University of Illinois
Spring 2006

Department of Economics
Roger Koenker

Economics 471: Applied Econometrics
Problem Set 1

> October. This is one of the peculiarly dangerous months to speculate in stocks in. The others are July, January, September, April, November, May, March, June, December, August and February. [Mark Twain, Pudd'nhead Wilson's Calendar (1900)]

Due date: Thursday, February 2.

The software link of the class website contains a small R program designed to download daily data on US stock prices from the Yahoo finance website. The program then transforms this price data, which has been adjusted for dividends, into monthly returns. And in a final step the returns series are transformed into "excess returns" be subtracting off returns on a risk-free asset chosen to be the 13 week US T-bill return. By default the series begin in January 1998, but this could be altered, if you would like to experiment.

Useage of the function is, I hope, quite simple. You begin by pasting the function into your R session, or reading in the file using the command

```
source("capm.R")
```

verify that it is available by simply typing the name of the function, `capmRead` and you should see the text of the function on your screen. Now you need to select a small number, 3-5, of stocks for analysis and find their stock exchange symbols. This is quite conveniently done on the Yahoo finance webpage `http://finance.yahoo.com/lookup` or browsing the link indicated as "stocks by industry" given on that page.

To illustrate, suppose that we have selected IBM, KO, and GE. We can then download the data for these companies by typing:

```
d <- capmRead(c("IBM","KO","GE"))
```

This command should send an inquiry off to Yahoo and return with an array of excess return data for the three companies as well as data on the excess returns of holding the market portfolio, which for our purposes will be defined as the Standard and Poor's 500 index. You can see the first 10 observations of the array by simply typing, `d[1:10,]`. To begin to do something constructive with the data it is convenient to do:

```
attach(d)
```

This has the effect of putting the variables included in `d` into the current search path of R, so that you can now reference them by their original symbols. Thus, if you would like to plot IBM excess returns vs. the market excess returns, you can do,

```
plot(market,IBM)
```

and superimposing the classical CAPM regression line on the plot can be done by,

```
f <- lm(IBM ~ market)
abline(f)
summary(f)
```

The capital asset pricing model suggests that there should be a linear relationship between the risk premium on a stock j, that is the difference between stock j's return and the risk free rate, and the overall market risk premium, that is

$$r_{jt} - r_{ft} = \alpha + \beta(r_{mt} - r_{ft}) + u_{jt} \tag{1}$$

where r_{jt}, is the return of the jth firm at time t, r_{ft} is the risk free return at time t, and r_{mt} is the market return at time t. Note that these returns are monthly, so if you compute a mean monthly return for 12 months you can approximate the annual return as

$$R = (1 + \bar{r})^{12} - 1.$$

For example, if $\bar{r}$ is .0049 then R is about .06, or 6 percent per year.

1. Choose 4 firms and estimate the model (1) given above based on the full sample. Try to choose two firms that would expect a priori to be risky and two others you expect to be relatively safe. Do the results of the regressions support your original viewpoint? Explain why, or why not.

2. Try plotting the the firm risk premium and the market risk premium first as superimposed time series, then as a scatter plot to see whether there are anomalies that you think would be problematic.

3. Test the hypothesis that $\alpha = 0$ vs. $\alpha > 0$ for each of the firms. Interpret the results of these tests economically.

4. Construct a 95 percent confidence interval for the parameter β for each of the firms. Interpret the intervals emphasizing whether they include the value one.

5. The R^2 of your regressions is usually interpreted, following Sharpe, to be the proportion of the total risk of the firm that can be attributed to the market, and $1 - R^2$ as the proportion of risk that is undiversifiable. How do these values correspond to the $\hat{\beta}$'s that you computed?

6. On the class website in the data link there is a file called `gold.d` that contains data the return of gold and an associated market and risk free rate for the period January 1970 through 2005. Try estimating the CAPM model for gold for several subperiods and see whether you can find a period for which the β coefficient is negative. This procedure is intended to show that assuming that β is fixed over time is often a questionable. Compute a 95 percent confidence interval for β, for several of these periods and interpret these intervals and their variability.