LECTURE NOTES 1

ARIEL M. VIALE

I. Portfolio Choice and Asset Pricing

Definition 1. (Gross return) Let $\widetilde{x}_k$ denote the one period payoff of asset $k \in K$ with price $p_k \geq 0$. If $p_k > 0$ the gross return of asset k is defined as,

$$\widetilde{R}_k = \frac{\widetilde{x}_k}{p_k}. \tag{1}$$

Remark. The rate of return of asset $k \in K$ is defined as $\widetilde{r}_k = \widetilde{R}_k - 1 = \frac{\widetilde{x}_k - p_k}{p_k}$.

Definition 2. (Risk premium) If there exists a riskless asset with return $\overline{R}_f$, then the risk premium of the kth risky asset is defined by $\mathbb{E}\left[\widetilde{R}_k\right] - \overline{R}_f$.

Intuition: The extra average return that investors ask for in order to hold the risky asset instead of the riskless asset. The main goal of asset pricing is to explain why different assets carry different risk premia.

Claim. Assume a one period economy with a single consumption good (i.e., the numeraire or unit in which prices are measured so the price of the single consumption good is equal to 1). Let w_0 denote the amount of the single consumption good postponed at the beginning of the period, which is invested in the risky asset k and let η_k denote the number of shares the investor chooses to hold of the risky asset k. Assume the end of period random endowment $\widetilde{\omega}$ (e.g., labor income), which we assume is totally consumed in addition of the value of the end of period portfolio of assets. If we assume that the beginning of period consumption c_0 has been already optimally chosen so that $w \longmapsto v(c^*, w)$, then the typical investor's behavioral problem is,

$$\begin{aligned} &\underset{\{\eta_1,\ldots,\eta_K\}}{Max} E\left[u\left(\widetilde{w}\right)\right] \\ &s.t. \quad \textstyle\sum_k \quad \eta_k \cdot p_k = w_0 \\ &and\ \widetilde{w} \quad = \quad \widetilde{\omega} + \sum_k \eta_k \cdot \widetilde{x}_k \end{aligned} \tag{2}$$

Alternatively, the portfolio choice problem can be stated in terms of the amount $\phi_k = \eta_k p_k$ of the consumption good invested in each risky asset,

$$\underset{\{\phi_1,\ldots,\phi_K\}}{Max} E\left[u\left(\widetilde{w}\right)\right] \tag{3}$$

Date: 07/13/11 - Rev.: 09/10/13.

$$s.t. \quad \sum_k \quad \phi_k = w_0$$

$$and\ \widetilde{w} \quad = \quad \widetilde{\omega} + \sum_k \phi_k \cdot \widetilde{R}_k,$$

and in terms of the fraction $z_k = \frac{\phi_k}{w_0}$ of the beginning of period wealth invested in each asset,

$$\underset{\{z_1,\ldots,z_K\}}{Max} E\left[u\left(\widetilde{w}\right)\right] \tag{4}$$

$$s.t. \quad \sum_k \quad z_k = 1$$

$$and\ \widetilde{w} \quad = \quad \widetilde{\omega} + w_0 \cdot \sum_k \left(z_k \cdot \widetilde{R}_k\right)$$

Note. $\eta_k, \phi_k, z_k < 0$, $\forall k$. Investors can borrow and sell assets short. If the asset has a positive payoff then the term $\eta_k \cdot \widetilde{x}_k$ has negative sign reflecting the fact that the investor must pay back the value of the asset borrowed. Short selling is restricted by an implicit solvency contraint on $\widetilde{w}$ i.e., it must be in the domain of the definition of the utility function with probability 1. When introducing them one has to follow Kuhn-Tucker and check if the constraints bind at the optimum.

Assumption 1. There exists an interior solution so that the partial derivatives w.r.t. η_k are equal to zero at the optimum.

Assumption 2. Differentiation and expectation can be interchanged.

Substituting the second constraint in the objective function, the Lagrangian function can be written as,

$$\mathcal{L} \equiv E\left[u\left(\widetilde{\omega} + \sum_k \eta_k \cdot \widetilde{x}_k\right)\right] - \lambda\left(\sum_k \eta_k \cdot p_k - w_0\right) \tag{5}$$

with first order necessary condition (F.O.N.C.) w.r.t. η_k,[1]

$$(\forall k) \qquad \frac{\partial \mathcal{L}}{\partial \eta_k} \equiv E\left[u'\left(\widetilde{w}\right)\widetilde{x}_k\right] \quad - \quad \lambda p_k = 0$$

$$\Rightarrow (\forall k) \qquad E\left[\frac{u'\left(\widetilde{w}\right)}{\lambda}\widetilde{x}_k\right] \quad = \quad p_k, \qquad rearranging\ terms$$

$$\Rightarrow (\forall k) \qquad E\left[\frac{u'\left(\widetilde{w}\right)}{\lambda}\widetilde{R}_k\right] \quad = \quad 1, \qquad if\ p_k \neq 0$$

$$\Rightarrow (\forall k) \qquad E\left[u'\left(\widetilde{w}\right)\left(\widetilde{R}_k - \overline{R}_f\right)\right] \quad = \quad 0, \qquad if\ \exists \overline{R}_f \neq 0$$

[1] $E\left[\frac{u'(\widetilde{w})}{\lambda}\widetilde{R}_k\right] - E\left[\frac{u'(\widetilde{w})}{\lambda}\widetilde{R}_j\right] = 1 - 1, \forall k, j; k \neq j. If\, \widetilde{R}_j = \bar{R}_f \Rightarrow E\left[\frac{u'(\widetilde{w})}{\lambda}\left(\widetilde{R}_k - \overline{R}_f\right)\right] = 0.$

Remark. Asset pricing theory focuses in the F.O.N.C. of the portfolio choice problem.

Definition 3. (Excess return) The random variable $\left(\widetilde{R}_k - \overline{R}_f\right)$ is the excess return or the payoff of a zero-cost portfolio that is long in a unit of consumption invested in the risky asset k and short the same amount in the riskless asset.

Intuition: At the optimum, the marginal utility of adding any zero-cost portfolio should be equal to zero, otherwise the optimum portfolio wouldn't be an optimum as the utility of the investor will improve.

Proof. See Back (2010, pages 24-25: Existence, concavity, and continuity of utility function plus the monotone convergence theorem guarantees proof by contradiction) □

II. Stochastic Discount Factors and Representative Agent (RA) Asset pricing

Definition 4. (Stochastic discount factor) A stochastic discount factor (SDF) is any random variable $\widetilde{m}$ such that,

$$(\forall k) \text{ } and \text{ } (\forall s \in S) \qquad \sum_{s=1}^{S} \widetilde{m}(s)\, \widetilde{x}_k(s)\, prob_s = E\left[\widetilde{m}\widetilde{x}_k\right] = p_k \tag{6}$$

Remark. that pays one unit of the consumption good in a particular state s and zero in all other states. Then the price of such security $q_s = \widetilde{m}(s)\, prob_s$ is a state price, and the SDF in any particular state of the world is the ratio between the state price and the probability of the state. Because the SDF sets the price of a unit of consumption good in each state is also called "state price density" or "pricing kernel". Technically, $\widetilde{m}$ is the Radon-Nikodym derivative of the set function that assigns prices to sets of states or events relative to the probabilities of these events. If each p_k is positive then,

$$(\forall k) \qquad E\left[\widetilde{m}\widetilde{R}_k\right] = 1 \tag{7}$$

$$\Rightarrow (\forall k, j) \qquad E\left[\widetilde{m}\left(\widetilde{R}_k - \widetilde{R}_j\right)\right] = 0 \tag{8}$$

if $\exists \overline{R}_f$ then,

$$E\left[\widetilde{m}\right] = \frac{1}{\overline{R}_f} \tag{9}$$

Note. (7)-(8) hold for portfolios as well as individual assets.

Definition 5. (Asset pricing theory) a set of hypotheses that implies some particular form for $\widetilde{m}$ is an asset pricing theory.

Claim. Explaining risk premia of different assets and deriving a SDF is equivalent.

Proof. Let $E\left[\widetilde{m}\widetilde{R}\right]=1$. Using the definition of covariance between two random variables as equal to the difference between the expectation of their product and the product of their expectations then,

$$cov\left(\widetilde{m},\widetilde{R}\right)+E\left[\widetilde{m}\right]E\left[\widetilde{R}\right] = 1 \qquad \textit{given } (7)$$

and assuming that $\exists\overline{R}_f$, we substitute (9) in the previous equation,

$$cov\left(\widetilde{m},\widetilde{R}\right)+\frac{1}{\overline{R}_f}E\left[\widetilde{R}\right] = 1$$

$$\Rightarrow \overline{R}_f\cdot cov\left(\widetilde{m},\widetilde{R}\right)+E\left[\widetilde{R}\right] = \overline{R}_f, \qquad \textit{rearranging terms}$$

$$\Rightarrow E\left[\widetilde{R}\right]-\overline{R}_f = -\overline{R}_f\cdot cov\left(\widetilde{m},\widetilde{R}\right) \qquad \textit{rearranging terms}$$

□

Remark. Risk premia is determined by the covariances of returns with some SDF.

Note. Given the concavity of $u(\cdot)$, the F.O.N.C. implies that optimal wealth is inversely related to the SDF. That is investors consume less in states that are more expensive. The covariance between the discount factor and the return must have negative sign reflecting the positive tradeoff between risk and return.

Definition 6. (Gorman aggregation) Gorman (1961) showed using duality that utility functions with affine expenditure functions lead to indirect utility functions with linear and parallel "Engel" or iso-income curves. Hence, prices in equilibrium are independent of the initial distribution of wealth across investors. This is a necessary and sufficient condition to generalize society's behavior to that of a representative agent (RA) in economics.

1. Aggregation based on scale invariance (CRRA utility function).

Definition 7. A preference $\succ^*$ is said to be scale invariant (SI) if $w'\succ^* w \Rightarrow \theta w'\succ^*\theta w \quad \forall\theta\in(0,\infty)$.

Definition 8. A representative agent (RA) is the SI intertemporal trader with endowment equal to $\sum_i w_{si}=w \qquad \forall s$ and absolute risk aversion $\theta=\left(\sum_{i=1}^{I}\frac{1}{\theta_i}\right)^{-1}$.

2. Aggregation based on translation invariance (CARA utility function).

Definition 9. A preference $\succ^*$ is said to be translation invariant (TI) if $w'\succ^* w\Rightarrow w'+\alpha\cdot 1\succ^* w+\alpha\cdot 1 \quad \forall\alpha\in\mathbb{R}$.

Definition 10. A representative agent (RA) is the TI intertemporal trader with endowment equal to $\sum_i w_{si}=w \qquad \forall s$ and absolute risk aversion $\theta=\left(\sum_{i=1}^{I}\frac{1}{\theta_i}\right)^{-1}$.

Note. If preferences are also assumed to admit an additive-separable representation like e.g., is the case with the standard expected utility model, then the restrictions above imply specific parameterizations (e.g., power, logarithmic, exponential). However, Gorman aggregation is more general than additive-separability as it is possible if and only if investors have linear risk tolerance (equivalently optimal investments are affine in wealth - Mossin, 1968) with the same cautiousness parameter α, that is encompasses a more general class of utility functions with hyperbolic risk aversion (HARA) and the recursive utility model of Kreps-Porteus.

Example. (Preference-based equilibrium pricing)
We now re-state the investor's behavioral problem as one of a representative agent (RA) that seeks to solve the problem of choosing an optimal consumption plan for "date 0" (beginning of the period) and "date 1" (end of period) at the beginning of the period in addition to choose the optimal portfolio formed from investing the amount postponed in consumption. Let A_t be the RA's information set. Preferences $\succeq^*$ are represented by the expected utility function (hence they satisfy SI and TI and leads to an additive-separable representation),

$$E\left[\sum_{t=0}^{\infty}\beta^t U\left(c_t\right)|A_0\right],$$

where $\beta \in (0,1)$ is the RA's subjective discount factor. If at time t the RA's optimal asset holding is interior to the set of allowable portfolios, then in equilibrium asset prices can be determined by the F.O.N.C. (Euler equation),

$$p_t\left(\widetilde{x}_\tau\right) = E\left[\phi_\tau^{\tau-t}\widetilde{x}_T\,|A_t\right],$$

where $\phi_\tau^{\tau-t} = \beta\frac{U'(c_\tau)}{U'(c_t)}$ is the SDF and the ratio of marginal utilities is the intertemporal marginal rate of substitution of consumption (IMRS) between dates t and τ. Note that the approach requires a parametric assumption about the RA's utility function and the underlying economic structure to determine the joint distribution of $\phi_\tau^{\tau-t}$ and p_t. In continuous time additive separable utility leads to the Consumption-based Capital Asset Pricing Model (CCAPM) of Breeden (1979).

III. Empirical Asset Pricing

The set of plausible asset pricing models is defined by a set of pricing kernels indexed by some $N-$dimensional vector of parameters θ in the admissible parameter space $\Phi \in \mathbb{R}^N$. Note that the parameter vector must satisfy all constraints and functional relations determined by theory (e.g., variances have to be positive and risk aversion parameters have to be bounded by some minimum value). The basic premise of empirical asset pricing is that there exists a unique parameterization $\theta_0 \in \Phi$ that is consistent with the population distribution of the observed vector of asset prices or returns and defines the economy's SDF. The main goal of empirical asset pricing is to construct the best estimator of θ_0.

The general estimation strategy for θ_0 involves the choice of:

- A sample of size T of observed asset prices or returns $\vec{y}_T \equiv \left(y_T,\ldots,y_1\right)^\top$.
- An admissible parameter space $\Phi \in \mathbb{R}^N$.
- A N-vector of functions $D\left(y_t;\theta\right)$ with the property that θ_0 is the unique element of Φ satisfying $E\left[D\left(y_t;\theta_0\right) = 0\right]$.

Note that although the estimation strategy relies on the asset pricing theory of interest, theory does not preclude a unique D. There are multiple plausible admissible choices for D. Thus, we reinterpret D as the first order condition for maximizing the non-stochastic criterion function $Q_0(\theta) : \Phi \to \mathbb{R}$ with θ_0 as its solution,

$$\frac{\partial Q_0}{\partial \theta}(\theta_0) = E\left[D\left(y_t; \theta_0\right)\right] = 0. \tag{10}$$

Thus, the choice of the estimation strategy reduces to the problem of choosing a good criterion function Q_0. As long as the function is well behaved, there will be a global maximum or minimum (depending on the function) θ^* that is unique and equal to the population parameter vector of interest θ_0. A necessary step in verifying that $\theta^* = \theta_0$ is to verify that D satisfies $E\left[D\left(y_t; \theta_0\right) = 0\right]$.

To construct the estimator of θ_0 we work with the sample version of the criterion function $Q_T(\theta)$ which is a known function of $\vec{y}_T$. The sample estimator θ_T that maximizes or minimizes $Q_T(\theta)$ over Φ is the solution to the first order condition,

$$\frac{\partial Q_T}{\partial \theta}(\theta_T) = \frac{1}{T}\sum_{t=1}^{T} D\left(y_t; \theta_T\right) = 0. \tag{11}$$

Under some weak regularity conditions, $Q_0(\theta)$ should converge to its population counterpart $Q_T(\theta)$, so we expect $\theta_T \xrightarrow{p} \theta_0$ as $T \to \infty$ (consistency property). Moreover, we choose a D that gives the estimator with smallest asymptotic variance (efficiency property).

Intuition: the estimator is more efficient if it uses more information. A consistent and efficient estimator is our best estimator.

Equiped with the empirical model and an estimation strategy we proceed with the empirical study. At this stage, the financial econometrician might still have to:

(1) choose a computational method to find the global optimum in $Q_T(\theta)$.
(2) choose a set of statistics to test hypothesis of interest (given the derivation of their large-sample properties including its asymptotic distribution).
(3) Assess the small-sample distributions of the test statistics and the reliability of the inference procedures used in the empirical exercise (using e.g., Monte Carlo methods, bootstrapping, etc.).

References

[Breeden (1979) Journal of Financial Economics]
[Gorman (1961) Metroeconomica]
[Mossin (1968) Journal of Business]