

Martingale Theory in Economics and Finance

Yongmiao Hong
Cornell University & Xiamen University

April 28, 2009

School of Economics, SKK University

Outline

- I. Concepts of Martingale
- II. A Motivating Example: Efficient Market Hypothesis
- III. A General Framework: Rational Expectations
- IV. Applications
 - a. Consumption Smoothing
 - b. Martingale Theory of Stock Price
 - c. Dynamic CAPM
 - d. Derivatives Pricing
- V. Conclusion

1 What is the Martingale Process?

Definition [Martingale] Suppose P_t is a time series process and $I_{t-1} = \{P_{t-1}, P_{t-2}, \dots\}$ is the information set available at time $t - 1$. Then $\{P_t\}$ is called a martingale process with drift α if

$$P_t = \alpha + P_{t-1} + X_t,$$

where X_t is a stationary time series process such that

$$E(X_t|I_{t-1}) = 0.$$

The process $\{X_t\}$ is called a martingale difference sequence (MDS).

Question: How important is the concept of martingale?

2 A Motivating Example: Efficient Market Hypothesis

The predictability of price changes has become a major theme of the financial research.

Definition [Efficient Market Hypothesis] Suppose P_t is an asset price (e.g., stock price) and

$$\begin{aligned} X_t &= \ln(P_t/P_{t-1}) \\ &= \ln\left(1 + \frac{P_t - P_{t-1}}{P_{t-1}}\right) \\ &\simeq \frac{P_t - P_{t-1}}{P_{t-1}}, \end{aligned}$$

is the rate of return on the asset from period $t - 1$ to period t . Let I_{t-1} be an information set available at period $t - 1$. We say that the efficient market hypothesis (EMH) holds if

$$E(X_t|I_{t-1}) = E(X_t) \text{ almost surely.}$$

In other words, the return $X_t - \mu$ is an MDS with respect to the information set I_{t-1} .

Question: What is the interpretation for $E(X_t)$?

The buy-and-hold trading strategy

$$T^{-1} \sum_{t=1}^T X_t \xrightarrow{p} E(X_t)$$

as $T \rightarrow \infty$.

Remarks:

(i) When the efficient market hypothesis holds, the expected return is zero. Thus, no systematic trading strategy that exploits the conditional mean dynamics can be more profitable in the long run than holding the market portfolio—of course, one can still temporarily “beat” the market through sheer luck.

To gain a deep insight into the EMH, let us consider some complex financial strategy. Suppose there are a riskfree asset with constant price 1, and a risky asset with price P_t that follows a martingale process. Let w_0 be an initial endowment. It can be invested in both assets, and the portfolio allocations can be regularly updated without any transaction costs. The portfolio value at time t is

$$w_t = \alpha_{0t} + \alpha_t P_t,$$

where α_{0t} and α_t are the quantities invested at time t in the risk-free and risky assets respectively. The portfolio is self-financed if, at each date, the exact portfolio value is being invested. The self-financing condition can be written as

$$w_{t+1} = \alpha_{0t} + \alpha_t P_{t+1}.$$

It follows that the updating of the portfolio value is

$$w_{t+1} - w_t = \alpha_t (P_{t+1} - P_t).$$

Whenever the decisions on allocations are based on the information P_t , the updating is such that

$$\begin{aligned} E_t(w_{t+1} - w_t) &= E_t[\alpha_t (P_{t+1} - P_t)] \\ &= \alpha_t [E_t(P_{t+1}) - P_t] \\ &= 0 \end{aligned}$$

given $E_t(P_{t+1}) = P_t$. Therefore, the martingale property is satisfied for any strategy of portfolio allocations no matter how complex it is. It is not possible to find a strategy ensuring a strictly positive net return with probability 1 at some fixed horizon. Indeed, the market is efficient if even a skilled investor has no sure advantage.

(ii) Under the efficient market hypothesis, the log of price P_t follows a martingale process:

$$\ln P_t = \ln P_{t-1} + X_t,$$

where $\{X_t - \mu\}$ is an MDS (i.e., $E(X_t|I_{t-1}) = \mu$ a.s.), and α is the drift parameter of the martingale process.

Question: What is the distinction between the martingale hypothesis and the random walk hypothesis?

Recall that $\{P_t\}$ follows a geometric random walk if

$$\ln P_t = \ln P_{t-1} + X_t,$$

where $\{X_t - \mu\}$ is an IID random sequence with $E(X_t) = 0$.

The notion of independence implies that the current return does not depend on past returns. Consequently, it is impossible to predict the future return using past returns. Bachelier (1900,1964) develops an elaborate mathematical theory of speculative prices. Roberts (1959) presents a largely heuristic argument for why successive price changes should be independent. Osborne (1959) develops the proposition that it is not absolute price changes but the logarithmic price changes which are independent of each other. If the auxiliary assumption is that the changes themselves are normally distributed, this implies that prices are generated as a Brownian motion.

Clearly, the random walk hypothesis implies the martingale hypothesis but not vice versa (unless $\{X_t\}$ is a Gaussian process). In other words, if $\{P_t\}$ follows a geometric random walk, then the market is efficient. However, P_t may not follow a random walk when the market is efficient. A random walk is more restrictive than a martingale. The martingale only rules out serial dependence in conditional mean, whereas the random walk rules out not only this but also serial dependence involving the higher order conditional moments of X_t . This distinction is rather important, because the random walk model may be inconsistent with optimizing economic models (e.g., the rational expectations theory).

Example 1: Suppose the return $\{X_t\}$ follows an autoregressive conditional heteroskedasticity (ARCH) process:

$$\begin{cases} X_t &= h_t^{1/2} z_t, \\ h_t &= \alpha_0 + \alpha_1 X_{t-1}^2, \\ \{z_t\} &\sim \text{i.i.d. } (0,1). \end{cases}$$

Here, $E(X_t|I_{t-1}) = 0$ a.s., so the market is efficient. However, $\{X_t\}$ is not IID, because its conditional variance $\text{var}(X_t|I_{t-1}) = h_t$ depends on the square of the previous return X_{t-1} . Such serial dependence in variance does not help predict the level of the asset return, although it implies that the future volatility of asset return is predictable using the currently available information.

Remark: Most high frequency financial time series have strong volatility clustering (e.g., Mandelbrot 1963). That is, a high volatility tends to be followed by another high volatility; and a low volatility tends to be followed by a low volatility. Statistically speaking, there exists positive autocorrelation in the squares of returns or the absolute returns.

(iv) There are three notions of market efficiency. They differ in the definition of the information set I_{t-1} (see, e.g., Campbell, Lo and MaKinlay 1997, pp.):

1. The weak form of efficient market hypothesis: I_{t-1} only contains the past history of $X_s, s < t$.

2. The semi-strong form of efficient market hypothesis: I_{t-1} contains the past history of $\{X_s, s < t\}$ and other public information available at period $t - 1$.
3. The strong form of efficient market hypothesis: I_{t-1} contains all public information and some private information, available at period $t - 1$.

Remark: The knowledge of new trading strategies and new modelling techniques can be viewed as part of private information.

Remark: The above definition of EMH is a classical version. It may be more appropriate to be termed as “non-predictability hypothesis”. There is a modern version of definition of market efficiency in the rational expectations framework, where the market can be efficient when asset returns are still predictable using past information. However, the predictable component of the asset return is the time-varying risk premium. The ultimate risk-adjusted excess asset return is not predictable using the past information.

3 Rational Expectations

A General Framework: The Optimizing Approach

Cochrance, John (2001): *Asset Pricing*, Princeton University Press: Princeton.

One important task in economics is to predict how economic agents’ behavior changes when their environment changes, e.g., as a result of changes in government policies or external shocks. For this purpose, we need estimates of parameters describing economic agents’ preferences and firms’ technologies, i.e., structural parameters. Within the rational expectation macroeconomic school these parameters are estimated based on dynamic equilibrium theories where economic agents solve constrained optimization problems.

Suppose a representative consumer maximizes the expected present value of his/her intertemporal utility

$$\max_{\{C_t\}} E_t \sum_{j=0}^{\infty} \beta^j u(C_{t+j}), \quad 0 < \beta < 1,$$

subject to the intertemporal budget constraint

$$A_{t+1} = R_{t+1}(A_t + Y_t - C_t),$$

with the initial endowment A_0 given. Here, C_t is the consumption of the agent at time t , Y_t is the agent’s “labor income” at time t , A_t is the amount of a single earning asset valued in units of the consumption good, held at the beginning of period t , R_{t+1} is the real gross rate of return on the asset between periods t and $t + 1$, and β is the constant discount factor following from the subjective rate of time preference. We assume that $\{R_{t+1}\}$ is a stochastic process and that R_{t+1} becomes known to the agent only at the beginning of period $t + 1$. We also assume that $\{Y_t\}$ is an uncontrollable random process,

and there are no taxes, transaction costs or other frictions, and financial markets are complete. We further assume that the utility function $u(\cdot)$ of the economic agent is concave, strictly increasing, and twice continuously differentiable. These, among other things, imply that the consumer is risk averse.

Examples of $u(\cdot)$: Commonly used utility functions include:

1. Quadratic utility:

$$u(C_t) = a + bC_t + cC_t^2, \quad b > 0, c < 0.$$

2. Constant relative risk aversion (CRRA) utility

$$\begin{aligned} u(C_t) &= (C_t^\gamma - 1)/\gamma, \quad \gamma > 0. \\ \gamma &= -\frac{u''(C_t)}{u'(C_t)}, \end{aligned}$$

3. Logarithmic utility

$$u(C_t) = \ln(C_t).$$

4. Exponential utility

$$u(C_t) = -\exp(-\gamma C_t).$$

For the exponential utility function, the absolute risk aversion $-u''(C_t)/u'(C_t) = \gamma$ is independent of consumption C_t . For this reason, this utility is called the constant absolute risk aversion (CARA) utility.

Question: How to solve for the intertemporal utility optimization problem?

Denote

$$V_t(A_t) = \max_{\{C_t\}} E_t \left[\sum_{j=0}^{\infty} \beta^j u(C_{t+j}) \right]$$

subject to the intertemporal budget constraint. Then

$$V_t(A_t) = \max [u(C_t) + \beta E_t [V_{t+1}(A_{t+1})]],$$

where

$$R_{t+1}(A_t + Y_t - C_t).$$

By differentiating with respect to C_t , we obtain

$$u'(C_t) - \beta E_t [V'_{t+1}(A_{t+1})R_{t+1}] = 0.$$

Because $V'_{t+1}(A_{t+1}) = u'(C_{t+1})$, we have

$$E_t[\beta u'(C_{t+1})R_{t+1}] = u'(C_t).$$

The Euler equation associated with the above intertemporal utility optimization problem is

$$E[\beta R_t u'(C_t) | I_{t-1}] = u'(C_{t-1}), \quad t = 1, 2, \dots,$$

where I_{t-1} is the information set available to the economic agent at time $t - 1$. Note that the econometrician may only know a subset of I_{t-1} .

Question: What is the economic interpretation for the Euler equation?

Answer: The RHS is the utility value of consuming one dollar today. The LHS is the expected utility value (discounted to the present value of utility) of investing one dollar today, receiving dividend D_t and then selling it tomorrow. In an equilibrium, the two strategies will give the same present value of utility.

Remarks:

(i) This model is very simplified, due to many strong assumptions (e.g., one representative agent who is living forever, constant time discount factor, time-invariant utility function, no frictions, complete markets, etc). However, it is also very general, because it covers many interesting economic examples, as we will illustrate below.

(ii) It implies a direct and unique relationship between the variation in asset returns and the variation in the intertemporal marginal rate of substitution (IMSE); that is, there is a strong relationship between financial market prices and consumptions.

4 Applications

4.1 Hall's (1978) Martingale Theory of Consumption

Purpose: To examine the dynamics of the consumption series given the expected asset returns.

Suppose the rate of return on the asset $R_t = R > 0$ is a constant for all t , then the Euler equation becomes

$$E[u'(C_t) | I_{t-1}] = (\beta R)^{-1} u'(C_{t-1}).$$

This can be equivalently written as

$$u'(C_t) = (\beta R)^{-1} u'(C_{t-1}) + \varepsilon_t,$$

where

$$E(\varepsilon_t | I_{t-1}) = 0 \text{ a.s.}$$

Remarks:

(i) The conditional expectation of the marginal utility of consumption $u'(C_t)$ given I_{t-1} only depends on the one-lagged marginal utility of consumption $u'(C_{t-1})$. No other variables (e.g., asset prices) in the information set I_{t-1} help predict the expected value

of $u'(C_t)$, once $u'(C_{t-1})$ has been included. We call that $\{u'(C_t)\}$ is a first order Markov process in mean.

Definition [Markov process in mean]: Put $I_{t-1} = \{X_{t-1}, X_{t-2}, \dots\}$. We call a time series $\{X_t\}$ as a p -th order Markov process in mean if

$$E(X_t|I_{t-1}) = E(X_t|X_{t-1}, \dots, X_{t-p}) = g(X_{t-1}, \dots, X_{t-p}).$$

(ii) It is possible that the conditional variance (or other higher order conditional moments) of $u'(C_t)$ given I_{t-1} depends on I_{t-1} rather than only on $u'(C_{t-1})$. Therefore, $u'(C_t)$ may not be a Markov process. [Recall the definition of a Markov process in Lecture 02.]

(iii) When $u(C_t) = (C_t^\gamma - 1)/\gamma$ is a constant relative risk aversion utility function, then $u'(C_t) = C_t^{\gamma-1}$, and the Euler equation becomes

$$E[C_t^{\gamma-1}|I_{t-1}] = (\beta R)^{-1} C_{t-1}^{\gamma-1}.$$

(iv) If $\beta R = 1$, then $\{u'(C_t)\}$ is a martingale process:

$$u'(C_t) = u'(C_{t-1}) + \varepsilon_t.$$

(v) If $u(C_t) = a + bC_t + cC_t^2$ and $\beta R = 1$, then $\{C_t\}$ is a martingale process:

$$u'(C_t) = (\beta R)^{-1} u'(C_{t-1}) + \varepsilon_t,$$

$$C_t = C_{t-1} + \varepsilon_t,$$

$$E(C_t|I_{t-1}) = C_{t-1}.$$

$$\text{where } E(\varepsilon|I_{t-1}) = 0 \text{ a.s.}$$

Question: How to test the martingale hypothesis for consumption?

Existing methods in macroeconomics use an auxiliary regression

$$C_t = \alpha + \beta C_{t-1} + \gamma' Z_{t-1} + \varepsilon_t$$

and test whether $\beta = 1$ and $\gamma = 0$ jointly, where $Z_{t-1} \in I_{t-1}$ is a vector of some macroeconomic variables such as asset returns.

Remarks:

(i) This is a traditional approach to testing omitted variables. Statistical inference here is complicated because C_t is $I(1)$ and Z_t may be $I(0)$. Standard asymptotic theory may not be useful here.

(ii) An alternative way is to consider

$$\Delta C_t = \alpha + \beta' Z_t + \varepsilon_t$$

and test whether $\beta = 0$. Because both ΔC_t and Z_t are stationary, we can use the conventional asymptotic normal distribution theory.

4.2 Martingale Theory of Stock Prices

Purpose: To examine the dynamics of asset prices given the expected IMRS.

Suppose the asset is a share of an enterprise that sells for price P_t per share at period t and pays nonnegative random dividends of D_t consumption goods to the owner of the share at the beginning of t .

Let S_t be the number of shares owned by the consumer at the beginning of t , and letting $A_t = (P_t + D_t)S_t$, we have the budget constraint

$$\begin{aligned} & (P_{t+1} + D_{t+1})S_{t+1} \\ = & \frac{P_{t+1} + D_{t+1}}{P_t} [(P_t + D_t)S_t + Y_t - C_t]. \end{aligned}$$

Alternatively, we can write the budget constraint as

$$P_t S_{t+1} = (P_t + D_t)S_t + Y_t - C_t.$$

Remark: The gross yield on shares from time t to time $t+1$ is $R_{t+1} = (P_{t+1} + D_{t+1})/P_t = P_{t+1}/P_t + D_{t+1}/P_t$.

Question: How to interpret the budget constraint equation?

Rewrite

$$\beta E \left[R_t \frac{u'(C_t)}{u'(C_{t-1})} | I_{t-1} \right] = 1, \quad t = 1, 2, \dots$$

Using the formula that $\text{cov}(X, Y) = E(XY) - E(X)E(Y)$, the

$$\begin{aligned} X &= R_t = \frac{P_t + D_t}{P_{t-1}}, \\ Y &= \frac{u'(C_t)}{u'(C_{t-1})} \end{aligned}$$

Euler equation becomes

$$\begin{aligned} 1 &= \beta E \left[\frac{P_t + D_t}{P_{t-1}} \frac{u'(C_t)}{u'(C_{t-1})} | I_{t-1} \right] \\ &= \beta E \left(\frac{P_t + D_t}{P_{t-1}} | I_{t-1} \right) E \left[\frac{u'(C_t)}{u'(C_{t-1})} | I_{t-1} \right] \\ &\quad + \beta \text{cov} \left[\frac{u'(C_t)}{u'(C_{t-1})}, \frac{P_t + D_t}{P_{t-1}} | I_{t-1} \right]. \end{aligned}$$

If the conditional expectation of the intertemporal marginal rate of substitution $E[u'(C_t)/u'(C_{t-1}) | I_{t-1}] = \gamma$ is a constant, then the second term

$$\text{cov} \left[\frac{u'(C_t)}{u'(C_{t-1})}, \frac{P_t + D_t}{P_{t-1}} | I_{t-1} \right] = 0,$$

and so we have

$$1 = \gamma\beta E\left[\frac{P_t + D_t}{P_{t-1}} \middle| I_{t-1}\right].$$

$$\begin{aligned} (\gamma\beta)E[(P_t + D_t)|I_{t-1}] &= P_{t-1}. \\ P_t &= (\gamma\beta)E[(P_{t+1} + D_{t+1})|I_t] \end{aligned}$$

Remark: Imposing the no-bubble transversality condition $\lim_{j \rightarrow \infty} \beta^j E_t(P_{t+j}) = 0$, we can obtain the following present value model for the asset price at time t :

$$P_t = \sum_{j=1}^{\infty} (\gamma\beta)^j E_t[D_{t+j}].$$

Thus, the price of the asset today is determined as the present discounted value of all expected future dividends.

Remarks:

- (i) What are the interpretations for the above two conditions?
- (ii) When the share price is adjusted for dividends and discounting in time. it follows a first order univariate Markov process in mean. No other variables help predict the share prices.

Question: When will P_t follow a martingale?

If $\gamma\beta = 1$ and $D_t = 0$, then we obtain a martingale model for the asset price:

$$\begin{aligned} E(P_t|I_{t-1}) &= P_{t-1} \text{ or} \\ P_t &= P_{t-1} + \varepsilon_t, \end{aligned}$$

where

$$E(\varepsilon_t|I_{t-1}) = 0 \text{ a.s..}$$

4.3 Dynamic Capital Asset Pricing

CAPM = Capital Asset Pricing Model

Purpose: To examine asset allocations and their dynamic cross-sectional restrictions given the expected IMRS.

Now suppose there are n assets in the economy. At the beginning of period t , the owner of asset i receives a “dividend” of D_{it} units of the consumption good. The consumption good is not storable, so that the equilibrium condition holds for all t , namely $C_t = D_t = \sum_{i=1}^n D_{it}$ for all t . The budget constraint becomes

$$\sum_{i=1}^n P_{it} S_{it+1} = \sum_{i=1}^n (P_{it} + D_{it}) S_{it} - C_t,$$

Question: How to interpret the budget constraint equation?

The Euler equation now consists of a system of equations:

$$E \left[\beta \frac{u'(C_t)}{u'(C_{t-1})} (P_{it} + D_{it}) | I_{t-1} \right] = P_{it-1},$$
$$i = 1, \dots, n.$$

Question: What is the economic interpretation for the Euler equation?

Substituting the equilibrium condition $C_t = D_t$ gives

$$P_{it-1} = E_{t-1} \left[\beta \frac{u'(D_t)}{u'(D_{t-1})} (P_{it} + D_{it}) \right].$$

Remark: This can be used to calculate the equilibrium asset prices. The simultaneous Euler equations impose a cross-sectional restriction on all asset prices in each period and such a cross-sectional restriction is time-varying (i.e. dynamic).

Stochastic Discount Factor

Define the stochastic discount factor

$$M_t = \beta \frac{u'(C_t)}{u'(C_{t-1})}.$$

Note that M_t depends on the economic agent's preference and does not depend on asset prices. Then the Euler equation becomes

$$E [M_t (P_{it} + D_{it}) | I_{t-1}] = P_{it-1} \text{ for all } i, t.$$

Question: What is the interpretation for the stochastic discount factor M_t ? How to interpret the Euler equation?

Answer: SDF is the aggregate measure of the market expectation about future uncertainty.

Remark: $M_t \geq 0$.

By a Taylor series expansion, we obtain

$$M_t = \beta \frac{u'(C_{t-1} + \Delta C_t)}{u'(C_{t-1})}$$
$$= \beta + \beta \frac{u''(C_{t-1})}{u'(C_{t-1})} \Delta C_t + \text{higher order term,}$$

where the first term corresponds to the case when the economic agent is risk neutral, and the second term $\beta[u''(C_{t-1})/u'(C_{t-1})]\Delta C_t$ depends on the consumption growth and

the risk attitude of the economic agent. [Recall $u''(\cdot)/u'(\cdot)$ is the curvature of the utility function, which determines the risk aversion of the economic agent.]

Remark: The Euler equation characterization is consistent with the perhaps more familiar expected return-beta representation. Let

$$R_{it} = (P_{it} + D_{it})/P_{it-1}$$

be the rate of gross return on asset i from period $t - 1$ to period t . Then the Euler equation can be written as

$$\begin{aligned} 1 &= E(M_t R_{it} | I_{t-1}) \\ &= \text{cov}(M_t, R_{it} | I_{t-1}) \\ &\quad + E(M_t | I_{t-1}) E(R_{it} | I_{t-1}). \end{aligned}$$

It follows that

$$\begin{aligned} E(R_{it} | I_{t-1}) &= \frac{1}{E(M_t | I_{t-1})} \\ &\quad - \frac{\text{cov}(M_t, R_{it} | I_{t-1})}{E(M_t | I_{t-1})} \\ &= \frac{1}{E(M_t | I_{t-1})} \\ &\quad + \frac{\text{cov}(M_t, R_{it} | I_{t-1})}{\text{var}(M_t | I_{t-1})} \left[\frac{-\text{var}(M_t | I_{t-1})}{E(M_t | I_{t-1})} \right] \\ E(R_{it} | I_{t-1}) &= \alpha_{t-1} + \beta_{it-1} \lambda_{t-1}, \text{ say,} \end{aligned}$$

where λ_{t-1} is the market risk factor at $t - 1$, β_{it-1} is the conditional covariance coefficient between the asset return R_{it} and the stochastic discount factor M_t , and α_{t-1} is the rate of return on riskfree asset.

Question: Why is α_{t-1} the riskfree rate at time t ?

Answer: Note that the Euler equation also applies to the (unit price) riskfree asset that pays α_{t-1} in period t which is known at the end of period $t - 1$:

$$1 = E_{t-1}(M_t \alpha_{t-1}) = \alpha_{t-1} E_{t-1}(M_t).$$

It follows that

$$\alpha_{t-1} = \frac{1}{E_{t-1}(M_t)}.$$

We can equivalently write

$$R_{it} = \alpha_{t-1} + \beta_{it-1} \lambda_{t-1} + \varepsilon_{it}, \quad i = 1, \dots, n.$$

where

$$E(\varepsilon_{it} | I_{t-1}) = 0 \text{ a.s.}$$

$$\varepsilon_{it} \sim i.i.d.(0, \sigma^2),$$

$$\{\varepsilon_{it}\} \text{ and } \{\varepsilon_{jt}\} \text{ independent.}$$

Remarks:

(i) $\text{cov}(\varepsilon_{it}, \varepsilon_{jt-l}) = 0$ for all $l > 0$ and all i, j . However, it is possible that there exist spatial correlation across different assets: $\text{cov}(\varepsilon_{it}, \varepsilon_{jt}) \neq 0$ for all i, j . This may occur because asset prices may respond to the same shocks and move together.

(ii) For each asset i , $\text{var}(\varepsilon_{it}|I_{t-1})$ may be a time-varying function. That is, there may exist volatility clustering for the return on each asset.

(iii) It is important to note that both α_{t-1} and β_{it-1} are functions of I_{t-1} and are so time-varying.

Dynamic Asset Pricing Model Specification

(i) Any dynamic asset pricing model $\mathcal{M} = \{M_t(\theta) : \theta \in \Theta\}$ can be viewed as a specification (i.e., model) for the stochastic discount factor M_t , where θ is the unknown model parameter vector.

Example 1: [Capital Asset Pricing Model (CAPM)]:

The CAMP model is equivalent to assume

$$M_t(\theta) = a + bR_t^m,$$

where $\theta = (a, b)'$ and R_t^m is the return on the market portfolio.

Remark: The CAMP model, established independently by Sharpe (1964), Lintner (1965) and Mossin (1966), is based on the assumption that either all utility functions are quadratic, or the returns are normally distributed, or both.

Example 2 [Fama and French's (1993) Linear Factor Model] A linear asset pricing model is equivalent to assume that the stochastic discount factor is a linear function of some factors:

$$M_t(\theta) = b_0 + b_1' f_t,$$

where $\theta = (b_0, b_1)'$ and f_t is a vector of time-varying risk factors.

Question: How to model the stochastic discount factor M_t ?

Answer: Any macroeconomic variables that can help predict consumption growth and/or risk aversion.

(ii) We say that an asset pricing model is correctly specified if there exists $\theta_0 \in \Theta$ such that

$$M_t(\theta_0) = M_t \text{ almost surely for all } t.$$

To check correct asset pricing, it suffices to check whether the Euler condition holds:

$$E [X_{it+1}(\theta_0)|I_t] = 0 \text{ for some } \theta_0 \text{ and all } i, t,$$

where

$$X_{it+1}(\theta) = M_{t+1}(\theta)(P_{it+1} + D_{it+1}) - P_{it}$$

is the model pricing error for asset i .

Cochrane's (2001) equivalent characterization: The Euler condition

$$E[X_{it}(\theta_0)|I_{t-1}] = 0 \text{ for some } \theta_0$$

is equivalent to

$$E[X_{it}(\theta_0)Z_t] = 0 \text{ for some } \theta_0$$

where $Z_t \in I_{t-1}$ is any measurable function of I_{t-1} and is called an instrument. This equivalent characterization is interesting, but it is impossible to check all possible Z_t given any sample of finite number of observations.

(iii) Most empirical studies on testing asset pricing models in finance focus on the unconditional implication of the Euler condition:

$$E[X_{it}(\theta_0)] = 0 \text{ for } i = 1, \dots, n$$

That is, they check whether the average pricing error is zero for each asset i . This is only one of an infinite number of implications of the Euler condition. Such a test will miss many possible model misspecifications whose price error has zero unconditional mean but nonzero conditional mean.

Example 1: AR(1) Pricing Error

$$X_{it}(\theta_0) = \alpha X_{it-1}(\theta_0) + u_{it},$$

where $\{u_{it}\}$ is i.i.d. $(0, \sigma_{iu}^2)$. Here,

$$E[X_{it}(\theta_0)] = 0$$

but

$$E[X_{it}(\theta_0)|I_{t-1}] = \alpha X_{it-1}(\theta_0).$$

Example 2: Nonlinear Moving Average Pricing Error

$$X_{it}(\theta_0) = \alpha z_{it-1} z_{it-2} + z_{it},$$

where $\{z_{it}\} \sim$ i.i.d. $(0, \sigma_{iu}^2)$. Again, we have

$$E[X_{it}(\theta_0)] = 0$$

but

$$E[X_{it}(\theta_0)|I_{t-1}] = \alpha u_{it-1} u_{it-2}.$$

Question: For what kinds of asset pricing models, will the pricing errors have the forms as given in these two examples?

Question: Why is it important to test the conditional implications of correct asset pricing?

When the average model pricing error is zero but the pricing error exhibits clustering, the asset pricing model cannot capture the dynamics of asset price movements.

(iv) It is a daunting job to check the Euler condition, mainly because I_{t-1} may not be available to econometricians. The so-called generalized method of moments (GMM) is used to estimate and test Euler equation. Commonly used instruments are lagged interest rate levels, lagged stock price changes, lagged consumption growth, and so on. However, these are just a finite number of instruments which constitute a very tiny subset of I_{t-1} .

Question: What is the GMM method?

Define the moment condition

$$m(X_t, \theta) = \left[\beta \frac{u'(C_t)}{u'(C_{t-1})} R_t - 1 \right] \otimes Z_t,$$

where R_t is a $n \times 1$ asset return vector, Z_t is a $k \times 1$ instrumental vector contained in I_{t-1} , $X_t = (C_t, C_{t-1}, R'_t, Z'_t)'$, and θ is a $d \times 1$ model parameter vector containing β and preference parameters in $u(\cdot)$. Then the Euler equation implies

$$E[m(X_t, \theta_0)] = 0,$$

where $m(X_t, \theta)$ is a $lk \times 1$ vector.

Given a random sample $\{X_t\}_{t=1}^T$, the GMM estimator is defined as

$$\hat{\theta} = \arg \min_{\theta \in \Theta} \hat{m}(\theta)' \hat{V}^{-1} \hat{m}(\theta)$$

where

$$\hat{m}(\theta) = T^{-1} \sum_{t=1}^T m(X_t, \theta)$$

and \hat{V} is a pre-specified $lk \times lk$ symmetric positive definite matrix that may also depends on the sample.

When the Euler equation is correctly specified, the following so-called J test statistic

$$J \equiv T \hat{m}(\hat{\theta})' \hat{V}^{-1} \hat{m}(\hat{\theta}) \rightarrow^d \chi_{lk-d}^2.$$

Question: What are the advantages of GMM and the J -test?

(i) GMM does not have to specify the data generating processes for related economic variables (such as the consumption and asset prices).

(ii) Model parameters can be consistently estimated and the estimators are asymptotically normal (with proper scaling) when the processes are stationary. The time series processes can be unknown conditional heteroskedastic with non-Gaussian distributions.

It is not as restrictive as MLE. The latter requires knowledge of the full conditional distribution of the economic structural model, rather than only the Euler equation.

Question: What are the disadvantages or drawbacks of the GMM method?

(i) How to choose instruments? Or how to make full/efficient use of the available information? There is no general rule for this.

(ii) There has been a vast empirical literature on the GMM method for macroeconomic and financial applications. Often some important model parameter estimates are insignificant, and the J -test statistic is insignificant, likely due to its low power. In fact, the J -test, which is based on the GMM objective function, rewards large sampling errors for GMM estimators. How to make the J -test more powerful? Or how to design more powerful tests?

(iii) The rejection of the J -test does not provide useful information about how to revise a macroeconomic/financial model.

4.4 Derivative Pricing

In modern finance, the pricing of contingent claims is of particular interest, given the phenomenal growth in turnover and volume of financial derivatives over the past decades. Derivative pricing formulas are highly nonlinear even when they are available in closed form.

The Euler equation can also be used to characterize derivative pricing.

In a standard dynamic exchange economy, the equilibrium price of a derivative at date t with a single liquidating payoff $Y(S_T)$ at date T that is a function of the underlying asset S_T is given by

$$P_t = E_t [M_{t,T} Y(S_T)],$$

where the conditional expectation is taken with respect to the information set available to the representative agent at time t , $M_{t,T}$ is the marginal rate of substitution between dates t and T .

Remark: If $Y(S_T) = r_f$, the riskfree compound interest rate. Then the Euler equation implies

$$\begin{aligned} 1 &= e^{r_f(T-t)} E_t[M_{t,T}], \\ E_t[M_{t,T}] &= e^{-r_f(T-t)} = \frac{1}{\alpha_{t-1}}. \end{aligned}$$

Remark: The Euler equation is the key condition for optimal derivative pricing.

Assuming that the conditional distribution of the future underlying asset price S_T

has a density representation $f_t(\cdot)$, then the conditional expectation can be expressed as

$$\begin{aligned}
 P_t &= E_t [Y(S_T)M_{t,T}] \\
 &= \int Y(s)M_{t,T}f_t(s)ds \\
 &= E_t(M_{t,T}) \int Y(s)\frac{M_{t,T}f_t(s)}{E_t(M_{t,T})}ds \\
 &= \exp(-r_t\tau) \int Y(s)f_t^*(s)ds \\
 &= \exp(-r_t\tau)E_t^*[Y(S_T)],
 \end{aligned}$$

where $\tau = T - t$ is the remaining time to maturity, and

$$f_t^*(S_T) = \frac{M_{t,T}f_t(S_T)}{\int M_{t,T}f_t(s)ds}$$

is called the risk neutral density (RND) function. This function is also called the risk-neutral pricing probability, or the equivalent martingale measure, or the state-price density. It contains rich information on the pricing and hedging of risky assets in an economy, and can be used to price other assets, or to recover the information about the market preferences and asset price dynamics. Obviously, the RND function differs from $f_t(S_T)$, the physical density function of C_T conditional on the information available at time t .

Example [Black-Scholes formula] $Y_t(S_T) = \max(S_T - K_t, 0)$, where S_T is the stock price at time T , and K_t is the strike price.

Remarks:

(i) $f_t^*(S_T)$ depends on $M_{t,T}$ and $f_t(S_T)$. If the latter are correctly specified, then $f_t^*(S_T)$ will give optimal pricing in the sense that $M_{t,T}Y(S_T) - P_t$ will have a zero conditional mean with respect to information set I_t .

(ii) **Question:** How to obtain $f_t^*(S_T)$?

Suppose $P_t = P_t(S_t, K_t, r_t, \tau)$ is the option pricing model. Then there is a close relation between the second derivative of G_t with respect to the strike price K_t and the RND function:

$$\frac{\partial^2 P_t}{\partial K_t^2} = \exp[-r_t\tau] f_t^*(K_t).$$

An Example: The Call Option Price Formula

The payoff at time T is

$$Y_t(S_T) = \max(S_T - K_t, 0),$$

where K_t is the strike price. The price of the call option can be represented by

$$\begin{aligned} C_t &= e^{-r_f(T-t)} \int_0^\infty f_t^*(S_T) Y(S_T) dS_T \\ &= e^{-r_f(T-t)} \int_0^\infty f_t^*(S_T) \max(S_T - K_t, 0) dS_T. \end{aligned}$$

For $S_T \geq K_t$, we have

$$C_t = e^{-r_f(T-t)} \int_0^\infty f_t^*(S_T) (S_T - K_t) dS_T.$$

The first derivative

$$\begin{aligned} \frac{\partial C_t}{\partial K_t} &= -e^{-r_f(T-t)} (S_T - K_t) |_{S_T=K_t} \\ &\quad + e^{-r_f(T-t)} \int_{K_t}^\infty f_t^*(S_T) \frac{\partial(S_T - K_t)}{\partial K_t} dS_T \\ &= -e^{-r_f(T-t)} \int_{K_t}^\infty f_t^*(S_T) dS_T. \end{aligned}$$

By further differentiation, we have

$$\frac{\partial^2 C_t}{\partial K_t^2} = e^{-r_f(T-t)} f_t^*(K_t).$$

(iii) If we know $M_{t,T}$, then we can predict $f_t(S_T)$ using $M_{t,T}$ and $f_t^*(S_T)$.

(iv) If we know $f_t(S_T)$, then we can obtain $M_{t,T}$ from $f_t^*(S_T)$ and $f_t(S_T)$. Recall that $M_{t,T}$ contains vital information about the market expectation of future uncertainty.

Remark: All derivative pricing models are equivalent to a specification of the RND $f_t^*(S_T)$. The Euler equation provides a basis to check the optimality of derivative pricing in a dynamic context.

5 Conclusion

This lecture provides a unified framework to via a number of economic theories including consumption smoothing, stock valuation and efficient market hypothesis, dynamic capital asset pricing, and derivatives pricing.