 2000	Dec 2007 – Current	Till date

Although the definition worked quite well, there are several problems with it. One is that it does not provide monthly dates of when recessions began or ended. For this purpose the National Bureau of Economic Research (NBER), whose chronology of recessions is widely accepted, uses monthly measures of production, employment, sales, and income, all expressed in real terms (after allowing for inflation).

Another problem with the two-consecutive-quarters definition is that there can be serious declines in economic activity even without two consecutive quarters of negative growth.

PREDICTORS OF RECESSION

Although there are no completely reliable predictors, the following are regarded to be possible predictors

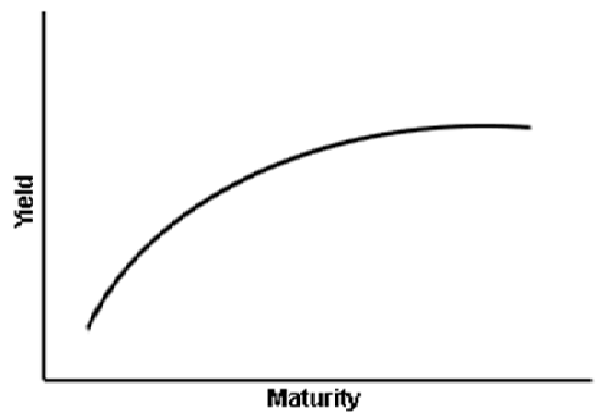
- **Inverted yield curve** - The model developed by economists uses yields on 10-year and three-month Treasury securities as well as the Fed's overnight funds rate. It is, however, not a definite indicator; it is sometimes followed by a recession 6 to 18 months later.
- **A Stock market drop** of a significant amount - of at least 10% of the overall value of the market - is considered by many to be one of the more reliable predictors of a recession.
- The three-month change in the **unemployment rate** and initial jobless claims.
- **Index of Leading (Economic) Indicators** (includes some of the above indicators) many experts rely on the Index of Leading Economic Indicators when it comes to predictors of a recession. This Index includes a number of different factors - including inflation, unemployment rates, consumer spending and so forth in order to make at least a basic determination pertaining to the state of the economy as a whole.
- **Lowering of Home Prices** - Lowering of home prices or value, too much personal debts, etc.

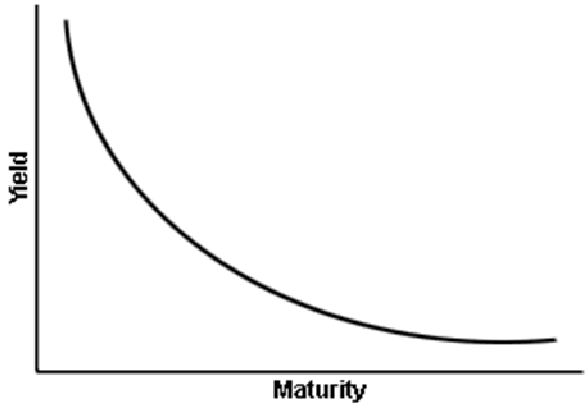
Yield curve is often considered as the sole predictor of recessions for the US economy.

Defining a Yield Curve

The U.S. Treasury issues many types of debt, from short-term to long-term. We can buy Treasury bonds that mature in three months or in thirty years, and quite a few in between. We can buy them straight from the Treasury or from others who are selling ones they bought from the Treasury. The interest rate on the bonds that is relevant to the yield curve is the “yield to maturity” or just “yield,” which accounts for both the coupon payments from the bonds and the price we paid for the bond. The yield curve simply plots the yield on the bond against its time to maturity.

Usually, the yield curve slopes up: longer-term bonds have higher yields than do short-term bonds, as people feel those longer-term bonds have more risk, requiring a higher return. Putting differently, the spread (difference) between any given long rate and a short rate is usually positive. When short rates rise above long rates, the yield curve is said to be inverted (and the spread is negative).





The steepness of the yield curve should be an excellent indicator of a possible future recession for several reasons. Current monetary policy has a significant influence on the yield curve spread and hence on real activity over the next several quarters. A rise in the short rate tends to flatten the yield curve as well as to slow real growth in the near term. This relationship, however, is only one part of the explanation for the yield curve's usefulness as a forecasting tool.

An inverted yield curve normally signals a recession, which begins about six months later. The stock market usually begins to fall six months prior to any recession. So, the appearance of an inverted yield curve normally is followed very shortly by a falling stock market. The yield curve for U.S. Treasury debt certificates is the one that investors use to predict the economy. Investors assume that the Treasury is the safest lender -- the least likely to default -- and therefore the rates on Treasury debt are least affected by risk.

Yield curve predictions about future growth come in two general flavors. One tries to predict the rate of growth that can be expected at some point in the future, the other tries to predict the probability of a recession occurring. The first uses a term spread to predict future output, usually at a horizon of two, four, or six quarters. This approach has an advantage in that it reveals more of the information that is in the yield curve. Inversions are not the curve's only signal: While negative spreads precede negative growth, positive spreads precede positive growth. In general, steeper the curve, higher the expected growth.

So the most straightforward, and, arguably, robust way to use the yield curve is to use the statistical relationship between its slope and future economic growth, and then look where the current yield curve is pointing. This approach requires choosing first a specific spread to use as a measure of the yield curve's slope, and then choosing a measure of output.

One of the most reliable and most often watched spreads is the one between 10-year U.S. Treasury bonds and 3-month T-bills. The natural and probably the most popular choice for growth is real GDP growth, taken at a quarterly or on a year-over-year basis, and predicting out one year.

Since 1960, the economy has been in a recession 14 percent of the time. So while not predicting a recession for sure, the yield curve indicates that the odds are substantially greater than average.

An inverted yield curve has been viewed as an indicator of a pending economic recession. When short-term interest rates exceed long-term rates, market sentiment suggests that the long-term outlook is poor and that the yields offered by long-term fixed income will continue to fall. More recently, this viewpoint has been called into question as foreign purchases of securities issued by the U.S.

Despite the evidence linking the yield curve to economic growth, and even though yield-curve inversions preceded the two most recent recessions, many have suggested that the yield curve no longer reliably predicts economic growth. Noting that the economy is continually evolving, particularly the financial sector, they discount past successes. They point to two recent "near misses" in 1995 and 1998, when a flat yield curve did not presage slow growth. And indeed, evidence since the early 1990s suggests that the relationship between the yield curve and growth has shifted, if not disappeared. Thus, some people argue that it is inappropriate to use data before 1990 to measure the connection between the yield curve and economic growth. Even using this more modern sample, though, the yield curve's predictions have not fallen in line with the consensus

There are two scenarios when the **yield curve will steepen** (i.e. long-term rates rise faster than short-term rates or short-term rates fall faster than long-term rates):

1. In periods where the fixed income market believes inflation will be rising. This usually occurs when the economy is heating up. In such an environment, the yield curve will steepen because investors fear that a strong economy may lead to inflation.

   Thus, long-term investors demand to be compensated for the inflation risk. The same pattern tends to occur in an actual rising inflationary environment, which reflects the fixed income market's belief that inflation will erode the value of the long bond's yield.

2. In periods when the economy is actually weakening. When the economy is weakening, the Federal Reserve will normally lower short-term interest rates. As a result, short-term rates fall faster than long-term rates, thus causing the yield curve to steepen.

There are two **scenarios when the yield curve tends to flatten** (short-term rates rise faster than long-term rates, or long-term rates fall faster than short-term rates):

1. The Federal Reserve is raising short-term interest rates to restrict economic growth. Typically, during this scenario short- term rates will rise faster than long-term rates.
2. The market believes that inflation is under control or poised to decline. Under this scenario, long-term interest rates will decline faster than short-term rates because investors no longer feel they need the same compensation for committing money for the long-term.

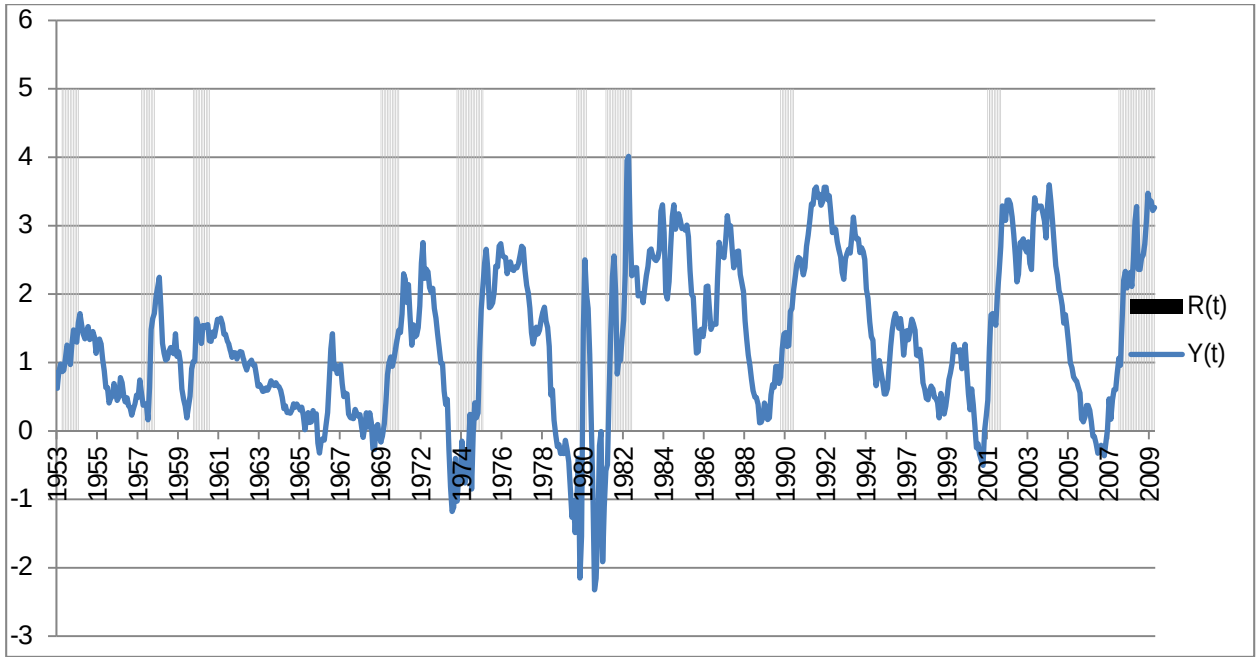
Taking it one step further, **a yield curve inversion** (i.e. short-term rates higher than long-term rates) suggests that investors may be viewing the economic environment from two different perspectives:

1. In the short end of the curve, investors are expecting short-term rates to rise (i.e.; the Fed may be raising the Fed Funds rate in an attempt to control inflation and/or economic growth).
2. For longer-term investors, the belief is that the sharply rising short-term rates will potentially cause a recession. Traditionally a recession would cause inflation to fall, which historically has caused long-term interest rates to fall. Thus, long-term investors actively seek to lock in the long-term interest rates before rates drop.

Economic Indicators Used

Yield Spread

It is the difference between yields on differing debt instruments, calculated by deducting the yield of one instrument from another. The higher the yield spread, the greater the difference between the yields offered by each instrument. The spread can be measured between debt instruments of differing maturities, credit ratings and risk.



Data: TB10 10 Year Treasury Bond Constant Maturity Rate
TB3 3 Month Treasury Bill Rate

Variable: Spread Percentage difference between 10 year and 3 month Treasury Bill Rates

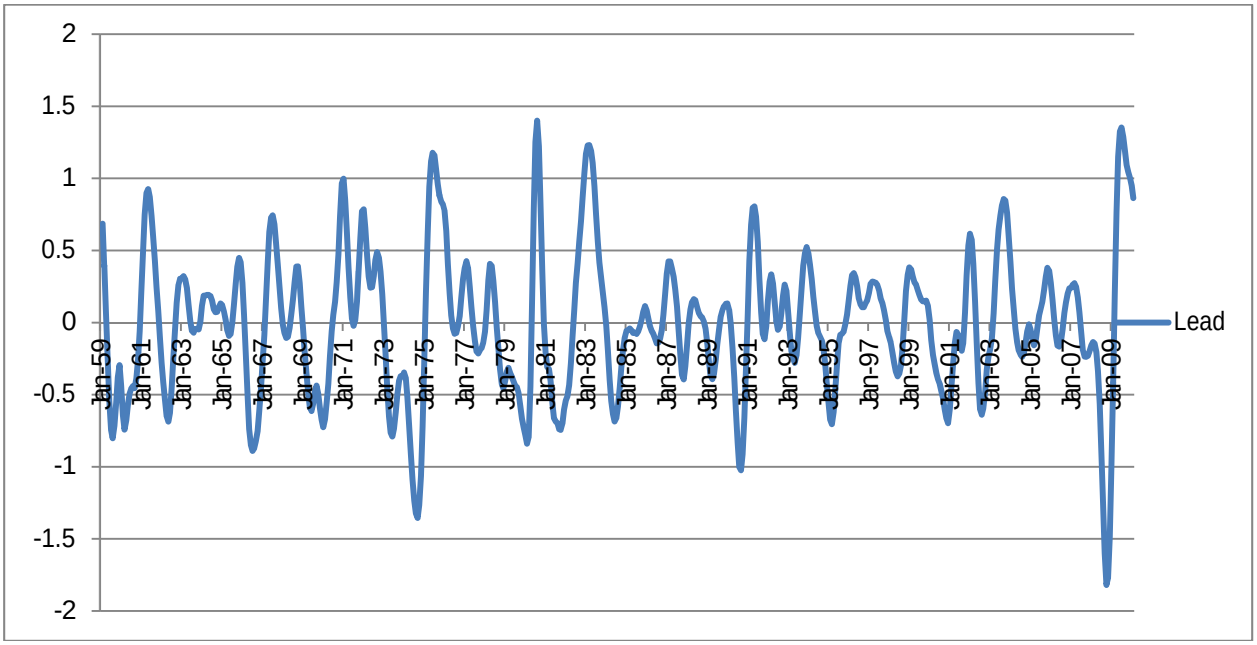
$$Spread = 100\left[\ln\left(1+\frac{TB10}{100}\right)-\ln\left(1+\frac{TB3}{100}\right)\right]$$

Conference Board's Index of Leading Economic Indicators

It is an index published monthly by the Conference Board used to predict the direction of the economy's movements in the months to come. The index is made up of 10 economic components, whose changes tend to precede changes in the overall economy. These 10 components include:

1. the average weekly hours worked by manufacturing workers
2. the average number of initial applications for unemployment insurance
3. the amount of manufacturers' new orders for consumer goods and materials
4. the speed of delivery of new merchandise to vendors from suppliers
5. the amount of new orders for capital goods unrelated to defense
6. the amount of new building permits for residential buildings
7. the S&P 500 stock index
8. the inflation-adjusted monetary supply (M2)
9. the spread between long and short interest rates
10. consumer sentiment

The Composite Index of Leading Indicators is a number that is used by many economic participants to judge what is going to happen in the near future. By looking at the Composite Index of Leading Indicators in the light of business cycles and general economic conditions, investors and businesses can form expectations about what's ahead, and make better-informed decisions.



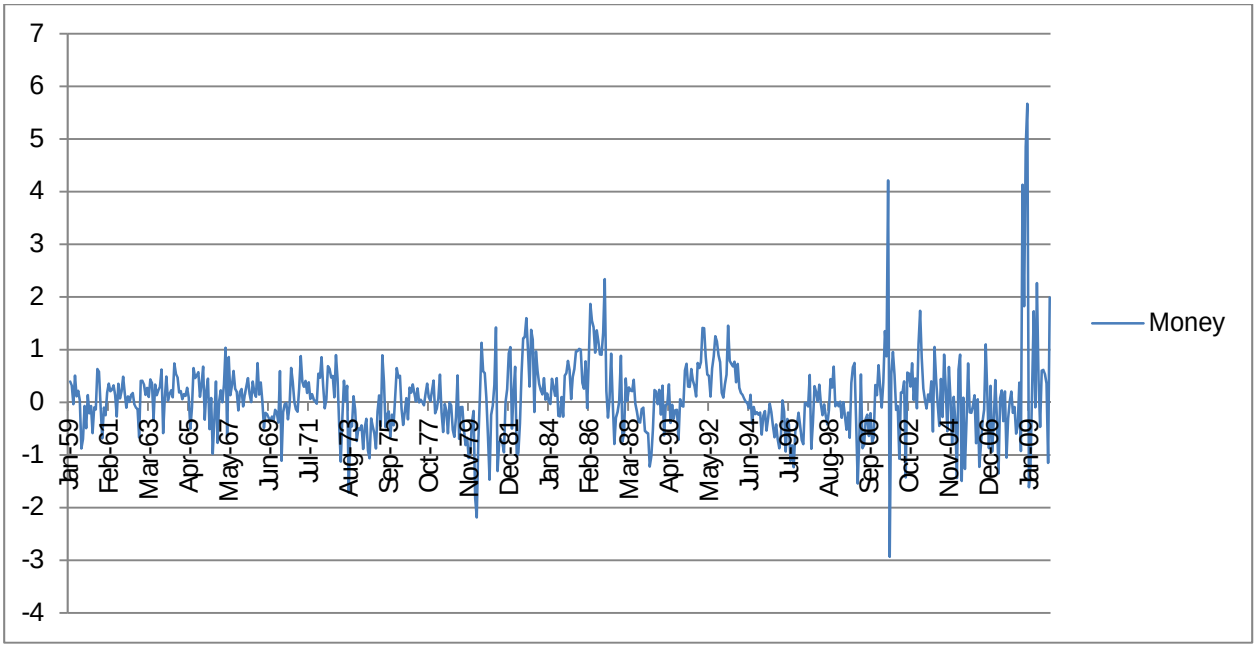
Data: LI Conference Board's Index of Leading Economic Indicators
Variable: Lead Percentage change in the Conference Board's Index of leading indicators

$$Lead = LI_t - LI_{t-1}$$

Money Supply M1

It is the entire quantity of bills, coins, loans, credit and other liquid instruments in a country's economy.

Money supply is divided into multiple categories - M0, M1, M2 and M3 - according to the type and size of account in which the instrument is kept. The money supply is important to economists trying to understand how policies will affect interest rates and growth.



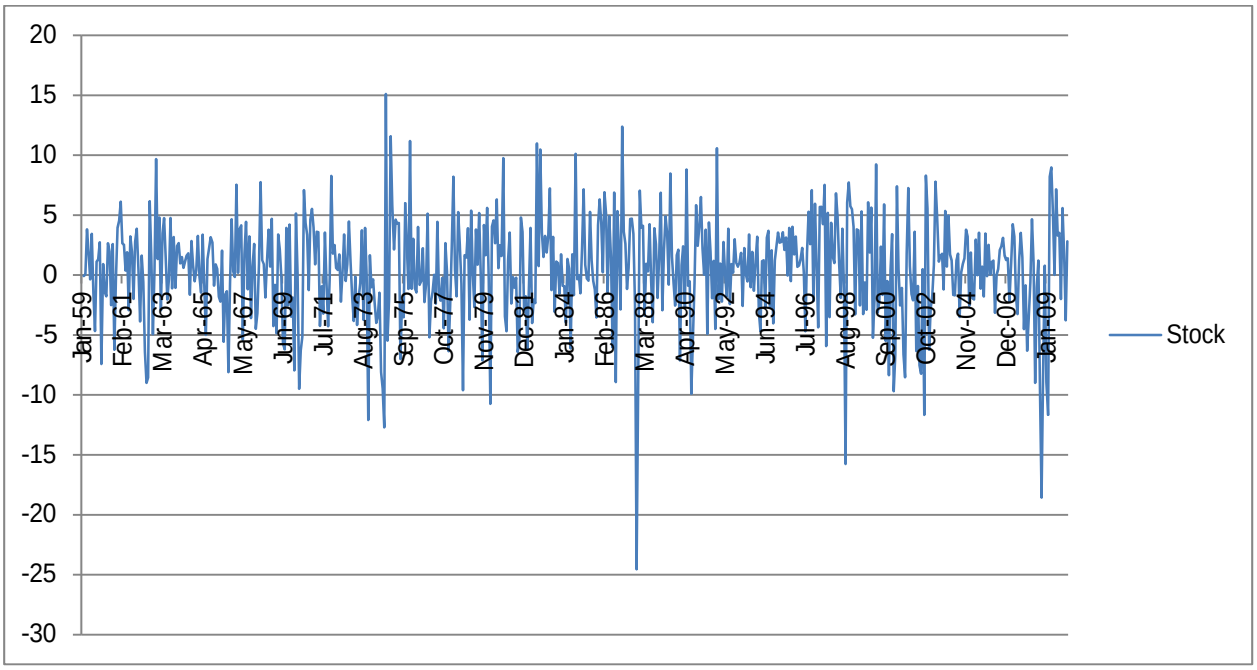
Data:	M1	M1 Money Stock
	CPI	Consumer Price Index
Variable:	Money	Monetary Growth measured using M1 Money Stock

$$Money = 100\left[\ln\left(\frac{M1_t}{CPI_t}\right) - \ln\left(\frac{M1_{t-1}}{CPI_{t-1}}\right)\right]$$

Standard & Poor's 500 U.S. Equity Index

It is an index of 500 stocks chosen for market size, liquidity and industry grouping, among other factors. The S&P 500 is designed to be a leading indicator of U.S. equities and is meant to reflect the risk/return characteristics of the large cap universe.

Companies included in the index are selected by the S&P Index Committee, a team of analysts and economists at Standard & Poor's. The S&P 500 is a market value weighted index - each stock's weight is proportionate to its market value.
The S&P 500 is one of the most commonly used benchmarks for the overall U.S. stock market. The Dow Jones Industrial Average (DJIA) was at one time the most renowned index for U.S. stocks, but because the DJIA contains only 30 companies, most people agree that the S&P 500 is a better representation of the U.S. market. In fact, many consider it to be *the* definition of the market.



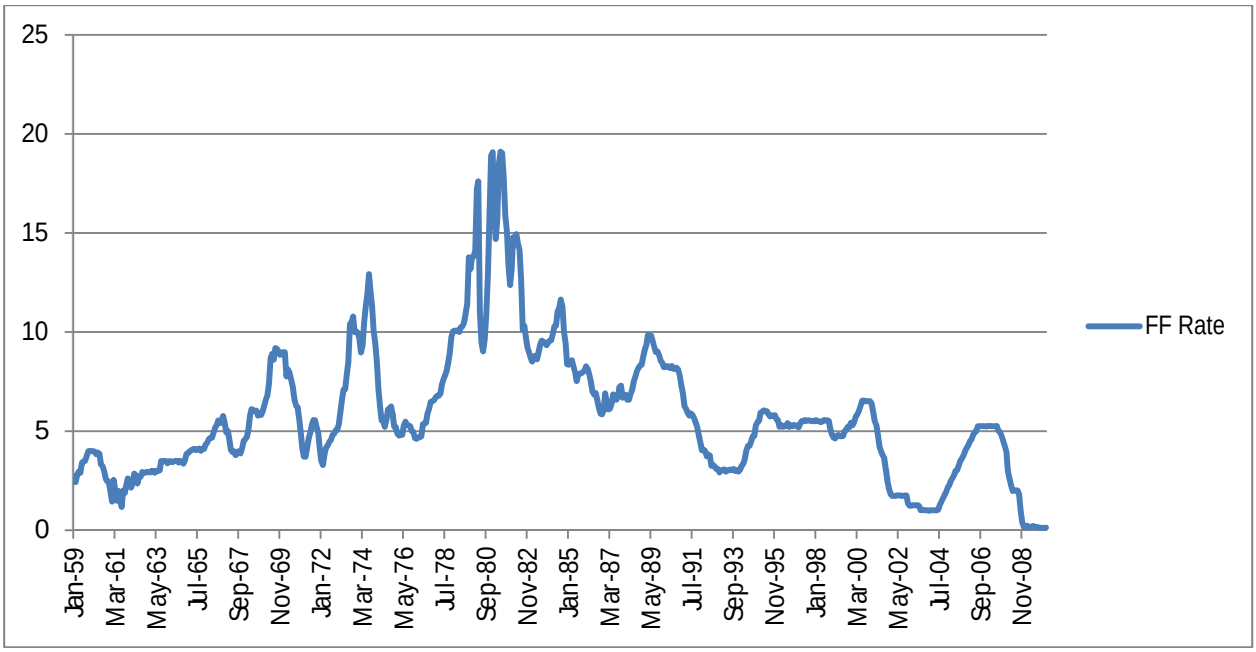
Data: S&P500 Standard & Poor's 500 U.S. Equity Index
Variable: Stock Percentage change in S&P500 Index of stock prices

$$Stock = 100\left[\ln \frac{S\,\&\,P500_t}{S\,\&\,P500_{t-1}}\right]$$

Federal Funds Rate

It is the interest rate at which a depository institution lends immediately available funds (balances at the Federal Reserve) to another depository institution overnight.

This is what news reports are referring to when they talk about the Fed changing interest rates. In fact, the FOMC (Federal Open Market Committee) sets a target for this rate, but not the actual rate itself (because it is determined by the open market).



Data:	FF	Federal Funds Rate
Variable:	FF	Federal Funds Rate in percentage

APPROACH

To assess how well the indicator variable predicts recessions, we use the so-called Probit model, which, in our application, directly relates the probability of being in a recession to a specific explanatory variable such as the yield curve spread. For example, one of the most successful models in our study estimates the probability of recession k months ahead in the future as a function of the current value of the yield curve spread between the ten-year Treasury note and the three-month Treasury bill.

The Probit Model

The objective is to examine recession probabilities using the simple probit model as used earlier. The Probit model is used to predict a recession dummy variable, R_t, where

$R_t = 1$ If the economy is in recession in period t and

$R_t = 0$ Otherwise

The dummy variables enable the isolation of recession forecasts. The equation is located below with the dependent variable defined as the probability of recession at time t.
The model is based on

Φ : Cumulative standard normal density function

k : The forecast horizon in months

The coefficient estimates are derived from the regression analysis

$$\Pr(R_t = 1) = \Phi(c_0 + c_1 Spread_{t-k})$$

$$\Pr(R_t = 1) = \Phi(c_0 + c_1 Spread_{t-k} + c_2 Lead_{t-k} + c_3 Money_{t-k} + c_4 Stock_{t-k} + c_5 FF_{t-k})$$

The Maximum Likelihood Estimates are obtained and there significance tested. The measure of fit for the probit model is based on and the Hosmer and Lemeshow Goodness-of-Fit Test. Model fit are better with a higher value of R^2 and greater significance of Chi-square. The lag period k with the most significant statistics is then used for prediction.

The value of studying the predictability of recessions is based on inputting available data into the model to forecast a recession in the future.

Data Sources

The data used is based on monthly time series ranging from January 1953 to October 2009.

1. NBER Recession Period
   Source: www.nber.org

2. 10 year Treasury Bond Constant Maturity Rate (Percent)
   Source: www.research.stlouisfed.org – TB10

3. 3 Month Treasury Bill Secondary Market Rate (Percent)
   Source: www.research.stlouisfed.org – TB3

4. Conference Board's Index of Leading Economic Indicators
   Source: Bloomberg, Symbol LEI CHNG – LI

5. M1 Money Stock (Seasonally Adjusted)
   Source: www.research.stlouisfed.org – M1

6. Consumer Price Index for All Urban Consumers: All Items (Seasonally Adjusted)
   Source: www.research.stlouisfed.org – CPI

7. Standard & Poor's 500 U.S. Equity Index
   Source: www.standardandpoors.com, FactSet – S&P 500

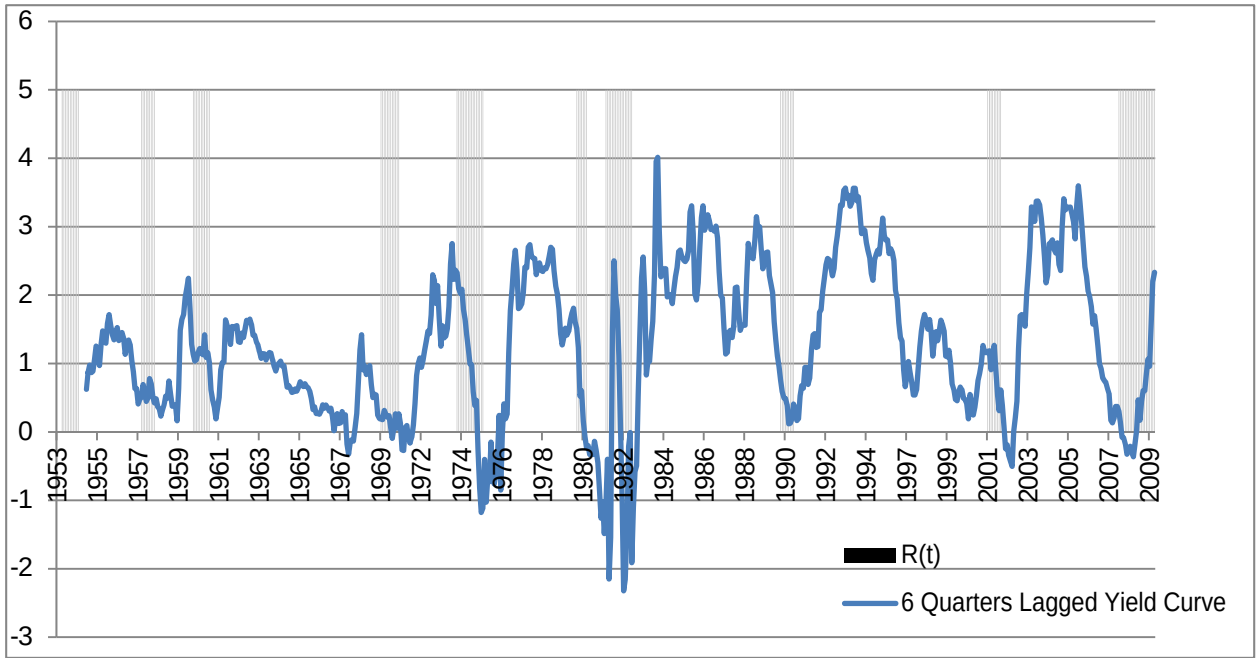
Total Data Set : January 1953 to October 2009 (679 observations)
Data Frequency : Monthly

Software Packages Used: SAS Enterprise Guide, SAS 9.1, Microsoft Excel + PHStat

The Yield Curve Model

Step 1: Preliminary Analysis of Dependence of Yield Curve on Recession Period

Theoretically, a yield curve inversion should lead to a recession. But such an effect is not immediate. For example, an 18 month lag in the yield curve inversion shows a better relation with the corresponding recession periods.



Also, every inversion in a yield curve does not predict recession. So, with an appropriate time lag in the yield spread and a significant Probit model, a suitable fit could be found.

Step 2: Applying Probit Regression to find a suitable lag

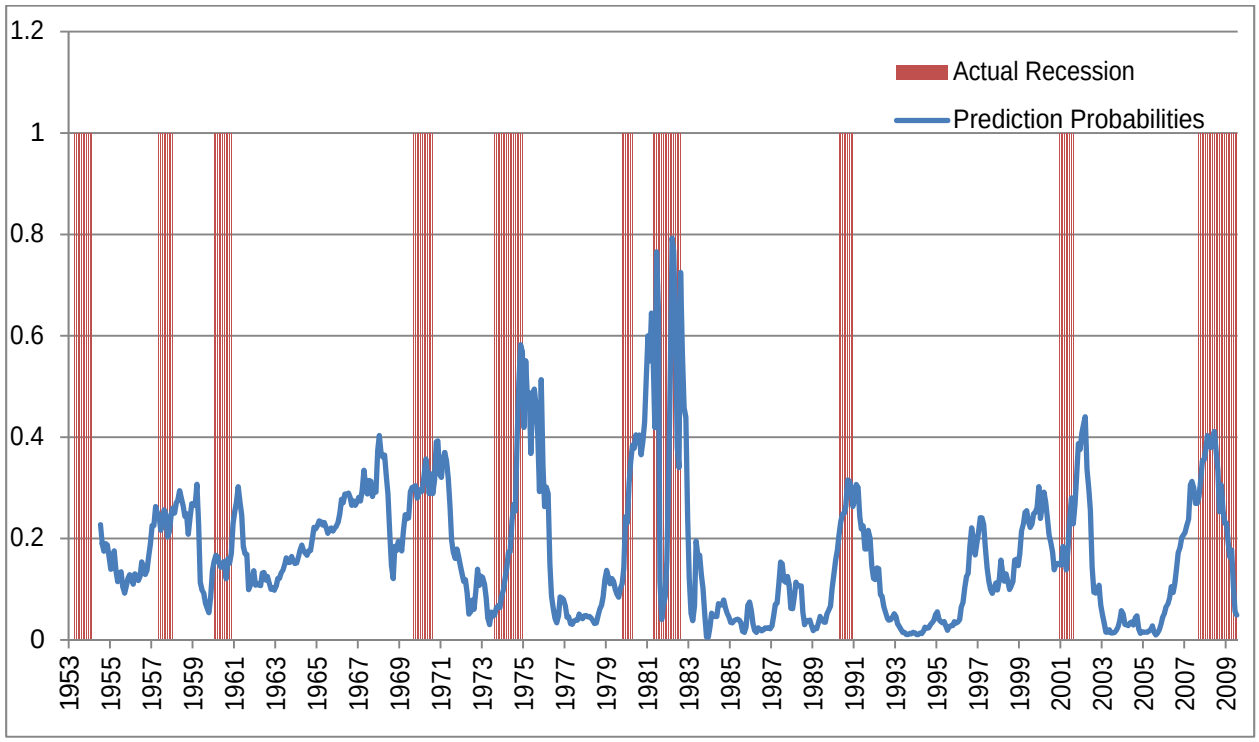
The Probit model is applied on the time series data with the yield curve lag ranging from 1 month to 24 months and corresponding statistics are shown.

	Maximum Likelihood Estimates				Hosmer and Lemeshow Goodness-of-Fit Test				
Yield Spread Lag	Intercept	Pr > ChiSq	Slope	Pr > ChiSq	Chi-Square	DF	Pr > ChiSq	R-Square	Max-rescaled R-Square
1	-0.7751	<.0001	-0.098	0.0515	20.1575	8	0.0098	0.0057	0.0093
2	-0.6682	<.0001	-0.1894	0.0002	12.9613	8	0.1132	0.0208	0.0338
3	-0.5713	<.0001	-0.2799	<.0001	9.5407	8	0.2988	0.0433	0.0702
4	-0.5071	<.0001	-0.3511	<.0001	11.7359	8	0.1634	0.0641	0.1042
5	-0.4569	<.0001	-0.4128	<.0001	18.9951	8	0.0149	0.084	0.1369
6	-0.4055	<.0001	-0.4787	<.0001	21.1615	8	0.0067	0.1065	0.174
7	-0.3559	<.0001	-0.5468	<.0001	27.6158	8	0.0006	0.13	0.2128
8	-0.3301	<.0001	-0.5909	<.0001	31.8904	8	<.0001	0.1452	0.2385
9	-0.3153	<.0001	-0.6216	<.0001	34.2434	8	<.0001	0.1553	0.2556
10	-0.3192	0.0001	-0.6274	<.0001	29.1675	8	0.0003	0.1564	0.258
11	-0.3218	0.0001	-0.6349	<.0001	20.3057	8	0.0092	0.1578	0.261
12	-0.3146	0.0002	-0.6589	<.0001	18.8218	8	0.0158	0.1626	0.2696
13	-0.3249	<.0001	-0.6555	<.0001	13.7243	8	0.0892	0.1591	0.2645
14	-0.3421	<.0001	-0.6405	<.0001	12.219	8	0.1417	0.1526	0.2544
15	-0.3515	<.0001	-0.6248	<.0001	17.0851	8	0.0292	0.1469	0.2448
16	-0.3599	<.0001	-0.6101	<.0001	18.5979	8	0.0172	0.1407	0.2342
17	-0.3914	<.0001	-0.5663	<.0001	14.2255	8	0.0761	0.1259	0.2094
18	-0.4169	<.0001	-0.5309	<.0001	12.458	8	0.1319	0.1143	0.19
19	-0.4459	<.0001	-0.4933	<.0001	13.7401	8	0.0888	0.1019	0.1694
20	-0.4734	<.0001	-0.4575	<.0001	14.6861	8	0.0655	0.0904	0.1502
21	-0.5087	<.0001	-0.4125	<.0001	12.4861	8	0.1308	0.0769	0.1277
22	-0.542	<.0001	-0.3723	<.0001	15.4599	8	0.0508	0.0649	0.1078
23	-0.5748	<.0001	-0.3348	<.0001	13.4154	8	0.0983	0.0541	0.0897
24	-0.6069	<.0001	-0.2993	<.0001	14.6517	8	0.0663	0.0442	0.0733

The statistics are most significant for an 18 month lag on yield spread.

Step 3: Analysis of the model

The predicted values for probability of recession with an 18 month lag obtained after fitting the model is used to compare with the real recession periods.



The model predicts recessions with a significant change in probability, with one false signal in 1968, up to 1990 recession but fails to predict further recessions, especially the 2007 recession, efficiently.

Step 4: Improving the Probit Model

Re-considering the yield curve, it can be found that yield curve has significantly flattened after the 1990s. So, separate Probit analysis for the significance of yield curve up to 1990 and after 1990 could be performed. The yield spreads could be grouped as follows

1. Group 1: Yield Curve up to 1990
2. Group 2: Yield Curve after 1990

The statistics for the Probit regression for the above two group is shown below

Group 1

	Maximum Likelihood Estimates				Hosmer and Lemeshow Goodness-of-Fit Test				
Yield Spread Lag	Intercept	Pr > ChiSq	Slope	Pr > ChiSq	Chi-Square	DF	Pr > ChiSq	R-Square	Max-rescaled R-Square
3	-0.4047	<.0001	-0.4538	<.0001	14.545	8	0.0686	0.0939	0.1473
6	-0.2693	0.0045	-0.6946	<.0001	15.2543	8	0.0544	0.1755	0.2775
7	-0.2261	0.019	-0.7807	<.0001	15.4553	8	0.0509	0.2018	0.3201
8	-0.2206	0.0232	-0.8151	<.0001	18.7829	8	0.0161	0.2132	0.3391
9	-0.2226	0.0226	-0.8321	<.0001	13.5596	8	0.094	0.2183	0.3483
10	-0.2464	0.0112	-0.8057	<.0001	10.3046	8	0.2443	0.2099	0.3359
11	-0.2681	0.0056	-0.7832	<.0001	4.4226	8	0.8171	0.2019	0.3241
12	-0.2842	0.0033	-0.7716	<.0001	2.8577	8	0.943	0.1964	0.3163
13	-0.3187	0.0009	-0.7199	<.0001	4.4968	8	0.8098	0.1796	0.2902
14	-0.3499	0.0002	-0.6756	<.0001	6.429	8	0.5993	0.1646	0.2668
15	-0.3678	0.0001	-0.6359	<.0001	7.8926	8	0.444	0.1514	0.2453
16	-0.3858	<.0001	-0.5961	<.0001	8.4364	8	0.392	0.1371	0.2219
17	-0.43	<.0001	-0.5236	<.0001	9.1644	8	0.3286	0.1126	0.182
18	-0.4672	<.0001	-0.4654	<.0001	7.7397	8	0.4593	0.0935	0.1511
19	-0.5121	<.0001	-0.4017	<.0001	8.3275	8	0.4022	0.0733	0.1184
20	-0.5517	<.0001	-0.3475	<.0001	8.4958	8	0.3866	0.0571	0.0922
21	-0.5985	<.0001	-0.2878	<.0001	7.1724	8	0.5181	0.041	0.0661
22	-0.6454	<.0001	-0.2323	0.0006	13.4621	8	0.0969	0.0275	0.0444
23	-0.6977	<.0001	-0.1745	0.0086	18.032	8	0.021	0.0159	0.0255
24	-0.7466	<.0001	-0.1232	0.0614	22.1785	8	0.0046	0.008	0.0128

The best lag period for recessions in the Group 1 is found out to be 12 months.

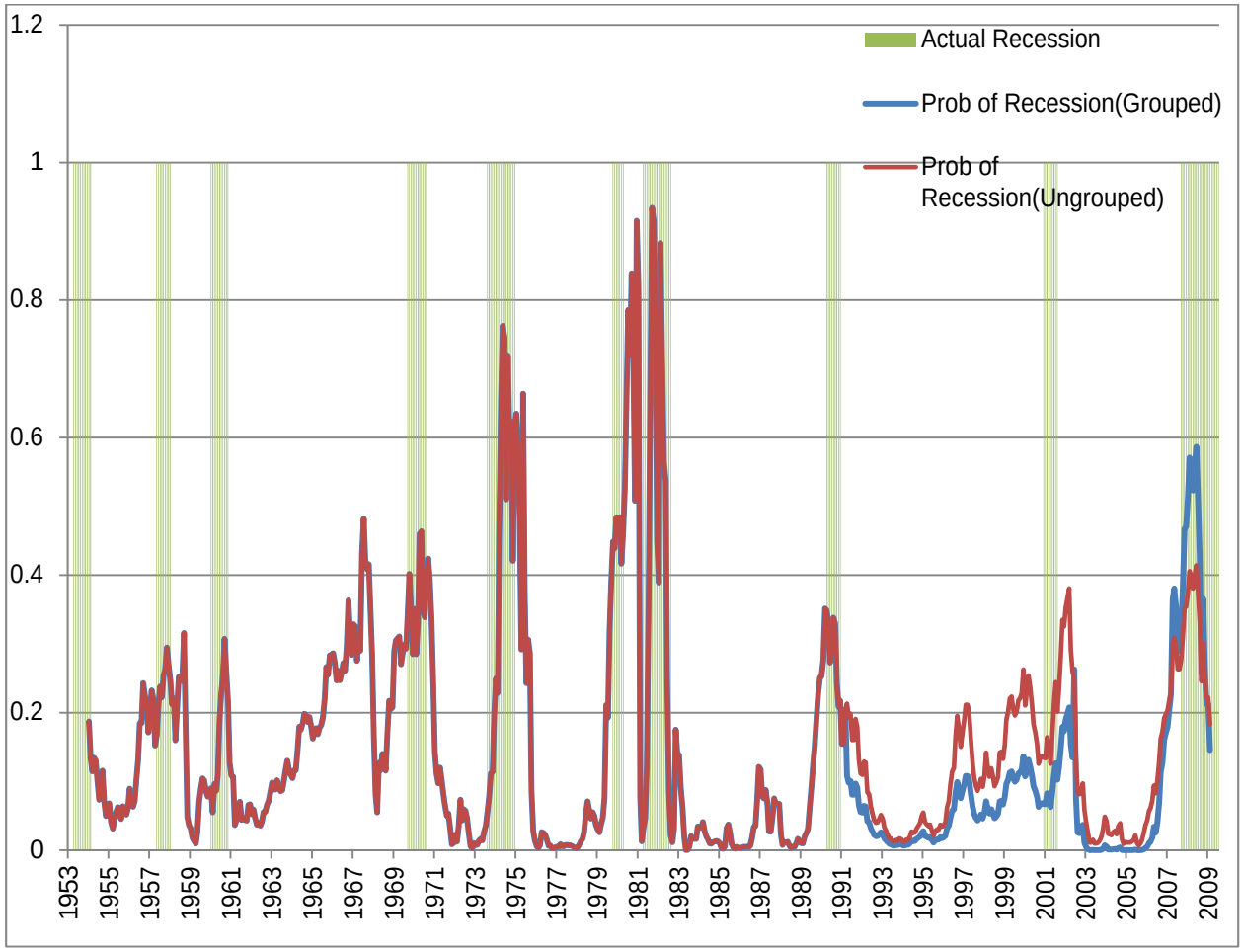
Group 2 analysis of the yield curve data has been done in Part 2 of the project.

Step 5: Model Forecasts

The results for the Group 2 data are taken from Part 1 of the Project.

Using the grouped model, the recession probabilities are forecasted using historical data only. For example 1982 recession is forecasted using data only up to 1982 and similarly for other recession periods.

The probability of recession using separate lag periods is plotted against the actual recession periods.



Results and Conclusion

Conclusions Drawn from Step 1

The simple graph of the yield curve depicts the Recession Periods and the Curve shows that almost all the recessions were preceded by an inverted yield curve.

The analysis of the yield curve with a 6 quarters lag and using the suitable Probit model depicts that it is a fairly better model as compared to the normal yield curve in predicting the Recession periods. As it predict the Recessions with a fairly better probability approximately 6 quarters before.

Conclusions Drawn from Steps 2 and 3

The Probit model is applied on the time series data with the yield curve lag ranging from 1 month to 24 months and corresponding statistics suggest that the yield curve model with a lag of 18 months is the most significant model for predicting of Recessions out of all the others for the entire data of spread from 1954 to 2009(Ungrouped data).

This model predicts:

- **Recession of 1982 with a probability of 76%**
- **Recession of 2001 with a probability of 41%**
- **Recession of 2007 with a probability of 37%**

The model predicts recessions with a significant change in probability, with one false signal in 1968, up to 1990 recession but fails to predict further recessions, especially the 2007 recession, efficiently.

Conclusions Drawn from Steps 4 and 5

Separate Probit analysis for the significance of yield curve up to 1990 and after 1990 was performed. The yield spreads was grouped as follows

- Group 1: Yield Curve up to 1990
- Group 2: Yield Curve after 1990

The statistics for the Probit regression for the above two group suggest that:

- For group 1 - The yield curve model with a lag of 12 months is the most significant model for predicting of Recessions of before 1990
- For group 2 - The yield curve model with a lag of 19 months is the most significant model for predicting of Recessions of before 1990

This model predicts:

- **Recession of 1982 with a probability of 91%**
- **Recession of 2001 with a probability of 41%**
- **Recession of 2007 with a probability of 60%**

The current Grouped Probit Model improves the probability of forecasting a recession.

The Five Variable Model

Step 6: Applying Probit Regression to find a suitable lag

Now we apply the following three models and test the significance of each model with various lead periods (in months):

1. Model A: $\Pr(R_t = 1) = \Phi(c_0 + c_1 Spread_{t-k})$
2. Model B: $\Pr(R_t = 1) = \Phi(c_0 + c_1 Spread_{t-k} + c_2 FF_{t-k})$
3. Model C: $\Pr(R_t = 1) = \Phi(c_0 + c_1 Spread_{t-k} + c_2 Lead_{t-k} + c_3 Money_{t-k} + c_4 Stock_{t-k} + c_5 FF_{t-k})$

Model A and B have been chosen because such models have been used widely in previous research. Model C is based on all the economic indices significant theoretically.

				Pr > Chi Sq						Hosmer and Lemeshow Goodness-of-Fit	
		R-Square	Max-rescaled R-Square	Intercept	Yield	FF	Lead	Money	Stock	Chi Sq	Pr > Chi Sq
Lead 0	Model A	0.0011	0.0018	<.0001	0.4174					27.5441	0.0006
	Model B	0.0334	0.0562	<.0001	0.1317	<.0001				47.5227	<.0001
	Model C	0.0931	0.1569	<.0001	0.0052	0.0002	<.0001	0.3066	0.9029	15.3727	0.0523
Lead 3	Model A	0.0531	0.0892	<.0001	<.0001					10.8592	0.2098
	Model B	0.0869	0.1461	<.0001	0.0059	<.0001				44.0553	<.0001
	Model C	0.2621	0.4409	<.0001	0.8983	0.001	<.0001	0.8333	0.4862	23.6052	0.0027
Lead 6	Model A	0.1216	0.2042	<.0001	<.0001					20.8793	0.0075
	Model B	0.145	0.2434	<.0001	<.0001	<.0001				36.036	<.0001
	Model C	0.3174	0.5329	<.0001	0.0002	0.0166	<.0001	0.3928	0.0407	5.826	0.6667
Lead 9	Model A	0.1777	0.2977	0.0002	<.0001					19.6206	0.0119
	Model B	0.1911	0.3203	<.0001	<.0001	0.002				41.1228	<.0001
	Model C	0.2623	0.4396	<.0001	<.0001	0.0979	<.0001	0.9692	0.3857	12.5145	0.1297
Lead 12	Model A	0.1854	0.3101	0.0005	<.0001					15.7136	0.0467
	Model B	0.1904	0.3185	0.001	<.0001	0.0561				5.6973	0.6811
	Model C	0.1931	0.3229	0.0008	<.0001	0.0959	0.2651	0.9356	0.6732	13.4611	0.0969
Lead 15	Model A	0.1673	0.2803	<.0001	<.0001					8.9622	0.3455
	Model B	0.1707	0.2861	0.0014	<.0001	0.1095				7.8403	0.4492
	Model C	0.1787	0.2995	0.0028	<.0001	0.0554	0.0342	0.9573	0.7077	15.6791	0.0472
Lead 18	Model A	0.1308	0.2212	<.0001	<.0001					10.2063	0.2508
	Model B	0.1337	0.226	0.0005	<.0001	0.1557				23.3673	0.0029
	Model C	0.1545	0.2614	0.0011	<.0001	0.0564	0.0007	0.556	0.3739	8.0659	0.4271

Model C with a 6 months lead period shows the best fit but the coefficient estimate of Money is not significant.
Detailed statistics for this model is shown in the SAS generated output.

SAS Output for Lead 6 Model C Full Model Regression

Model Information

Data Set	WORK.SORTTEMPTABLESORTED	
Response Variable	NBER Recession	NBER Recession
Number of Response Levels	2	
Model	binary probit	
Optimization Technique	Fisher's scoring	

Response Profile

Ordered Value	NBER Recession	Total Frequency
1	1	102
2	0	505

Probability modeled is NBER Recession=1.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	551.655	329.839
SC	556.064	356.290
-2 Log L	549.655	317.839

R-Square 0.3174 **Max-rescaled R-Square** 0.5329

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	231.8170	5	<.0001
Score	200.3037	5	<.0001
Wald	132.3893	5	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.5173	0.2252	45.3747	<.0001
Yield Curve2	1	-0.3103	0.0844	13.5105	0.0002
FF Rate2	1	0.0585	0.0244	5.7402	0.0166
Lead2	1	-2.0539	0.2426	71.6746	<.0001
Money2	1	0.1032	0.1208	0.7301	0.3928
Stock2	1	-0.0358	0.0175	4.1884	0.0407

Partition for the Hosmer and Lemeshow Test

		NBER Recession = 1		NBER Recession = 0	
Group	**Total**	**Observed**	**Expected**	**Observed**	**Expected**
1	61	0	0.01	61	60.99
2	61	0	0.18	61	60.82
3	61	0	0.62	61	60.38
4	61	0	1.33	61	59.67
5	61	3	2.26	58	58.74
6	61	5	3.83	56	57.17
7	61	4	6.79	57	54.21
8	61	16	13.75	45	47.25
9	61	31	27.21	30	33.79
10	58	43	44.73	15	13.27

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
5.8260	8	0.6667

Generated by the SAS System (Local, XP_PRO) on 24APR2010 at 5:58 PM

Step 7: Applying Step-Wise Probit Regression

A step-wise regression is applied to take only the significant variables in the model.

Lead	R-Square	Max-rescaled R-Square	Pr > Chi Sq						Hosmer and Lemeshow Goodness-of-Fit	
			Intercept	Yield	FF	Lead	Money	Stock	Chi Sq	Pr > Chi Sq
0	0.0914	0.154	<.0001	0.001	0.0002	<.0001			26.7477	0.0008
3	0.2614	0.4397	<.0001		0.0005	<.0001			27.4454	0.0006
6	0.3166	0.5316	<.0001	0.0004	0.0197	<.0001		0.0411	3.2825	0.9154
9	0.258	0.4323	<.0001	<.0001		<.0001			12.1198	0.1459
12	0.1854	0.3101	0.0005	<.0001					15.7136	0.0467
15	0.1735	0.2908	0.0104	<.0001		0.037			7.3405	0.5004
18	0.1476	0.2496	0.0028	<.0001		0.0008			11.8423	0.1584

The statistics for step-wise regression on various lead periods are summarized in the above table.
The detailed statistic as generated by SAS for the lead 6 model is shown below.

SAS Output for Lead 6 Step-Wise Regression

Model Information

Data Set	WORK.SORTTEMPTABLESORTED	
Response Variable	NBER Recession	NBER Recession
Number of Response Levels	2	
Model	binary probit	

Number of Observations Read	614
Number of Observations Used	607

Response Profile

Ordered Value	NBER Recession	Total Frequency
1	1	102
2	0	505

Probability modeled is NBER Recession=1.

Note: 7 observations were deleted due to missing values for the response or explanatory variables.

Stepwise Selection Procedure

Step 0. Intercept entered:

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
200.3037	5	<.0001

Step 1. Effect Lead2 entered:

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	551.655	359.154
SC	556.064	367.971
-2 Log L	549.655	355.154

R-Square	0.2742	**Max-rescaled R-Square**	0.4603

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	194.5012	1	<.0001
Score	155.8216	1	<.0001
Wald	114.1193	1	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
37.5609	4	<.0001

Note: No effects for the model in Step 1 are removed.

Step 2. Effect Yield Curve2 entered:

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	551.655	334.042
SC	556.064	347.268
-2 Log L	549.655	328.042

R-Square	0.3059	**Max-rescaled R-Square**	0.5135

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	221.6132	2	<.0001
Score	177.0403	2	<.0001
Wald	129.5457	2	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
10.8844	3	0.0124

Note: No effects for the model in Step 2 are removed.

Step 3. Effect FF Rate2 entered:

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	551.655	330.915
SC	556.064	348.549
-2 Log L	549.655	322.915

R-Square	0.3117	Max-rescaled R-Square	0.5233

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	226.7405	3	<.0001
Score	192.6393	3	<.0001
Wald	133.9164	3	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
5.2924	2	0.0709

Note: No effects for the model in Step 3 are removed.

Step 4. Effect Stock2 entered:

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	551.655	328.555
SC	556.064	350.598
-2 Log L	549.655	318.555

R-Square	0.3166	Max-rescaled R-Square	0.5316

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	231.1001	4	<.0001
Score	198.5225	4	<.0001
Wald	133.1445	4	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
0.6547	1	0.4184

Note: No effects for the model in Step 4 are removed.

Note: No (additional) effects met the 0.05 significance level for entry into the model.

Summary of Stepwise Selection

Step	Effect Entered	Effect Removed	DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq	Variable Label
1	**Lead2**		1	1	155.8216		<.0001	Lead2
2	**Yield Curve2**		1	2	25.2282		<.0001	Yield Curve2
3	**FF Rate2**		1	3	5.3637		0.0206	FF Rate2
4	**Stock2**		1	4	4.6906		0.0303	Stock2

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.5304	0.2259	45.8989	<.0001
Yield Curve2	1	-0.2885	0.0807	12.7616	0.0004
Lead2	1	-2.0511	0.2371	74.8164	<.0001
Stock2	1	-0.0357	0.0175	4.1737	0.0411
FF Rate2	1	0.0565	0.0242	5.4372	0.0197

Association of Predicted Probabilities and Observed Responses

Percent Concordant	91.9	**Somers' D**	0.840
Percent Discordant	7.9	**Gamma**	0.841
Percent Tied	0.1	**Tau-a**	0.235

Association of Predicted Probabilities and Observed Responses

Pairs	51510	**c**	0.920

Partition for the Hosmer and Lemeshow Test

		NBER Recession = 1		NBER Recession = 0	
Group	**Total**	**Observed**	**Expected**	**Observed**	**Expected**
1	62	0	0.01	62	61.99
2	61	0	0.18	61	60.82
3	61	0	0.64	61	60.36
4	61	0	1.38	61	59.62
5	61	2	2.21	59	58.79
6	61	5	3.80	56	57.20
7	61	7	6.77	54	54.23
8	61	16	14.10	45	46.90
9	61	29	27.43	32	33.57
10	57	43	43.99	14	13.01

Hosmer and Lemeshow Goodness-of-Fit Test

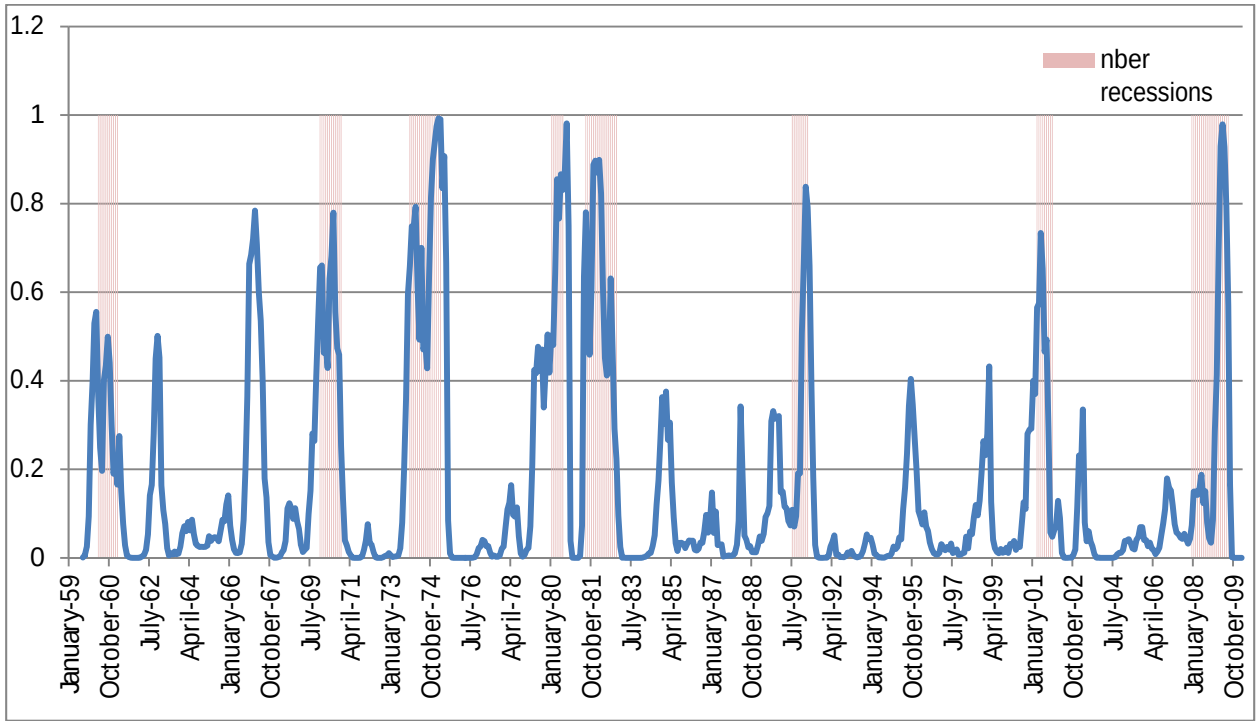
Chi-Square	**DF**	**Pr > ChiSq**
3.2825	8	0.9154

Generated by the SAS System (Local, XP_PRO) on 24APR2010 at 6:16 PM

The variable Money is not entered into the model.
So the final model which can be used for predictions is:

$$\Pr(R_t = 1) = \Phi(c_0 + c_1 Spread_{t-k} + c_2 Lead_{t-k} + c_3 Stock_{t-k} + c_4 FF_{t-k})$$

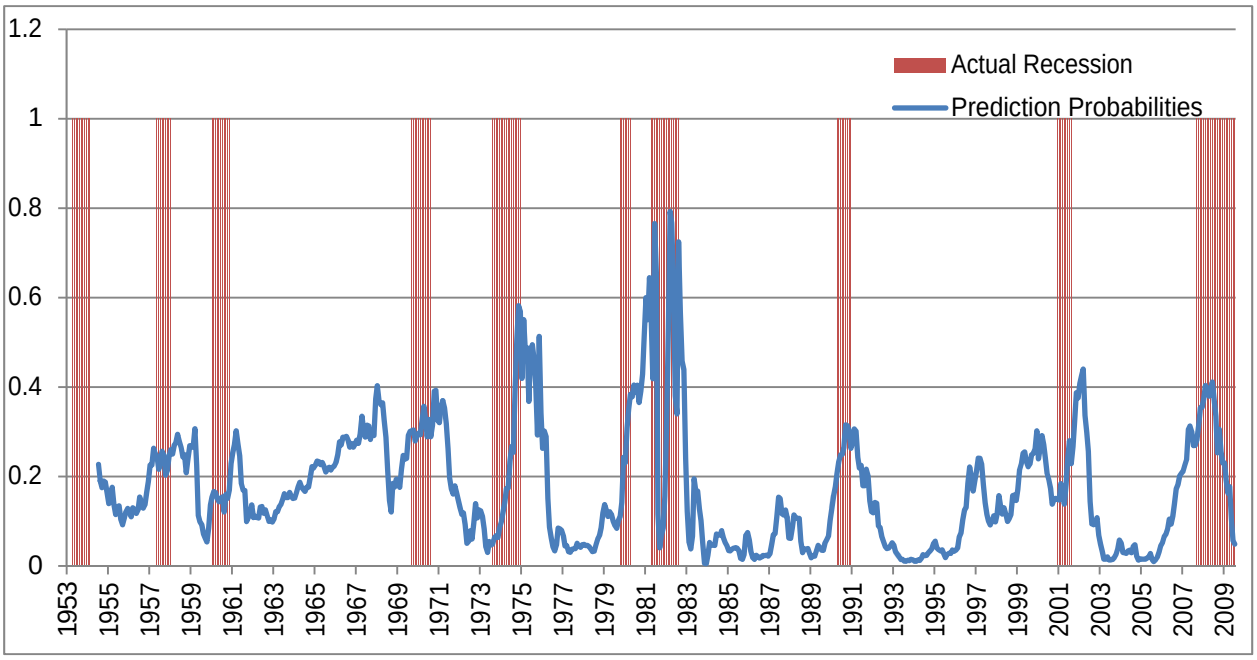
Step 8: Model Predictions



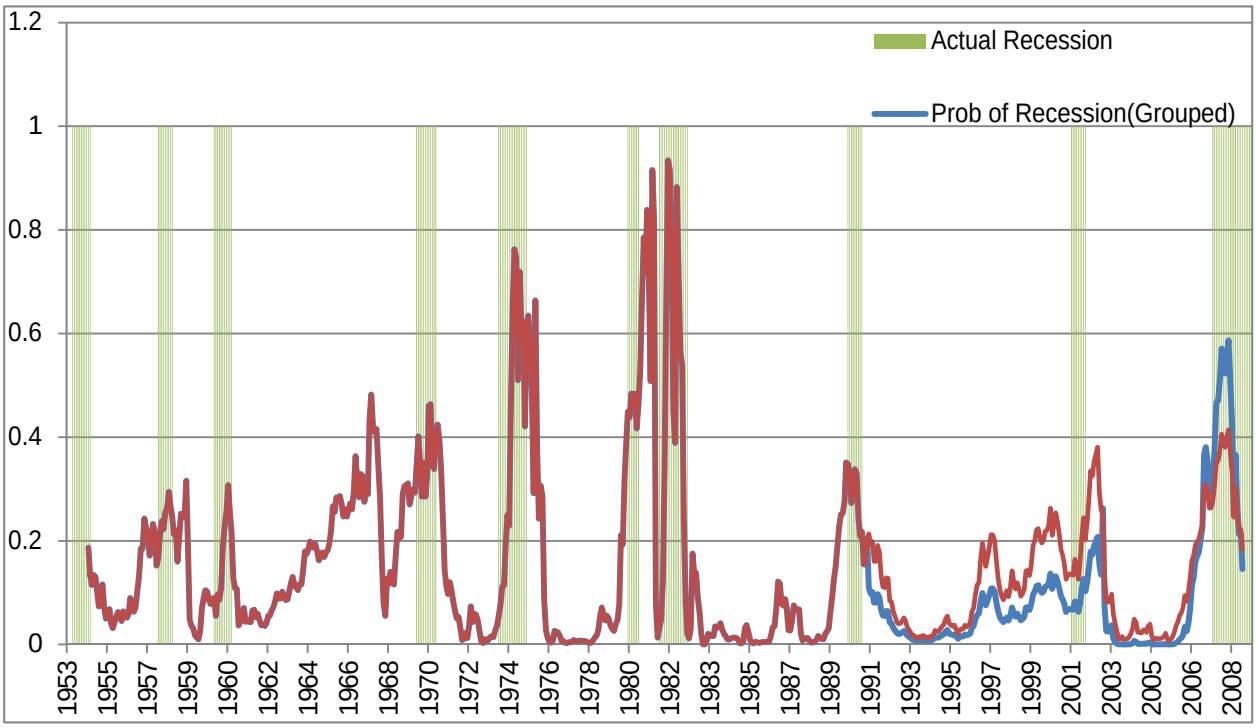
The predicted probabilities of recessions for the final four variable model is plotted above.
It is evident that the probabilities of recession have increased giving a much result fit than the previous models.
A brief comparison of all the three models studied is shown and tabulated in the following pages.

Improvement in Result

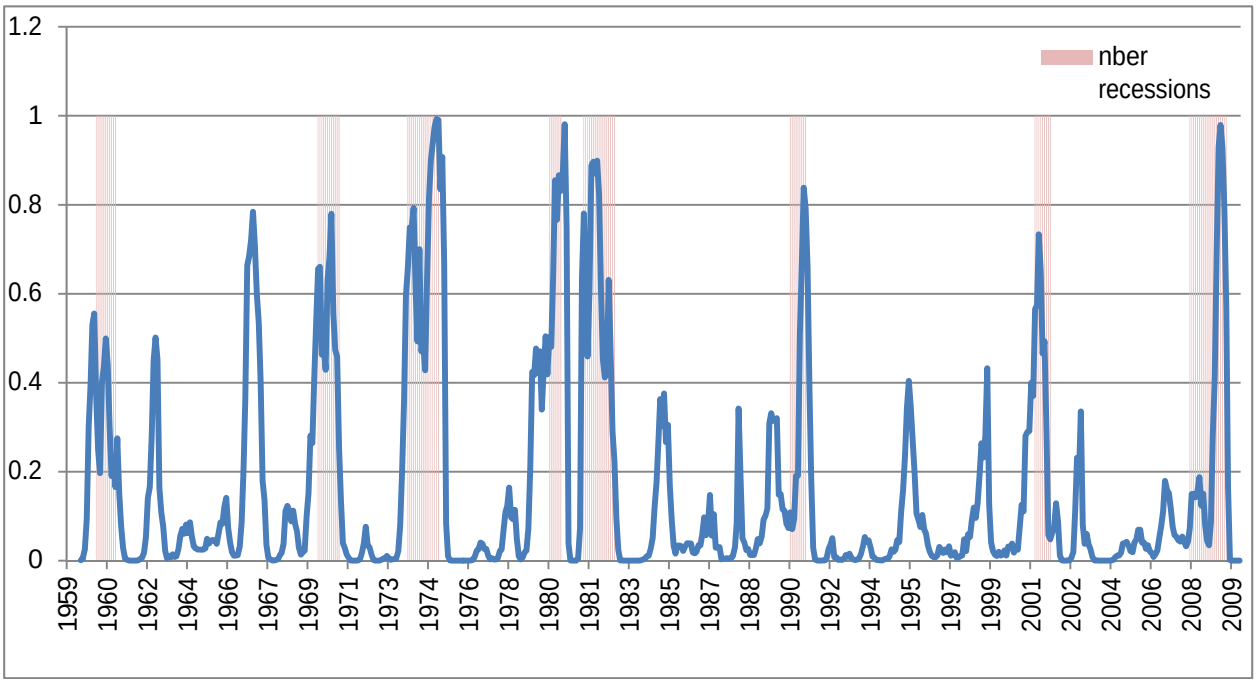
Basic Model



Grouped Model



Final Model



DATE	Model 1	Model 2	Model 3
Aug 1957 –April 1958	26.3	31.57	--
Apr 1960 –Feb 1961	25.42	30.74	55.56
Dec 1969 –Nov 1970	35.67	46.04	77.94
Nov 1973 –Mar 1975	58.21	74.54	97.47
Jan–July 1980	35.09	52.66	98.1
July 1981 – Nov 1982	79.26	91.53	89.89
July 1990 – Mar 1991	26.53	35.14	79.16
Mar – Nov 2001	21.19	21.19	73.36
Dec 2007 – Current	37.42	57.08	97.9

REFERENCES

1. Hardeep Bamara -"*The Predictive Power of Yield Curve in Forecasting U.S. Recessions*", A Dissertation for MBA, Simon Fraser University Library, Spring 2006.
2. Jonathan H. Wright -*"The Yield Curve and Predicting Recessions"* – Finance and Economics Discussion Series, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington, D.C., May 2003.
3. John Slvia, Sam Bullard and Huiwen Lai -*"Forecasting U.S. Recessions with Probit Stepwise Regression Models"*– Journal of Business Economics, January 2008.
4. Michael Dotsey "*The Predictive Content of the Interest Rate Term Spread for Future Economic Growth*", Federal Reserve Bank of Richmond Economic Quarterly, Volume 84/3, Summer 1998.
5. www.stloiusfed.org, Federal Reserve Bank of St Louis
6. www.nber.org, National Bureau of Economic Research
7. www.ustreas.gov, United States Department of Treasury
8. www.wikipedia.com
9. www.investopedia.com
10. www.recession.org