TrainExer31

This is a stock market data set for the United States for 1927-2013 (yearly data). The source of the data is the updated version of the Goyal and Welch (2008)[1] data. The data are available from the website of Prof Amit Goyal, http://www.hec.unil.ch/agoyal/, where the updated data through 2013 (which are used in the lecture) are downloadable from http://www.hec.unil.ch/agoyal/docs/PredictorData2013.xlsx. Particularly the application in lecture 3.5 follows the aforementioned paper.

The variables are:

- Year
- Index: The S&P500 index
- Dividends: Dividends on the index ("D12" in the Goyal and Welch [GW] file)
- Riskfree: Riskfree rate ("Rfree" in GW)
- LogEqPrem: Log of the equity premium (calculated following GW[2])
- BookMarket: Book to market ratio ("b/m" in GW)
- NTIS: Equity issued ("ntis" in GW)
- DivPrice: Dividend to price ratio (calculated following GW[3])
- EarnPrice: Earnings to price ratio (calculated following GW[4])
- Inflation: Inflation rate ("infl" in GW)

[1] "A comprehensive look at the empirical performance of equity premium prediction", Review of Financial Studies 21(4), p1455-1508.
[2] Calculated as: log((Index+D12)/Index(-1))-log(1+Rfree), where x(-1) denotes value from previous period, log is the natural logarithm, D12 dividends and Rfree the riskfree rate.
[3] Calculated as: log(D12)-log(Index), where D12 are dividends.
[4] Calculated as: log(E12)/log(Index), where E12 are earnings.