# 4 The APT

The no-arbitrage results that we just derived are very general. They are so general that they are very difficult to use. There is a notion among economists that ”you get what you pay for” in terms of assumptions and results. We made very weak assumptions in the previous section, and we got fairly weak results. What would an MBA say to you if you told her ”there is some positive pricing kernel out there but I have no idea what it is?”

We will turn now to the APT. The APT is an older model (published by Ross in 1976) that has been used quite a bit by both economists and practitioners. It makes slightly stronger assumptions and gets, as a result, stronger predictions. Like the no-arbitrage results of the last section, the APT does not rely on much economic argument. It only assumes that returns are driven by a linear factor model and that there are no “asymptotic arbitrage opportunities.”

The stuff we will talk about today is covered in chapter 7 of Ingersoll.

## 4.1 Notation and Assumptions

Suppose that returns are driven by a set of factors, $f_1, f_2, ..., f_k$, such that

$$r_{it} = \alpha_i + \beta_{i1} f_{1t} + ... + \beta_{ik} f_{kt} + \varepsilon_{it}, \tag{36}$$

or, in vector notation,

$$\mathbf{r} = \alpha + \beta \mathbf{f} + \varepsilon. \tag{37}$$

Notice that this looks very much like a regression equation. The APT is loosely based on multivariate regression analysis. Notice also that the factors are macroeconomic aggregates rather than firm-specific characteristics. This is what is called a “factor model.” The APT assumes that returns are generated by a factor model and then it derives the expected returns relation that follows from that assumption. Figuring out what the APT really means is a little bit tricky. At first glance, it seems that we are assuming a factor model and then doing a lot of math to arrive at the same factor model. The APT assumes that returns are generated by a factor model and then it shows that, with no arbitrage, each asset's expected return is a linear function of the asset's "sensitivities" to the factors or its "factor loadings."

To make the derivation of the APT simple we need to make some assumptions about the factor model. Let

$$\begin{aligned} E(\varepsilon) &= \mathbf{0}, \\ E(\mathbf{f}) &= 0, \\ E(\mathbf{f}\mathbf{f}^{'}) &= \mathbf{I}, \\ E(\varepsilon\mathbf{f}^{'}) &= 0, \\ E(\varepsilon\varepsilon^{'}) &= \mathbf{G} \end{aligned} \tag{38}$$

where $\mathbf{G}$ is a diagonal matrix with bounded diagonal elements, $\{G_{ii}\} \equiv s_i^2 < S^2$. These assumptions are mostly innocuous. The only assumptions with teeth are that a factor structure describes the returns generating process and that $\mathbf{G}$ is diagonal.

## 4.2 Ideas in APT

The APT is derived by combining two separate ideas. The first idea is the law of large numbers, an asymptotic statistical concept.

**The Law of Large Numbers.** *Let $z_i$ represent a sequence of iid. random variables with finite expectation, $\mu_z$. Then for any $\epsilon > 0$*

$$\lim_{n\to\infty} \text{Prob}\left(\left|\frac{1}{n}\sum_{i=1}^{n} z_i - \mu_z\right| < \epsilon\right) = 1, \tag{39}$$

*or, equivalently,*

$$\text{plim}\left(\frac{1}{n}\sum_{i=1}^{n} z_i\right) = \mu_z. \tag{40}$$

You may have run across the law of large numbers in an econometrics class before - it is frequently used to prove consistency, for example. Here we will use it to argue that if we have enough assets in our portfolios, we can create portfolios that have no residual variance. In other words, we can use the LLN to make our portfolios have returns that look like

$$\mathbf{r} = \alpha + \beta\mathbf{f}, \tag{41}$$

which is just like the form that we assumed above except that it does not contain $\varepsilon$, the error term.

The second idea in the APT is that in the absence of arbitrage opportunities, securities that satisfy the exact factor structure outlined above must have expected returns that are linear in $\beta$. To see this, it is useful to think of the one factor case. Suppose that there is only one factor and we can find two different well diversified portfolios that are both sensitive to this factor,

$$\begin{aligned} r_1 &= \alpha_1 + \beta_1 f, \\ r_2 &= \alpha_2 + \beta_2 f \end{aligned} \tag{42}$$

By our assumptions, $E(r_1) = \alpha_1$, and likewise for asset 2. We will invest $w$ in portfolio 1 and $1 - w$ in portfolio 2. This gives us a return of

$$r_p = w\alpha_1 + (1-w)\alpha_2 + [w\beta_1 + (1-w)\beta_2]f. \tag{43}$$

We can construct a risk-free portfolio by setting the term in brackets to zero. We choose

$$w^* = \frac{\beta_2}{\beta_2 - \beta_1}, \tag{44}$$

and we know that the resulting portfolio must have a return equal to the risk-free rate since it is riskless. Thus,

$$w^*\alpha_1 + (1 - w^*)\alpha_2 = r_f, \tag{45}$$

or

$$\left(\frac{\beta_2}{\beta_2 - \beta_1}\right)(\alpha_1 - \alpha_2) = r_f - \alpha_2, \tag{46}$$

or

$$\frac{\alpha_1 - \alpha_2}{\beta_1 - \beta_2} = \frac{\alpha_2 - r_f}{\beta_2}. \tag{47}$$

By symmetry, the same thing holds if we interchange the 1 and 2 subscripts. Notice that

$$\frac{\alpha_1 - \alpha_2}{\beta_1 - \beta_2} = \frac{\alpha_2 - \alpha_1}{\beta_2 - \beta_1}, \tag{48}$$

so that

$$\frac{\alpha_2 - r_f}{\beta_2} = \frac{\alpha_1 - r_f}{\beta_1}. \tag{49}$$

Finally, this implies that

$$\alpha_i = r_f + \lambda\beta_i = E(r_i). \tag{50}$$

The APT math holds for any number of securities, of course. Suppose we try it with two factors and three assets, so that

$$\begin{aligned} r_1 &= \alpha_1 + \beta_{11} f_1 + \beta_{12} f_2, \\ r_2 &= \alpha_2 + \beta_{21} f_1 + \beta_{22} f_2, \\ r_3 &= \alpha_3 + \beta_{31} f_1 + \beta_{32} f_2. \end{aligned} \tag{51}$$

Now if we find weights such that $\sum_{i=1}^{3} w_i\beta_{i1} = 0$ and $\sum_{i=1}^{3} w_i\beta_{i2} = 0$ then we will have formed a riskless portfolio again. This portfolio will have to satisfy $\sum_{i=1}^{3} w_i\alpha_i = r_f$. In matrix form,

$$\begin{pmatrix} \alpha_1 - r_f & \alpha_2 - r_f & \alpha_3 - r_f \\ \beta_{11} & \beta_{21} & \beta_{31} \\ \beta_{12} & \beta_{22} & \beta_{32} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{52}$$

The vector of weights cannot equal a zero vector because it must sum to one. The matrix in this problem must be singular (that's from a matrix algebra theorem), so the first row must be a linear combination of the last two rows, which we assume are linearly independent. Thus,

$$E(r_i) = \alpha_i = r_f + \lambda_1\beta_{i1} + \lambda_2\beta_{i2} \tag{53}$$

in this case, and

$$E(r_i) = r_f + \sum_{j=1}^{k} \lambda_j\beta_{ij} \tag{54}$$

in the more general case of $k$ factors.

## 4.3 Derivation

To derive the APT, we need to define an "asymptotic arbitrage." We will then assume that no asymptotic arbitrage exists and derive the implication of that assumption.

**Asymptotic Arbitrage.** *An asymptotic arbitrage is a sequence of portfolios,* $\theta^n =$

$\{\{\theta_1^2, \theta_2^2\}, \{\theta_1^3, \theta_2^3, \theta_3^3\}, ...\{\theta_1^n, \theta_2^n, ..., \theta_n^n\}\}$, *that satisfy:*

$$\begin{gathered} \textstyle\sum_{i=1}^n \theta_i^n = 0, \\ \textstyle\sum_{i=1}^n \theta_i^n E(r_i) \geq \delta > 0, \\ \textstyle\sum_{i=1}^n \sum_{j=1}^n \theta_i^n \theta_j^n \sigma_{ij} \to 0. \end{gathered} \tag{55}$$

Why should no asymptotic arbitrage exist? How could you exploit such an arbitrage?

**The Arbitrage Pricing Theory (APT).** *If returns are generated by the factor model defined in (37) and (38) and there are no asymptotic arbitrage opportunities, then there exists a linear pricing model that gives expected returns with a mean squared error of zero,*

$$\alpha_i - \lambda_0 - \sum_{j=1}^k \lambda_j \beta_{ij} = \nu_i \tag{56}$$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n \nu_i^2 = 0. \tag{57}$$

Think of a linear projection (or a least squares regression) of the vector $\alpha$ onto the space spanned by the matrix $\beta$ and a vector of ones. We could express the result of this "regression" as

$$\alpha_i = \lambda_0 + \sum_{j=1}^k \lambda_j \beta_{ij} + \nu_i \tag{58}$$

Since the error terms in regressions are orthogonal to the regressors, the conditions:

$$\sum_{i=1}^n \nu_i = 0, \tag{59}$$

$$\sum_{i=1}^n \nu_i \beta_{ik} = 0, \quad \forall k, \tag{60}$$

must hold. Now consider the arbitrage portfolio (a portfolio with zero investment),

$$\theta_i = \frac{\nu_i}{\sqrt{n \sum_{i=1}^{n} \nu_i^2}} = \psi \nu_i, \tag{61}$$

where $\psi = \frac{1}{\sqrt{n \sum_{i=1}^{n} \nu_i^2}}$ for notational convenience. The profit on this portfolio is

$$\psi \sum_{i=1}^{n} \nu_i r_i = \psi \sum_{i=1}^{n} \nu_i [\alpha_i + \sum_{j=1}^{k} \beta_{ij} f_j + \varepsilon_i] = \psi \sum_{i=1}^{n} \nu_i (\alpha_i + \varepsilon_i), \tag{62}$$

while the expected profit is,

$$\psi \sum_{i=1}^{n} \nu_i \alpha_i = \psi \left[ \lambda_0 \sum_{i=1}^{n} \nu_i + \sum_{j=1}^{k} \lambda_j \sum_{i=1}^{n} \nu_i \beta_{ij} + \sum_{i=1}^{n} \nu_i^2 \right] = \frac{1}{\sqrt{n}} \sqrt{\sum_{i=1}^{n} \nu_i^2}. \tag{63}$$

The variance of the profit is

$$\frac{\sum_{i=1}^{n} \nu_i^2 s_i^2}{n \sum_{i=1}^{n} \nu_i^2} \leq \frac{S^2}{n} \tag{64}$$

Now suppose that the APT theorem above is false. If it is false then the expected profit of this arbitrage portfolio is non-zero while the variance goes to zero with $n$. This would be an arbitrage opportunity, so it can't exist in equilibrium. Therefore, the expected profit term must also vanish, meaning that $\frac{\sum_{i=1}^{n} \nu_i^2}{n} \to 0$ and the theorem holds. This concludes our proof of the APT theorem.

There is quite a large literature about the APT. People have worried about whether or not the APT is testable, how precisely it is likely to fit, and other issues. We will not review all of these developments here. You can read about them in Ingersoll if you want to.

## 4.4 Connections

The APT can be made a little more transparent with the tools that we have developed in previous sections. For example, consider the result of applying our pricing kernel result to the APT equations without error,

$$
\begin{aligned}
\text{APT}: \; r_{it} &= \alpha_i + \beta_{i1} f_{1t} + ... + \beta_{ik} f_{kt}, \\
E[M(1+r_{it})] = 1 &= E[M](1+\alpha_i) + \beta_{i1} E[M f_{1t}] + ... + \beta_{ik} E[M f_{kt}], \\
1 + r_f &= (1+\alpha_i) + \beta_{i1} E[M f_{1t}](1+r_f) + ... + \beta_{ik} E[M f_{kt}](1+r_f), \\
\alpha_i &= r_f - \beta_{i1} E[M f_{1t}](1+r_f) - ... - \beta_{ik} E[M f_{kt}](1+r_f), \\
\alpha_i &= r_f + \lambda_1 \beta_{i1} + ... + \lambda_k \beta_{ik}
\end{aligned} \tag{65}
$$

where $\lambda_j = -E[M f_{jt}](1+r_f), \;\; j = 1, \; 2, \; ... \; k$. Thus, using our previously derived arbitrage results, we can prove the APT result much more easily.

It is also useful to me to think of what the APT assumptions mean in terms of the finite state world that we have assumed in previous sections. The APT assumptions are basically equivalent to the condition that the $N \times S$ matrix of gross returns in various states, $\mathbf{R}$ can be decomposed as follows:

$$
\mathbf{R} = \sum_{j=1}^{k} \beta_j f_j^{'} + \epsilon \tag{66}
$$

where $\beta_j$ is an $N$-vector of security betas on factor $j$, $f_j$ is an $S$-vector that contains factor $j$'s realizations in each state, and $\epsilon$ has the property that $\epsilon^{'}\epsilon$ is a diagonal matrix. The outer product terms above, $\beta_j f_j^{'}$, each consist of an $N \times S$ matrix of rank 1. Thus, the middle term of equation (66) is (at most) of rank $k$.

This representation illustrates the truly important assumption of the APT. If $k$ is reasonably small, the APT assumes that the almost all returns can be described by a returns matrix of low rank. According to the APT, all assets have returns that are essentially linear combinations of a few important factors. But there is really nothing special about the way that the APT decomposes the matrix $\mathbf{R}$. So if a simpler way to decompose $\mathbf{R}$ is available, it may be possible to develop a more parsimonious APT. There is a literature on a "nonlinear" APT that attempts to exploit this fact.[3]

## 4.5 Homework Problems

1. Suppose you found that an APT model with the market return and an unexpected inflation factor fits expected returns quite well:

$$\begin{aligned} r_{it} &= \alpha_i + \beta_{im} r_{mt} + \beta_{i\iota}\tilde{\iota} + \varepsilon_{it} \\ E[r_{it}] &= \alpha_i = r_f + \lambda_m \beta_{im} + \lambda_\iota \beta_{i\iota}. \end{aligned} \tag{67}$$

   Suppose, furthermore, that you found $\lambda_\iota = -0.4$. What does the fact that $\lambda_\iota < 0$ mean? What sign do you expect $\lambda_m$ to have?

2. What is the stochastic discount factor implied by the APT model if the APT has only two factors? You should be able to relate the stochastic discount factor to parameters of the APT model.

[3] Bansal, R. and S. Viswanathan, 1993, No Arbitrage and Arbitrage Pricing: A New Approach, *Journal of Finance*, XLVIII, 1231-1262.