Stat 107: Quantitative Methods for Economics

Homework 8: Due Monday, December 5

Homework policy

Homework is due by 11:59pm of the assigned date. Homework is to be handed in electronically via the Canvas system. Late homework will not be accepted.

You are encouraged to discuss homework problems with other students (and with the instructor and TFs, of course), but you must write your final answer in your own words. Solutions prepared "in committee" or by copying someone else's paper are not acceptable.

Note: All R code needed for this assignment is in file mypairspack.txt
Cut and paste this code into R once-see the file to read what the different functions are.

The purpose of this problem is to test two pairs trading strategy. There are many methods to perform pairs trading; for this exercise we will use a method detailed in the paper *Evaluation of Pairs Trading Strategy at the Brazilian Financial Market,* available on the course web site, and the method used by pairslog.com (which unfortunately is currently down).

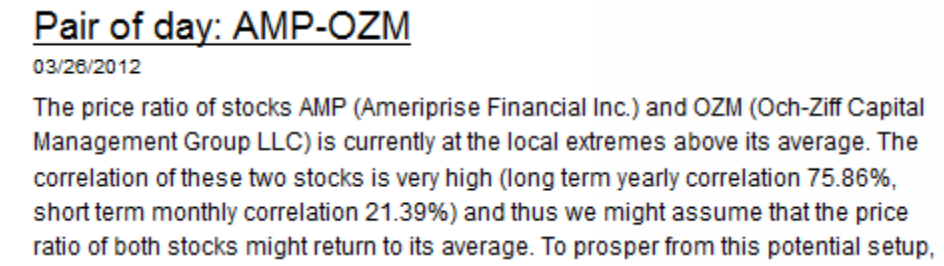
The most important parts of the Brazilian paper are on pages 6 and 7. The trading idea in the paper is as follows.

1) Choose two stocks that you think are related to each other
2) Normalize both stock prices
3) Trade when the absolute value of difference of the normalized value is larger than some cut-off value d (see his Table 2 for values of d).
4) Close out the trades when the difference crosses 0.

For the website pairslog.com they do the following.

1) Choose two stocks that you think are related to each other
2) Calculate the normalized ratio of closing prices.
3) Trade when the normalized value is larger than 2.
4) Close out the trades when the ratio crosses 0.
5) Note that if you try to match their numbers exactly, pairslog.com divides by the standard deviation for the population, so they divide by n instead of n-1 in their standard deviation formula.

Note that there are many variations of these two techniques.
Let's now implement this in R. From pairslog.com we find that AMP/OZM is potentially a good pair of stocks to trade (look the symbols up on Yahoo so you know what there are):



The website pairslog.com uses a 14 day window to calculate statistics such as means and variance so will use that in our calculations.

The following R code shows how to implement the two methods

Brazilian Method

```
stock1="AMP"
stock2="OZM"
s1=getSymbols(stock1,from = "2010-01-
01",auto.assign=FALSE)
s2=getSymbols(stock2,from = "2010-01-
01",auto.assign=FALSE)
p1 = Ad(s1)
p2 = Ad(s2)
np1 = (p1-runMean(p1,n=14))/runSD(p1,14)
np2 = (p2-runMean(p2,n=14))/runSD(p2,14)
```

Now that we have the normalized date, we form the difference.

```
ndiff=(np1-np2)
```

We now are ready to implement the pairs trading strategy. When the difference gets too far away from its mean value of 0, a pairs trade is entered as follows:

- If diff < -2, we go long stock1 and short stock2 in equal dollar amounts
- If diff > 2, we short stock1 and go long stock2 in equal dollar amounts
- The trade is closed when the absolute value of the difference crosses zero.

All this would be calculated with end of day prices, or in practice, with prices obtained just before the close of the market.

Here are some example trades:

```
           AMP.Close OZM.Close    ndiff
2012-03-08     55.32      9.25  0.63672607
2012-03-09     56.08      9.41  1.07945215
2012-03-12     55.85      9.25  1.23936695
2012-03-13     57.60      9.44  2.31156372
2012-03-14     56.75      9.42  1.21812781
2012-03-15     58.02      9.49  1.73595062
2012-03-16     57.72      9.67  0.08564145
2012-03-19     58.09      9.69  0.08201246
2012-03-20     57.27      9.67 -0.40673004
2012-03-21     57.39      9.69 -0.44928937
2012-03-22     56.49      9.48 -0.11995505
2012-03-23     57.10      9.46  0.35458872
2012-03-26     58.25      9.33  2.15412855
2012-03-27     57.70      9.24  2.00528559
2012-03-28     57.82      9.15  2.43452438
```

All trades are at the close of market (end of day trades they are called).

On 3/13 we would short AMP at 57.60 and go long OZM at 9.44
On 3/20 we would cover AMP at 57.27 and sell OZM at 9.67

On 3/26 we would short AMP at 58.25 and go long OZM at 9.33

Pairslog.com Method

```
stock1="AMP"
stock2="OZM"
s1=getSymbols(stock1,from = "2010-01-01",auto.assign=FALSE)
s2=getSymbols(stock2,from = "2010-01-01",auto.assign=FALSE)
rat=Cl(s1)/Cl(s2)
nrat=(rat-runMean(rat,14))/runSD(rat,14)
```

We now are ready to implement the pairs trading strategy. When the normalized ratio gets too far away from its mean value of 0, a pairs trade is entered as follows:

- If nrat < -2, we go long stock1 and short stock2 in equal dollar amounts
- If nrat > 2, we short stock1 and go long stock2 in equal dollar amounts
- The trade is closed when the normalized ratio zero.

All this would be calculated with end of day prices, or in practice, with prices obtained just before the close of the market.

Here are some example trades:

	AMP.Close	OZM.Close	ndiff
2012-03-05	55.54	9.53	0.5865
2012-03-06	54.17	9.02	2.2101
2012-03-07	54.55	9.38	0.2194
2012-03-08	55.32	9.25	1.5825
2012-03-09	56.08	9.41	1.1919
2012-03-12	55.85	9.25	1.6408
2012-03-13	57.60	9.44	1.8302
2012-03-14	56.75	9.42	1.0439
2012-03-15	58.02	9.49	1.5472
2012-03-16	57.72	9.67	0.1981
2012-03-19	58.09	9.69	0.3532
2012-03-20	57.27	9.67	-0.4018
2012-03-21	57.39	9.69	-0.4996
2012-03-22	56.49	9.48	-0.1749
2012-03-23	57.10	9.46	0.6187
2012-03-26	58.25	9.33	2.3221
2012-03-27	57.70	9.24	1.9796
2012-03-28	57.82	9.15	2.0226

All trades are at the close of market (end of day trades they are called).

On 3/6 we would short AMP at 54.17 and go long OZM at 9.02
On 3/20 we would cover AMP at 57.27 and sell OZM at 9.67
We lost 5.4% on AMP but made 7.02% on OZM

On 3/26 we would short AMP at 58.25 and go long OZM at 9.33

There are four separate deliverables for this problem.

1. Run the two pairs trading strategies on BKEP/DSX from October 2013 to present. To make the coding easier, we did the coding for you! The R function is pairs.trade in the file mypairspack.txt. You will have to slightly modify the code to keep trade of the individual trades though.

    a) Using the last 180 days of data, what is the correlation between the prices of these two securities and are the two securities co-integrated? (report the appropriate p-value). The R function myvals in the supplied R code file will be useful for this.
    b) How much money total (today) would someone have made if they invested $10000 in each short and long position each time ? That is, what was the final profit, assuming they could invest $10,000 each time? We are assuming not compounding our growth but investing the same fixed amount each time. Note that it is possible your last trade is not closed out by the end of our time period- decide how to handle that if needed. The default in the code is to close out the trade.
    c) How many pair trades triggered? How many were profitable?
    d) What was the best pair trade? The worst pair trade?
    e) What was the average of the losing trades and average of the winning trades?
    f) Finally, which pairs trading method would you prefer? Explain.

2. Repeat the steps above with the pair FAS and FAZ. These are ETFs that are designed to have a correlation of -1. and JNJ. See more information about pairs trading ETFs here: http://etfdb.com/etf-trading-strategies/how-to-use-a-pairs-trading-strategy-with-etfs/

3. Repeat the steps above with the pair DPW and EVM. The results should be pretty good for this pair. Look up DPW on finance.yahoo.com. Does it look like a stock that would be easy to trade in a pairs trading strategy (that is, could you short it, or purchase it in volume?).

4. Repeat the steps above with the randomly chosen pair BECN and GOV. This pair is not highly correlated and hence not a good candidate for pairs trading. What do the trading results look like for a pair that is probably not mean reverting? Report the same items as above.

Technical Note for those curious.

Pairslog.com reports 1-pvalue from the adf.test when they talk about cointegration. So for them the higher their "cointegration value" the better. A bit weird.

As an example, here is data from their website. Under the CI column is their cointegration value.

Dates calculated for 04/02/2015, unifo

Pair	Last 1	Last 2	Corr.	Corr. 30	Corr. 180	CI▼	Ratio	%mean	Delta	#	APT	#P
LMT - MOG-A Aerospace/De..	198.72	73.98	0.02	0.86	0.43	0.99	2.7	-0	-0.1	39	-50.4	21
NEWP - TRMB Scientific &..	19.1	24.68	0.37	0.71	0.43	0.99	0.8	-0.2	-0.2	35	28.2	19
HNT - UNH Health Care ..	60.17	117.36	0.66	0.92	0.8	0.99	0.5	0.8	0.7	39	-59.9	24
NVDA - SLAB Semiconducto..	21.06	50.9	0.87	0.47	0.89	0.99	0.4	-4.2	-1.3	42	166.2	33
DCT - SSS REIT - Indus..	34.98	93.95	0.47	0.55	0.73	0.99	0.4	0.8	0.8	37	808.6	25
DCT - PSA REIT - Indus..	34.98	197.31	0.35	0.81	0.55	0.99	0.2	-0.1	-0.1	34	925.2	23

We will examine the pair NEWP/TRMB. Their CI value is "0.99".

```
> getSymbols("TRMB",from="2014-04-01")
[1] "TRMB"
> lm(Cl(NEWP)~-1+Cl(TRMB))

Call:
lm(formula = Cl(NEWP) ~ -1 + Cl(TRMB))

Coefficients:
Cl(TRMB)
  0.5866

> diff=Cl(NEWP)-.5866*Cl(TRMB)
> adf.test(diff)

        Augmented Dickey-Fuller Test

data:  diff
Dickey-Fuller = -3.7923, Lag order = 6, p-value = 0.01988
alternative hypothesis: stationary
```

Our time period is slightly different but we would report (suing their method) 1-0.01988 = 0.98. But you get the idea.

Note-the adf.test is from library(tseries).