Time Series Analytics Notes
The Market Model

The objective of the **market model** is to measure the riskiness of a particular stock relative to the riskiness of a broadly based market average. Relative risk is an important concept in securities analysis in finance. The riskiness of a stock is usually assessed in terms of its price volatility. If the price of a stock is relatively more volatile than the general market, then the stock is riskier than the typical stock. If the price of a stock is relatively less volatile than the general market, then the stock is less risky than the typical stock. If the price of a stock is just as volatile as the general market, then the stock is just as risky as the typical stock.

Modern investment theory focuses on choosing stocks to maximize expected returns for a given level of portfolio risk. The idea is that there is a trade-off between risk and return. The investor can have higher expected returns, but at the cost of higher risk. So the investor chooses the level of risk (price volatility) that can be tolerated and then searches for a set of stocks to maximize expected returns, subject to the constraint that their collective risk (termed *portfolio risk*) not exceed the tolerable level.

Figure 1. Portfolio concepts

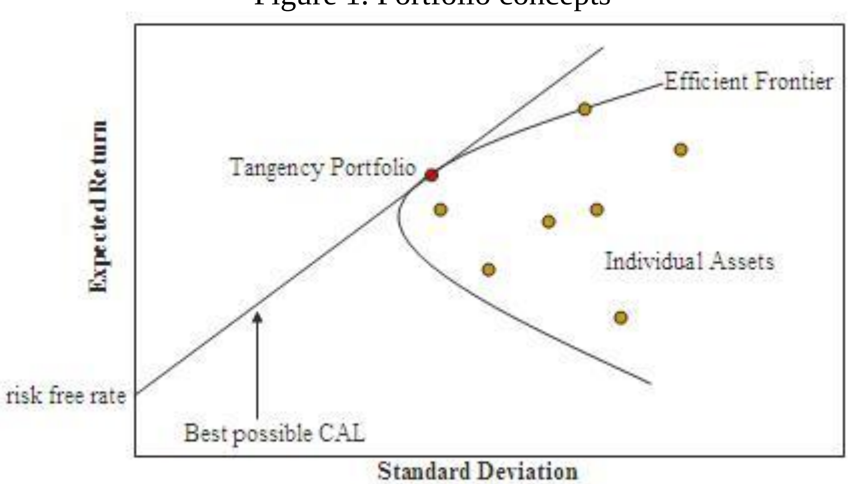

Figure 1 illustrates these concepts. The vertical axis represents expected percentage returns obtainable from investments in assets. The horizontal axis represents the risk of those assets (here called "standard deviation", but could be any risk measure). The risks and returns of individual assets are points in the plane, here shown as small circles. An asset A lying above (higher return) and to the left (lower risk) of another asset B is preferred to asset B. By combining individual risky assets (excluding risk-free assets like Treasury bills) in various proportions, asset portfolios can be generated to yield any (risk, return) combination inside the hyperbola shown in Figure 1. This hyperbolic region of feasible portfolios is sometimes called the "Markowitz bullet". The upper left edge of the Markowitz bullet is the "Efficient Frontier". Every interior point is inferior in risk and/or return to some point on the Efficient Frontier. So only the Efficient Frontier portfolios need to be considered as investment strategies. If a risk-free asset like Treasury bills is added to the asset mix, then the feasible portfolio set expands: If all funds are invested in the risk-free asset, then the return shown on the vertical axis labeled "risk-free rate" is achievable – and by dividing investment funds appropriately between the risky investments and the risk-free asset, so is any point on a line connecting the intercept with any point in the Markowitz bullet. By rotating that line upward to the left, we obtain preferable (risk, return) combinations. The rotation must stop at the CAL (capital allocation line) shown in Figure 1, as there are no feasible risky portfolios to the left of the CAL. The point of tangency of the CAL with the Efficient Frontier identifies the optimal portfolio.

A portion of the portfolio risk can be eliminated through diversification among several stocks. But the part of risk due to the general market cannot be diversified away.

In this document, I explain the ideas behind relative price volatility. This leads to the concept of **beta** as a measure of relative volatility. I show how beta can be obtained from a simple linear regression of time series data. I also show how that regression tells the investor how much of the stock's risk can be diversified away. First, however, I explore some alternative, superficially plausible ways to measure relative price volatility and show why they are deficient.

Initial Attempts at Measuring Relative Price Volatility

Quantifying the idea of relative price volatility requires us to compare the variability of a stock price to the variability of a general market average. There are several ways we might implement the comparison. Since the concept of standard deviation naturally comes to mind when we think about volatility, we might first think of comparing the standard deviation of a stock to the standard deviation of the market. We might divide the standard deviation of the stock price by the standard deviation of a general market average. For example, if we are interested in Dell, we might divide the standard deviation of Dell stock price by the standard deviation of the S&P 500 index: σ(Dell) $\div$ σ(S&P 500). To implement this idea, we would collect price data for Dell and the S&P 500 index for a period of time. The calculation for this and subsequent attempts to assess relative price volatility would be subject to the period of time and the frequency of sampling – for example, daily for 90 days, or monthly for 10 years.

But reflection suggests a couple of reasons why this idea may not be good:

1) For one thing, it may be misleading to compare the riskiness of different stocks on the basis of the standard deviation ratio. The standard deviation of a high-priced stock (like Berkshire Hathaway) is usually larger than the standard deviation of a low-priced stock, even though the variability of the high-price stock may be smaller as a percent of its price than the variability of the low-priced stock as a percent of its price. [A one percent change in the price of Berkshire Hathaway is more than the entire price of Dell! So by this measure, Berkshire may appear to be more volatile than Dell only because of the difference in scale. σ(Berkshire) $\div$ σ(S&P 500) would be vastly larger than σ(Dell) $\div$ σ(S&P 500).]
2) The second problem with the standard deviation ratio is that the ordinary standard deviation may be an inherently defective measure of stock price variability. The reason is that standard deviation measures variability around one fixed mean level – the overall mean. But both stock prices and market averages are usually Random Walks, so they meander around, rather than staying close to one mean level. The variability around the meandering trend is usually much smaller than the variability around the overall mean level. Thus, the standard deviation is not an especially meaningful statistic for a Random Walk. For the same reason, the mean is also not an especially meaningful statistic for a RW. Although the mean tells you where the RW has been (at least on average), the future of the RW has no memory of its past and will meander where it pleases. So the historical mean gives no reliable indication of where the RW will be in the future.[1]

[1] The upshot of this line of thought may be shocking to you: The most commonly used statistic in the universe – the mean – may be meaningless, depending upon the model that the data follow!

We could fix the first of these two problems by scaling the standard deviation in proportion to the mean. For example, calculate (or estimate) [σ(Dell)/μ(Dell)] ÷ [σ(S&P 500)/μ(S&P 500)].[2] However, the second problem remains: Neither the mean nor the standard deviation is very meaningful for Random Walks because RWs lack stable mean levels.

So let's try again: What about the period-to-period *changes* in stock prices as a basis for a relative volatility measure? The *changes* in a Random Walk are a Random Sample. Thus the changes have a constant mean level. So the standard deviation of the changes, which measures the variability around that mean level, is very meaningful. It would then make sense to calculate σ(Dell *change*) ÷ σ(S&P 500 *change*). This solves problem two above. But not problem one. The ratio for Berkshire, σ(Berkshire *change*) ÷ σ(S&P 500 *change*), would be much larger than the ratio for Dell, σ(Dell *change*) ÷ σ(S&P 500 *change*), because even a very small percentage change in Berkshire is greater than the entire price of a share of Dell – we still have differences in scale.

However, both problems can be solved if we calculate σ(Dell *percent* change) ÷ σ(S&P 500 *percent* change). For example, if Dell typically varies by ±1.5% and the S&P 500 typically varies by ±1%, then the volatility measure is about 1.5% ÷ 1% = 1.50. If Berkshire typically varies by 0.75%, then the volatility measure for Berkshire is about 0.75% ÷ 1% = 0.75 – even though a 0.75% change in Berkshire's price might dwarf Dell's entire price. This method provides a solution to the problem of measuring relative volatility *as I posed it*. However, this method merits only Honorable Mention!

The problem is that investors also want a solution that *ties* the volatility of Dell to market volatility so that they can tell how much market volatility affects Dell volatility. The Honorable Mention solution, σ(Dell percent change) ÷ σ(NYSE percent change), does not tell us whether volatility in the S&P 500 *leads to* volatility in Dell, or whether Dell varies independently of S&P 500. It is possible to have two stocks with exactly the same value for this ratio, and one stock marches in lock step with the S&P, up and down, but the other goes down and up in opposite directions.

The Market Model

Fortunately, the problem of tying the volatility of Dell to the volatility of the Market has a simple fix: Just multiply the Honorable Mention ratio by the correlation coefficient (ρ) between the two percentage changes:

relative riskiness measure = ρ(Dell percent change, S&P 500 percent change) x
σ(Dell percent change) ÷ σ(S&P 500 percent change).

The more the percentage changes in Dell march in lockstep with the percentage changes in the S&P 500, the higher the correlation and the larger this risk measure.

Astonishingly, this final expression turns out to be the formula for the slope in a regression – the regression of the percentage changes in the stock (as response variable Y) against the percentage changes in the market average (as predictor X). This regression model is referred to as the **Market Model**. So the Market Model for a stock is the regression model:

Percentage change in stock = α + β * percentage change in market average

[2] The standard deviation divided by the mean is a statistic called the **coefficient of variation**. This ratio expresses the variation as a proportion (or percentage) of the overall level. E.g., if the mean level is about 10 and the standard deviation is about 2, then the coefficient of variation is about 20%. In a context in which the mean is stable, the coefficient of variation can be a very meaningful and informative statistic.

The slope of this regression equation is our relative riskiness measure and is called **beta**.[3] Beta tells how much the stock is expected to go up (or down) in response to a one percent increase (or decrease) in the market average.[4] For example, if the beta of a stock is 2, then the stock is expected to go up 2% if the market goes up 1%. If the beta of a stock is 0.5, then the stock is expected to go up 0.5% if the market goes up 1%. Stocks with betas exceeding 1 are riskier than the market. Stocks with betas less than 1 are less risky than the market. If the beta of a stock is 1, then the stock has the same risk as the market. Having a high beta is not necessarily good. In bear markets, stocks with high betas go down more rapidly than most stocks, but stocks with low betas go down more slowly.

Aside: CAPM and the Market Model.

Finance makes extensive use of a theory called the Capital Asset Pricing Model (CAPM) to evaluate fair prices for risky assets. I want to warn you that CAPM is not the same as the Market Model, although it is related. The major statistical difference is that the Market Model relates the *returns* of an asset to *market returns*, whereas CAPM relates *excess returns* of an asset to *excess returns of the market*. An excess return is the difference between a return (percentage change) and a risk-free return, such as the yield on a government bond. It is called an excess return because it is the premium that an investor can earn for assuming the extra risk of owning an asset with inherent risk.

- To estimate the Market Model, one runs a regression: $R_i = \alpha + \beta R_m + \varepsilon$, where R_i is the % change (return) on asset *i* in a given time period and R_m is the % change (return) in the market.
- To estimate CAPM, one runs a regression $(R_i - R_f) = \alpha + \beta(R_m - R_f) + \varepsilon$, where the excess returns are calculated by subtracting the risk-free rate R_f from the asset return and from the market return before running the regression.

If the risk-free rate R_f is constant over the time period, then the beta of CAPM is the same as the beta of the Market Model; otherwise the two betas are different.

End aside.

The Market Model is an example of the use of time series regression *to explain* rather than *to predict*.[5] In the Market Market, we regress the percentage changes of a stock on the percentage changes in a market index in order to understand how stock returns are related to market returns. We are not trying to *forecast* stock returns from market returns. In fact, the Market Model cannot be used for forecasting. The reason is that the Market Model regresses stock returns on *contemporaneous* market returns. If we plug market returns into the Market Model equation, we will get a "forecast" of stock returns *for the same time period* – not a real *forecast* at all! We would need to use lagged market returns as predictor variable in order to have

[3] Another equivalent formula for beta is the covariance between the percentage changes in the stock and the percentage changes in the market average, divided by the variance of the percentage change in the market average. This calculation of beta is used in finance courses. In general, the slope in a simple regression of Y on X equals cov(Y,X) / Var(X), and this also equals corr(Y,X) * stdev(Y) / stdev(X).

[4] To be precise, the percentage change in the stock involves both beta and alpha – i.e., we should use the entire regression equation. But for reasons explained after the Aside, it is reasonable to expect alpha to be approximately zero.

[5] Recall that the two primary uses of time series models are to predict the future and to explain the present.

a legitimate forecasting model. The reason that we want to explain the stock returns in terms of a market index is to evaluate the (relative) risk of the stock so that we can choose stocks of appropriate risk for our portfolios.

The intercept of the regression equation market model is called **alpha**. Alpha indicates how much the stock is expected to change if the market average remains the same. Ordinarily, we expect almost all stocks to have alphas that are zero to within a statistical margin of error. This is because the market is efficient. If stock A and stock B are otherwise equal (including in risk), but A has a positive alpha and B does not, then an investor can get a free lunch by buying A. Stock A tends to increase even if the market remains level. So investors will bid up the price of A relative to B until there is no longer a free lunch. Similarly, if A has a negative alpha, then investors will sell A until its price declines enough to eliminate the negative alpha.

In order to rely on inferences based on the market model, the market model regression residuals should satisfy L, H, I, and N. It is not necessary that either the stock percentage changes or the general market percentage changes be random samples by themselves (and hence, the stock and market average be random walks), but they usually will be.

Sidebar. Use of logarithms in stock market models.

I presented the Market Model using percentage changes in prices for the response and predictor variables. I also presented the Random Walk model for stock prices in terms of changes in stock prices. In fact, for most analyses of securities prices, the logarithm of prices is generally used instead of actual prices. It is worth a couple of paragraphs to explain exactly how logs are used and why.

Let us start with the RW. A security price is a RW if its changes are a RS. Suppose we analyze the logarithm[6] of a security price and find that the changes in the log price are a RS. What does that mean? Well, by the definition of RW, it means that the log price is a RW. But what does that *really* mean?

Let Y_t denote the price of the security. Our premise is that $\log Y_t - \log Y_{t-1}$ is a RS. But $\log Y_t - \log Y_{t-1} = \log \frac{Y_t}{Y_{t-1}} = \log \frac{Y_{t-1} + (Y_t - Y_{t-1})}{Y_{t-1}} = \log\left(1 + \frac{(Y_t - Y_{t-1})}{Y_{t-1}}\right) \approx \frac{(Y_t - Y_{t-1})}{Y_{t-1}}$ since log (1 + x) $\approx x$ if x is small.[7] So $\log Y_t - \log Y_{t-1} \approx \frac{(Y_t - Y_{t-1})}{Y_{t-1}} \approx \varepsilon_t$ (a RS). Now $\frac{(Y_t - Y_{t-1})}{Y_{t-1}}$ is the percentage change in Y_t. But the percentage changes in Y_t are just the returns of Y_t. So this says that if $\log Y_t$ is a RW, then the returns in Y_t are approximately a RS. You can easily check that the steps of the argument can also be repeated in reverse. So $\log Y_t$ is a RW if and only if the returns in Y_t are approximately a RS. Since investors focus on returns (yields), using the logs of securities prices provides a natural scale for securities analysis.

[6] Throughout this discussion, log means log to the base *e*.
[7] *x* is the leading term in the Taylor's series expansion of log (1 + *x*).

Now for the Market Model. From the immediately preceding discussion, we have that $\log Y_t - \log Y_{t-1} = \log \frac{Y_t}{Y_{t-1}}$ is approximately the percentage changes (returns) of Y_t. Thus, if Y_t is a security price and X_t is a market index, then the Market Model is approximately $\log \frac{Y_t}{Y_{t-1}} = \alpha + \beta \log \frac{X_t}{X_{t-1}} + \varepsilon_t$. Thus logarithms also have a natural role in the Market Model as well.

End sidebar.

Interpreting R-square in the Market Model

In general, we can divide the influences on the price of a stock approximately into three categories: (1) general market influences, (2) industry-specific influences, (3) company-specific influences. The market model assesses the importance of the first of these categories. In general, the R-square of a regression tells what proportion of the variability of Y can be explained by the predictors. So the R-square of the market model tells what proportion of the variability of the stock returns is due to general market influences. This is risk that cannot be diversified away.[8] But it can be managed by choosing stocks with R-squares appropriate for the investor's risk tolerance. 1 – R-square assesses how important categories (2) and (3) are together. This risk can be diversified away by choosing different companies and different industries.[9] The market risk is *systematic*; the company risk is *specific*.

How Diversification Works

It is worth a couple of paragraphs to explain the intuition for why diversification works for the 1 – R-square (specific) part of stock price volatility but not for the R-square (systematic) part. The basic reason is our old friend *averaging*! When we average data, the high values tend to cancel the low values. The application of the averaging principle to stock diversification is as

[8] Except partially, by investing in other markets, like foreign stocks or real estate or gold, which may be somewhat less dependent on domestic stock market averages. Another option is to invest in synthetic securities like an ETF hedged to yield returns opposite the general market. Practical experience with the latter has been mixed, at best; the point of such synthetics is to reduce portfolio beta without eliminating the return as well.

[9] Technical Aside: This can be said somewhat more precisely: If you have a portfolio of 100 stocks in different industries and each stock has a market model R-square of 25%, then the portfolio as a whole has no less market risk than any of its individual stocks. The volatility of the whole portfolio remains just as determined in an absolute sense by general market influences as is any one of its constituent stocks. But since the 100 stocks are in different industries and companies, the influence of company and industry developments on the portfolio is reduced 100-fold in the overall portfolio variance (from 75% to 0.75%) and 10-fold in overall portfolio standard deviation (assuming the industry and company influences are independent of each other). So the importance of general market developments relative to company and industry developments is much greater for the overall portfolio than for individual stocks. But this is due to driving out the effect of individual company influences and not to increasing market influences. In large, diversified portfolios, general market conditions are more important than individual company or industry news. Small, undiversified portfolios are much more sensitive to idiosyncratic micro-scale developments in their constituent companies.

follows: When we own a portfolio of stocks, some stocks go up while others go down. The positive price changes *tend* to cancel the negative price changes. The resulting average change is much less volatile than the individual stock price changes. In fact, if the standard deviation of individual changes is σ, then the standard deviation of the portfolio change is $\sigma / \sqrt{n}$, where *n* is the number of stocks in the portfolio. [10] So by owning more stocks (which is what is meant by diversification), we can drive the portfolio volatility toward zero.

But not so fast! That theory applies if the price changes of a group of stocks are a Random Sample. The trouble is – they are not! Each stock's price change on a given day is correlated with the General Market change on that day and therefore with the price changes of all other stocks on that day. All stocks feel the effect of the General Market in a similar way, although by different amounts. Part of every stock "tries" to move in the same direction as the General Market. This violates the independence requirement for a RS. When the General Market goes up, part of the change in every stock will be an (implicit) up change. When the General Market goes down, part of the change in every stock will be an (implicit) down change. This prevents the up changes from canceling the down changes in a portfolio. Because of this correlation, when the General Market goes up, the probability is high that the portfolio will also go up; when the General Market goes down, the probability is high that the portfolio will also go down. The part of the stock price change that is *not* tied to the General Market varies randomly up or down. It is the *untied* parts of the price changes among the stocks in the portfolio that can tend to cancel each other out. The untied parts are a Random Sample. Their collective volatility can be driven toward zero by owning more stocks. But the whole of the change cannot. R-square is the part of the change that is tied to the General Market; 1 – R-square is the part that is not tied. No matter how many stocks you buy, the R-square part will not go away. The 1 – R-square part will. The smaller the R-square, the better diversification works.

[10] If the portfolio is equally weighted among the stocks. If not equally weighted, the formula is more complicated, but the basic principle remains valid.

Example: Dell and S&P 500

This example is worked out in detail in the Excel workbook "Dell Market Model.xlsx". I will only sketch the development here. The example covers three months in the summer of 2012. For Market price, I take the S&P 500 index. I compute the returns for Dell and the S&P 500 by computing the fractional change from day to day = (current price – previous price) / previous price. I regress Dell's daily returns as Y on the S&P 500's daily returns as X. Table 1 shows the regression output:

Table 1. Market model for Dell, summer 2012.

Summary	Multiple R	R-Square	Adjusted R-Square	StErr of Estimate		
	0.5956	0.3547	0.3442	0.0144		
ANOVA Table	Degrees of Freedom	Sum of Squares	Mean of Squares	F-Ratio	p-Value	
Explained	1	0.0070	0.0070	33.5365	< 0.0001	
Unexplained	61	0.0127	0.0002			
					Confidence Interval 95%	
Regression Table	Coefficient	Standard Error	t-Value	p-Value	Lower	Upper
Constant	-0.0022	0.0018	-1.1849	0.2407	-0.0058	0.0015
S&P Return	1.0972	0.1895	5.7911	< 0.0001	0.7183	1.4760

The regression equation is *Estimated Dell return* = -0.0022 + 1.0972 *S&P return*. Dell's beta over this time period is 1.0972 – only a little bigger than the average stock (which has a beta of 1.00, by definition.) In this time period, Dell was about as risky as the average stock. The R-square is 0.3547. This says that the Market is responsible for about 35.47% of Dell's price volatility. The remaining 64.53% is due to industry or Dell-specific conditions. 64.53% of Dell's price volatility can potentially be diversified away by purchasing other stocks. The alpha is only 1.1849 standard deviations below zero. So it is readily believable that Dell's alpha is really zero.

SUMMARY

The **market model** measures the riskiness of a particular stock relative to the riskiness of a broadly based market average. In this framework, risk is equated with price volatility. The model provides a numerical measure of the stock's price volatility called **beta**, which is the slope of a simple linear regression: Regress the percentage changes in the stock price (as Y) on the percentage changes in a broad market average (as X). Beta is the slope.

The Market Model:

Percentage change in stock = α + β * percentage change in market average

- $\beta = 1$ means that the stock has as much risk (price volatility) as the general market.
- $\beta > 1$ means that the stock has more risk (price volatility) than the general market.
- $\beta < 1$ means that the stock has less risk (price volatility) than the general market.

- The R-square of this regression indicates how much the stock price depends on the general market.
- R-square is the proportion of price volatility that cannot be diversified away.
- 1 – R-square indicates how much the stock price depends upon the company and its industry.
- 1 – R-square is the proportion of price volatility that can be diversified away by averaging with other stocks.

Portfolio diversification works by virtue of the averaging principle: In any set of stocks, the high price changes tend to cancel the low price changes on a given day. But the principle works imperfectly for stocks because the price changes are not independent of each other. On any given day, the high price changes do *not* balance the low price changes because of the correlation between all stocks and the Market. Part of each stock's price change represents the effect of the Market trying to move each stock in the same direction as the Market is going.