Solution
Assignment 1

In the recent finance literatures, it is suggested that asset prices are well described by a so-called factor model. Excess returns are linearly explained from excess returns on a number of factor portfolios. According to CAPM, the intercept term should be zero, just like the coefficient for any other variable included in the model the value of which is known in advance (e.g. a January Dummy). The data set ***project1.dta*** contains excess returns on three-factor portfolio model. Extended from CAPM, the three-factors model includes two additional variables into the model. The portfolios are re-formed each month on the basis of the most recent available information on firm size and book-to-market value of equity. In addition to risk premium of market portfolio (r_m), the r_{smb} factor is based on the firm size and reflects the difference in returns between a portfolio of stocks with a small firm size and a portfolio of stock with a big firm size. The r_{hml} factor classifies the stocks based on the ratio of book value to market value of equity and computes the difference in returns between a portfolio of stocks with a high book-to-market ratio (value stocks) and a portfolio of stocks with a low book-to-market ratio (growth stocks).

CAPM: $$r_{jt} = \alpha_j + \beta_{j1} r_{mt} + \varepsilon_{jt} \quad (1)$$

Fama- French: $$r_{jt} = \alpha_j + \beta_{j1} r_{mt} + \beta_{j2} r_{smbt} + \beta_{j3} r_{hmlt} + \varepsilon_{jt} \quad (2)$$

Where:
- r_{jt} = excess return on portfolio *j* at time *t* and
- *j* = 1, 2, …., 20 portfolios categorized by four groups of firm size and five groups of value of the stock.
- r_{mt} = excess return on market portfolio at time *t* – representing market risk premium.
- r_{smbt} = return on a small-stock portfolio minus the return on a large-stock portfolio (Small Minus Big) at time *t* – representing size premium.
- r_{hmlt} = return on a value-stock portfolio minus the return on a growth-stock portfolio (High Minus Low) at time *t* – representing value premium.

(1) Regress the excess returns on portfolio upon the excess return on the market portfolio with constant term, noting that this corresponds to the CAPM (model (1)).

```
. reg rj r_mf

      Source |       SS           df       MS      Number of obs   =     1,048
-------------+----------------------------------   F(1, 1046)      =  17487.59
       Model |  37136.0054         1  37136.0054   Prob > F        =    0.0000
    Residual |  2221.24787     1,046  2.12356393   R-squared       =    0.9436
-------------+----------------------------------   Adj R-squared   =    0.9435
       Total |  39357.2533     1,047  37.5904998   Root MSE        =    1.4572

------------------------------------------------------------------------------
          rj |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        r_mf |   1.104191   .0083499   132.24   0.000     1.087806    1.120575
       _cons |   .3707874    .045301     8.18   0.000     .2818963    .4596786
------------------------------------------------------------------------------

. est store capm
```

(2) Make interpretation of the estimated β coefficients, t-test, F-test, R^2 and Adjusted-R^2.

(3) Perform hypothesis testing whether your estimated portfolio has the same risk as the market.

```
. test r_mf=1

 ( 1)  r_mf = 1

       F(  1,  1046) =  155.70
            Prob > F =    0.0000
```

(4) Test the validity of the CAPM by testing whether the constant terms in regression is zero. What are the differences between regressing the models with and without constant term.

```
     _cons |   .3707874    .045301     8.18   0.000     .2818963    .4596786
```

(5) Test whether there exists significant January effect in the regression models.

```
. g jan=(month==1)

. reg rj r_mf jan

      Source |       SS           df       MS      Number of obs   =     1,048
-------------+----------------------------------   F(2, 1045)      =   8755.66
       Model |  37140.8477         2  18570.4239   Prob > F        =    0.0000
    Residual |  2216.40555     1,045  2.12096225   R-squared       =    0.9437
-------------+----------------------------------   Adj R-squared   =    0.9436
       Total |  39357.2533     1,047  37.5904998   Root MSE        =    1.4564

------------------------------------------------------------------------------
          rj |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        r_mf |   1.103777   .0083492   132.20   0.000     1.087394     1.12016
         jan |   .2465016   .1631396     1.51   0.131    -.0736168    .5666201
       _cons |    .350576    .047208     7.43   0.000     .2579428    .4432092
------------------------------------------------------------------------------
```

(6) Regress the three-factor model (2) by using OLS. Compare the estimated results with the one-factor (CAPM) model (1).

```
. reg rj r_mf

      Source |       SS           df       MS      Number of obs   =     1,048
-------------+----------------------------------   F(1, 1046)      =  17487.59
       Model |  37136.0054         1  37136.0054   Prob > F        =    0.0000
    Residual |  2221.24787     1,046  2.12356393   R-squared       =    0.9436
-------------+----------------------------------   Adj R-squared   =    0.9435
       Total |  39357.2533     1,047  37.5904998   Root MSE        =    1.4572
------------------------------------------------------------------------------
          rj |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        r_mf |   1.104191   .0083499   132.24   0.000     1.087806    1.120575
       _cons |   .3707874    .045301     8.18   0.000     .2818963    .4596786
------------------------------------------------------------------------------

. est store capm

. reg rj r_mf r_smb r_hml

      Source |       SS           df       MS      Number of obs   =     1,048
-------------+----------------------------------   F(3, 1044)      =   8657.06
       Model |   37836.295         3  12612.0983   Prob > F        =    0.0000
    Residual |  1520.95827     1,044  1.45685658   R-squared       =    0.9614
-------------+----------------------------------   Adj R-squared   =    0.9612
       Total |  39357.2533     1,047  37.5904998   Root MSE        =     1.207

------------------------------------------------------------------------------
          rj |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
```

```
       r_mf |   1.042735   .0074665   139.66   0.000     1.028084    1.057386
      r_smb |   .2281417   .0120183    18.98   0.000     .2045589    .2517244
      r_hml |   .1124214    .010856    10.36   0.000     .0911193    .1337235
      _cons |    .307938   .0376921     8.17   0.000     .2339772    .3818989
------------------------------------------------------------------------------

. est store ff

. estimates table capm ff, stat(r2 F rss) star(0.1 0.05 0.01)

--------------------------------------------
    Variable |     capm            ff        
-------------+------------------------------
        r_mf |  1.1041908***     1.042735***  
       r_smb |                  .22814165***  
       r_hml |                  .11242138***  
       _cons |  .37078743***    .30793805***  
-------------+------------------------------
          r2 |  .94356192        .96135507    
           F |  17487.585        8657.0624    
         rss |  2221.2479        1520.9583    
--------------------------------------------
           legend: * p<.1; ** p<.05; *** p<.01
```

(7) Test whether there exists a significant first order autocorrelation in the estimated regression model. In case of autocorrelation, what are the consequences of the problem? And how can we solve the problem.

```
. reg rj r_mf r_smb r_hml

      Source |       SS           df       MS      Number of obs   =     1,048
-------------+----------------------------------   F(3, 1044)      =   8657.06
       Model |   37836.295         3  12612.0983   Prob > F        =    0.0000
    Residual |  1520.95827     1,044  1.45685658   R-squared       =    0.9614
-------------+----------------------------------   Adj R-squared   =    0.9612
       Total |  39357.2533     1,047  37.5904998   Root MSE        =     1.207

------------------------------------------------------------------------------
          rj |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        r_mf |   1.042735   .0074665   139.66   0.000     1.028084    1.057386
       r_smb |   .2281417   .0120183    18.98   0.000     .2045589    .2517244
       r_hml |   .1124214    .010856    10.36   0.000     .0911193    .1337235
       _cons |    .307938   .0376921     8.17   0.000     .2339772    .3818989
------------------------------------------------------------------------------

. estat dwatson

Durbin-Watson d-statistic(  4,  1048) =  1.829536

. prais rj r_mf r_smb r_hml, corc
Iteration 0:  rho = 0.0000
Iteration 1:  rho = 0.0837
Iteration 2:  rho = 0.0855
Iteration 3:  rho = 0.0855
Iteration 4:  rho = 0.0855
Cochrane-Orcutt AR(1) regression -- iterated estimates

      Source |       SS           df       MS      Number of obs   =     1,047
-------------+----------------------------------   F(3, 1043)      =   8573.85
       Model |  37240.7344         3  12413.5781   Prob > F        =    0.0000
    Residual |  1510.09858     1,043   1.4478414   R-squared       =    0.9610
-------------+----------------------------------   Adj R-squared   =    0.9609
       Total |   38750.833     1,046  37.0466855   Root MSE        =    1.2033

------------------------------------------------------------------------------
          rj |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        r_mf |   1.045358   .0074378   140.55   0.000     1.030764    1.059953
       r_smb |   .2268022   .0119303    19.01   0.000     .2033921    .2502122
       r_hml |   .1094813   .0109429    10.00   0.000     .0880086    .1309539
       _cons |   .3081707     .04105     7.51   0.000     .2276208    .3887206
-------------+----------------------------------------------------------------
         rho |    .085518
------------------------------------------------------------------------------
Durbin-Watson statistic (original)    1.829536
Durbin-Watson statistic (transformed) 2.004839
```

(8) Test whether your estimated portfolio has size premium.

```
r_smb |   .2268022   .0119303   19.01   0.000    .2033921    .2502122
```

(9) Test whether your estimated portfolio has value premium.

```
r_hml |   .1094813   .0109429   10.00   0.000    .0880086    .1309539
```

(10) Perform *F*-test for the hypothesis that the coefficients for the two new factors are jointly equal to zero.

```
. test r_smb r_hml

 ( 1)  r_smb = 0
 ( 2)  r_hml = 0

       F(  2,  1043) =  237.12
            Prob > F =    0.0000
```

(11) Perform Chow-test whether January and other month share the same structure of the Fama-French model (2).

```
. g jan_rmf=r_mf*jan

. g jan_rsmb=r_smb*jan

. g jan_rhml=r_hml*jan

. reg rj r_mf jan_rmf r_smb jan_rsmb r_hml jan_rhml jan

      Source |       SS           df       MS      Number of obs   =     1,048
-------------+----------------------------------   F(7, 1040)      =   3766.88
       Model |  37863.8479         7  5409.12114   Prob > F        =    0.0000
    Residual |  1493.40533     1,040  1.43596667   R-squared       =    0.9621
-------------+----------------------------------   Adj R-squared   =    0.9618
       Total |  39357.2533     1,047  37.5904998   Root MSE        =    1.1983

------------------------------------------------------------------------------
          rj |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        r_mf |   1.041944   .0076958   135.39   0.000     1.026843    1.057045
     jan_rmf |  -.0270275   .0301801    -0.90   0.371    -.0862484    .0321934
       r_smb |   .2427885   .0126738    19.16   0.000     .2179194    .2676576
    jan_rsmb |  -.0565731   .0484615    -1.17   0.243    -.1516666    .0385205
       r_hml |   .1238979   .0115231    10.75   0.000     .1012868    .1465091
    jan_rhml |  -.0379402   .0379196    -1.00   0.317    -.1123479    .0364674
         jan |  -.2750941   .1702533    -1.62   0.106    -.6091732    .0589849
       _cons |   .3439995   .0388991     8.84   0.000     .2676699     .420329
------------------------------------------------------------------------------

. test jan jan_rmf jan_rsmb jan_rhml

 ( 1)  jan = 0
 ( 2)  jan_rmf = 0
 ( 3)  jan_rsmb = 0
 ( 4)  jan_rhml = 0

       F(  4,  1040) =    4.80
            Prob > F =    0.0008
```

However, some other literatures have claimed that macroeconomic factors play important role in determining the return of the portfolio. Then, the model should be arbitrage pricing model as

APT: $$r_{jt} = \alpha_j + \beta_{j1} r_{mt} + \gamma_{j2} r_{\text{int}\, t} + \gamma_{j3} r_{fxt} + \gamma_{j4} r_{goldt} + \varepsilon_{jt} \qquad (3)$$

Where: r_{intt} = interest rate at time *t* – expected to have negative relationship.
r_{fxt} = change in exchange rate at time *t* – expected to have negative impacts.
r_{goldt} = change in gold price at time *t* – expected to have negative impacts.

(12) Regress the APT (model (3)) of your portfolio. Make interpretation of the estimated result.

```
. reg rj r_mf r_int r_fx r_gold

      Source |       SS           df       MS      Number of obs   =     1,048
-------------+----------------------------------   F(4, 1043)      =   4372.19
       Model |  37142.1595         4  9285.53988   Prob > F        =    0.0000
    Residual |  2215.09376     1,043  2.12377158   R-squared       =    0.9437
-------------+----------------------------------   Adj R-squared   =    0.9435
       Total |  39357.2533     1,047  37.5904998   Root MSE        =    1.4573

------------------------------------------------------------------------------
          rj |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        r_mf |   1.104281   .0083563   132.15   0.000     1.087884    1.120678
       r_int |   .1588066   .1384347     1.15   0.252    -.1128356    .4304488
        r_fx |  -.0210917   .0222047    -0.95   0.342    -.0646627    .0224793
      r_gold |  -.0031087   .0036353    -0.86   0.393     -.010242    .0040245
       _cons |    .063137    .293473     0.22   0.830    -.5127278    .6390017
------------------------------------------------------------------------------

. est store apm
```

(13) Whether there exists serious multicollinearity problem in model (3). Should any independent variables in model (3) be dropped? Give explanation of your decision.

```
. pwcorr rj r_mf r_int r_fx r_gold

             |       rj     r_mf    r_int     r_fx   r_gold
-------------+---------------------------------------------
          rj |   1.0000
        r_mf |   0.9714   1.0000
       r_int |  -0.0018  -0.0103   1.0000
        r_fx |  -0.0312  -0.0250   0.0082   1.0000
      r_gold |   0.0196   0.0265   0.0148  -0.0041   1.0000

. estat vif

    Variable |       VIF       1/VIF
-------------+----------------------
        r_mf |      1.00    0.998568
      r_gold |      1.00    0.999060
        r_fx |      1.00    0.999299
       r_int |      1.00    0.999602
-------------+----------------------
    Mean VIF |      1.00

. reg rj r_mf r_fx

      Source |       SS           df       MS      Number of obs   =     1,048
-------------+----------------------------------   F(2, 1045)      =   8743.23
       Model |  37137.8735         2  18568.9367   Prob > F        =    0.0000
    Residual |  2219.37982     1,045  2.12380844   R-squared       =    0.9436
-------------+----------------------------------   Adj R-squared   =    0.9435
       Total |  39357.2533     1,047  37.5904998   Root MSE        =    1.4573

------------------------------------------------------------------------------
          rj |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        r_mf |   1.103995    .008353   132.17   0.000     1.087604    1.120385
        r_fx |  -.0208242   .0222041    -0.94   0.349    -.0643939    .0227454
       _cons |   .3903649   .0498815     7.83   0.000     .2924855    .4882442
------------------------------------------------------------------------------

. est store apm2

. reg rj r_mf
```

```
      Source |       SS           df       MS      Number of obs   =     1,048
-------------+----------------------------------   F(1, 1046)      =  17487.59
       Model |  37136.0054         1  37136.0054   Prob > F        =    0.0000
    Residual |  2221.24787     1,046  2.12356393   R-squared       =    0.9436
-------------+----------------------------------   Adj R-squared   =    0.9435
       Total |  39357.2533     1,047  37.5904998   Root MSE        =    1.4572

------------------------------------------------------------------------------
          rj |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        r_mf |   1.104191   .0083499   132.24   0.000     1.087806    1.120575
       _cons |   .3707874    .045301     8.18   0.000     .2818963    .4596786
------------------------------------------------------------------------------

. est store capm

. est table apm apm2 capm, stat(r2 F rss) star(0.1 0.05 0.01)

--------------------------------------------------------------
    Variable |      apm            apm2            capm
-------------+------------------------------------------------
        r_mf |  1.1042805***    1.1039949***    1.1041908***
       r_int |   .1588066
        r_fx |  -.0210917      -.02082424
      r_gold | -.00310872
       _cons |  .06313695       .39036486***    .37078743***
-------------+------------------------------------------------
          r2 |  .94371828       .94360938       .94356192
           F |  4372.1933       8743.2258       17487.585
         rss |  2215.0938       2219.3798       2221.2479
--------------------------------------------------------------
                              legend: * p<.1; ** p<.05; *** p<.01
```

(14) Which of the following hypotheses about the coefficients can be tested using a *t*-test? Which of them can be tested using an *F*-test? In each case, state the number of restrictions.

(i) $H_0 : \beta_1 = 1$

(ii) $H_0 : \gamma_2 + \gamma_3 = 1$

(iii) $H_0 : \gamma_2 + \gamma_3 = 1 \; and \; \gamma_4 = 1$

(iv) $H_0 : \beta_1 = 0 \; and \; \gamma_2 = 0 \; and \; \gamma_3 = 0 \; and \; \gamma_4 = 0$

(v) $H_0 : \gamma_2 \gamma_3 = 1$

(15) In question (14) which would you expect to be bigger – the unrestricted residual sum of squares or the restricted sum of squares, and why?

(16) From CAPM model (1), what would be the conceptually differences if we (a) regress the excess returns on portfolio upon the excess return on the market portfolio with constant term or (b) regress the excess return on the market portfolio upon the excess returns on portfolio with constant term? Make comparison of (a) and (b).

```
. reg rj r_mf

      Source |       SS           df       MS      Number of obs   =     1,048
-------------+----------------------------------   F(1, 1046)      =  17487.59
       Model |  37136.0054         1  37136.0054   Prob > F        =    0.0000
    Residual |  2221.24787     1,046  2.12356393   R-squared       =    0.9436
-------------+----------------------------------   Adj R-squared   =    0.9435
       Total |  39357.2533     1,047  37.5904998   Root MSE        =    1.4572
------------------------------------------------------------------------------
          rj |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        r_mf |   1.104191   .0083499   132.24   0.000     1.087806    1.120575
       _cons |   .3707874    .045301     8.18   0.000     .2818963    .4596786
------------------------------------------------------------------------------

. reg r_mf rj

      Source |       SS           df       MS      Number of obs   =     1,048
-------------+----------------------------------   F(1, 1046)      =  17487.59
```

```
       Model |  28739.3768         1  28739.3768   Prob > F        =    0.0000
    Residual |   1719.0131     1,046  1.64341596   R-squared       =    0.9436
-------------+----------------------------------   Adj R-squared   =    0.9435
       Total |  30458.3899     1,047  29.0911079   Root MSE        =     1.282

------------------------------------------------------------------------------
        r_mf |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          rj |    .854528   .0064619   132.24   0.000     .8418482    .8672078
       _cons |  -.2824635   .0401698    -7.03   0.000    -.3612861   -.2036409
------------------------------------------------------------------------------
```