2 General Properties of Asset Pricing Models

We are interested in pricing assets. Assets are either financial securities or actual means of production (e.g. land, machines) that people buy in order to earn a profit. The form of profit that we will analyze is called a return, and is generally defined as the profit earned on an asset divided by the original price of the asset. We want to explain the fact that some assets earn high returns and others earn negative returns. We want to be able to write down an equation that says $E(r_i) = f(x_i)$, where E is the expectations operator, r_i is the return on asset i, and x_i is a vector of characteristics of asset i. To figure out which asset characteristics should matter and which function of characteristics makes the most sense, we make up asset pricing models.

In this section we describe a generic asset pricing model. It turns out that all asset pricing models can be described as special cases of one single equation. In general, we want a way to describe the expected return for any asset. The equation we discuss in this section describes expected returns.

2.1 $1 = E(RM)$

The "mother of all asset pricing models" (my term) is a convenient way to summarize all of asset pricing. We will derive it in a number of ways, with a number of different assumptions. The model is:

$$E(R_i M) = 1, \quad i = 1, 2, \ldots, n. \tag{1}$$

where E stands for expectation, R_i is the gross return on asset i (gross returns are equal to net returns plus one, $R = r + 1$), and M is known as either the **pricing kernel**, the **stochastic discount factor**, or the **intertemporal marginal rate of substitution**. Almost all asset pricing models can be expressed as a special case of $E(R_i M) = 1$. Models are distinguished by their specification of M. Notice that while

R_i is a ¯rm-speci¯c quantity, a single pricing kernel works for all assets. If we assume that there is a risk free asset, then the model can be expressed as:

$$E(R_i^e M) = 0; \quad R^e = R_i - R_f; \tag{2}$$

since equation 1 must hold for the risk free rate, R_f, as well as R_i.

Much of the class will be spent motivating various forms for the pricing kernel, M. For now, think of M as an aggregate measure of discomfort. In many of the models we will consider, M is the marginal utility of consumption. In bad states of the world, M is high; M is low in good states. Viewing M this way, think about a security that pays o® relatively well in bad states. Such a security will have a high value when M is high and a low value when M is low. This is a low risk security, perhaps even an insurance contract of some sort. Now think of a security that pays o® well when M is low, but does poorly when M is high. This second security is fairly risky. In order for equation 2.1 to hold, the second security will have to have higher payo®s on average than the ¯rst security. This is the intuition behind all of asset pricing. Assets that help to smooth consumption are relatively safe and hence do not pay high expected returns.

2.2 Alternative Representations

There are several di®erent ways to express the $E(RM) = 1$ result. Which representation is preferred depends on the application at hand.

2.2.1 Factor-Based Models

One common (and traditional) representation is the factor-based model. Demonstrating the link between factor-based models and the pricing kernel model should help you see that the fairly risky security described above must pay higher returns than the low risk security. You may be familiar with factor-based models like the CAPM or the APT.

Using the fact that

$$COV(A, B) = E(AB) - E(A)E(B), \tag{3}$$

a simple relation between factor-based models and stochatic discount models can be derived. The fact that

$$E(R_i M) = 1 = COV(R_i, M) + E(R_i)E(M) \tag{4}$$

implies that

$$E(R_i) = \frac{1}{E(M)} - \frac{COV(R_i, M)}{E(M)} \tag{5}$$

$$= \frac{1}{E(M)} - \frac{VAR(M)}{E(M)} \left[\frac{COV(R_i, M)}{VAR(M)} \right] \tag{6}$$

$$= \gamma_0 - \gamma_M \beta_{iM}. \tag{7}$$

Thus, any stochastic discount factor model can be expressed as a factor-based model by realizing that the factor in the factor-based model is the pricing kernel, M. Now we can verify the intuition of the previous section. An asset with returns that are positively correlated with M pays relatively low returns.

The risk premium in the factor-based model is equal to the variance of M divided by the mean of M, and the intercept equals one over the mean of M. Interpretting these terms requires some thought about the expected and permissible values of the pricing kernel, which happens to be discussed in section 2.4.

2.2.2 State Prices

If we are willing to assume that there are a finite number of possible future states of the world, we can express the pricing kernel in other useful ways. Assume that there are S possible states of the world, indexed by s, and that there are N assets, indexed by i. Let the $N \times S$ matrix D represent the payoffs of the N securities in the S states, and let security prices be given by the N-vector q. Denoting portfolio weights by the N-vector θ, the portfolio θ costs $q \cdot \theta$ and it pays off $D'\theta$.

Within this framework, a **state-price** vector is a strictly positive S-vector, ψ, with $q = D\psi$. We will interpret state prices, ψ_j, later. If state prices exist, then the $1 = E(RM)$ results described in section 2 hold. Remember that all uncertainty in this model is about which state will be revealed to be true next period. If we use the symbol π_s to mean the probability of state s then the stochastic discount factor result is the same as

$$\sum_{s=1}^{S} \pi_s M_s R_{is} = 1. \tag{8}$$

The state price relation to prices can be written as

$$q_i = \sum_{s=1}^{S} D_{is}\psi_s \quad \text{or} \quad \sum_{s=1}^{S} R_{is}\psi_s = 1. \tag{9}$$

So, all we need to do to get the pricing kernel result from the state-price representation of the model is to define

$$M_s = \frac{\psi_s}{\pi_s}. \tag{10}$$

So the existence of state prices means that there is a pricing kernel that makes the mother of all asset-pricing models fit.

2.2.3 Risk-neutral Probabilities

We can perform the same trick that produced a pricing kernel in the previous section with any sort of “probability” that we care to consider. Many results use what is sometimes called “risk-neutral probabilities” which are constructed as

$$\phi_0 = \sum_{s=1}^{S} \phi_s; \tag{11}$$

$$\hat{\phi}_s = \frac{\phi_s}{\phi_0}. \tag{12}$$

The elements of $\hat{\phi}$ can be considered probabilities because they sum to one. With these risk-neutral “probabilities”, the price of any security is the discounted expected payoff with these “probabilities,”

$$q_i = \phi_0 \hat{E}(D_i) = \phi_0 \sum_{s=1}^{S} \hat{\phi}_s D_{is}. \tag{13}$$

Note that this construction requires that $\phi_0 = \frac{1}{R_f}$.

Sometimes in dynamic contexts people refer to a “change of measure” or “equivalent martingale measure” result that employs probabilities like these. We will discuss these results in more detail later.

Each of these representations (the pricing kernel model, the factor-based model, state prices and risk-neutral probabilities) is essentially equivalent, and each will appear as an implication that we derive from some asset pricing model.

2.3 When Does 1 = E(RM) Hold?

Why is 1 = E(RM) such a convenient model? Because we can expect it to hold under very general circumstances. Supposing again that all the uncertainty about the economy can be summarized by a discrete set of S states, some sort of pricing kernel will always exist as long as the problem

$$q = D\psi; \tag{14}$$

or equivalently,

$$R\psi = 1_S \tag{15}$$

has a solution. Again, the $N \times S$ matrices R and D represent the gross returns and the payoffs of the N securities in the S states, repsectively. Within this framework, we can argue that some sort of pricing kernel will always exist.

First of all, consider the case in which $S = N$ and D is of full rank. In this case, we say that markets are complete. In a world with complete markets, we can construct Arrow-Debreu securities or state contingent claims. A state contingent claim pays off one unit of consumption if a particular state is realized, and it pays zero otherwise. It is easy to create Arrow-Debreu securities from ordinary securities simply by inverting the payoff matrix. To create Arrow-Debreu securities, we need an $S \times S$ matrix of weights, W, such that

$$WD = I: \tag{16}$$

Of course, the matrix $W = D^{-1}$ fits this description.

Knowing this, we can see now why the state-price vector has its name. If we apply the weight matrix that we just found to the definition of a state-price vector given in section 2.2.2,

$$q = D\psi \quad \Rightarrow \quad D^{-1}q = \psi; \tag{17}$$

we see that the cost of one unit of payoff in state j is just ψ_j. Furthermore, the price of any security is just the sum of its payoffs in particular states times the cost of a unit of payoff in each state. Agents can use state contingent claims to insure against all future states of the world, so they should smooth consumption perfectly in complete market worlds.

What happens when $N > S$ and D is of full rank? Then we have what are called redundant securities. If we assume that a very weak form of no-arbitrage holds then redundant securities will be priced correctly by a pricing kernel in complete markets.

The Law of One Price. Two bundles with exactly the same characteristics (e.g. portfolios with equal payo®s in all states) have to sell for the same price.

In linear algebra terms, the N ¡ S rows of redundant asset payo®s have to be linear combinations of the ¯rst S rows of D. In other words, redundant securities are basically portfolios of state contingent claims. As such, they must be correctly priced by state prices for the law of one price to hold.

Now consider the case where $N < S$. In this case, we say that markets are not complete. People cannot insure against all possible future states because a complete set of contingent claims is not avaiable. When this is the case, relying on simple linear algebra results, we know that there are an in¯nite number of solutions to problem 15.

Thus, under fairly general circumstances, we know that a pricing kernel exists. This is not a very interesting result, however, since we have no way of identifying pricing kernels and there is not a unique pricing kernel. We need to know more about the pricing kernel's properties in order to get anything useful out of the $1 = E(RM)$ type of model. We can get more speci¯c properties for M by making stronger asumptions to set up our model.

2.4 Empirical Restrictions on M

Without any additional assumptions, we can derive some empirical restrictions on the pricing kernel, M. First of all, since $E(R_i M) = 1$ holds for the risk-free asset, it is true that

$$E(M) = \frac{1}{R_f}: \tag{18}$$

This implies that the intercept in the factor-based model implied by $E(R_i M) = 1$ is the risk-free rate. This is consistent with the CAPM and other factor-based models, and it satisfies the requirement that $\psi_0 = \frac{1}{R_f}$ implied by the risk-neutral probability representation. Equation 18 also implies that the risk premium term, $\gamma_M = \frac{VAR(M)}{E(M)}$, is positive. Since the risk premium is positive, it must be the case that riskier securities have negative betas on M in the factor-based model. This is consistent with the description of M in section 2.1. Assets with payoffs that are positively correlated with M could actually have negative expected returns (e.g. insurance contracts).

Second, we can make statements about the variability of M. We will use the fact that

$$COV(A; B) = \rho(A; B)\sigma(A)\sigma(B); \tag{19}$$

where $\rho(A; B)$ is the correlation coefficient between A and B, and $\sigma(j)$ is the standard deviation of j. Applying this fact to the excess returns version of our stochastic discount factor model,

$$0 = E(MR_i^e) = E(M)E(R_i^e) + COV(M; R_i^e) \tag{20}$$

implies that

$$\sigma(M) = \frac{-E(M)E(R_i^e)}{\rho(M; R_i^e)\sigma(R_i^e)}. \tag{21}$$

Since $|\rho| \leq 1$,

$$\sigma(M) \geq \frac{E(M)E(R_i^e)}{\sigma(R_i^e)}. \tag{22}$$

Furthermore, since $E(m) = \frac{1}{R_f}$,

$$\sigma(M) \geq \frac{1}{R_f}\left[\frac{E(R_i^e)}{\sigma(R_i^e)}\right]. \tag{23}$$

Notice that the term in brackets is security i's Sharpe Ratio. This relation is used to construct the Hansen-Jaganathan bound, a region of permissible values for the moments of M.[1] The HJ bound is often used by empiricists to examine whether

[1] Hansen, L. P., and R. Jaganathan, 1991, "Implications of security market data for models of dynamic economies," Journal of Political Economy, 99, 225-261.

particular models for M are reasonable.

2.5 Conditional Models

The models discussed above have all involved unconditional expectations. Many of the models we will consider will be conditional models, taking the form

$$E_t(R_{i,t+1}M_{t+1}) = 1, \tag{24}$$

where E_t means "the expected value conditional on all information available at time t." Conditional expectations are defined carefully in Duffie (pg. 224) and elsewhere.

In general, conditional and unconditional models have different empirical implications. The conditional model above can be written as a factor-based model:

$$E_t(r_{t+1}) = r_f + \gamma_t \beta_t \tag{25}$$

where the i and M subscripts have been dropped to avoid confusion, and the net return, r, replaces the gross return, R. The t subscripts that appear here denote conditional moments. So, for example, the β_t in this model consists of a conditional covariance divided by a conditional variance. The t subscripts mean that both betas and risk premiums can change with time and information.

People often convert conditional models to unconditional models using something called the **law of iterated expectations**. The law of iterated expectations states that if the set of possible events B is a subset of the set of possible events A (B represents more information than A) then $E[E(x|B)|A] = E(x|A)$. Applying this law to our conditional model, the expected value of $E_t(r)$ is just $E(r)$. Applying it to the right hand side is more complicated - γ_t and β_t are both random variables at time t. The unconditional model turns out to be,

$$E(r_{t+1}) = r_f + E(\gamma_t)E(\beta_t) + COV(\gamma_t; \beta_t): \tag{26}$$

2.6 Homework Problems

1. Write the CAPM as a stochastic discount factor model. The CAPM is usually written as

$$E(r_i) = r_f + \gamma\beta_i; \tag{27}$$

where β_i is security i's beta on the market portfolio and γ is some positive risk premium. Verify that you can get a relation for expected returns in the same form as (27) by using the stochastic discount factor,

$$M = a + bR_m; \tag{28}$$

where R_m is the gross return on the market portfolio. You probably want to start with,

$$0 = E(R_i^e M)\ )\ ? \tag{29}$$

2. Show mathematically that if the law of one price holds, all redundant assets are priced by the same pricing kernel as the primitive assets. Hint: Use the discrete-state notation and partition the payoff matrix into primitive and redundant security payoffs.

3. If the law of one price holds, when will there be a unique pricing kernel? (markets not complete, exactly complete, or more than complete?) Carefully explain why.