Chapter 4

PAIRS TRADING

4.1 Developing and Testing Strategies

Before we start thinking about pairs trading as a strategy, it is wise to give some thought to some fundamental issues that underly the development of trading strategies.

A fundamental activity in quantitative finance is the development and refinement of trading strategies. For example, in *equity pairs trading* or other such *convergence strategies*, one needs to determine the securities to be traded along with a whole set of trading parameters including the entry and exit conditions for the trade and the amount of each security to be bought or sold. All strategies must be developed and extensively tested on historical data before they can "go live." These notes provide some initial discussion of the issues that arise in this strategy development process, that come up in the finance papers we will be discussing in this course, and will be relevant to your homework and final projects. The issues addressed include:

- Finding, downloading and cleaning an appropriate data set
- Survivorship bias
- Implementation shortfall
- "Past performance is no guarantee of future success"

4.1.1 Specifying potential strategies

The first step is to list the potential characteristics of the strategies that you are considering. For example, if you are considering a pairs trading or convergence strategy, you will need to answer the following sorts of questions.

- Which securities are you going to consider trading? For example, stocks only, commodities, options, futures, stocks vs. commodities, currencies, etc.
- What criteria will you use to choose a pair? For example, will you consider market capitalization, recent volume, statistical features such as correlation or cointegration, actual trading profit, etc.
- What and how much past data will you use to construct the criteria? For example, fixed past period, x-day moving average, etc.
- Will you hedge positions using out-of-the-money options? Will you hedge only the short side?
- For a chosen pair, what criteria will you use to open, close, or modify positions?
- What will your strategy do if a day arrives when part of an open position suddenly has no price recorded?

Generally, one will be comparing a collection of strategies that may use different, but overlapping, datasets.

4.1.2 Identifying the appropriate dataset

After you have narrowed down the strategies you are considering, you will want the historical data you use for backtesting to be as representative as possible of the type of data that will arise when your strategy "goes live." You wouldn't want to develop a strategy using index data and then apply it to individual stocks or vice-versa since indices tend to be less variable than the individual stocks that comprise that index. A strategy developed and tested with bull-market data might perform very differently in a bear market. Strategies developed during time periods with high interest rates may perform much differently in periods with low interest rates. So, you must give some thought to what historical data you wish to use, not only in terms of the actual securities but also the market conditions.

You should choose the types of securities that you are interested in trading and gather data at a frequency that matches the trading frequency you anticipate for your strategy. For example, if you plan to trade at most once per day, then daily data could be used. If you are planning a strategy with monthly time steps, then monthly data might be appropriate. If you plan to trade multiple times during a single day, tick-by-tick data might be needed. Where can you obtain such data? For equity pairs trading, daily and montly data are available for individual securities at finance.yahoo.com. Large databases of daily and monthly data exist at Wharton Research Data Services (*WRDS*). These include high/low/closing prices, volumes, options, fundamentals, etc. Intraday data can be collected at Interactive Brokers. Large databases of intraday data can be purchased from the NYSE. Datafeed subscriptions can also be purchased. Presumably, such data will be supplied by one's employer.

4.1.3 Obtaining and cleaning the data

Once you have identified the data you need and where you can get them, you need to download them. Suppose that you obtain a file of daily data on a large number of NYSE-traded stocks. The form of the data file must be understood. Typically, each line (or record) of the file will contain all information (data fields) for one stock on one day. The information might include a ticker, other identifying information, the date, the daily high, low, close, and volume. If the data concern options, one might obtain some of the greeks, strike, call/put flag, exercise style, issue date, expiry, etc.

Now, suppose that you have a data file that you obtained from some reliable source. You should never take for granted that all data fields for all the records are exactly what you expect them to be. It is quite common for some fields to be missing in some records, or for some to be miscoded, or to contain special characters that indicate that the item means something different from what was specified. If this latter is the case, there should be some documentation from the data provider that tells you what these special codes mean. Sometimes the name of a data field is not particularly informative, or (worse yet) means something different from what you expect. Missing values are easy to find in comma-separated files: you will find two consecutive commas with no information between. Miscoded values can be harder to find. Badly miscoded values are the easiest to find. For example, if 181.03 is coded as 18103.00, one should notice strange statistical behavior for any procedure that uses this data field. The problem is that there are typically many thousands of data fields, and one rarely has time to look at them all individually. Before one does an analysis and tries to draw conclusions, one should look at summaries of the data values that are going to be used in the analysis just to verify that nothing stands out. For example, if several series of prices are going to be used, one could compute means and standard deviations for each series and see if any of them show anomalies. This could indicate the presence of badly miscoded values. For example, if the standard deviation is several times the size of the mean, there are probably some extreme values.

Other sorts of "cleaning" of the data file may be needed to produce the final data set. When analyzing data over long periods of time, stocks may split, and that will cause big jumps in price that can be handled if we know when they occurred. One needs to decide how one will deal with dividends, and so on. Finally, you want to store the data in a form that you can easily manage. It is possible that your firm may have people whose job it is to obtain and format data. If so, you will still need to identify what you will need and be able to recognize whether or not you have it.

4.1.4 Cross-Validation

Strategies are often specified up to some set of parameters, and those parameters are often chosen using historical data in the backtesting process. For example, in equity pairs trading strategies, there are a number of parameters (such as the entry and exit conditions and the size of the trade) to be chosen. In such cases, it is tempting to use the entire dataset to optimize these parameters.

Generally, this is a very bad approach, one that leads to overfitting (read "overconfidence"). If the parameters are going to be chosen in the backtesting process, then it is important to divide the historical data set into two parts, the training set and the validation set. One uses the training set to find the choices of parameters, but then cross-validates on the remaining data. This can prevent overfitting and will give some idea of how the strategy will perform when it is used for actual trading.

4.1.5 Missing Data - and survivorship bias

Missing data can be a significant nuisance in backtesting, and it is tempting to restrict the data to avoid such problems. Unfortunately this can cause significant biases in the testing protocol. Suppose one were to notice that a certain security has a number of values missing in the dataset. It is tempting to drop that stock from consideration. Some data are missing because they were lost or were entered incorrectly, but sometimes data are missing because a stock did not trade. In an extreme case, an entire block of data can be missing because the stock went out of existence, perhaps because of a merger or delisting. In cases of a trading halt or the stock going out of existence, eliminating these data from the dataset corresponds to using future information that would not be available for the strategy to use under "live" (non-anticipative) conditions. There is no way to know that at some point in the future, this security would be delisted, or would have its trading suspended.

There is a second problem that can occur with the elimination of selected stocks from the dataset, namely *survivorship bias*. If one systematically eliminates stocks that went out of existence in the relevant time period from the dataset, then the resulting stocks are no longer representative, but tend to be biased. Generally, for medium and large capitalization companies, the companies that do not survive (i.e. go bankrupt or are merged out of existence) tend to be relatively weak. Consequently, eliminating those companies tends to bias the sample toward the stronger, better performing companies. Again, one is using future data to make this decision. Introducing survivorship bias into the testing process results in adding anticipative elements to the testing process, and this should be avoided.

Training data are always available during training, so choices about training data can be made based on what is available. This will generally not cause problems so long as one is careful to (i) test choices with data that were not used in training, and (ii) repeat the process that led to the choices with the relevant set of data that are available before taking a procedure live. We will be more specific about what this last point means with particular examples later.

4.1.6 Implementation shortfall and transaction costs

In his 1988 paper, Andre Perold (1988) addressed the issue that trading strategies always do much better on paper than they do in real trading. He introduced the concept of *implementation shortfall*.

At a crude level, implementation shortfall refers to the difference between the value of a trade *on paper* at the moment of the decision to make the trade minus the actual proceeds from the trade. The former is almost always larger (sometimes significantly so) because the cost of trading reduces the actual proceeds. Some of these costs such as commissions, bid-ask spreads and taxes are obvious and relatively easy to measure. Others are more difficult to measure, for example *market impact*. The simplest example of market impact is the following. Suppose that you decide to open a position long 150 shares of a particular equity. But, when you attempt to make the purchase, you find that only 100 shares are available at the price that you used in your decision. The remaining 50 shares will need to be purchased at the next highest available price, or you might decide not to open the position because it isn't as advantageous as it appeared.

Over the years, the ideas originated by Perold have been greatly refined. The paper by Kissell (2006) contains a detailed analysis of the components of transactions costs and implementation shortfall. These ideas are nicely laid out in the ITG publication *Below the Waterline*. This will be a topic of discussion later in the course, but the major point to be made is that transaction costs must be taken into consideration. On paper, a trading strategy can look very profitable, but in reality, it may lose money because the gains are insufficient to cover the transactions costs.

4.1.7 Other kinds of data

In addition to data about equities, there are foreign exchange rates, commodity prices, option prices, quarterly/annual fundamentals, etc. There are a number of procedures that require data from multiple sources. These data come in multiple formats. Usually, you would need to "line them up," which will mean different things in different examples. We will see some later in this course. Options data are particularly cumbersome because of the number of different strike prices and expiration dates that are available at any given time. If your trading strategy requires an option with a particular strike, that strike might not be available. One might be able to replicate a particular option with other options.

Market microstructure data can be very large. Many securities have over a million orders per day. Some data sets record orders to the millisecond or even the microsecond. Despite the precision, there are issues about how accurate such time stamps really are. The times at which events are recorded for each stock will typically vary, and it becomes a task to line them up if one wishes to study joint behavior.

4.1.8 "Your performance may vary"

It is always important to remember that security markets are not stationary, so the performance exhibited by a trading strategy over the time period used for backtesting may not be representative of its performance in the future. Stocks chosen for a convergence trade based on historical data may well not produce the profits expected when the strategy goes live. It is always vital to monitor future performance to determine whether the the market conditions or the relationships that held during the backtesting period have changed. When this happens it is “back to the drawing board.”

4.2 Introduction to Pairs Trading

The idea behind pairs trading is basically the following. Suppose that two stock price series tend to move (from day to day) together, that is, they follow similar patterns over moderate stretches of time. Occasionally, there will be times when the common pattern is broken, but eventually the two series tend to get back “in synch” (or converge) again. When the prices get sufficiently far “out of synch” (or diverged), one can buy the one that is currently low and sell the one that is currently high. Then one waits until the prices converge and unwinds the position at a profit. This strategy will no longer work if the reason that the stocks diverged was related to a fundamental change in the relationship between the two stocks.

There are many questions that must be answered before one can implement such a procedure.

- Which pair of stocks should be traded?
- How do we tell when they have diverged?
- How much do we buy/sell of each stock?
- How do we tell when they have converged?
- When and how do we cut our losses when things go badly?

There are empirical methods for answering the first question, but one must use common sense also. One method is to examine the historical paths of all stocks that one would consider trading, and choose those pairs that move together most closely. To be more precise, one can standardize all price series to put them on a common scale (to make the interpretation of visual effects easier) and then compute the “distances” between all pairs of the resulting standardized series. Those pairs that have the smallest distances are the ones that have moved most closely together. The same calculation can be done with the logarithms of the price series, since these are the processes that are most commonly modeled statistically. Indeed, some pairs traders advocate using logarithms instead of raw prices to see how closely pairs move together.

There are multiple strategies for choosing how much to invest in a pairs trade. One strategy is to always invest \$0. That is, purchase the same dollar amount of the long stock as you sell of the short stock. Another strategy is to try to make the trade “market neutral.” More specifically, one tries to make the position uncorrelated with a broader market index. A third strategy is to try to maximize some measure of how much return we expect from the trade.

Typically, a pairs trader opens a position when the absolute value of the difference between the standardized series gets larger than some multiple of its historical standard deviation. Some traders use a multiple of 2. The larger the multiple, the less frequently a trader will open a position.

Closing positions can be an art. Some traders suggest waiting until the two standardized series cross again. Others will close when they have reached a predefined profit target on the trade.

It is a good idea to have rules about when to bail out if things are going against you. You need to know how much of your investment you can afford to lose before you get started.

4.3 An Example

4.3.1 General Calculations

Consider the stocks in SIC industry group 283, Drugs. We obtained daily stock prices from the CRSP (Center for Research in Security Prices) database. There were 291 stocks that had closing values for every trading day of 2003. In the pairs formation phase of the example, we consider the logarithms of closing prices of those 291 stocks for the year 2003, a period with 252 trading days. We work with logarithms because many of the models for stock prices say that log prices are like Brownian motions. Also, the upper tails of the distributions of the logarithms are not so heavy as those of the original series. This results in 291 time series of 252 observations each. We next standardized each of these series. That is, suppose that one of our stock-price series is denoted by $X_1, \ldots, X_{252}$. Define

$$\begin{aligned} A &= \frac{1}{252}\sum_{i=1}^{252}\log(X_i), \text{ and} \\ S &= \sqrt{\frac{1}{251}\sum_{i=1}^{252}[\log(X_i) - A]^2}, \end{aligned}$$

then create the standardized series $Z_1, \ldots, Z_{252}$ with $Z_i = [\log(X_i) - A]/S$. We next compute the Euclidean distance between each pair of standardized series. In general, for two standardized series $Z_1, \ldots, Z_n$ and $Z'_1, \ldots, Z'_n$, $\sum_{i=1}^n Z_i = \sum_{i=1}^n Z'_i = 0$ and $\sum_{i=1}^n (Z_i)^2 = \sum_{i=1}^n (Z'_i)^2 = n-1$. Consequently, the sample correlation coefficient r between the two series is given by $r = \frac{1}{n-1}\sum_{i=1}^n Z_i Z'_i$. (In our example, $n = 252$.)

The squared Euclidean distance between the Z and Z' series is given by

$$\begin{aligned} d^2 &= \sum_{i=1}^n (Z_i - Z'_i)^2 \\ &= \sum_{i=1}^n Z_i^2 + \sum_{i=1}^n Z_i'^2 - 2\sum_{i=1}^n Z_i Z'_i \end{aligned}$$

$$
\begin{aligned}
&= n-1+n-1-2(n-1)r \\
&= 2(n-1)(1-r).
\end{aligned}
$$

Consequently, the squared Euclidean distance between two series is linearly related to their sample correlation coefficient, and this distance is minimized by the pair with the largest sample correlation coefficient.

For each series, one can find one other series from among the other 290 series which is closest to it in Euclidean distance. This yields at most 291 pairs to work with. There anything magical about starting with these 291 pairs. For example, we could start with those pairs that have the largest correlations amongst all pairs. The reason for the particular choice we made was to avoid having too many pairs that involve the same stock. After eliminating duplicate pairs we are left with 255 distinct pairs. We consider the 50 with the smallest Euclidean distances. There is nothing magical about 50.

For each pair, we compute the mean and standard deviation of the logarithm of the ratio of the stock series for the year 2003. If a pair of series is denoted by $\{X_i\}_{i=1}^{252}$ and $\{X_i'\}_{i=1}^{252}$, then we compute

$$
\begin{aligned}
M &= \frac{1}{252}\sum_{i=1}^{252}\log\left(\frac{X_i}{X_i'}\right), \\
D &= \sqrt{\frac{1}{251}\sum_{i=1}^{252}\left[\log\left(\frac{X_i}{X_i'}\right)-M\right]^2}.
\end{aligned}
$$

Starting in January 2004, we begin trading each of the 50 pairs as follows. Each day $i = 253, \ldots, 504$, compute

$$
LR_i = \frac{\log\left(\frac{X_i}{X_i'}\right)-M}{D}, \tag{4.1}
$$

until the first day that $LR_i \notin [-2, 2]$. (One could use $[-3, 3]$ or some other interval depending on how aggressive one wants to be.) If $LR_i > 2$, wait until day $i+1$ and go short \$1 in shares of X and long \$1 in shares of X'. That is, buy $1/X_{i+1}'$ shares of X' and sell $1/X_{i+1}$ shares of X on day $i+1$. On the other hand, if $LR_i < -2$, go short \$1 in shares of X' and long \$1 in shares of X on day $i+1$. Once a position is open, we continue to monitor LR_i to see when to unwind the position. The common strategy is to wait until the first time it crosses 0. This is an indication that, whatever temporary anomaly separated the two stocks has now been forgotten. One could stop when $|LR_i|$ gets less than some prespecified value instead of waiting for it to cross 0. If any positions were open on the last trading day of the year, they were closed regardless of whether this resulted in profit or loss. The reason that we wait a day before trading is to remove an artifact of bid/ask spread. When a stock rises, it is reasonable to think that it sold close to an ask price. When it falls, it might have sold near a bid. Waiting a day helps to reduce the effect of the bid/ask spread on the size of any profits that we compute. Of course, in actual trading, a trader would not need to wait a day to trade, nor would he/she be working with daily closing prices.

The reason for trading based on the difference of logarithms, rather than the difference of prices is the following. Suppose that a position opens on day i long in stock priced X_i and short in stock priced X'_i. Then the value of the position on day $j > i$ will be

$$\frac{X_j}{X_i} - \frac{X'_j}{X'_i}. \tag{4.2}$$

This is more or less than 0 according as $\log(X_j/X'_j)$ is more or less than $\log(X_i/X'_i)$, the initial log-ratio. If the positions had been opened with equal numbers of shares, then the value of the position would be a constant times the difference of the prices, and it would have been appropriate to trade based on the difference of the prices. But equal numbers of shares would mean that the net investment is not zero.

Another thing that we did not consider, but would need to be taken into account, is transaction costs. Pairs trading can get very expensive if one has to open and close positions frequently. Also, when you own or short sell a dividend-paying stock you should include the income or expense of the dividend in the profit calculation. The data used in the example discussed here were prices adjusted for dividends and splits.

Another important consideration is what to do if it looks like a trade is going badly. A stop-loss for such trades could be useful. One possibility is to unwind a position if $|LR_i| > 5$ or use some more or less conservative cutoff than 5. Another possibility is to bail out when you have lost some predetermined amount or fraction of the position. Clearly, "fraction" is relative to something other than the inition 0 investment, for example, the size of the long (short) position. To do this, one must monitor the value of the position (**??**) rather than the log-ratio. In addition, we shall discuss hedging as a way to bound losses in Section **??**.

4.3.2 Using Bid and Ask Prices

Alternatively, one could monitor the ratios of bid and ask prices and open a position on day i rather than day $i+1$. For example, let

$$LR_{S,i} = \frac{\log\left(\dfrac{X_{i,b}}{X'_{i,a}}\right) - M}{D}, \text{ and } LR_{L,i} = \frac{\log\left(\dfrac{X_{i,a}}{X'_{i,b}}\right) - M}{D},$$

where the extra subscripts b and a indicate bid and ask prices. Note that $LR_{L,i} \geq LR_{S,i}$ for all i because ask prices are never lower than bid prices. If $LR_{S,i} > 2$, we go short $1/X_{i,b}$ shares of X and long $1/X'_{i,a}$ shares of X' on day i. If $LR_{L,i} < -2$, we go short $1/X'_{i,b}$ shares of X' and long $1/X_{i,a}$ shares of X'.

If $LR_{S,i} > 2$, we would then monitor $LR_{L,j}$ for $j > i$ until $LR_{L,j} \leq 0$. Similarly, if $LR_{L,i} < -2$, we would monitor $LR_{S,j}$ until is goes above 0. The reason that we monitor the opposite series from the one that caused the position to open is that, when we close a position, we have to buy what we sold and sell what we bought.

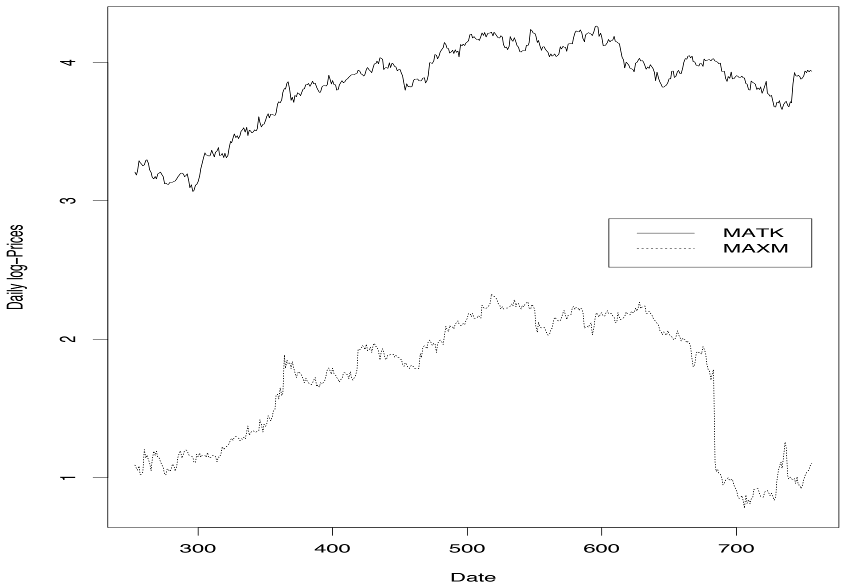

Figure 4.1: Time series of daily share prices of MATK and MAXM in 2003–2004.

4.3.3 Specific Numerical Calculations

As a specific example, consider the following two stocks. The first (X) is MATK, Martek Biosciences Corp., a commercial microbe developer in Columbia, MD. The second (X') is MAXM, Maxim Pharmaceuticals appears no longer to be traded. Figure **??** is a plot of the two series of daily prices for the 504 trading days in 2003 and 2004. Clearly, the stocks have quite different sizes and variations. A more informative plot is the standardized daily prices in Figure **??**. The two stocks seem to move together rather closely in the first year. After that, there are a couple of divergences and convergences of the two series. We chose these two stocks based on the data from the first year and then acted as if we traded during the next year.

Figure **??** shows the series of LR_i values from (**??**) for 2004 using MATK and MAXM. The trading boundaries of ± 2 and ± 3 are shown to illustrate the differences between the two strategies. If the position will eventually open, one will generally make more profit or lose less by waiting for a more extreme boundary to be crossed (e.g., ± 3 rather than ± 2). But, sometimes only the less extreme boundary is crossed, in which case one will not trade if waiting for the more extreme boundary. Because we either wait one day before trading or use the bid and ask prices, there are rare cases in which the profit is lower even when the more extreme boundary is reached. This would occur, for example, if the series of LR_i values included $LR_i = 2.6$, $LR_{i+1} = 3.1$, $LR_{i+2} = 2.5$. A trader using ± 2 would open the position on day $i + 1$, while the trader using ± 3 would open the position on day $i + 2$.

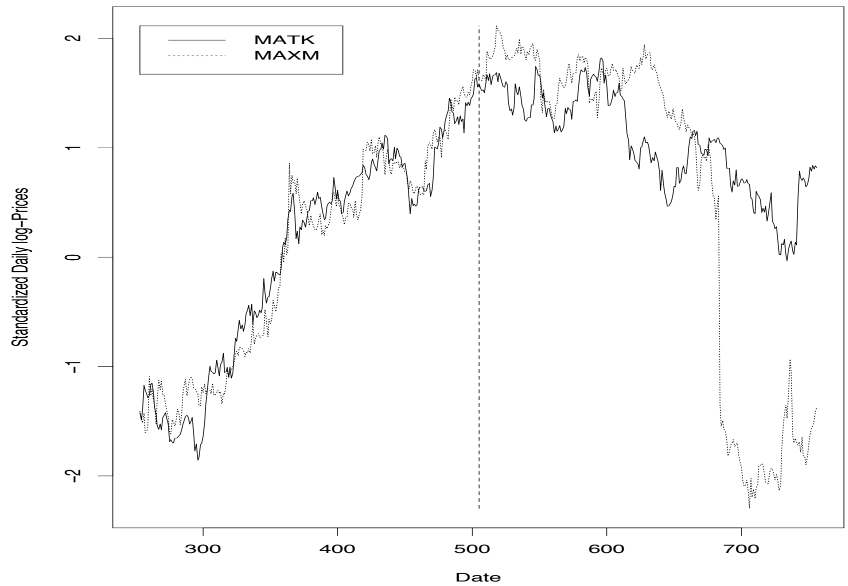

Figure 4.2: Standardized daily share prices of MATK and MAXM. The vertical line appears at the start of 2004

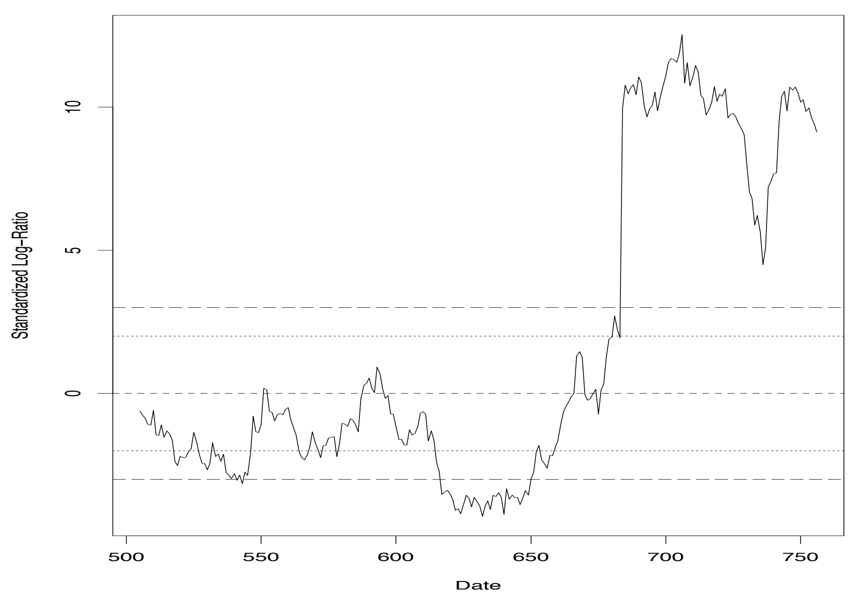

Figure 4.3: Standardized daily log-ratio of MATK and MAXM. Horizontal lines indicate possible trading boundaries

Table **??** compares the profits of the two different cutoffs for the pair we have been examining. The

Table 4.1: Comparison of two strategies for trading MATK and MAXM. Trades are based on bid and ask prices. "Short" is the stock in which we go short.

Boundary ±2				Boundary ±3			
Open	Close	Short	Profit	Open	Close	Short	Profit
January 22	March 10	MAXM	0.1793	February 25	March 10	MAXM	0.2581
March 30	May 3	MAXM	0.2080				
June 10	August 25	MAXM	0.2572	June 15	August 25	MAXM	0.3485
September 14	September 20	MATK	-0.4587	September 20	September 21	MATK	-0.0576

last position that was opened with both boundaries had to bail out, but with boundary equal to 3, it lost less because it didn't open so early. The big drop in MAXM happened during September 2004 when a class action suit was filed against the company. If we had seen that coming, we would not have opened a position in mid-September.

Table **??** compares the two cutoffs for all of the 50 pairs that traded, this time using bid/ask prices and not waiting a day to open/close positions. Using the more extreme cutoff made less profit but also traded less.

To illustrate the variation in trading strategies, such as the one described above, we performed the same operations starting the formation period at each month (21 trading days) from January 2002 through December 2007. This gave us 47 repetitions of the above strategy with different but overlapping formation and trading periods. Some of the same pairs would trade in adjacent repetitions, but we would expect differences from one repetition to the next. Figure **??** shows the series of profits. There is some clear serial dependence, and there are quite a few repetitions in which profit is negative. The average profits were 2.215 and 1.631 with cutoffs of 2 and 3 respectively. The first-order serial correlations were 0.3888 and 0.3268 respectively. The sample standard deviations were 3.990 and 2.740 respectively. The approximate variance of the average of n consecutive terms in an AR(1) process is $\sigma^2(1+\rho)/[n(1-\rho)]$, where ρ is the first-order autocorrelation and σ^2 is the common variance, which can be estimated by the sample variance to order $1/n$. The estimated standard errors of the two sample means are then 0.887 and 0.567 respectively.

4.3.4 General Considerations

The methods illustrated in this chapter as well as those in the papers are examples of methods that might be used to conduct pairs trading. There are many other possibilities. For example, instead of looking for pairs that are close together throughout the formation period, one might look for pairs that diverge somewhat but for which the moving average of the log-ratio is relatively stable. This would mean that there is evidence of both opportunity to trade and tendency to reconverge. For example, one could actually pretend to trade selected pairs and see which ones lead to trading

Table 4.2: Positions and profits for 50 pairs in 2004. Trading was based on bid/ask prices without waiting a day. A bail-out point of 4.5 was used for the log-ratio.

	Boundary ±2				Boundary ±3			
Stocks	Open	Bail	Closed	Profit	Open	Bail	Closed	Profit
DEPO/APPX	2	0	1	0.3011	1	0	0	0.0843
KV.2/CHIR	1	0	1	0.1997	1	0	1	0.3110
PARD/RPRX	0	0	0	0.0000	0	0	0	0.0000
MLNM/TPTX	1	0	0	-0.2481	1	0	0	0.0046
USNA/ESPR	1	0	0	-0.0070	0	0	0	0.0000
CARN/VMSI	0	0	0	0.0000	0	0	0	0.0000
IMCL/ANPI	0	0	0	0.0000	0	0	0	0.0000
CARN/DEPO	1	0	0	-0.3381	1	0	0	-0.0433
DUSA/MICU	1	1	0	-0.6283	1	1	0	-0.3266
OSIP/ATRX	2	1	1	0.3734	1	1	0	0.0787
EPIX/ANPI	1	1	0	-0.2438	1	1	0	-0.1831
DEPO/INGN	2	0	1	0.7673	0	0	0	0.0000
WPI/SKYE	2	1	1	0.0213	1	1	0	-0.0963
DNA/AVAN	2	1	1	0.0314	1	1	0	-0.0949
MATK/QLTI	5	0	5	1.1707	1	0	1	0.2994
PPHM/ASTM	1	0	0	-0.3476	1	0	0	-0.0723
PPHM/ZILA	1	1	0	-0.3507	1	1	0	-0.2131
AMAG/MTEX	2	0	1	0.2896	1	0	0	0.0468
ZILA/GORX	1	0	1	0.3828	0	0	0	0.0000
VRX/SGEN	1	1	0	-0.2370	1	1	0	-0.1188
DNA/DSCO	1	0	1	0.4538	0	0	0	0.0000
CYPB/SUPG	2	0	1	0.3220	2	0	1	0.8931
BSX/CTE	0	0	0	0.0000	0	0	0	0.0000
ZILA/WPI	1	0	0	-0.0252	1	0	0	0.3436
DNA/GENR	0	0	0	0.0000	0	0	0	0.0000
DIGE/FLML	0	0	0	0.0000	0	0	0	0.0000
DNA/OXGN	0	0	0	0.0000	0	0	0	0.0000
ZILA/DEPO	2	1	1	0.0876	1	1	0	-0.2079
KOSP/IVGN	3	0	3	0.5644	1	0	1	0.2921
INKP/WCRX.1	0	0	0	0.0000	0	0	0	0.0000
CVM/ZILA	1	0	0	0.0216	1	0	0	0.2926
ANPI/BSX	0	0	0	0.0000	0	0	0	0.0000
PRX/ONXX	1	1	0	-0.6874	1	1	0	-0.4065
ONXX/IVGN	1	0	0	0.0291	1	0	0	0.2839
BCRX/ONXX	1	0	0	0.2841	0	0	0	0.0000
ANIK/ONXX	2	0	2	0.7909	1	0	1	0.5084
INSM/CRIS	1	0	0	-0.0926	1	0	0	0.4199
INKP/ENCY	1	0	0	0.1236	0	0	0	0.0000
MATK/MAXM	4	1	3	0.1857	3	1	2	0.5490
ZILA/INFI	1	0	0	-0.0748	1	0	0	0.1267
TEVA/IPXL	1	0	0	0.2028	0	0	0	0.0000
TEVA/INSM	0	0	0	0.0000	0	0	0	0.0000
TARO/USNA	1	1	0	-0.5897	1	1	0	-0.3750
CYTO/INFI	1	0	0	0.0298	1	0	0	0.2014
IDEV/LGND	2	1	1	0.2880	2	1	1	0.3209
TGEN/INSM	1	0	1	0.7676	0	0	0	0.0000
IDEV/IVD	3	0	2	0.5115	1	0	1	0.3847
INKP/NUVO	1	0	1	0.4100	1	0	1	0.4286
IDEV/INSM	1	0	1	1.1359	0	0	0	0.0000
KV.1/KV.2	1	1	0	-0.0209	1	1	0	-0.0137
Totals	61	13	30	5.8549	35	13	10	3.7182

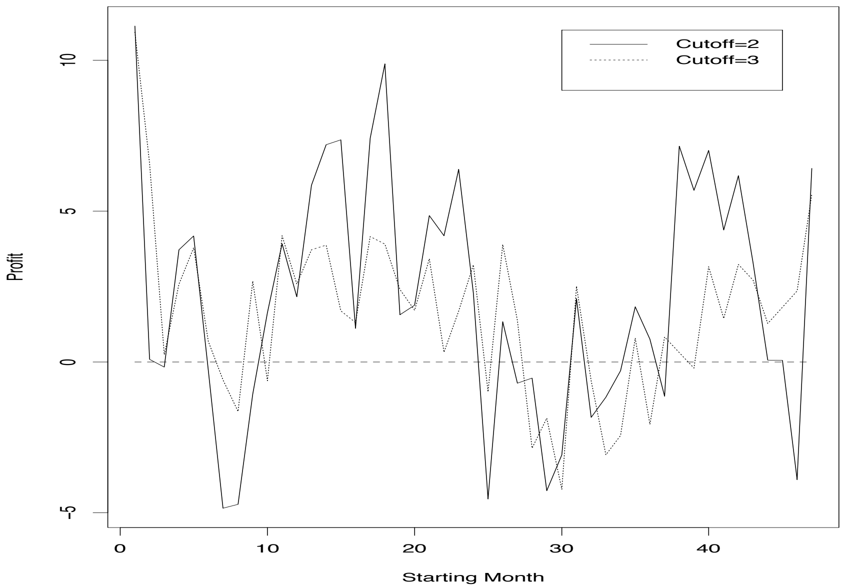

Figure 4.4: Time series of profits from trading 50 pairs starting each month from January 2002 through December 2005

opportunities or positive returns during the formation period, and then choose the pairs that did best. One could change the monitoring system to use a moving average and standard deviation of recent values for standardization rather than the mean and standard deviation fromt the formation period. There are many variations that might make more sense. You will get to try some on the homework.

Rules for choosing pairs and for opening and closing positions can be quite arbitrary. The important thing to keep in mind is that, whatever appears to work when experimenting with past data *must* be further tested with other data that were not used in the original experiment. The plot in Figure **??** is one attempt to do this sort of "out-of-sample" checking with the drug industry data.

4.4 Hedging Positions

Instead of simply setting a bail-out point in order to cut losses, one could hedge positions in order to mitigate some of the losses. Suppose that we are about to open a position that is long in X and short in X'. The position degrades in at least three possible scenarios:

- if a decline in X is not offset by a larger decline in X',
- if a rise in X' is not offset by a larger rise in X,

- if simultaneously X declines and X' rises .

The key components of all three of these scenarios are a decline in X and/or a rise in X'. These are the two contingencies against which we might consider hedging. One might have fundamental information suggesting that either X is undervalued or X' is overvalued, but not both. In such a case, one might want to hedge only one part of the position against further movement in the same direction. For example, if X is undervalued, we might want to hedge against X continuing to decline. Alternatively, we could hedge both parts of the position. The most straightforward hedge would be to buy an out-of-the-money (OTM) put option for X and an OTM call option for X'. On any given day, there are typically several options available at different strikes and with different expiration dates. Since one does not typically expect pairs trading positions to remain open for long periods of time, one might choose options that expire between one and two months in the future. Alternatively, one could use options that expire near the end of whatever trading period one desires. Or one could choose the trading period based on when potential hedging options expire. For the option strikes, one needs to weigh how much one is willing to risk losing versus how much one is willing to pay for such insurance. We shall indicate how to do this next.

In order to keep our investment at \$1 long and \$1 short, we need to take into account the prices of both stocks and options. Let the ask prices of the two options be c and c'. Then we must buy $a = 1/(X_0 + c)$ shares of X and the same number of put options. We must go short $a' = -1/(-X_0' + c')$ shares of X' and buy the same number of call options. In this case, the "cost" of the options is approximately $ac + a'c'$. If the strikes are s and s' respectively, the maximum loss that could occur is $-as + a's'$. The maximum loss would occur if X drops below s and we have to exercise our put to unload our a shares at s each, while X' simultaneously rises above s' and we have to cover our a' short shares at s' each. For all pairs of available options, one could plot the maximum losses against the costs and see if there is a reasonable combination.

A more direct approach parallels the choice of a bail-out point. One could compute how far each stock would have to move in order to reach the bail-out point. For example, in (**??**), think of X_i' being fixed at it's starting value X_0' and imagine how far X_i would have to fall in order for LR_i to fall to the lower bail-out point $-B$. A little algebra shows that X_i would have to reach the value

$$H = X_0' \exp(M - D * B). \tag{4.3}$$

Similarly, if we think of X_i as fixed at its starting value X_0, we can calculate how far X_i' would have to move in order for LR_i to reach the upper bail-out point B, namely $H' = X_0 \exp(-M + D * B)$. We could then buy a put option on X with strike close to H and a call option on X' with strike close to H'. If we choose our bail-out based on amount lost, the calculation of the strikes for hedging is simpler. If we are willing to lose proportion B of our investment, the strikes should be $X_0(1 - B)$ and $X_0'(1 + B)$.

As a specific example, consider the two stocks that traded in Figure **??**. This pair opens a position on September 14, 2004, and bails out on September 20, 2004. Suppose that we had hedged this position from the start. The position is short in MATK X and long in MAXM X', the opposite of the general discussion above. So, all of our calculations will be the reverse of what we just saw.

The bid price of MATK when the position opens is 55.66, and the ask price of MAXM is 5.67. Choosing the hedge to correspond to hitting the bail-out boundaries leads to strikes of 68.53 and 4.80 respectively. MATK has options with strikes of 65 and 70 available on September 14, 2004 with a wide range of expirations. MAXM has options with strikes of 2.5 and 5 available. For illustration, we choose expirations in December 2004 and declare this to be the end of the trading period. All the options are American style. Table **??** shows the various costs and maximum losses for the different combinations of OTM options that are available. The second row corresponds to

Table 4.3: Avalable OTM options for MATK/MAXM trading.

MATK call strike	MATK call ask	MAXM put strike	MAXM put ask	Maximum loss	Cost
60	3.50	2.5	0.80	0.7639	0.1907
60	3.50	5.0	2.05	0.5026	0.3326
65	1.95	2.5	0.80	0.8238	0.1599
65	1.95	5.0	2.05	0.5625	0.3018
70	1.05	2.5	0.80	0.8954	0.1428
70	1.05	5.0	2.05	0.6341	0.2847
75	0.50	2.5	0.80	0.9732	0.1327
75	0.50	5.0	2.05	0.7120	0.2746
80	0.25	2.5	0.80	1.0573	0.1281
80	0.25	5.0	2.05	0.7961	0.2700

the choice based on hitting the bail-out boundaries. Figure **??** plots the maximum loss against cost. The point that corresponds to hitting the bail-out boundary is in the lower right corner of the plot, being the most conservative and most costly. We can now see what would have happened if we had hedged. We will use the stikes of 70 and 2.5 for illustration. After the position opens, one can calculate the value of the position on each day. The value is what we would take in if we unwound the position. For the unhedged position, this would equal

$$-\frac{X_t}{X_0} + \frac{X'_t}{X'_0},$$

on day t, where X_t is the ask price of X and X'_t is the bid price of X'. For the hedged position, the value is

$$\frac{\min\{X_t, s\}}{-X_0 + c} + \frac{\max\{X'_t, s'\}}{X'_0 + c'},$$

plus the bid prices of whichever options we did not exercise, if any. Figure **??** shows these values from the day the positions open until the options expire. The hedged position always has lower value than the unhedged position, mainly because neither stock price crosses the strike of its corresponding option. If one had been willing to wait past the bail-out cue, one would have recovered some of the losses. On the other hand, if one had been willing to lose as much as 0.9 on the position, one could have chosen a larger bail-out boundary. The problem with the bail-out boundaries in terms of the log-ratio is that they put no bound whatsoever on the amount that you can lose. That is, it is possible for the log-ratio to approach a bail-out boundary without crossing it while simultaneously incurring an arbitrarily large loss.

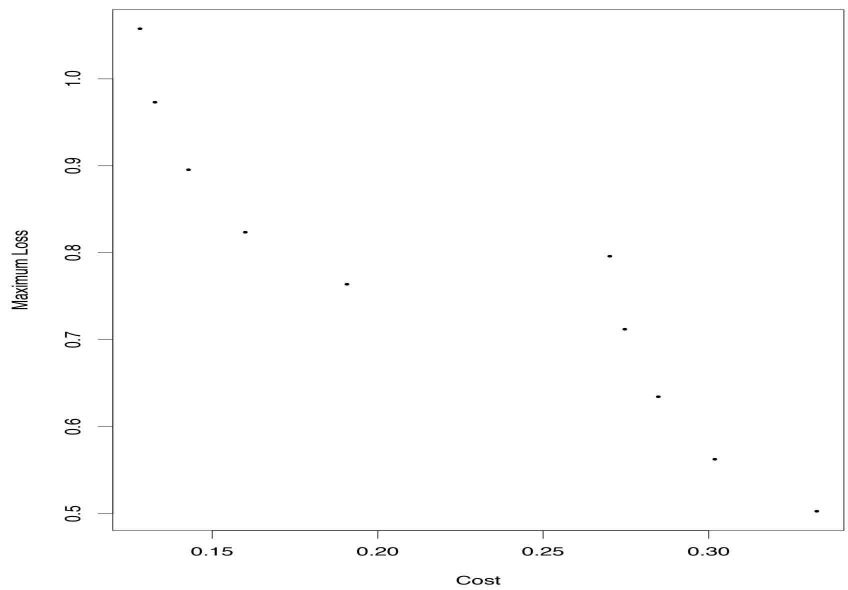

Figure 4.5: Plot of maximum loss versus cost of hedging.

There are several things that could happen to the prices and that would have an impact on whether the hedging pays off.

- The two stock prices could converge without LR_i reaching either bail-out point and without either option being exercised. In this case, the hedging was not needed. We do get to sell off the options, if they have any value.
- The two stock prices could continue to diverge until at least one of the options could be exercised. Whether bailing at this point causes a greater or smaller loss than using a bail-out point depends on what the other stock price does. It is possible to hold the position a while longer to see if the prices start to converge again. Once one of the options enters the money, the loss from that part of the position is bounded. In this case, hedging gives you more time to decide whether or not to bail out.
- The two stock prices could converge while one of the two options ends up in the money. In this case, the hedging actually increases the profit from the position.

As a final note, a position that is short c shares and long c call options on the same underlying is equivalent (in terms of immediate exercise value) to having c put options on the same underlying minus some cash, except for the cost of the cash (interest). Similarly a position that is long equal numbers of shares and put options is equivalent to having that number of call options. The difference in price between the two equivalent options should be very small. This alternative representation gives a different way to hedge a pairs trade. It has the advantage of needing half

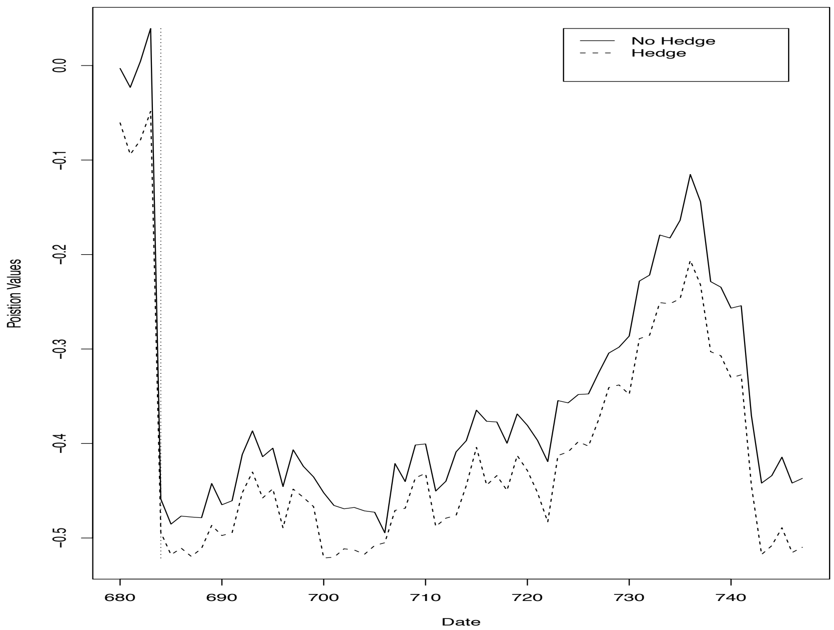

Figure 4.6: Values of hedged and unhedged positions. The vertical line indicates where the unhedged position would have bailed out.

as many transactions (two instead of four). In order to make the net investment equal to \$0, the investor would need to borrow the total price of the two options, The two options are deep in-the-money, so that there is high probability that both will still be in-the-money when the position unwinds. If the log-ratio moves toward 0, the options would move deeper into the money, and exercising them would provide enough cash to pay off the loan at a profit. In the example above, a put option on MATK with strike of 70 costs 15.20, and a call on MAXM with strike of 2.50 costs 4.1. To replicate the hedged portfolio, we would need $-1/(-55.66+1.05) = 0.01831$ of the MATK and $1/(5.67+0.80) = 0.15456$ of the MAXM, for a total cost of

$$0.01831 \times 15.2 + 0.15456 \times 4.1 = 0.912.$$

The amount of cash needed to complete the equivalence (ignoring interest) would be

$$-0.01831 \times 70 + 0.15456 \times 2.5 = -0.895.$$

Figure **??** adds the value of the position that uses only options to Figure **??** over the same time period. The value of this "equivalent hedge" is calculated as the sum of two parts, one for each option. The parts are the maximum of the value of selling the option and the value of exercising the option.

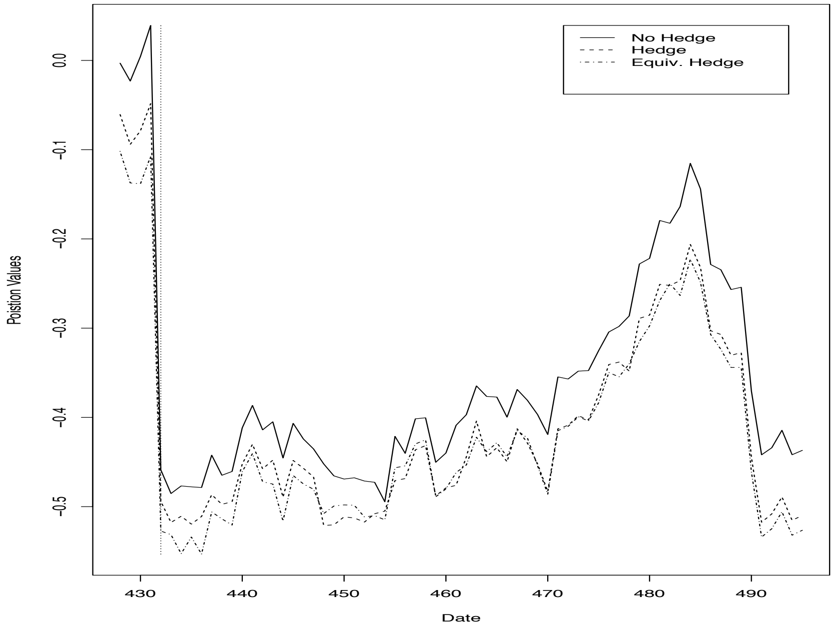

Figure 4.7: Values of hedged and unhedged positions along with equivalent hedging using options only. The vertical line indicates where the unhedged position would have bailed out.

4.5 Different Purchase Amounts

In the example above, the strategy called for trading equal dollar amounts of each stock. This made the trade a zero dollar investment trade. Except for transaction costs, the cost of the long purchase equaled the receipts from the short sale. Another popular choice for how much to go long and short is to make the trades market neutral. This means that what becomes of the trade does not depend on whatever general trend the market as a whole may follow. Our ability to do this is limited by our ability to model the relationship between individual stocks and the broader market. In the CAPM model, each stock price X equals $\alpha + \beta Y + \epsilon$ where Y is the broader market index, and α and β are peculiar to the particular stock. If another stock price, X', is given by $X' = \alpha' + \beta' Y + \epsilon'$, then

$$\frac{c}{\beta}X - \frac{c}{\beta'}X' = c\left(\frac{\alpha+\epsilon}{\beta} - \frac{\alpha'+\epsilon'}{\beta'}\right),$$

which does not depend on Y. Of course, this is a gross idealization, but the idea should remove some of the dependence on overall market movements. If one invests in this market neutral manner, one must compute profits compared to the risk free rate, because each trade will have a non-zero net investment. To avoid this, one could choose to trade only pairs whose β's were very close.

Another alternative is to choose the amount to invest in each stock according to the degree of cointegration between the stocks. Cointegration is a property that two time series have when a linear combination of the series is stationary even when neither of the individual series is stationary.

You may recall the meaning of stationary time series from an earlier time series course. The idea is that a series is stationary if the joint distribution of every collection of values depends only on how far apart they are in time (their relative positions) but not on their absolute positions in time. The key feature of stationary series that makes them useful for trading is the following. If a stationary process moves away from its mean value, it tends to move back eventually. Such behavior by a log-ratio series is exactly what makes a pairs trade profitable. We will discuss cointegration and its implications for pairs trading in Chapter **??**

4.6 Relation to "Pairs Trading: performance of a relative value arbitrage rule" by Gatev, Goetzmann, and Rouwenhorst

The authors compare a number of trading rules like the ones introduced above. Indeed, the illustrations given above were modeled on the strategies described in this paper. The authors did an extensive study with daily prices from 1962–1997. They used 12-month periods to form their pairs and the ensuing 6-month periods for trading. Each month they started a new 12-month formation period. They opened positions with 0 net investment as in the trading rules introduced above, however, they did not claim to trade based on the logarithm of the price ratio (difference of the logarithms). The reason for trading based on the difference of logarithms, rather than the difference of prices is the following. Suppose that a position opens on day i long in stock priced X_i and short in stock priced X'_i. Then the value of the position on day $j > i$ will be

$$\frac{X_j}{X_i} - \frac{X'_j}{X'_i}.$$

This is more or less than 0 according as $\log(X_j/X'_j)$ is more or less than $\log(X_i/X'_i)$, the initial log-ratio. If the positions had been opened with equal numbers of shares, then the value of the position would be a constant times the difference of the prices, and it would have been appropriate to trade based on the difference of the prices.

The authors of this paper introduced the idea of waiting a day after stock prices have diverged before opening a position in order to remove bias based on the bid/ask bounce. In referring to a "winner" as the stock whose price is high relative to the "loser" in the pair, they say

> Part of any observed price divergence is potentially due to price movements between bid and ask quotes: conditional on divergence the winner's price is more likely to be an ask quote and the loser's price a bid quote. Since we have used these same prices for the start of trading, our returns may be due to the fact that we are implicitly buying at bid quotes (losers) and selling at ask quotes (winners). The opposite is true at the second crossing (convergence): part of the drop in the winner's price can reflect a bid quote, and part of the rise ie of the loser's price an ask quote.

They use this reasoning to justify the one-day wait before opening and closing positions in their study of excess returns. They also claim that the drop in profits that they observe when waiting one day to trade, compared to trading with the same day's closing prices, provides an estimate of transaction costs.

One possible explanation for why pairs trading strategies seem to earn excess returns is that they are more risky. The authors regress their excess returns on various risk factors (see Fama and French, 1996[1]). They find that, although some of the risk factors do explain some of the excess returns, there are still significant excess returns not explained by known risk factors. They also conclude that their pairs portfolios are not particularly risky as measured by VaR.

In forming the pairs to trade, the authors used all stocks, not just those in a particular industry. They found that utility stocks figured prominently in their pairs. They also separated stocks into broad industry categories and found significant excess returns in each category, with utilities providing the largest excess returns.

4.7 Some Computer Code

4.7.1 R Code for Pairs Trading Example

The function below trades a single pair using a bail-out limit based on the log-ratio. It opens positions on the same day that the trading signal occurs and uses bid and ask prices to open and close positions rather than daily closing prices.

```
pair.trade1=function(dates,price,bid,ask,
  dayst,dayet,cutoff,bail,closet=0,norms){
#
# Trade one pair with a bail-out point based on log-ratio
# price,bid,ask all have 2 columns, one for each stock
# norms is a vector containing the mean
#   and standard deviation with which to normalize the series of log-ratios
# dayst and dayet are start and end days for trading
#
  profit=0
  nopen=0
  nclose=0
  nbail=0
  baillist=0
  tradelist=rep(0,2)
  openlist=rep(0,2)
  closelist=rep(0,2)
```

[1]Fama, E.F. and French, K., "Multifactor explanations of asset pricing anomalies," *Journal of Finance*, **51**, 1996, 55-84.

```
  stock1=price[,1]
  stock2=price[,2]
  ratio.mean=norms[1]
  ratio.std=norms[2]
  stock1b=bid[,1]
  stock2b=bid[,2]
  stock1a=ask[,1]
  stock2a=ask[,2]
  fna<-which(is.na(stock1a[dayst:dayet])|is.na(stock2a[dayst:dayet]))
  fnb<-which(is.na(stock1b[dayst:dayet])|is.na(stock2b[dayst:dayet]))
# Stop trading when any bid or ask price is missing
  endday<-dayet-dayst
  if(length(fna)>0){
    endday<-fna[1]-2
  }
  if(length(fnb)>0){
    endday<-min(c(endday,fnb[1]-2))
  }
  open.trade=0
  if(endday>0){
    for(j in 1:endday ){
      daynow<-dayst+j-1
# If a position is open, check to see if it should close
      if(open.trade!=0){
        if(open.trade==-1){
           value=n1*stock1b[daynow]+n2*stock2a[daynow]
         }else{
           value=n1*stock1a[daynow]+n2*stock2b[daynow]
         }
         bailout=(open.trade*ratio.trade[j-startt]>=bail)
   if((open.trade*ratio.trade[j-startt]<=closet)|bailout){
            open.trade=0
            nclose=nclose+1
            profit=c(profit,value)
            closelist=rbind(closelist,c(dates[daynow],daynow))
          }
# If we ever bail out of a position, never trade this pair again
        if(bailout){
          nbail=nbail+1
          nclose=nclose-1
          baillist=c(baillist,daynow)
          break
        }
# If no position is open, check to see if we should open one
      }else{
```

```
# Don't trade penny stocks
if((stock1b[daynow]>1)&(stock2b[daynow]>1)){
          ratio.tradeab=(log(stock1a[daynow:dayet]/stock2b[daynow:dayet])-
      ratio.mean)/ratio.std
          ratio.tradeba=(log(stock1b[daynow:dayet]/stock2a[daynow:dayet])-
      ratio.mean)/ratio.std
          if(ratio.tradeba[1]>=cutoff){
            open.trade=1
            n1=(-open.trade)/stock1b[daynow]
            n2=open.trade/stock2a[daynow]
            ratio.trade<-ratio.tradeab
          }
          if(ratio.tradeab[1]<=-cutoff){
            open.trade=-1
            n1=(-open.trade)/stock1a[daynow]
            n2=open.trade/stock2b[daynow]
            ratio.trade<-ratio.tradeba
          }
          if(open.trade!=0){
#  startt is the day before the position opens so that ratio.trade[j-startt]
#    will be the normalized log ratio on day j while the position is open
    startt=j-1
            nopen=nopen+1
            tradelist=rbind(tradelist,c(n1,n2))
            openlist<-rbind(openlist,c(dates[daynow],daynow))
          }
}
     }
  }
}
# If a trade is open on the last day, close it
  if(open.trade!=0){
    if(open.trade==-1){
      profit=c(profit,n1*stock1b[endday+dayst]+n2*stock2a[endday+dayst])
    }else{
      profit=c(profit,n1*stock1a[endday+dayst]+n2*stock2b[endday+dayst])
    }
    closelist=rbind(closelist,c(dates[endday+dayst],endday+dayst))
  }
# The output is the stocks in the pair, the numbers of positions opened, bailed, and
# closed normally, the total profit, a list of the coefficients for each position,
# the dates positions were opened, the dates they closed, which ones were bailed,
# and the mean and s.d. of the log-ratio during the
# formation period.
  if(nopen>0){
```

```
   openlist=openlist[-1,]
   closelist=closelist[-1,]
   tradelist=tradelist[-1,]
   profit=profit[-1]
  }
  if(nbail>0) baillist=baillist[-1]
  list(trans=c(nopen,nbail,nclose),profit=profit,
       trades=tradelist,opens=openlist,closes=closelist,
       bails=baillist,normalize=norms)
}
```

The inputs and outputs are described in the comments.

If one wishes to trade the top pairs, as in the example presented earlier, one can use the program above for each pair. To find the top pairs, one could use code like the following.

```
trade.cor=function(days,price.mat,bid.mat,ask.mat,ticks,daysf=1,dayef=126,
  dayet=dayef+126,npairs=20,cutoff=2,bail,closet=0){
  if(missing(bid.mat)|missing(ask.mat)|missing(price.mat))
    stop("You must supply bid and ask and closing prices")
#
# days is a list of dates
# price.mat is a matrix with one row for each date and one column for each stock
# bid.mat and ask.mat are in the same form as price.mat and contain
#   bid and ask prices
# ticks is a list of ticker symbols
# daysf is the index (in the days list) of the day on which to start the
#   formation period
# dayef is the day to end the formation period. Trading period starts
#   the next day.
# dayet is the day to stop trading
# npairs is how many pairs to consider for trading
# cutoff is the absolute log-ratio that triggers the start of a trade
# bail is the absolute log-ratio for bailing out.
# closet is the absolute log-ratio that triggers closing a trade normally
#
# Remove stocks with missing data during the formation period
#
  INDp = !is.na(apply(price.mat[daysf:dayef,],2,prod))
  INDa = !is.na(apply(ask.mat[daysf:dayef,],2,prod))
  INDb = !is.na(apply(bid.mat[daysf:dayef,],2,prod))
  IND = as.logical(INDp*INDa*INDb)
  price=price.mat[,IND]
  bid=bid.mat[,IND]
```

```
  ask=ask.mat[,IND]
  ticks=ticks[IND]
  nstock=dim(price)[2]
#
# Compute correlations
#
  price.cor=cor(log(price[daysf:dayef,]))
  pairm=c(1:nstock)
#
# For each stock, find the other stock with largest correlation
#
  for(i in 1:nstock){pairm[i]=sort.list(price.cor[i,])[nstock-1]}
  pairmn=t(apply(cbind(c(1:nstock),pairm),1,sort))
#
# Remove duplicate pairs
#
  pairmu=pairmn[!duplicated.data.frame(as.data.frame(pairmn)),]
  napairs=dim(pairmu)[1]
  numpairs=min(napairs,npairs)
#
# Choose the pairs with highest correlations
#
  toppairs=pairmu[sort.list(price.cor[pairmu]),][napairs:(napairs-numpairs+1),]
#
# Initialize outputs
#
  profit=rep(0,numpairs)
  nopen=rep(0,numpairs)
  nclose=nopen
  nbail=nopen
#
# Loop through the top pairs
#
  for(i in 1:numpairs){
    prices=price[,toppairs[i,]]
    bids=bid[,toppairs[i,]]
    asks=ask[,toppairs[i,]]
#
# normalize based on entire formation period
#
    norms=c(mean(log(prices[daysf:dayef,1]/prices[daysf:dayef,2])),
        sqrt(var(log(prices[daysf:dayef,1]/prices[daysf:dayef,2]))))
#
# Call the program that trades a single pair
#
```

```
    result=pair.trade1(days,prices,bids,asks,dayef+1,dayet,cutoff,
      bail,closet,norms)
#
# Store the output
#
    nopen[i]=result$trans[1]
    nbail[i]=result$trans[2]
    nclose[i]=result$trans[3]
    profit[i]=sum(result$profit)
  }
#
# Output includes the inputs plus the tickers for pairs that traded,
#   the numbers of transactions, the profits, and the stocks that
#   were included in the formation period.
#
  list(nums=c(nstock,numpairs),dates=days[c(daysf,dayef,dayet)],
       crits=c(cutoff,bail,closet),
       tickers=cbind(ticks[toppairs[1:numpairs,1]],
ticks[toppairs[1:numpairs,2]]),
       pairs=toppairs[1:numpairs,],trans=cbind(nopen,nbail,nclose),
          profit=profit,include=IND)
}
```

4.8 Gatev, Goetzmann, and Rouwenhorst (2006)

Reading Notes
"Pairs Trading: Performance of a Relative-Value Arbitrage Rule"

The paper by Gatev, Goetzmann, and Rouwenhorst presents an introduction to pairs trading (with some historical commentary on its origin) and reports on a large study that documents the effectiveness of this trading strategy. The authors implemented a pairs trading strategy using daily data from the U.S. equity market over the 1962 to 2002 time period. The bottom line of the study is that pairs trading gives excess returns of 11% per year for the top portfolios, although this figure does not properly account for many effects such as transactions costs or holding costs. The authors do go on to explore the robustness of their results to factors such as "bid-ask bounce," short selling costs, and transaction costs. We summarize below the significant details of the paper. Of course, these notes are just highlights and are a very poor substitute to reading the actual paper.

4.9 Study Methodology

4.9.1 Pairs Selection and Trading Period

The authors consider a pairs trading strategy over a basic 18 month period. The first 12 months are called the formation period, while the last 6 months are the trading period. All the stocks from the CRSP daily price data base are considered as candidates for a pair, but stocks that had one or more days with **no** trades were eliminated. The authors consider this to be a way to identify relatively liquid stocks. This does not really create a survivorship bias, because the evaluation of the trading strategy is done over the next 6 months, and the authors handle the case in which a member of a pair that is being traded disappears.

Once the eligible stocks have been identified, stock price and returns series are derived for each. The stock price series are normalized to standardize their location and scale. The authors consider all pairs of stocks in the database and "... then choose a matching partner for each stock by finding the security that minimizes the sum of squared deviations between the two normalized price series." Thus, if there are n stocks under consideration, there will be n pairs selected, some of which may be duplicates (e.g. stock B is closest to stock A, while stock A is the closest to stock B). The pairs can be ranked in order of closeness with the pair having the smallest least squares difference ranked number 1. In a separate analysis, the authors also divided the stocks into four industry categories: Utilities, Transportation, Financial, and Industrial and restricted the pairs to be in the same industry category. Table 3 on page 812 indicates that there were 156 utility stocks, 61 transportation stocks, 371 financial stocks and 1729 industrials for a total of 2317.

In the unrestricted case, the authors rank order all the pairs selected in the formation period according to the least squares criterion and use them for trading. Results are reported for the top 5, the top 20, numbers 101-120 of the top 120, and all pairs. It was noted (see the text at the bottom of Figure 1 page 807) that the average number of pairs in the "all" portfolio was 2057.

It is interesting to note that section 1.4 of the paper describes the underlying concept of pairs trading in terms of cointegration. Nevertheless, they carry out their pairs formation using correlation.

Q1[2]: How would the performance change if one used a cointegration-based measure rather than a correlation-based measure?

Q2: How could the performance of a "pairs trade" using a set of 3 or more stocks using cointegration vectors compare with a strict pairs trading strategy?

After the pairs are chosen, there is a six month trading period. For each pair, the trade is opened when there is a two-standard deviation gap between the two stocks. One dollar of the stock with the relatively high price is shorted, while one dollar of the stock with the relatively low price is purchased. This trade is self-financing with an initial portfolio value of 0, and any returns are excess

[2] We will occasionally provide sample questions that might be investigated in a statistical arbitrage project.

returns over the market. A position is closed when the two stock prices converge or when the six month trading period ends. A pair that closes can reopen again if the two standard deviation criterion is met. At the end of the six month period, a trade that is still open will be closed out. Finally, the authors state that "If a stock is delisted from CRSP, the position is closed using the delisting return or the last available price."

Figure 1 is taken from the paper. It illustrates the opening and closing of an illustrative pairs trade (Kennecott and Uniroyal in this case). In this case there are 5 distinct trades with these two stocks, and the final trade closes at the end of the trading period. Note that the first, second and fourth trades are long Uniroyal, while the third and fifth are short Uniroyal.

Q3: How does the performance change if the two sigma rule were modified?

Q4: How does the performance change as a function of the lengths of the formation and trading periods?

Q5: How can the trading strategy used by Gatev *et. al.* be improved and how much better do the improved strategies perform. For example, at the end of section 3.3 (page 811) the authors indicate that it is foolish to open a trade at the end of the trading interval, yet their trading rules would do so. Are there other improvements that can be introduced?

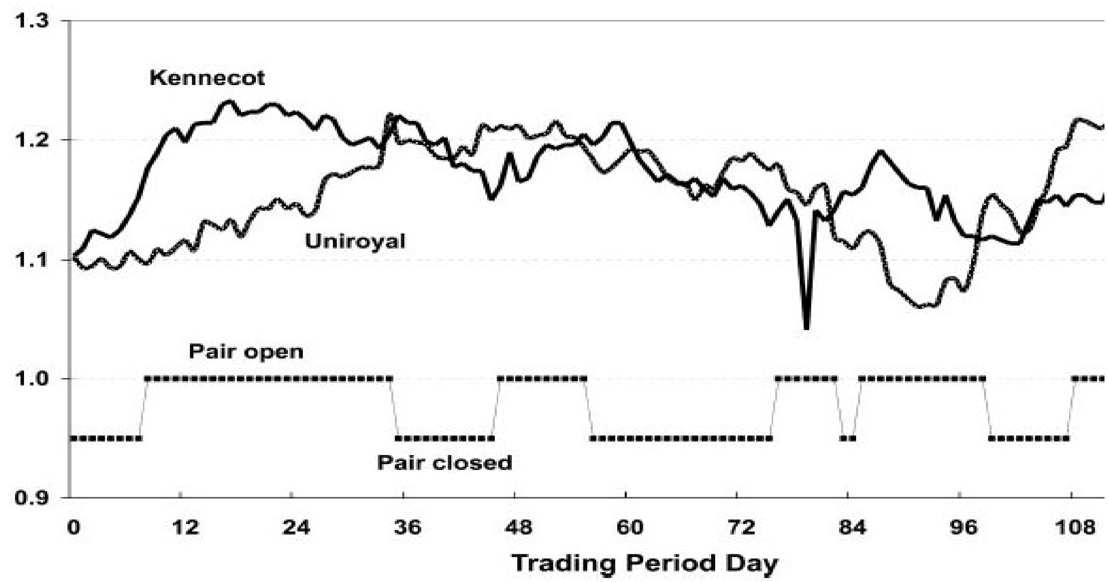

Figure 1
Daily normalized prices: Kennecott and Uniroyal (pair 5)
Trading period August 1963–January 1964.

Figure 4.8: Kennecott/Uniroyal pairs trading period, page 804

4.9.2 Returns evaluation

Some pairs may never open, but those that do will have one or more cash flows during the six month trading period. Many will have a cash flow at the end of six months if they have a trade that is still open at the end. The authors state that "the excess return on a pair during a trading interval is computed as the reinvested payoffs during the trading interval. In particular, the long and short portfolio positions are marked-to-market daily. The authors assume that cash derived from closing positions (which must be positive except possibly at the end of the trading period) earns a 0% rate of return, and is assumed to be reinvested during the trading interval. Thus the returns are value-weighted and are calculated as follows:

$$r_{P,t} = \frac{\sum_{i \in P} w_{i,t} r_{i,t}}{\sum_{i \in P} w_{i,t}},$$

where $w_{i,t} = \Pi_{j=1}^{t-1}(1 + r_{i,j})$, r denotes the returns, w denotes the weights, and P refers to the set of pairs that is being traded.

The authors use two measures of total excess return on a portfolio of pairs: 1) the return on committed capital (counting $1 for every pair whether open or not), and 2) the return on the actual employed capital (counting $1 only for open positions).

The study begins in 1962 (call this time 0). Then for each k, $(k \geq 0)$ there is a 12 month formation period followed by a 6 month trading period, i.e. after 1 year there will be 12 portfolios with partially overlapping formation periods and 6 with partially overlapping trading periods. The time series of returns indexed by k will thus be correlated. The 40 year (480 month) study yielded a time series of 474 returns.

4.9.3 Buying and selling prices

It is important to ask how the the stock purchases and sales are priced. In the empirical results presented in section 3, the authors consider two approaches. In one (Method A), they assume that a pair is opened at the end of the day on which the prices satisfy the two-standard deviation criterion and is closed at the end of the day on which the prices converge. The authors note the bias in this approach (see line 7 from the bottom, page 807). When an "open pair" signal occurs, the price difference has exceeded the two-sigma threshold. This means that either one stock has had a gain, the other has had a loss or both. Price divergence on that day is likely to mean that the stock that is shorted is priced near the "ask," while the stock that is purchased is likely to be priced near the "bid." Furthermore, when prices converge, just the opposite is true. Consequently, it is likely that the "bid-ask" bounce will increase the apparent profitability of the pairs trade. To combat this bias, the authors also use a second approach (Method B) in which the positions are initiated and liquidated on the date following the open and closing signal events.

4.10 Empirical Results

Section 3 of the paper presents the empirical results. We summarize each of the major points in the paper.

4.10.1 Profits from the pairs trading strategy

Table 1 of the paper (see Figure **??** below) presents some of the major results. Panel A presents

Table 1
Excess returns of unrestricted pairs trading strategies

Pairs portfolio	Top 5	Top 20	101–120
A. Excess return distribution (no waiting)			
Average excess return (fully invested)	0.01308	0.01436	0.01081
Standard error (Newey-West)	0.00148	0.00124	0.00094
t-Statistic	8.84	11.56	11.54
Excess return distribution			
Median	0.01194	0.01235	0.00955
Standard deviation	0.02280	0.01688	0.01540
Skewness	0.62	1.39	1.34
Kurtosis	7.81	10.54	10.30
Minimum	−0.10573	−0.06629	−0.03857
Maximum	0.14716	0.13295	0.12684
Observations with excess return < 0	26%	15%	21%
Average excess return on committed capital	0.00784	0.00805	0.00679
B. Excess return distribution (one day waiting)			
Average monthly return (fully invested)	0.00745	0.00895	0.00795
Standard error (Newey-West)	0.00119	0.00096	0.00085
t-Statistic	6.26	9.29	9.40
Excess return distribution			
Median	0.00699	0.00690	0.00694
Standard deviation	0.02101	0.01527	0.01438
Skewness	0.34	1.45	0.98
Kurtosis	10.64	16.13	7.78
Minimum	−0.12628	−0.08218	−0.04266
Maximum	0.14350	0.13490	0.10464
Observations with excess return < 0	35%	23%	28%
Average excess return on committed capital	0.00463	0.00520	0.00503

Summary statistics of the monthly excess returns on portfolios of pairs between July 1963 and 2002 (474 observations). We trade according to the rule that opens a position in a pair at the day that prices of the stocks in the pair diverge by two historical standard deviations (Pan results in Panel B correspond to a strategy that delays the opening of the pairs position by on pairs are ranked according to least distance in historical price space. The "top n" portfolios i n pairs with least distance measures, and the portfolio "101–120" studies the 20 pairs after th The average number of pairs in the all-pair portfolio is 2057. The t-statistics are computed usi West standard errors with six-lag correction. Absolute kurtosis is reported.

Figure 4.9: Table 1, page 807 of Gatev, et al.

the results assuming purchases and sales are made on the day the open pairs trade signal or the close trade signal is given. Panel B presents the results assuming the purchases and sales are made on the next day. The top and bottom lines of both Panels A and B give the two methods for computing the average excess return. The four columns give 5 different portfolios - the top 5, top 20, pairs 101-120, and all pairs. The results in Panel A show the average monthly excess ranging from 1.436% (top 20) to 1.081% (pairs 101-120). The standard errors are approximately 10% of the averages, and the results are highly statistically significant. Features of the excess returns distribution are presented with a range for the monthly of -2.721% to 17.178% (All pairs). Note that the standard deviation for the top 5 is larger than the other portfolios, presumably because fewer pairs results in more variable results. The minimum monthly return is also the smallest for this portfolio, -10.573% compared with -2.721% for the All portfolio. An interesting statistic is the fraction of observations that are negative. These range from 15% (Top 20) to 26% (Top 5). For example, the authors report that "During the full sample period of 474 months, a portfolio of 20 pairs experienced 71 monthly periods with negative payoffs, compared to 124 months for a portfolio of 5 pairs," (line 3 from the bottom of page 806). Overall, the results are strongly supportive of pairs trading as a method to produce excess monthly returns, and the overall returns can be quite large.

The results in Panel B are not as strong as those in Panel A, possibly because of the "bid-ask bounce" phenomenon. Nevertheless, they range from .715% (All) to .895% (Top 20). As a consequence of this difference, the authors indicate that for the rest of the paper they report results for the second (sale and purchase on the following day) strategy.

Section 3.4 and Table 3 (pp811-812, but not included in this set of notes) discuss the returns when the pairs are restricted to both be in one of the four industry categories. Again, they rank all the pairs on the least squares metric, then construct the top 5, top 20, to 101-120, and all portfolios (except for the transportation sector which has fewer than 100 and therefore doesn't have an entry for 101-120). Considering just the top 20 portfolio, the monthly excess returns are 1.084% for the utilities, .775% for the financials, .607% for the industrials, and .577% for the transportation sector, each is economically and statistically significant. Nevertheless, with the exception of the utilities sector, negative results are more likely in the top 20 portfolio. There were 23% negative returns reported in Table 1 for the overall. In contrast there were 19% for the utilities portfolio, but 33% for the financials, 36% for the industrials, and 42% for the transportation sector. While pairs trading still has excess returns, restricting to a smaller, if more homogeneous, subset of stocks appears to reduce the returns.

Q6: Is there an effective way to better match stocks using a combination of cointegration/correlation information and fundamental information?

4.10.2 Trading statistics and portfolio composition

Panel A of Table 2 of the paper (see Figure **??** below) gives information about the trades that are made, and it also gives information about the composition of the top portfolios.

Table 2
Trading statistics and composition of pairs portfolios

Pairs portfolio	Top 5	Top 20	101–120	
A. Trading statistics				
Average price deviation trigger for opening pairs	0.04758	0.05284	0.07560	
Average number of pairs traded per six-month period	4.81	19.30	19.41	1'
Average number of round-trip trades per pair	2.02	1.96	1.78	
Standard deviation of number of round trips per pair	0.62	0.40	0.27	
Average time pairs are open in months	3.75	3.76	3.98	
Standard deviation of time open, per pair, in months	0.80	0.45	0.38	
B. Pairs portfolio composition				
Average size decile of stocks	2.54	2.71	3.41	
Average weight of stocks in top three size deciles	0.78	0.74	0.58	
Average weight of stocks in top five size deciles	0.91	0.91	0.79	
Average weight of pairs from different deciles	0.66	0.69	0.75	
Average decile difference for mixed pairs	0.97	0.97	0.97	
Average sector weights				
Utilities	0.72	0.71	0.32	
Transportation	0.02	0.02	0.02	
Financials	0.11	0.13	0.26	
Industrials	0.15	0.14	0.40	
Mixed sector pairs	0.20	0.22	0.44	

Trading statistics and composition of portfolios of pairs portfolios between July 1963 and Decer (474 months). Pairs are formed over a 12-month period according to a minimum-distance crit then traded over the subsequent 6-month period. We trade according to the rule that opens a po pair on the day following the day on which the prices of the stocks in the pair diverge by two standard deviations. The "top n" portfolios include the n pairs with least distance measures portfolio "101–120" includes the 20 pairs after the top 100. Panel A summarizes the trading char of a pairs strategy. Pairs are opened when prices diverge by two standard deviations. Average de trigger opening of pair is the cross-sectional average of two standard deviations of the pair prices c Panel B contains information about the size and industry membership of the stocks in the var portfolios.

Figure 4.10: Table 2, page 809, of Gatev et al.

Trading Statistics

Table 2 gives information about the nature of the pairs and the frequency of trades. The two standard deviation criterion translated to a 4.76% (top 5), 5.28% (top 20), to a 16.89% (all) difference in the prices. The authors indicate that figures around 5% are small, and it may indicate that a larger number of standard deviations than 2 should be used. Nearly all of the pairs have at least one open trade during their trading period (19.3 out of 20 for the top 20). The average number of round-trip trades is slightly less than 2 (1.96 for the top 20, 1.62 for all). A trade is open on average from 3.75 to 3.97 months.

Portfolio composition

Panel B of Table 2 gives some information about the composition of the top portfolios in terms of firm size and industry category. For the top portfolios (5, 20, 101-120) the firms tend to be large, e.g. 74% in the top 3 size deciles and 91% in the top 5 size deciles for the Top 20 portfolio. Indeed

62% of the entire set of pairs lie in the top 5 size deciles, perhaps indicating that the method of pair selection favors liquid stocks which are often in the higher size deciles. Perhaps most interesting was the breakdown by industrial category. In the top 20 category 71% were utility stocks even though those stocks comprise only 8% of the entire set of stocks.

4.10.3 Transaction costs

Section 3.3 on transaction costs is one of the least satisfying in the paper, and an area that could be substantially improved. The authors conceptualize transaction costs in terms of the "bid-ask bounce," and they estimate this through the reduction in returns in the two different purchase and sale price algorithms (same day purchase/sale versus one day later purchase/sale). The differences in the excess returns of these two are the result of a mixture of pairs trading a differing number of times. They use average quantities to try to reconstruct trading costs instead of creating a measure on a trade-by-trade basis. The authors come up with a 81bp per trade (162bp per round trip) transaction cost figure and compare this with published number of 37bp by Peterson and Fialkowski. The authors gratuitously say that since their pairs are over-weighted by large stocks, the transaction costs may be lower, but do not justify this. When they translate the results to a 6 month trading period, they argue that the transaction cost per pair per 6 months might be 324bp (162bp times 2.02 round trips/6 months). This would reduce the 437bp to 540bp excess returns per 6 months (multiply the monthly returns by 6) to a final value of 113bp to 225bp. These are still both economically and statistically significant, but certainly one would want to conduct a more accurate accounting of the transaction costs.

Note also that section 3.9 (Robustness to short-selling costs) discusses certain technical issues associated with shorting a stock. The discussion is framed in terms of the work of D'Avolio (2002). The abstract to the D'Avolio paper states:

> To short a stock, an arbitrageur must first borrow it. This paper describes the market for borrowing and lending U.S. equities, with an emphasis on the conditions generating and sustaining short sales constraints. Eighteen months (4/2000 - 9/2001) of daily data provided by a large institutional lending intermediary establish a rich set of facts on loan supply ("shortability"), loan fees ("specialness"), and loan recalls. The data suggests that while loan market specials and recall are rare on average, the incidence of these short sales constraints is increasing in the divergence of opinion among investors. An implication is that beyond some threshold, investor optimism itself can limit arbitrage via the loan market mechanism.

Since pairs trading requires shorting stocks, the above abstract indicates that there are risks in so doing. Section 3.9 provides two tests of these concerns. The results are displayed in Table 9 (page 825), and the authors argue that they show that the above concerns do not significantly reduce the effectiveness of the pairs trading strategy studied in this paper.

4.10.4 Risk characteristics

The two remaining substantive sections (3.5 on the risk characteristics of the pairs trading strategy and 3.8 on a subperiod analysis and the possibility of a dormant risk factor) look closer into the excess returns and their decomposition into market risk factors. We defer a discussion of this until later in the course as the reader must first learn about the Fama French factors and their extension by Carhart to include momentum. This will be given later in the course.

