A Proposal About a New Approach to Asset Pricing Theory: Expected Anomaly Model♣

WOONGKI LEE♦

ABSTRACT

An integrated economic view of asset pricing models including the pricing kernel model, cross-sectional model, time-series model, and expected anomaly model is presented. Several perspectives on the cross-section of expected anomalies are proposed. Modifications to the expected anomaly model are suggested. Asymptotic adjustments for errors-in-variables biases are discussed. Finally, an earlier concern against using spread portfolios is resolved.

♣ I thank Bumjean Sohn at the Korea University for helpful comments. All errors and omissions remain my responsibility.

♦ Business School, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul, 02841, Korea.
Tel.: +82 10 9759 8400.
E-mail address: w_lee@korea.ac.kr.

Disclosure Statement

Disclosure:

In accordance with the *Journal of Finance*'s disclosure policy and my ethical obligation as a researcher, I declare that I have no relevant financial interests that relate to the research described in this paper.

(Introduction)

In a long history of financial economics, there have been considerable debates on anomalies (anomalous patterns in the cross-section of expected returns) – whether the anomalies are supporting or opposing evidences for the efficient market hypothesis. These debates result in an enduring tension between two philosophies – the efficient market hypothesis versus behavioral theory. Leaving these philosophical debates aside, every time a new anomaly is discovered, a new set of empirical factors is devised to explain the anomalies. For example, Fama and French (1993) and Carhart (1997)'s multi-factor models originate from some of the most empirically powerful anomalies.

I have presented the expected anomaly model (a new model) providing different perspectives on the cross-section of expected anomalies as well as on the economy. I have proved that the capital asset pricing model (CAPM) holds for the anomalies so that an investor who attempts to exploit the discovered anomaly would fail to earn an extra profit exceeding the risk premia from the systematic risks. Next, I have suggested more general tests of market efficiency introducing the systematic risk measured on the anomaly, which is a prominent concept of the expected anomaly model.

The expected anomaly model is closely linked to the well-known cross-sectional model and time-series model. The expected anomaly model is a special case of the cross-sectional model (an expected anomaly is the sum of a risk-neutral payoff and risk premia of the anomaly, while an asset's expected return is the sum of the risk-neutral payoff and risk premia of the asset). Furthermore, the time-series model is a special case of the expected anomaly model (an alpha in the time-series model behaves like a time-invariant anomaly in the expected anomaly model). Thus, the expected anomaly model is in the middle of the two models.

Ferson et al. (1999) worry that the use of spread portfolios may lead to some spurious test results of whether the cross-section of expected returns is explained by the systematic risks. I have summarized their concern with a notion of beta-alignment. In more detail, betas in the cores-sectional model may be aligned to asset-characteristics via two sources – herding behavior or construction of the spread portfolio. No one knows the sources from which the beta-alignment is derived. If the true cause of beta-alignment is the construction, the cross-sectional model doesn't imply the risk-reward principle. The expected anomaly model, however, can avoid such problem related to the alignment so that researchers believe the test results of whether the cross-section of expected anomalies is explained by the systematic risks.

Moreover, the expected anomaly model enables researchers to test whether an empirical factor serves as the unknown true factor related to the anomaly. Through the model, one can find opposing evidences for the empirical factor if the factor doesn't seem to describe the cross-section of the anomaly from which the factor is motivated. Yet, I have reserved such empirical applications of the expected anomaly model for subsequent series of researches as its inclusion will be overly burdensome for this current paper.

I. Derivation of expected anomaly model

As argued by Ross (1978) and Harrison and Kreps (1979), lots of asset pricing models have each unique alternative form, called the pricing kernel model, which holds using a pricing kernel. The expected anomaly model featured in this paper starts with the following pricing kernel model:

$$Price\left[R_t^i\right] = E\left[m_t \cdot R_t^i\right], \tag{1}$$

where R_t^i is a payoff of asset i, $Price[R_t^i]$ is a price of asset i, and m_t is the pricing kernel that prices the payoff of all the assets. The price is equal to one if the payoff is a return and zero if the payoff is an excess return. The pricing kernel is sometimes called the stochastic discount factor or state price deflator.[1] One appropriate interpretation of the kernel is the change of macroeconomy, which affects investor's marginal utility growth.[2] Because Eq. (1) is the central alternative form of asset pricing models, the time-series model, cross-sectional model, and expected anomaly model can be started by this equation.

To further clarify, divide an asset payoff into two parts:

$$R_t^i = \mathcal{R}_t^i + \mathcal{R}_t^{i^*}. \tag{2}$$

Note that the second term on the right-hand side, $\mathcal{R}_t^{i^*}$, is uncorrelated with the kernel, m_t. Now, plug Eq. (2) into Eq. (1) and derive an expected payoff below:

$$\begin{aligned} E[R_t^i] &= \frac{Price[R_t^i]}{E[m_t]} - \frac{Cov[m_t, \mathcal{R}_t^i]}{E[m_t]} \\ &= \gamma_0^i - \frac{Cov[m_t, \mathcal{R}_t^i]}{E[m_t]}, \end{aligned} \tag{3}$$

where $\gamma_0^i \equiv Price[R_t^i]/E[m_t]$ is a risk-neutral payoff of asset i and equals the risk-free rate, R^f, if the payoff is a return and zero if the payoff is an excess return.[3]

[1] The pricing kernel is also closely related to the state price density, risk-neutral probability, equivalent martingale measure, and Green's function.

[2] The pricing kernel, $m_t = \rho \cdot u'[C_{t+1}]/u'[C_t]$, is induced by maximizing two-period utility, $U[C_t, C_{t+1}] = u[C_t] + \rho \cdot u[C_{t+1}]$, where C_t denotes an optimal consumption at time t, $u[\cdot]$ is a utility function, and ρ is called the subjective discount.

[3] The terminology, "risk-neutral payoff", comes from the fact that γ_0^i equals the product of the asset's price and the risk-free rate, given $R^f = E[m_t]^{-1}$.

The starting point of Ross (1976)'s arbitrage pricing theory (APT) is a statistical characterization of a return by factors, $\boldsymbol{f_t}$, which are also used for specifying the kernel:

$$R_t^i = \alpha^i + \boldsymbol{\beta^{i\prime}} \cdot \boldsymbol{f_t} + \xi_t^i, \tag{4}$$

$$m_t = b - \boldsymbol{f_t'} \cdot \boldsymbol{b}. \tag{5}$$

This return characterization implies that $\mathcal{R}_t^i \equiv \alpha^i + \boldsymbol{\beta^{i\prime}} \cdot \boldsymbol{f_t}$ in Eq. (2) is fully explained by the factors and $\mathcal{R}_t^{i^*} \equiv \xi_t^i$ is orthogonal to the factors and kernel.[4] Applying Eq. (3) to the APT characterization, Eq. (4), the expected return is expressed as:

$$E\left[R_t^i\right] = \gamma_0^i + \boldsymbol{\beta^{i\prime}} \cdot \boldsymbol{\gamma}, \tag{6}$$

where each element of $\boldsymbol{\beta^i}$ is the systematic risk (the amount of risk exposure of the asset to each factor) and $\boldsymbol{\gamma} \equiv Var[\boldsymbol{f_t}] \cdot \boldsymbol{b}/E[m_t]$ denotes a vector of risk premia (prices of such risk exposures). Eq. (6) means that the expected return equals the sum of the risk-neutral payoff and risk premia of the asset. This equation derived from the pricing kernel model is called the cross-sectional model.

Deriving the expected anomaly model is quite similar in the manner of the cross-sectional model. Characterize a return as:

$$R_t^i = \mathcal{A}_t^i + \boldsymbol{\theta^{i\prime}} \cdot \boldsymbol{f_t} + \varepsilon_t^i, \tag{7}$$

where $\mathcal{A}_t^i$ denotes an anomaly of asset i and ε_t^i is a residual uncorrelated with the anomaly and factors and has zero mean. In this return characterization, $\mathcal{R}_t^i \equiv \mathcal{A}_t^i + \boldsymbol{\theta^{i\prime}} \cdot \boldsymbol{f_t}$ is a function of not only the factors but also the anomaly of the return and $\mathcal{R}_t^{i^*} \equiv \varepsilon_t^i$ is orthogonal to the anomaly,

[4] ξ_t^i is uncorrelated with the factors and has zero mean.

factors, and kernel. Applying Eq. (3) to Eq. (7), the expected return is repacked as:

$$E\left[R_t^i\right] = \gamma_0^i + \boldsymbol{\theta}^{\boldsymbol{i}\prime} \cdot \boldsymbol{\gamma} - \frac{Cov\left[m_t, \mathcal{A}_t^i\right]}{E[m_t]}$$
$$= \gamma_0^i + \boldsymbol{\theta}^{\boldsymbol{i}\prime} \cdot \boldsymbol{\gamma} + \boldsymbol{\delta}^{\boldsymbol{i}\prime} \cdot \boldsymbol{\gamma}, \tag{8}$$

where each element of both $\boldsymbol{\theta^i}$ and $\boldsymbol{\delta^i} \equiv Var[\boldsymbol{f_t}]^{-1} \cdot Cov\left[\boldsymbol{f_t}, \mathcal{A}_t^i\right]$ is systematic risk measured on the factors and anomaly, respectively. Note that the deltas are designed to have the same scale of systematic risks as the thetas for two reasons, even though these are measured on the different terms.[5] First, the designed deltas are easily obtained by running a regression of the factors on the anomaly. Second, if the thetas and deltas have the same scales of systematic risks, a unit theta and delta also have the same prices of measured risks, thereby facilitating the comparison.[6]

Take expectations in both sides of Eq. (7) and plug it into Eq. (8). Then the expected anomaly model is derived as:

$$E\left[\mathcal{A}_t^i\right] = \gamma_0^i - \boldsymbol{\theta}^{\boldsymbol{i}\prime} \cdot \frac{Price[\boldsymbol{f_t}]}{E[m_t]} + \boldsymbol{\delta}^{\boldsymbol{i}\prime} \cdot \boldsymbol{\gamma}$$
$$= \gamma_0^i - \boldsymbol{\theta}^{\boldsymbol{i}\prime} \cdot \boldsymbol{\gamma_o^f} + \boldsymbol{\delta}^{\boldsymbol{i}\prime} \cdot \boldsymbol{\gamma}, \tag{9}$$

where $\boldsymbol{\gamma_o^f} \equiv Price[\boldsymbol{f_t}]/E[m_t]$ is a vector of risk-neutral payoffs of the factors and each element of it equals the risk-free rate if the factors are returns and zero if the factors are excess returns.[7] The sum of the first two terms, $\gamma_0^i - \boldsymbol{\theta}^{\boldsymbol{i}\prime} \cdot \boldsymbol{\gamma_o^f}$, is the risk-neutral payoff of the anomaly, while the γ_0^i is the

---

[5] $\boldsymbol{\delta^i}[z] \equiv z \cdot Var[\boldsymbol{f_t}]^{-1} \cdot Cov\left[\boldsymbol{f_{t_t}}, \mathcal{A}_t^i\right]$ can have different scales of systematic risks according to the value of z.

[6] Total systematic risks contained in the asset equals the sum of the thetas and deltas: $\boldsymbol{\beta^i} = \boldsymbol{\theta^i} + \boldsymbol{\delta^i}$. Obviously, Eq. (6) and Eq. (8) are equivalent.

[7] $Price[\boldsymbol{f_t}] = E[m_t \cdot \boldsymbol{f_t}] = E[m_t] \cdot (E[\boldsymbol{f_t}] + Cov[m_t, \boldsymbol{f_t}]/E[m_t]) = E[m_t] \cdot (E[\boldsymbol{f_t}] - \boldsymbol{\gamma})$.

risk-neutral payoff of the asset.[8] The third term, $\boldsymbol{\delta}^{\boldsymbol{i}'} \cdot \boldsymbol{\gamma}$, represents the risk premia of the anomaly. Thus, the expected anomaly model indicates that the expected anomaly should be equal to the sum of the risk-neutral payoff and risk premia of the anomaly.

II. Expected anomaly model in the view of other models

A. In the view of cross-sectional model

Consider a simple and practical case where all payoffs are in a return form and factors are in an excess return form[9]. Then the vector of risk-neutral payoffs of the factors, $\boldsymbol{\gamma}_{\boldsymbol{o}}^{\boldsymbol{f}}$, in Eq. (9) becomes a vector of zeros and the equation is stylized as:

$$E\left[\mathcal{A}_t^i\right] = R^f + \boldsymbol{\delta}^{\boldsymbol{i}'} \cdot \boldsymbol{\gamma}. \tag{10}$$

In this chapter, the link between Eq. (10) and the cross-sectional model (especially the CAPM) is discussed.

As pointed out by Hansen and Richard (1987), a return on any asset can be decomposed into the three orthogonal ingredients:

$$R_t^i = r_t + w^i \cdot e_t + n_t^i, \tag{11}$$

where $r_t \equiv m_t / E[m_t^2]$ equals the return on the kernel which is mapped into the payoff space, e_t, equals the excess return produced by projecting a payoff vector of ones onto the space of excess returns, n_t^i, is a residual in the excess return form, and w^i is a parameter which sets the mean of the residual to

[8] The anomaly equals the return less the factor term and residual: $\mathcal{A}_t^i = R_t^i - \boldsymbol{\theta}^{\boldsymbol{i}} \cdot \boldsymbol{f}_{\boldsymbol{t}} - \varepsilon_t^i$. The risk-neutral payoff of $R_t^i - \boldsymbol{\theta}^{\boldsymbol{i}} \cdot \boldsymbol{f}_{\boldsymbol{t}}$ is equal to the risk-neutral payoff of $\mathcal{A}_t^i$ because the residual, ε_t^i, has zero price and risk-neutral payoff.

[9] It is traditional to use the spread portfolio as a proxy for the factor. The spread portfolio is in the excess return form – has zero price and risk-neutral payoff.

zero. From the fact that the residual excess return, n_t^i, has zero mean and the three ingredients are orthogonal to each other, the first and second moments of the return are expressed as:

$$E\left[R_t^i\right] = E[r_t] + w^i \cdot E[e_t], \tag{12}$$

$$E\left[R_t^{i^2}\right] = E[r_t^2] + w^{i^2} \cdot E[e_t^2] + E\left[n_t^{i^2}\right], \tag{13}$$

respectively. An asset' return with $n^i = 0$ has its minimized variance among other returns with the same mean. Therefore, the mean-variance frontier consists of the returns with $n^i = 0$ (minimum-variance returns). And the frontier is spanned by any two minimum-variance returns.

To address the similarity between the expected anomaly model and cross-sectional model, postulate a return adjusted by the market factor:

$$R_t^{i^*} = R_t^i - \theta^i \cdot \left(R_t^M - R^f\right), \tag{14}$$

where $R_t^{i^*}$ denotes the risk-adjusted return, R_t^M is a return on the market portfolio, $R_t^M - R^f$ is the market factor (an excess return on the market portfolio relative to the risk-free rate), and θ^i is a parameter which identifies the risk-adjusted return. The risk-adjusted return indicates a return on the achievable portfolio, long a unit asset and θ^i market portfolios and short θ^i risk-free assets.

By construction, r_t and n_t^i of the market factor in Eq. (11) vanish and consequently the market factor is exactly proportional to e_t. Therefore, subtracting the market factor from an asset's return doesn't affect its n_t^i. Any risk-adjusted return, $R_t^{i^*}$, in Eq. (14) has the same n_t^i as the corresponding return, R_t^i. This fact implies that the frontier of the entire set of risk-adjusted returns doesn't change and the CAPM holds for the risk-adjusted returns:

$$E\left[R_t^{i^*}\right] = R^f + \beta^{i^*} \cdot \left(E[R_t^M] - R^f\right), \tag{15}$$

where $\beta^{i^*} \equiv Cov\left[R_t^M, R_t^{i^*}\right]/Var[R_t^M]$ is the systematic risk contained in the risk-adjusted return.[10]

Recall that the residual, ε_t^i, in Eq. (7) is uncorrelated with the market factor and has zero mean. Then the CAPM also holds for the anomalies:

$$E\left[\mathcal{A}_t^i\right] = R^f + \delta^i \cdot \left(E[R_t^M] - R^f\right). \tag{16}$$

Even if the anomaly, $\mathcal{A}_t^i = R_t^{i^*} - \varepsilon_t^i$, is not in the return form, it has the same average and systematic risk as those of the risk-adjusted return.[11] Consequently, an investor holding the risk-adjusted asset would receive the reward, $E\left[\mathcal{A}_t^i\right]$, for bearing its systematic risk, δ^i.

Testing the expected anomaly model can provide an answer to the efficient market hypothesis as well as the cross-sectional model does. For example, someone who discovers an anomaly may want to make a profit from the anomaly-chasing strategy, long the high-return anomaly stocks and short the low-return anomaly stocks adjusting some risks related to the factors. If the markets are efficient (and all the investors immediately know the anomaly), however, he would fail to yield an extra profit exceeding the risk premia from the systematic risks. In other words, under circumstances in which markets are efficient, the anomaly, $\mathcal{A}_t^i$, in Eq. (7) becomes not anomalous in terms of the systematic risks.

B. In the view of time-series model

Start with a short review of the time-series model. Take expectations in both sides of Eq. (4) and plug it into Eq. (6). Then the time series model is derived as:

[10] The original CAPM is $E\left[R_t^i\right] = R^f + \beta^i \cdot \left(E[R_t^M] - R^f\right)$, where $\beta^i \equiv Cov\left[R_t^M, R_t^i\right]/Var[R_t^M]$ is the systematic risk contained in the asset.

[11] Recall that ε_t^i is uncorrelated with the market factor and has zero mean.

$$\alpha^i = \gamma_0^i - \boldsymbol{\beta}^{\boldsymbol{i}'} \cdot \frac{Price[\boldsymbol{f_t}]}{E[m_t]}$$
$$= \gamma_0^i - \boldsymbol{\beta}^{\boldsymbol{i}'} \cdot \boldsymbol{\gamma_o^f}, \tag{17}$$

where γ_0^i is the risk-neutral payoff of the asset and each element of $\boldsymbol{\gamma_o^f}$ is the risk-neutral payoffs of the factors. If factors are excess returns, α^i equals the risk-free rate (if the payoff is a return) or zero (if the payoff is an excess return). Thus, the alphas are restricted to the same values across the assets.

The alpha seems like the risk-neutral payoff of the anomaly, $\gamma_0^i - \boldsymbol{\theta}^{\boldsymbol{i}'} \cdot \boldsymbol{\gamma_o^f}$, in Eq. (9). To clarify the link between the expected anomaly model and time-series model, specify an anomaly as:

$$\mathcal{A}_t^i = a_0^i + \boldsymbol{a}^{\boldsymbol{i}'} \cdot \boldsymbol{x_t^i}, \tag{18}$$

where $\boldsymbol{x_t^i}$ denotes a vector of asset-characteristics (e.g., the size, book-to-market equity ratio, earnings-price ratio, price-dividend ratio of a firm or group, etc.) and each element of $\boldsymbol{a^i}$ is the amount of risk exposure of the asset to each asset-characteristic. Eq. (18) is called the linear specification and it is introduced to make the problem easier – Eq. (7) turns out to be a multiple linear regression of the asset-characteristics and factors on the return.[12]

Then the systematic risks, $\boldsymbol{\delta^i}$, contained in an anomaly are repacked as:

$$\boldsymbol{\delta^i} = W^i \cdot \boldsymbol{a^i}, \tag{19}$$

where each element of $\boldsymbol{a^i}$ represents risk exposure of the asset to each asset-characteristic and $W^i \equiv Var[\boldsymbol{f_t}]^{-1} \cdot Cov\left[\boldsymbol{f_t}, \boldsymbol{x_t^{i'}}\right]$ is the risk-scaling matrix which scales $\boldsymbol{a^i}$ in order to measure $\boldsymbol{\delta^i}$. The systematic risks, $\boldsymbol{\delta^i}$, vanish under either of two conditions. First, the risk-scaling matrix, W^i, shrinks (but

[12] One may posit a non-linear specification, such as $ln\left(\mathcal{A}_t^i - 1\right) = a_0^i + \boldsymbol{a}^{\boldsymbol{i}'} \cdot \boldsymbol{x_t^i}$. Then Eq. (7) turns out to be a semi-parametric regression.

it is a matter out of control). Second, the risk exposures, $\boldsymbol{a^i}$, become minuscule. If the specified anomaly less co-moves with the asset-characteristics (i.e., $\boldsymbol{a^i} \approx \mathbf{0}$), the anomaly would rarely varies over time. It means that the deltas are close to zeros and the anomaly becomes the alpha.[13] In this context, the time-series model is a special case of the expected anomaly model.

III. Tests of market efficiency

A. Basic ideas

Consider a cross-sectional regression of the betas and expected asset-characteristics on the expected returns:

$$E\left[R_t^i\right] = \gamma_0 + \boldsymbol{\beta^{i'}} \cdot \boldsymbol{\gamma} + E\left[\boldsymbol{x_t^{i'}}\right] \cdot \boldsymbol{c}, \tag{20}$$

where $\boldsymbol{\beta^i}$ is an estimated vector of the betas, each element of $\boldsymbol{x_t^i}$ denotes the asset-characteristic, and γ_0 equals the risk-free rate if all the payoffs are returns and zero if all the payoffs are excess returns. Credited to Fama and MacBeth (1973), Black and Scholes (1974), and Banz (1981), Eq. (20) provides descriptions of where the cross-section of expected returns comes from and whether the markets are efficient.

More generally, the following cross-sectional regression,

$$E\left[R_t^i\right] = \gamma_0 + \boldsymbol{\theta^{i'}} \cdot \boldsymbol{\gamma_\theta} + \boldsymbol{\delta^{i'}} \cdot \boldsymbol{\gamma_\delta} + E\left[\boldsymbol{x_t^{i'}}\right] \cdot \boldsymbol{c}, \tag{21}$$

is formed on more information about the discriminable systematic risks, $\boldsymbol{\theta^i}$ and $\boldsymbol{\delta^i}$.[14] Note that the risk premia, $\boldsymbol{\gamma_\theta}$ and $\boldsymbol{\gamma_\delta}$, in this regression are differently marked even though these would be the same if the markets are

[13] If the delta is zero, the thetas equals the betas because $\boldsymbol{\beta^i} = \boldsymbol{\theta^i} + \boldsymbol{\delta^i}$. Then the expected anomaly model, Eq. (9), is equivalent to the time-series model, Eq. (17).

[14] Recall that the thetas and deltas are measured on the factors and anomaly, respectively.

efficient. Traditional empirical work cares much about the statistical significance and directions of the slope estimates (i.e., the risk premia), while the magnitude of the estimates is often ignored.[15] Likewise, researchers may be interested in the questions of which estimate drives out the others and whether the directions are desired in the economic sense. The ingredients contained in Eq. (21) are useful for describing the economy. $\boldsymbol{\theta^i}$ represents co-movements in business cycles and market values, $\boldsymbol{\delta^i}$ shows correlations between the business cycles and asset-specific information, and $E[\boldsymbol{x_t^i}]$ is the expected information about the specific asset.

One can test the efficient market hypothesis in a different manner, by running a cross-sectional regression of the deltas and expected asset-characteristics on the expected anomalies:

$$E[\mathcal{A}_t^i] = \gamma_0 + \boldsymbol{\delta^{i'}} \cdot \boldsymbol{\gamma} + E[\boldsymbol{x_t^{i'}}] \cdot \boldsymbol{c}. \tag{22}$$

A set of anomalies is discovered in the history of empirical studies, thereby having provoked the huge controversy over the efficient market hypothesis and sparked a lot of theoretical and empirical literature on new factors. Hence, the cross-section of expected anomalies is just as important as the cross-section of expected return in the aspect of the empirical work. Eq. (22) focuses on the expected anomalies rather than the expected returns – researchers can arrive at some notable conclusions answering the question of how the expected anomalies are distributed across assets.

B. Errors-in-variables biases

A practical version of the expected anomaly model organized on the linear specification, Eq. (18), is the following three-stage procedure:

[15] There are, however, many researches to seek plausible explanations for the equity premium puzzle − why investors are too much afraid of holding risky assets.

The first stage: N *time-series regressions,*

$$R_t^i = a_0^i + \boldsymbol{a}^{\boldsymbol{i}'} \cdot \boldsymbol{x}_{\boldsymbol{t}}^{\boldsymbol{i}} + \boldsymbol{\theta}^{\boldsymbol{i}'} \cdot \boldsymbol{f}_{\boldsymbol{t}} + \varepsilon_t^i. \tag{23}$$

The second stage: N *time-series regressions,*

$$\mathcal{A}_t^i = \alpha^i + \boldsymbol{\delta}^{\boldsymbol{i}'} \cdot \boldsymbol{f}_{\boldsymbol{t}} + \eta_t^i. \tag{24}$$

The third stage: *a cross-sectional regression,*

$$E_{(T)}[\mathcal{A}_t^i] = Int. + \boldsymbol{\delta}^{\boldsymbol{i}'} \cdot \boldsymbol{\gamma} + v^i, \tag{25}$$

where $E_{(T)}[\cdot]$ is the sample mean operator, $Int.$ captures possible misspecifications in the expected anomaly model, and v^i denotes the pricing error.[16] In accordance with the linear specification, the anomaly, $\mathcal{A}_t^i$, is constructed from a_0^i and $\boldsymbol{a}^{\boldsymbol{i}}$ estimated in the first stage. Note that the alpha in Eq. (24) is equivalent to the alpha in the time-series model.

A simple and intuitive way to implement this procedure is to run each regression in order of the stages (referring to this method as the stage-by-stage approach). However, the stage-by-stage approach causes the errors-in-variables (EIV) biases because the anomalies and deltas are measured with error. The EIV biases (misleading standard errors of the estimates) might be a crucial issue of empirical implementations of the expected anomaly model – as compared with the cross-sectional model.

There have been several suggestions on how to fix the EIV biases (see Black et al., 1972; Litzenberger and Ramaswamy, 1979; MacKinlay and Richardson, 1991; Shanken, 1992; Jagannathan and Wang, 1998). Among other things, the generalized method of moments (GMM) approach is strongly

[16] In the perspective of the generalized method of moments developed by Hansen (1982), the pricing error equals the difference between the sample and popular means of the anomaly.

advocated for its simplicity and universality.[17] In addition, the GMM approach enables additional tests on the moment conditions (e.g., the J-test).

IV. Alignment issues

A. Beta-alignment

Many empirical studies use the spread portfolio, long the high-return characteristic stocks and short the low-return characteristic stocks, as if it represents the true factor in their models. Ferson et al. (1999) have cautioned against using the spread portfolio in the cross-sectional model because the spread portfolio can be just the other form of anomalous relation to returns. Their concern is summarized with a notion of beta-alignment discussed in this chapter. The terminology, "beta-alignment", stands for the pattern that the betas in the cross-sectional model are aligned to some asset-characteristics (e.g., B/Ms).

Since stattman (1980) and Rosenberg et al. (1985)'s researches, many supporting evidences for the B/M effect (expected returns are aligned to B/Ms) have accumulated. Given the data pretested for the B/M effect, empirical success of the cross-sectional model is guaranteed if the betas are aligned to the B/Ms.[18]

Denote by $\bar{f}_t \equiv R_t^H - R_t^L$ the spread portfolio, where R_t^H is an averaged return on high-B/M stocks and R_t^L is an averaged return on low-B/M stocks.[19] Then Eq. (4) is modified for the spread portfolio:

---

[17] One may need bias-adjusted GMM estimates, which have the same value as the estimates measured through the stage-by-stage approach. MATLAB code needed for this GMM estimation is available upon request.

[18] This is a simple syllogism. The B/M effect (expected returns are aligned to B/Ms) and the beta-alignment (the betas are aligned to B/Ms) lead to guaranteed empirical success of the cross-sectional model (expected returns are aligned to the betas).

[19] This spread portfolio is nothing less than Fama and French (1993)'s HML.

$$R_t^i = \overline{\alpha}^i + \overline{\beta}^i \cdot \bar{f}_t + \bar{\xi}_t^i, \tag{26}$$

where $\overline{\alpha}^i$ and $\overline{\beta}^i$ are measured parameters and $\bar{\xi}_t^i$ is a residual. By construction, measured betas of high-B/M stocks are approximately equal to one, and measured betas of low-B/M stocks are roughly equal to negative one.

In fact, the measured betas are strongly aligned to the B/Ms from about negative one to about positive one in the data. The puzzle is from where the beta-alignment is derived. There are probably two sources of the beta-alignment. First, it may be due to herding behavior (stocks with similar B/Ms have similar time-series of returns). Second, it may be due to the construction of the spread portfolio (the spread portfolio consists of weighted returns of many assets).

If the herding behavior is the true cause of the beta-alignment, the similar values of measured betas imply the similar amounts of systematic risks − this is a desired case in the economic sense. Yet, if the main cause of the beta-alignment is the construction, the measured betas provide almost no information on the systematic risks, thereby provoking confusions about whether the cross-section of expected returns is indeed associated with the systematic risks.

B. Delta-alignment

It is an advantage of the expected anomaly model to avoid the trap of the alignment issue mentioned above. To further formulate this advantage, plug Eq. (7) into the construction of the spread portfolio, $\bar{f}_t \equiv R_t^H - R_t^L$, and repack the spread portfolio as:

$$\bar{f}_t = (\mathcal{A}_t^H - \mathcal{A}_t^L) + (\theta^H - \theta^L) \cdot f_t + (\varepsilon_t^H - \varepsilon_t^L)$$

$$= \overline{g}_t + g_t^*, \tag{27}$$

where $\overline{g}_t \equiv \mathcal{A}_t^H - \mathcal{A}_t^L$ denotes a spread in true anomalies, $\mathcal{A}_t^i$, and g_t^* is the remaining. Note that f_t is the unobservable true factor, while $\overline{f}_t$ is the spread portfolio.

Next, modify Eq. (7) for the spread portfolio:

$$R_t^i = \overline{\mathcal{A}}_t^i + \overline{\theta}^i \cdot \overline{f}_t + \overline{\varepsilon}_t^i, \tag{28}$$

where $\overline{\mathcal{A}}_t^i$ denotes a measured anomaly and $\overline{\theta}^i$ and $\overline{\varepsilon}_t^i$ are a measured theta and residual, respectively. The measured anomaly should be close to the true anomaly if both the spread portfolio and true factor have similar distributions.

Assume that the true factor is a good proxy for the true factor. Then the measured anomalies, which are close to the true anomalies, are highly correlated with $\overline{g}_t$ in Eq. (27) because, by definition, the spread in true anomalies, $\overline{g}_t$, contains the weighted true anomalies. For example, under the assumption, the correlation between $\overline{\mathcal{A}}_t^H$ and $\overline{g}_t$ is about one and the correlation between $\overline{\mathcal{A}}_t^L$ and $\overline{g}_t$ is about negative one. As a result, the deltas containing the correlations between $\overline{\mathcal{A}}_t^i$ and $\overline{g}_t$ tend to be aligned to the B/Ms − this is called the delta-alignment related to the construction (in the similar manner of the beta-alignment).

Conversely, the delta-alignment related to the construction implies that the spread portfolio is a good proxy for the true factor. Thus, the delta-alignment from either source − the herding behavior versus construction of the spread portfolio − indicates that the systematic risks are aligned to the asset characteristics. Researchers need not distinguish the sources from which the delta-alignment is derived. They can accept test results of whether the cross-section of expected anomalies is explained by the systematic risks, without worrying about the data-pretested biases.

V. Conclusions

Each asset pricing model seems to be part of the economy. The pricing kernel model discusses consumption states, the time-series model focuses on risk-neutral payoffs, the cross-sectional model addresses the whole risk premia from the business cycle, and the anomaly expected anomaly model investigates the subdivided risk premia contained in the risk-adjusted return. There's a proverb saying "the blind man touches the elephant." It means a blind man who has only touched parts of the elephant can never grasp the entire picture of the elephant. Likewise, researchers should base their conclusions on test results not just of few models but of various models.

If the economy is considered to be an elephant, the asset pricing models can be compared to parts of the elephant as described below.

Elephant's body (the pricing kernel model): *All kinds of asset pricing models are summarized with the pricing kernel model, making the model seems like organs of the economy.*

Elephant's head (the cross-sectional model): *The cross-sectional model is most popular and has been developed from the 1950s (e.g., various versions of the CAPM). This model symbolizes the risk-reward principle.*

Elephant's tail (the time-series model): *The time-series model looks like an appetizer served before the cross-sectional model. However, the time-series model overly rejects empirical factors (see Lo and MacKinlay, 1990), while the cross-sectional model rarely rejects the factors (see Ferson et al., 1999).*

Elephant's legs (the expected anomaly model): *The expected anomaly model is motivated by the empirical ground of the anomalies. This model is robust to the data-pretested biases such as the B/M effect.*

The expected anomaly model provides different perspectives on not only the economy but also the cross-section of expected anomalies. There are a number of researches that can be empirically undertaken with this model because there remain many anomalies that are not yet clearly resolved. Researchers can test whether an anomaly is really anomalous in terms of the risk-reward principle or whether an empirical factor serves as the true factor related to the anomaly. Reserving further empirical applications of the expected anomaly model for subsequent papers, I hope this model will provoke the generation of other ideas to help explain asset pricing phenomena.

REFERENCES

Banz, Rolf W., 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3-18.

Black, Fischer, Michael C. Jensen, and Myron Scholes, 1972, The capital asset pricing model: Some empirical tests, in Michael Jensen, ed.: *Studies in the Theory of Capital Markets* (Praeger, New York).

Black, Fischer, and Myron Scholes, 1974, The effects of dividend yield and dividend policy on common stock prices and returns, *Journal of Financial Economics* 1, 1-22.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57-82.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33. 3-56.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607-636.

Ferson, Wayne E., Sergei Sarkissian, and Timothy Simin, 1999, The alpha factor asset pricing model: A parable, *Journal of Financial Market 2*, 49-68.

Hansen, Lar Peter, 1982, Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1029-1054.

Hansen, Lars Peter, and Scott F. Richard, 1987, The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models, *Econometrica* 55, 587-613.

Harrison, J. Michael, and David M. Kreps, 1979, Martingales and arbitrage in multiperiod securities markets, *Journal of Economic Theory* 20, 381-408.

Jagannathan, Ravi, and Zhenyu Wang, 1998, An asymptotic theory for estimating beta-pricing models using cross-sectional regression, *Journal of Finance* 53, 1285-1309.

Litzenberger, Robert H., and Krishna Ramaswamy, 1979, The effect of personal taxes and dividends on capital asset prices, *Journal of Financial Economics* 7, 163-195.

Lo, Andrew W., and A. Craig MacKinlay, 1990, Data-snooping biases in tests of financial asset pricing models, *Review of Financial Studies* 3, 431-467.

MacKinlay, A. Craig, and Matthew P. Richardson, 1991, Using generalized method of moments to test mean-variance efficiency, *Journal of Finance* 46, 511-527.

Rosenberg, Barr, Kenneth Reid, and Ronald Lanstein, 1985, Persuasive evidence of market inefficiency, *Journal of Portfolio Management* 11, 9-17.

Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341-360.

Ross, Stephen A., 1978, A simple approach to the valuation of risky streams, *Journal of Business* 51, 453-475.

Shanken, Jay, 1992, On the estimation of beta-pricing models, *Review of Financial Studies* 5, 1-33.

Stattman, Dennis, 1980, Book values and stock returns, *The Chicago MBA: A Journal of Selected Paper* 4, 25-45.