Problem Set 8

Part I reading questions

Brandt and Kavajecz

1.

    (a) How do Brandt and Kavajecz measure "orderflow?" You see a trade; how do you know if it's a "buy" or a "sell?" ("The market went up on a wave of buying" is a classic *fallacy* – for every buy, someone sold!) (Hint: see p. 2627- 2628)

    (b) Table IV: What does the number -0.72 in the top left corner of table IV mean? (This is a question about data and units – if x moves by what, what happens to y?)

    (c) Table IV: Which order flow is most important for movement in each yield?

2. In one sentence each, distinguish the "price discovery" vs "price impact" or "inventory" views of the correlation between order flow and price change.

3. Table VI. What pattern do you expect here if the phenomenon is simple "price pressure," demand for illiquid securities having price impact?

4. Table VII Same question.

  Krilenko et al. Don't pay a lot of attention to the regressions, but let's understand the basic facts and story.

1. Look at Figure 1 and 2. p.9, what according to Krilenko et al was the major spark? What event at least coincided with the end of the crash and quick recovery?

2. p. 16-19, Figure 4-5 and 6-7. Looking at profits together with net positions, do HFTs and intermediaries seem to be helping or hurting? (Note: to me the words and the figures do not match up. Get ready to discuss in class.)

3. p.20, Figure 10-11. How does the behavior of "opportunistic traders"contrast with those of HFT and intermediaries?

4. p.21. In their interpretation of regressions, how to Krilenko et al characterize HFT strategies?

    (p.35 discussion is their summary story. Connecting this to evidence is harder. The big question: do HFTs act as "market makers" and "increase liquidity?")

  Budish et al

1. How correlated are changes in the E-mini S&P 500 future (ES) and SPDR S&P 500 ETF (SPY)?

2. p.4. Do "snipers" add or subtract liquidity?

3. p.16. Has the pattern of correlation dropoff at high frequency changed over time, or has it remained constant?

4. To make money, we can't trade at the midpoints of Figure 1.1; we need a price spread that exceeds the bid-ask spread. Around p.20, Budish et al look at such arbitrage opportunities. The opportunity also has to last long enough that a signal can reach from New York to Chicago. What minimum time are they using here?

5. p. 22. Does New York seem to lag Chicago or vice versa?

6. p. 25. Has greater competition led to arbitrage opportunities that last for shorter time periods, or has it made arbitrage opportunities smaller, or both?

Duffie. Big picture. You know how a run on bank deposits works. Here we see how "shadow banking" contracts of broker dealers are also prone to runs. p. 54-60 are excellent on what dealer banks *do*. Bankruptcy happens when you run out of cash. Full stop. Forget fancy theories and present values! Runs happen if securities promise cash on demand, first come first serve, if the bank can't easily raise cash to satisfy demands, and if my getting my cash out makes the bank closer to failure so you have an incentive to get your cash out. Look for these characteristics in dealer banks. Obviously important to operating in markets.

1. p. 52. Alpha lost money, so is undercapitalized. Why not just sell more equity?

2. p. 59. What advantage is there, as a client, to investing in a hedge fund owned by a large dealer bank, rather than a standalone. After all, standalone management needs to work for you to keep their jobs!

3. p.61. Suppose you lend to a dealer bank, taking securities in repo as collateral. Why should you run if you suspect trouble? After all you have the collateral

4. Why not just borrow from the Fed?

5. p. 63. If your dealer is in trouble, why remove securities from the brokerage account? They're your securities after all. If your car dealer goes bankrupt, and your car is in for repair, you can go get it out even after bankruptcy.

6. p. 64 If you take your securities out, why does that make the dealer less liquid and more likely to go under

7. p. 66. If a dealer is downgraded, this alone drains cash. Why?

Gorton Why did the run bring down the system, not just a few banks?

1. p. 165. If everyone wants their money back from *every bank* at the same time, what must happen?

2. p. 167. The crisis originated in subprime. So why are other bond prices collapsing?

3. p. 169 Why do institutional investors favor repo over bank deposits?

4. p. 171 Central issues. How is the repo haircut like the money multiplier?

   p. 171. Read the e-coli analogy. For class, how is repo like e-coli?

Part II Hedge Fund data

This problem starts an exploration of the hedge fund index data used by Asness et al.

Get the hedge fund index returns from the class website. Ignore the last column of multi-strategy; it is missing a few data points and adapting the program to handle that is not worth the bother. Also get the factor returns.

1. If there is stale pricing, we should see returns that are correlated over time. Start by measuring the first-order correlation of the hedge fund excess returns by running $R^e_{t+1} = a + bR^e_t + \varepsilon_{t+1}$ Report b and the t statistic for each fund style. Compare to the same results for factors rmrf hml smb and umd. (The regression coefficient here is the same as the first order autocorrelation.) Do the hedge fund indices look different from the factors here? Is there a pattern to which funds have the worst (best if you could trade it!) correlation? (Note: my default factor is created by the difference in yield not a true rate of return. You will see how it doesn't work as it should here.)

2. Now, let's replicate Asness et al. Start by seeing how close we get to their results in their sample, up to September 2000.

   (a) Calculate the average return and Sharpe ratio of the overall hedge fund index, and compare it to the market average return and Sharpe. Are the means statistically significant?

   (b) But is it alpha or beta? Run the overall hedge fund index on the market return $R^e_t = \alpha + \beta \times rmrf_t + \varepsilon_t$ Report α, β. Compare the average return $E(R^e)$ with the alpha α, and check that $\beta \times E(rmrf)$ accounts for the reduction (increase?) in α. Is the alpha statistically significant? Is the alpha sharpe ratio $(\alpha/\sigma(\varepsilon))$ higher or lower than the raw Sharpe ratio? (A smaller alpha with a much smaller $\sigma(\varepsilon)$ can be a great opportunity!)

   (c) Run the hedge find index on current plus three lags of the market return and add up the coefficients,
   $$R^e_t = \alpha + \beta_0 R^{em}_t + \beta_1 R^{em}_{t-1} + \beta_2 R^{em}_{t-2} + \beta_3 R^{em}_{t-3} + \varepsilon_t.$$
   Report α and $\beta_0 + \beta_1 + \beta_2 + \beta_3$. Is this beta larger, and the alpha correspondingly smaller, as Asness found?

   (d) Cumulate both market and hedge fund return, and run the regression at an annual horizon. Do annual betas look more like monthly betas or like the betas with lags? Include the annual average return and annual alpha. (See note on annualizing returns in the programming hints.)

   (e) Back to monthly data, run the hedge fund index on "up and down" market movements separately. Create $R^{em+}_t = \{R^{em}_t \text{ if } R^{em}_t > 0, 0 \text{ if } R^{em}_t < 0\}$ and similarly R^{em-}_t. (Programming hints below) Then run
   $$R^e_t = \alpha + \beta_+ R^{em+}_t + \beta_- R^{em-}_t + \varepsilon_t.$$
   and report alpha and betas.

      i. Do you see a strong difference between up and down betas? Does it look like hedge funds have an index option component to their returns?
      ii. What do you make of the alphas in this regression?

   (f) Finally, mix these ideas. Run up/down betas with three lags.

   $$\begin{aligned} R^e_t = {}& \alpha + \beta_{0+} R^{em+}_t + \beta_{1+} R^{em+}_{t-1} + \beta_{2+} R^{em+}_{t-2} + \beta_{3+} R^{em+}_{t-3} \\ & + \beta_{0-} R^{em-}_t + \beta_{1-} R^{em-}_{t-1} + \beta_{2-} R^{em-}_{t-2} + \beta_{3-} R^{em-}_{t-3} + \varepsilon_t. \end{aligned}$$

      i. Report α, $\beta_+ \equiv \beta_{0+} + \beta_{1+} + \beta_{2+} + \beta_{3+}$ and $\beta_- \equiv \beta_{0-} + \beta_{1-} + \beta_{2-} + \beta_{3-}$
      ii. Report α, β_{0+}, $(\beta_{1+} + \beta_{2+} + \beta_{3+})$, β_{0-}, $(\beta_{1-} + \beta_{2-} + \beta_{3-})$

   Do lags help? Is there a difference between how much upside and downside lag there is? Does the pattern suggest "slow marking" or "illiquid assets?"

   (Note: you will get results similar to Asness but not exactly the same. You do not have to report standard errors, t stats. etc. However, you should look at them in all these regressions to get an idea of how precise our estimates are. I find standard errors of about 0.03-0.05 which means most of our conclusions about betas are statistically significant.)

3. Now repeat, using the post-Asness sample from 10/31/2000 to the end of your sample. How much of the analysis carries over to the latter period? If you find larger alphas, is it because the beta changed, or because the market premium changed? Does adding lags still matter? How do annual betas compare with monthly and monthly with three lags? Looking at betas and at R^2 are hedge funds more or less "alterative investments?" Is the up / down beta pattern still strong?

4. Now, let's look at the hedge fund styles (including the overall index). Use the full sample, and tabulate for each style the mean return, standard deviation, sharpe ratio; compute alpha and beta with three lags. Compute up/down betas also with three lags. Which styles have large betas? Which styles have differences between up and down betas?

5. Finally, let's try the Fama French factors. Motivating questions: Are hedge funds loading up on smb, hml, or following momentum strategies? Are they taking term structure risk (borrow short, lend long) or credit risk?

   So, run each index on all the factors, with three lags. Report the mean excess return, the alpha, and the fraction of mean return explained by each factor, $\beta_i E(rmrf)/E(R^{ei})$, $h_i E(hml)/E(R^{ei})$, etc. Why? The raw betas depend on the size of the factors - if we use $2 \times rmrf$, we get half the beta. So this is a nice way to say "how much of the expected return comes from which factor?"

   Now, look over the hedge fund styles. Overall, do they get a lot of expected return from exposures to these factors? If so, which ones? Are there specific styles that get a lot of expected return from specific factors, and does that pattern make sense?

Programming hints

The hedge fund data are backwards. You can reverse the rows with

```
T = size(x,1);
x = x((T:-1:1),:);                           % reverse order so oldest first
r = x(:,4:end) ;                             % returns; cut off date field.
```

x([2 1],:) produces the second row on top of the first row of x, and the second to last line extends this idea.

Since you only have to find regression coefficients, you can avoid a regression function and do them all at the same time with one big b = X\Y command.

To create the R^{em+} type of variables, you can use the logical expressions.

```
rx>0
```

creates a matrix of the same size as rx with ones where rx is positive and zeros where rx is negative. Thus,

```
rxplus = (rx>0).*rx
rxminus = (rx<0).*rx
```

produces the up/down returns we want. Then your right hand variable matrix for the regression with an up/down right hand variable is simply

```
X = [ ones(...) rxplus rxminus ]
```

I formed annualized returns quickly with this

```
 ra = 100*(exp(filter(ones(12,1),1,log(1+(rmrf+rf)/100)))-1)
rfa = 100*(exp(filter(ones(12,1),1,log(1+rf/100))) -1 )
rmrfa = ra-rfa
```

Look up filter and you'll see why this works.

Since you have to do the same thing over and over again it is a good idea to put the computations in a function that accepts different sizes of returns, factors, lags, and ups and downs. It's also a good idea to test the function to make sure it works. See my function test_asness on the class website for an example.

(Warning: Advanced and optional topic – for programming geeks only!) It would be awfully nice to print tables with words for headers identifying the strategy. Here's how I did it. First, I set up the labels,

```
lbls = {'Index','ConvArb','Short','EmergMkt','EqtyMktNeutral',...
        'EventDriven','Distressed','EDMultiStrat','RiskArb','BondArb',...
        'GlobalMacro','LongShortEqty','MgdFutures'};
```

This is a "cell array of strings" (Strings are vectors, i.e. ['a' 'b'] = 'ab' You can build a matrix of strings, but then they have to have the same length, e.g. ['ab'; 'cd'] but not ['ab'; 'cde'] The cell array lets you do that. Spaces and other characters in the strings are ok, I mashed mine together because I create filenames out of them.) Now you can refer to each label by index, for example,

```
for i = 1:13;
    disp(lbls(i))
end;
```

will write them out. The only hitch is that lbls(i) is a cell array, not a string. Thus, when functions want strings you have to convert. For example,

```
figure;
plot(dates,r(:,i));
title(lbls(i));
```

will return an error message, but if you change the data type of lbls(i) to a string with

```
title(char(lbls(i)));
```

or even fancier

```
title(['Graph of ' char(lbls(i))]);
```

this will work fine and give each graph its own title. Similarly,

```
for i = 1:14;
    fprintf('%20s %8.2f \n',char(lbls(i)),mean(r(:,i)));
end;
```

will make a nice table of the mean of each strategy with the strategy name next to it.

Annualizing. Annualizing correctly is not as easy as it sounds. If you take the sum or product of the returns you will get a good approximation, $R^e_{t \to t+12} = R^e_{t+1} + R^e_{t+2} + ..R^e_{t+12}$ or $R^e_{t \to t+12} = R^e_{t+1} \times R^e_{t+2} \times ..R^e_{t+12}$. But neither of these is really right. You can't compound excess returns because they are zero cost portfolios. If you make a dollar in month 1, how can you "reinvest" it in a zero cost portfolio? We need a strategy that makes sense. Here's what I did: I form $R_{t+1} = R^e_{t+1} + R^f_{t \to t+1}$ (i.e. rmrf+rf), and I use the already given total return for the hedge funds. This, you are allowed to cumulate, $R_{t \to t+12} = R_{t+1} \times R_{t+2} \times ..R_{t+12}$. Watch out, these R have to be numbers like 1.10 not 10 or 0.1. Then I form a new excess return $R^e_{t \to t+12} = R_{t \to t+12} - R^f_{t \to t+12}$ This is a genuine one-year horizon excess return.

When you're running the annual horizon regression you can use monthly observations of overlapping annual returns, i.e. $R_{t \to t+12}$ then $R_{t+1 \to t+13}$, then $R_{t+2 \to t+14}$, etc. This will ruin standard errors unless you use olsgmm to correct, but using overlapping data does not bias contemporaneous correlation or beta coefficients. And I didn't ask for standard errors.

Data

We're using the Credit Suisse "broad" Hedge Fund Index freely available at http://www.hedgeindex.com. From the website:

> "The Credit Suisse Hedge Fund Index was the industry's first—and remains the leading—asset-weighted hedge fund index. Asset-weighting, as opposed to equal-weighting, provides a more accurate depiction of an investment in the asset class.... The index universe is defined as funds with:

A minimum of U.S. $50 million assets under management ("AUM"),
A minimum one-year track record, and
Current audited financial statements."

Go read the selection rule for "Blue Chip" funds and you'll see some of my comments about selection bias at work, in order to produce data of interest to paying customers.

Since it's asset weighted, results, especially of the sector funds, will be representative of the largest funds. So, big changes in behavior as we saw in global macro can be the result of a single fund's decisions.

An interesting graph: The change of behavior of global macro – and the whole index – around 2000 coincides with the evaporation of global macro overall. I think there are only one or two funds behind this.

Historical Sector Weights

The Credit Suisse Hedge Fund Index rebalances quarterly. Below are the historical weights of the Index:

Multi-Strategy	Fixed Income Arbitrage	Dedicated Short Bias
Managed Futures	Event Driven	Convertible Arbitrage
Long/Short Equity	Equity Market Neutral	
Global Macro	Emerging Markets	