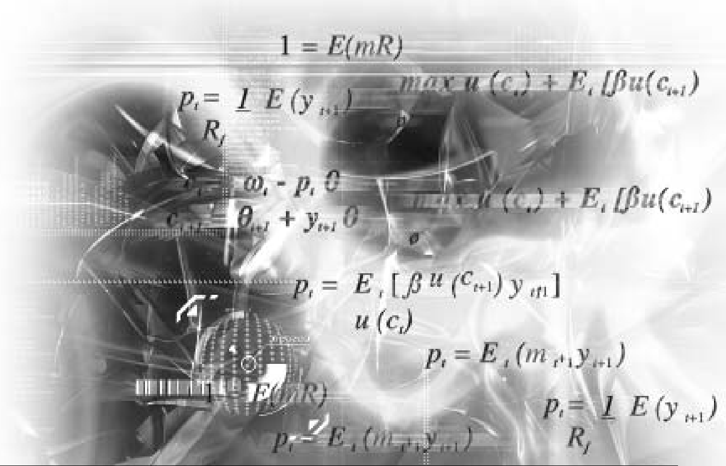
Pantisa Pavabutr, *Ph.D.*
Assistant Professor of Finance,
Faculty of Commerce and Accountancy,
Thammasat University



THE CHOICE OF FACTORS IN FACTOR ANALYSIS

DISSATISFACTION with the one factor model, in which excess return is a linear function of one factor or $r = f(F_1)$ has led to the pursuit of multi-factor models or $r = f(F_1, F_2, \ldots F_N)$. We are familiar with the special case of the one factor model under a few assumptions on investor utility and return distribution where the factor is market risk otherwise known as the capital asset pricing model (CAPM).

Computing the risk premiums from CAPM or multifactor models is straightforward. For the single factor model or CAPM, we run a time series regression from the model,

$$R_{it} - R_{ft} = a_i + ß_i \ (R_{mt} - R_{ft}) + \varepsilon_{it} \qquad t = 1, \ldots T \text{ for each } i.$$

After obtaining the betas, we run the cross-section regression,

$$E(R_i - R_{ft}) = ß_i \lambda_m$$

The multifactor case is similar. We start with the time series regression,

$$R_{it} - R_{ft} = a_i + ß_{1i} F_1 - ß_{2i} F_2 + \ldots + ß_{Ni} F_N + \varepsilon_{it} \qquad t = 1, \ldots T \text{ for each } i.$$

Then the cross-section regression has the form,

$$E(R_i - R_{ft}) = ß_{1i} \lambda_1 + ß_{2i} \lambda_2 + \ldots + ß_{Ni} \lambda_N$$

What is not so straightforward is how to deal with factor selection as empiricists often select factors which best fits data. The purpose of this article is to discuss why there is a need for multifactor models and the things we need to keep in mind to avoid adding irrelevant factors that have little economic foundation but happens to fit data well.

Why one factor isn't enough?

WHY are academics so disenchanted with one of the most classic and widely used theories in financial economics. If you are actually reading this article, I am assuming that chances are high that you have had experience or at least have an interest in trying to compute the market beta for markets in this region, at least for the Thai market. Although I have raised a couple of problems that occur when computing market beta for emerging markets in a related article titled "Note on Beta Estimation for Emerging Market Securitie", I have other observations to share here before I start discussing what literature on developed markets have to say. First, the beta estimates for stocks in emerging markets are extremely noisy, if not much noisier than beta estimates from stocks in developed markets. Not only because these markets lack the diversity in security listing in terms of the number of firms and industry variety, but also their short sample history is ridden with changes ie. companies being delisted, new companies joining the index, companies going through restructuring, policy changes, and the list goes on.

The origin of the CAPM and all factor models is based on the concept that the asset's price should be equal to the discounted value of the asset's future cashflow.

Grouping stocks together to compute portfolio betas does help reduce the noise but at the expense of losing information that by and large when we compute the cross-sectional regression of excess return on the security beta, the slope on the beta becomes insignificant. Besides, most emerging markets go through phases of extreme market correction. This turns out to be problematic when we try to understand the cross-section of stock returns because we will find that the loading on the market beta is negative. If we take this result literally, it means that high beta (risk) stocks gives lower return, which does not make much sense. However, after thinking about it a little bit, we will recognize that our analysis is based on realized return as opposed to expected returns. In other words, when we say that we want to invest in stocks with high beta because it will give us higher *"expected"* returns, it also means that we are willing to accept higher risk for *"realized"* returns that can be located on the upper end of the return distribution or the lower end of the return distribution[1]. When markets are bearish, high beta stocks have lower realized return than low beta stocks and hence the negative loading on the beta.

In the US market, the CAPM developed by Sharpe (1964), Lintner (1965), and Black (1972) received strong support prior to mid 1980's before anomalies started to emerge. For example, Banz (1981) shows that size is inversely related to expected return while Basu (1983) find that E/P is positively related to expected return. Later Fama and French (1993) find that book-to-market has strong explanatory power after controlling for ***ß***.

But academics are not ready to dismiss beta so easily. Chan and Chen (1988) show that when betas computed from portfolios formed on size have -0.988 correlation with the average size of stocks in the portfolio. This finding suggests that the true betas could be correlated to these variables.

Nevertheless, multifactor pricing tests survives and outperforms tests based on single factor market beta. Merton (1973) and Ross (1976) introduce multifactor asset-pricing models. The intuition behind Ross (1976) is expected returns of a security is determined by its covariance with common factors. Idiosyncratic movements in asset returns should not carry any risk prices. The implication of the model is that it opens up the opportunity for the empiricists to add as many forecasting variables without much economic structure. It is important then to place restrictions on the choice of factors to add. The next sub-section will describe the economic idea behind factor selection.

What are Consumption Based Models?

THE origin of the CAPM and all factor models is based on the concept that the asset's price should be equal to the discounted value of the asset's future cashflow. What deter-

[1] *Acctually, if we pick out the period of market crashes, for instnace when the NASDAQ plummeted you will also obtain this result. Besides, the weak role of the beta as reported in Fama and French (1991) in their study on NYSE, AMEX and NASDAQ for 1963-1990 and Lakonishok and Shapiro (1986) on NYSE for 1962-1981 could be a result of using realized returns for expectations.*

mines the discounted value is the marginal rate of substitution between future and current consumption levels. Economically speaking, the investor's decision on how much to consume and save for tomorrow will be such that the marginal gain from consuming more tomorrow be equal to the marginal loss of consuming less today.

To formalize this further[2], the investor's problem is,

$$\max_{\theta} u(c_t) + E_t\,[\beta u(c_{t+1})] \qquad (2.1)$$

subject to,

$$\begin{aligned} c_t &= \omega_t - p_t\theta \\ c_{t+1} &= \theta_{t+1} + y_{t+1}\theta \end{aligned} \qquad (2.2)$$

where c_t, c_{t+1}, ω_t represents consumption today, tomorrow and initial endowment or wealth. The investor maximizes his utility[3] by choosing, θ, the amount of asset he wants to buy. The asset has payoff worth y_{t+1} in the next period and the investor's discount rate is β.

Solving the first order condition of the objective function above we obtain the asset price,

$$p_t = E_t\left[\beta \frac{u'(c_{t+1})}{u'(c_t)} y_{t+1}\right] \qquad (2.3)$$

An alternative way to represent equation (2.3) is to define the stochastic discount factor or otherwise known as the pricing kernel as $m_{t+1} = \beta \frac{u'(c_{t+1})}{u'(c_t)}$, and write

$$p_t = E_t(m_{t+1} y_{t+1}) \qquad (2.4)$$

If there is no uncertainty, then we can think of the discount as the risk-free rate. Thus,

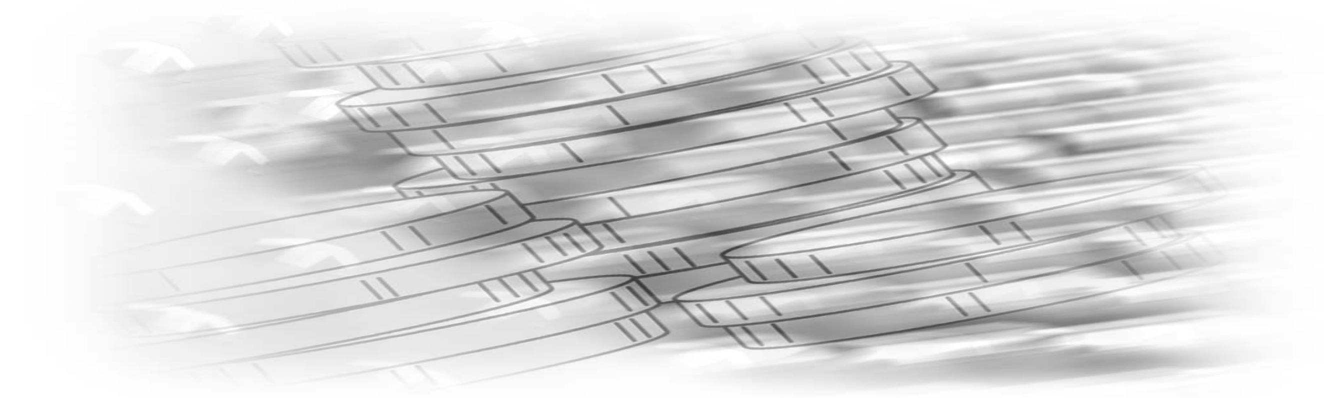
$$p_t = \frac{1}{R_f} E(y_{t+1}) \qquad (2.5)$$

Dividing equation (2.5) by price, p_t we have an alternative representation in terms of returns,

$$1 = E(mR) \qquad (2.6)$$

[2] *Much of this simplification is due to Cochrane, John H., Asset Pricing, Princeton. 2001*

[3] *The common utility form for the investor is* $u(C_t) = \frac{1}{1-\gamma} c_t^{1-\gamma}$



Since the covariance $cov(a,b) = E(ab) - E(a)E(b)$, we can write equations (2.4) as,

$$
\begin{aligned}
p_t &= E(m)E(y) + cov(m, y) \qquad (2.7)\\
&= \frac{E(y)}{R_f} + cov(m, y)\\
&= \frac{E(y)}{R_f} + \frac{cov\left[\beta u'(c_{t+1})y_{t+1}\right]}{u'(c_t)}
\end{aligned}
$$

We can see that the first term represents the discounted present value while the second term is additional risk adjustment. The equation (2.7) tells us that the asset price is lowered if its cashflow covaries positively with consumption as marginal consumption $\acute{u}(c_t)$ declines as c rises. After all, the investor is willing to pay more for a security that will stabilize consumption. Thus, the investor prefers a security that gives him high cashflow when he is not wealthy and vice versa. This is why investors care about covariance of the security cashflow with the stochastic discount factor rather than the overall volatility of individual security cashflow.

Then we apply the covariance decomposition on equation (2.6) and obtain

$$
\begin{aligned}
1 &= E(m)E(R) + cov(m, R) \qquad (2.8)\\
E(R) - R_f &= - R_f\, cov(m, R)\\
E(R) - R_f &= - \frac{cov[u'(c_{t+1}), R_{t+1}]}{E[u'(c_{t+1})]}
\end{aligned}
$$

[4] Derivations require one of the following additional assumptions: i) two-period quadratic utility, ii) two-period exponential utility and normally distributed returns, iii) infinite horizon, quadratic utility, and iid returns, iv) log utility.

The result of this exercise confirms that a security whose return covary positively with consumption must offer higher return to induce the investor to own them. In contrast, a security with return that negatively covary negatively with return will offer lower expected return.

With a little more algebra, we can write (2.8) as,

$$E(R) = R_f + \left[\frac{cov(R,m)}{\sigma^2_m}\right]\left[-\frac{\sigma^2_m}{E(m)}\right] \qquad (2.9)$$

or

$$E(R_i) = R_f + \beta_{i,m}\, \lambda_m \qquad (2.10)$$

Equation (2.10) is known as a beta pricing model where $\beta_{i,m}$ is the regression coefficient of return on the pricing kernel, m. The equation shows that expected return is a function of the quantity of risk, $\beta_{i,m}$ and the price of risk is λ_m. Here lies the idea behind the famous CAPM and the multifactor model. If the pricing kernel can be written as a linear function of the factor, then we have a factor model. In other words,

$$m_{t+1} = a + b_1 F^1_{t+1} + b_2 F^2_{t+1} + \dots b_N F^N_{t+1}$$

must hold. The famous one factor model, the CAPM is a special case where the discount factor is a linear function of the *"market portfolio"*.[4]

Conclusion on Factor Selection

SECTION 2 discusses the origin of factor pricing models. Using the consumption-based model as guideline we have a clear economic motivation in the choice of factors. The crux of it all is the potential factors added to our model should be those that are good proxies for marginal utility growth or consumption growth. These can be factors that forecast asset returns, ie. dividend yields, past security returns or factors that forecast future consumption ie. GDP growth, and interest rates.

For the case of emerging markets, size and turnover are often reported to have positive and significant impact on the cross-section of stock returns than the market beta itself. But unless we have a good theory why size and turnover are good proxies for consumption growth, we can never really dismiss the idea that these variables could just be highly correlated to the true market beta that is not observed empirically

REFERENCES :

Banz, Rolf W., (1981), "The Relationship between Return and Market Value of Common Stocks", **Journal of Financial Economics**, 9, 3-18.

Basu, Sanjoy, (1983), "Investment Performance of Common Stocks in Relation to their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis", **Journal of Finance**, 32, 663-692.

Black, Fischer, (1972), "Capital Market Equilibrium with Restricted Borrowing", **Journal of Business**, 45, 444-455.

Chan, K.C., and Nai-fu Chen, (1988), "An Unconditional Asset-pricing Test and the Role of Firm Size as an Instrumental Variable for Risk", **Journal of Finance**, 43, 309-325

Cochrane, John H., **Asset Pricing**, Princeton, 2001.

Fama, Eugene, and Kenneth French, (1993), "Common Risk Factors in Stock Returns and Bonds", **Journal of Financial Economics**, 33, 3-56.

Lintner, John, (1965), "The Valuation of Risk Assets and the Selection of Risky Investments in Sstock Portfolios and Capital Budget", **Review of Economics and Statistics**, 47, 13-37.

Merton, Robert, (1973), "An Intertemporal Capital Asset Pricing Model", **Econometrica**, 41, 467-487.

Ross, Stephen A., (1976), "The Arbitrage Theory of Capital Asset Pricing", **Journal of Economic Theory**, 13, 341-360.

Sharpe, William F., (1964), "Capital Asset Prices: A Theory of Market Equilibrium Under Condition of Risk", **Journal of Finance**, 19, 425-442.